arXiv:1505.04532v1 [cs.IT] 18 May 2015

Large System Analysis of Cognitive Radio Network via Partially-Projected Regularized Zero-Forcing Precoding

Jun Zhang, Chao-Kai Wen, Chau Yuen, Shi Jin, and Xiqi Gao

Abstract

In this paper, we consider a cognitive radio (CR) network in which a secondary multiantenna base station (BS) attempts to communicate with multiple secondary users (SUs) using the radio frequency spectrum that is originally allocated to multiple primary users (PUs). Here, we employ partially-projected regularized zero-forcing (PP-RZF) precoding to control the amount of interference at the PUs and to minimize inter-SUs interference. The PP-RZF precoding partially projects the channels of the SUs into the null space of the channels from the secondary BS to the PUs. The regularization parameter and the projection control parameter are used to balance the transmissions to the PUs and the SUs. However, the search for the optimal parameters, which can maximize the ergodic sum-rate of the CR network, is a demanding process because it involves Monte-Carlo averaging. Then, we derive a deterministic expression for the ergodic sum-rate achieved by the PP-RZF precoding using recent advancements in large dimensional random matrix theory. The deterministic equivalent enables us to efficiently determine the two critical parameters in the PP-RZF precoding because no Monte-Carlo averaging is required. Several insights are also obtained through the analysis.

Index Terms

Cognitive radio network, ergodic sum-rate, regularized zero-forcing, deterministic equivalent.

J. Zhang is with Jiangsu Key Laboratory of Wireless Communications, Nanjing University of Posts and Telecommunications, Nanjing 210003, P. R. China, and also with Singapore University of Technology and Design, Singapore 487372, Email address: zhang_jun@sutd.edu.sg.

C. K. Wen is with the Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan, E-mail address: ckwen@ieee.org.

C. Yuen is with Singapore University of Technology and Design, Singapore 487372, Email address: yuenchau@sutd.edu.sg.

S. Jin and X. Q. Gao are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, E-mail addresses: {jinshi, xqgao}@seu.edu.cn.

I. Introduction

The radio frequency spectrum is a valuable but congested natural resource because it is shared by an increasing number of users. Cognitive radio (CR) [1–4] is viewed as an effective means to improve the utilization of the radio frequency spectrum by introducing dynamic spectrum access technology. Such technology allows secondary users (SUs, also known as CR users) to access the radio spectrum originally allocated to primary users (PUs). In the CR literature, two cognitive spectrum access models have been widely adopted [4]: 1) the *opportunistic spectrum access* model and 2) the *concurrent spectrum access* model. In the opportunistic spectrum access model, SUs carry out spectrum sensing to detect spectrum holes and reconfigure their transmission to operate only in the identified holes [1, 5]. Meanwhile, in the concurrent spectrum access model, SUs transmit simultaneously with PUs as long as interference to PUs is limited [6, 7].

In this paper, we focus on the concurrent spectrum access model particularly when the secondary base station (BS) is equipped with multiple antennas. A desirable condition in the concurrent spectrum access model is for SUs to maximize their own performance while minimizing the interference caused to the PUs. Several transmit schemes have been studied to balance the transmissions to the SUs and the PUs [8–12]. In [8], a transmit algorithm has been proposed based on the singular value decomposition of the secondary channel after the projection into the null space of the channel from the secondary BS to the PUs. A spectrum sharing scheme has been designed for a large number of SUs [9], in which the SUs are pre-selected so that their channels are nearly orthogonal to the channels of the PUs. Doing so ensures that the SUs cause the lowest interference to the PUs.

In multi-antenna and multiuser downlink systems, a common technique to mitigate the multiuser interference is a zero-forcing (ZF) precoding [13–16], which is computationally more efficient than its non-linear alternatives. However, the achievable rates of the ZF precoding are severely compromised when the channel matrix is ill conditioned. Then, regularized ZF (RZF) precoding [17, 18] is proposed to mitigate the ill-conditioned problem by employing a regularization parameter in the channel inversion. The regularization parameter can control the amount of introduced interference. Several applications based on the RZF framework have been developed, such as transmitter designs for *non*-CR broadcast systems [19–22], security systems [23, 24], and multi-cell cooperative systems [25–28].

While directly applying RZF to CR networks, the secondary BS can *only* control the interference in inter-SUs. A *partially-projected* RZF (PP-RZF) precoding has been proposed [10, 11], which limits the interference from the SUs to the PUs by combining the RZF [17, 18] with the channel projection idea [8]. The PP-RZF precoding follows the classical RZF technique, although the former is based on the partially-projected channel, which is obtained by partially projecting the channel matrix into the null space of the channel from the secondary BS to the PUs. The amount of interference to the PUs decreases with increasing amounts of projection into the null space of the PUs, which can be achieved by tuning the projection control parameter. However, the search for the optimal regularization parameter and projection control parameter is a demanding process because it involves Monte-Carlo averaging. Therefore, a deterministic (or large-system) approximation of the signal-to-interference-plus-noise ratio (SINR) for the PP-RZF scheme has been derived [10, 11]. Unfortunately, only the CR channel with a *single* PU has been studied and the scenario where *multiple* PUs are present remains unsolved [10].

To apply the PP-RZF precoding scheme in a CR network with *multiple* PUs, a new analytical technique that deals with a *multi*-dimensional random projection matrix, which is generated by partially projecting the channel matrix into the null spaces of *multiple* PUs, is required. This paper aims to address the above mentioned challenge by providing analytical results in a more general setting than that in [10, 11]. Specifically, we focus on a downlink multiuser CR network (Fig. 1), which consists of a secondary BS with multiple antennas, SUs, and PUs as well as different channel gains. Our main contributions are summarized below.

- We derive deterministic equivalents for the SINR and the ergodic sum-rate achieved by the PP-RZF precoding under the general CR network. Unlike previous works [10, 11], our model considers *multiple* PUs and allows different channel gains from the secondary BS to each user. Owing to recent advancements in large dimensional random matrix theory (RMT) with respect to complex combinations of different types of independent random matrices [29], we identify the large system distribution of the Stieltjes transform for a new class of random matrix. Therefore, our extension becomes non trivial and novel.
- In the PP-RZF precoding, the regularization parameter and the projection control parameter can regulate the amount of interference to the SUs and the PUs, but a wrong choice of parameters can considerably degrade the performance of the CR network. However, the search for the optimal parameters is a demanding process because Monte-Carlo averaging

is required. We overcome the fundamental difficulty of applying PP-RZF precoding in the CR network. The deterministic equivalent for the ergodic sum-rate provides an efficient way of finding the asymptotically optimal regularization parameter and the asymptotically optimal projection control parameter. Simulation results indicate good agreement with the optimum in terms of the ergodic sum-rate.

- We provide several useful observations on the condition that the regularization parameter and the projection control parameter can achieve the optimal sum-rate. We also reveal the relationship between the parameters and the signal-to-noise ratio (SNR).

Notations—We use uppercase and lowercase boldface letters to denote matrices and vectors, respectively. An $N \times N$ identity matrix is denoted by $\mathbf{I}_N$, an all-zero matrix by $\mathbf{0}$, and an all-one matrix by $\mathbf{1}$. The superscripts $(\cdot)^H$, $(\cdot)^T$, and $(\cdot)^*$ denote the conjugate transpose, transpose, and conjugate operations, respectively. $\mathsf{E}\{\cdot\}$ returns the expectation with respect to all random variables within the bracket, and $\log(\cdot)$ is the natural logarithm. We use $[\mathbf{A}]_{kl}$, $[\mathbf{A}]_{l,k}$, or A_{kl} to denote the (k,l)-th entry of the matrix $\mathbf{A}$, and a_k denotes the k-th entry of the column vector $\mathbf{a}$. The operators $(\cdot)^{\frac{1}{2}}$, $(\cdot)^{-1}$, $\mathsf{tr}(\cdot)$, and $\det(\cdot)$ represent the matrix principal square root, inverse, trace, and determinant, respectively, $\|\cdot\|$ represents the Euclidean norm of an input vector or the spectral norm of an input matrix, and $\mathsf{diag}(\mathbf{x})$ denotes a diagonal matrix with $\mathbf{x}$ along its main diagonal. The notation "$\xrightarrow{a.s.}$" denotes the almost sure (a.s.) convergence.

II. System Model and Problem Formulation

A. System Model

As illustrated in Fig. 1, we consider a downlink multiuser CR network that consists of a secondary BS with N antennas (labeled as BS). The BS simultaneously transmits K independent messages to K single antenna SUs (labeled as $\mathsf{SU}_1, \ldots, \mathsf{SU}_K$). We assume that all the SUs share the same spectrum with L single antenna PUs (labeled as $\mathsf{PU}_1, \ldots, \mathsf{PU}_L$). Let $\mathbf{h}_k^H \in \mathbb{C}^{1\times N}$ be the fading channel vector between BS and SU_k, $\mathbf{f}_l^H \in \mathbb{C}^{1\times N}$ be the fading channel vector between BS and PU_l, and $\mathbf{g}_k \in \mathbb{C}^{N\times 1}$ be the precoding vector of SU_k. The received signal at SU_k can therefore be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{g}_k s_k + \sum_{j=1, j\neq k}^{K} \mathbf{h}_k^H \mathbf{g}_j s_j + z_k, \tag{1}$$

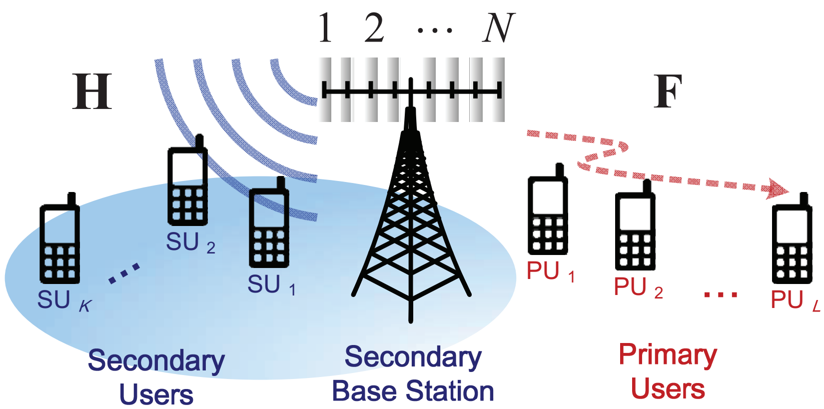

Fig. 1. A downlink multiuser cognitive radio network.

where s_k is the data symbol of SU_k, s_j's are independent and identically distributed (i.i.d.) data symbols with zero mean and unit variance, respectively, and z_k is the additive Gaussian noise with zero mean and variance of σ^2. For ease of exposition, we define $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times N}$, $\mathbf{F} \triangleq [\mathbf{f}_1, \ldots, \mathbf{f}_L]^H \in \mathbb{C}^{L \times N}$, $\mathbf{G} \triangleq [\mathbf{g}_1, \ldots, \mathbf{g}_K] \in \mathbb{C}^{N \times K}$, $\mathbf{y} \triangleq [y_1, \ldots, y_K]^T \in \mathbb{C}^K$, $\mathbf{s} \triangleq [s_1, \ldots, s_K]^T \in \mathbb{C}^K$, and $\mathbf{z} \triangleq [z_1, \ldots, z_K]^T \in \mathbb{C}^K$. The received signal of all the SUs in vector form is given by

$$\mathbf{y} = \mathbf{H}\mathbf{G}\mathbf{s} + \mathbf{z}. \tag{2}$$

We also assume that BS satisfies the average total transmit power constraint

$$\mathsf{E}\left\{\mathsf{tr}\left(\mathbf{G}\mathbf{G}^H\right)\right\} \leq N P_T, \tag{3}$$

where $P_T > 0$ is the parameter that determines the power budget of BS. Notably, if we consider the instantaneous transmit power constraint, i.e., $\mathsf{tr}(\mathbf{G}\mathbf{G}^H) \leq N P_T$, we can obtain the same constraint in a large-system regime, as shown in Appendix B-III.

The peak received interference power constraint or the average received interference power constraint is used to protect the PUs. Given that the latter is more flexible for dynamically allocating transmission powers over different fading states than the former [30, 31], we employ the average received interference power constraint and consider two cases: Case I—the average received interference power constraint at each PU and Case II—the total average received interference power constraint at all PUs[1]. These cases are respectively given by

$$\text{Case I (Per PU power constraint):} \quad \mathsf{E}\left\{\mathbf{f}_l^H \mathbf{G}\mathbf{G}^H \mathbf{f}_l\right\} \le P_l, \quad \text{for} \quad l = 1, \dots, L, \tag{4a}$$

$$\text{Case II (Sum power constraint):} \quad \mathsf{E}\left\{\mathsf{tr}\left(\mathbf{F}\mathbf{G}\mathbf{G}^H\mathbf{F}^H\right)\right\} \le P_{\text{all}}, \tag{4b}$$

where $P_l > 0$ denotes the interference power threshold of PU_l, and $P_{\text{all}} > 0$ represents the total interference power threshold of all PUs. We then set $P_l = \theta_l P_T$ and $P_{\text{all}} = \theta_{\text{all}} P_T$ with $\theta_l, \theta_{\text{all}}$ being positive scalar parameters to make a connection with the transmit power. Although we only consider equal power allocation for simplicity in this paper, our framework can be easily extended to arbitrary power allocation by replacing $\mathbf{G}$ with $\mathbf{G}\mathbf{P}^{\frac{1}{2}}$, where $\mathbf{P} = \mathsf{diag}(p_1, \dots, p_K)$ with $p_k \ge 0$ being the signal power of SU_k (see [21, 22] for a similar application).

Next, to incorporate path loss and other large-scale fading effects, we model the channel vectors by

$$\mathbf{h}_k^H = \sqrt{r_{1,k}}\,\tilde{\mathbf{h}}_k^H \quad \text{and} \quad \mathbf{f}_l^H = \sqrt{r_{2,l}}\,\tilde{\mathbf{f}}_l^H, \tag{5}$$

where $\tilde{\mathbf{h}}_k^H$ and $\tilde{\mathbf{f}}_l^H$ are the small-scale (or fast) fading vectors, and $r_{1,k}$ and $r_{2,l}$ denote the large-scale fading coefficients (or channel path gains), including the geometric attenuation and shadow effect. Using the above notations, the concerned channel matrices can be rewritten as

$$\mathbf{H} = \mathbf{R}_1^{\frac{1}{2}}\tilde{\mathbf{H}} \quad \text{and} \quad \mathbf{F} = \mathbf{R}_2^{\frac{1}{2}}\tilde{\mathbf{F}}, \tag{6}$$

where $\tilde{\mathbf{H}} \equiv [\frac{1}{\sqrt{N}}\tilde{h}_{ij}] \in \mathbb{C}^{K\times N}$ and $\tilde{\mathbf{F}} \equiv [\frac{1}{\sqrt{N}}\tilde{f}_{ij}] \in \mathbb{C}^{L\times N}$ consist of the random components of the channel in which $\tilde{h}_{ij}$'s and $\tilde{f}_{ij}$'s are i.i.d. complex random variables with zero mean and unit variance, respectively, and $\mathbf{R}_1 \in \mathbb{C}^{K\times K}$ and $\mathbf{R}_2 \in \mathbb{C}^{L\times L}$ are diagonal matrices whose diagonal elements are given by $[\mathbf{R}_1]_{kk} = r_{1,k}$ and $[\mathbf{R}_2]_{ll} = r_{2,l}$, respectively. In line with [10, 11], we assume that $\mathbf{H}$ is perfectly known to BS in this paper. Since BS needs to predict the interference power in (4), we further assume that perfect knowledge of $\mathbf{F}$ is available at BS [10, 11, 32]. To acquire perfect channel state information (CSI) for $\mathbf{H}$ and $\mathbf{F}$, transmission protocols need to incorporate certain cooperation among the PUs, the SUs, and BS [32]. Further research can focus on the case with imperfect CSI or estimation of channel [33, 34].

[1]Notably, multiple single-antenna PUs exist. These PUs can also be considered a single equivalent PU with multiple receive antennas.

In the downlink CR network (2), we consider the RZF precoding because this precoding's relatively low complexity compared with dirty paper coding [17, 18, 21, 27]. However, a direct application of the conventional RZF to the secondary BS will result in a very inefficient transmission because a large power back-off at the secondary BS is required to satisfy the interference power constraint (4). Therefore, following [10, 11], we adopt the RZF precoding based on the *partially-projected* channel matrix

$$\check{\mathbf{H}} = \mathbf{H}(\mathbf{I}_N - \beta\mathbf{W}^H\mathbf{W}), \tag{7}$$

where $\mathbf{W} \triangleq (\mathbf{F}\mathbf{F}^H)^{-\frac{1}{2}}\mathbf{F} \in \mathbb{C}^{L\times N}$, and $\beta \in [0, 1]$ is the projection control parameter. Note that the projected channel matrix $\check{\mathbf{H}}$ is obtained by *partially* projecting $\mathbf{H}$ into the null space of $\mathbf{F}$. Specifically, the RZF precoding matrix is given by

$$\mathbf{G} = \xi\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H, \tag{8}$$

where ξ is a normalization parameter that fulfills the BS transmit power constraint (3) and the interference power constraint (4), and $\alpha > 0$ represents the regularization parameter. We refer to this precoding as PP-RZF precoding.

Before setting each of the parameters in (8), two special cases of the PP-RZF precoding are considered first. On the one hand, if $\beta = 0$ then $\mathbf{G}$ degrades to the conventional RZF precoding. On the other hand, if $\beta = 1$ then $\check{\mathbf{H}}$ is completely orthogonal to $\mathbf{F}$ and we have $\mathbf{F}\check{\mathbf{H}}^H = \mathbf{0}$, i.e., no interference signal from the secondary BS will leak to the PUs. Therefore, the interference power constraint (4) is naturally guaranteed. Furthermore, the amount of the interference to the PUs decreases as the projection control parameter increases.

Now we return to the setting of the normalization parameter in (8). Considering Case I, from (3) and (4a), we have

$$\xi^2 \leq \xi_0^2 \triangleq \frac{P_T}{\mathsf{E}\left\{\frac{1}{N}\mathsf{tr}\left(\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}\right)\right\}}, \tag{9a}$$

$$\xi^2 \leq \xi_l^2 \triangleq \frac{\theta_l P_T}{\mathsf{E}\left\{\mathbf{f}_l^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}\mathbf{f}_l\right\}}, \quad \text{for } l = 1, \ldots, L. \tag{9b}$$

To satisfy (3) and (4a) simultaneously, we set $\xi^2 = \min\{\xi_0^2, \xi_l^2, l = 1, \ldots, L\}$. Then, the SINR of secondary user SU_k is given by

$$\gamma_k = \frac{\left|\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{h}}_k\right|^2}{\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{h}_k+\frac{\sigma^2}{\xi^2}}$$
$$= \frac{\rho\left|\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{h}}_k\right|^2}{\rho\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{h}_k+\nu}, \tag{10}$$

where $\check{\mathbf{H}}_{[k]} \triangleq [\check{\mathbf{h}}_1,\ldots,\check{\mathbf{h}}_{k-1},\check{\mathbf{h}}_{k+1},\ldots,\check{\mathbf{h}}_K]^H \in \mathbb{C}^{(K-1)\times N}$, $\check{\mathbf{h}}_k \triangleq (\mathbf{I}_N-\beta\mathbf{W}^H\mathbf{W})\mathbf{h}_k$, $\rho \triangleq P_T/\sigma^2$, and

$$\nu \triangleq \frac{P_T}{\xi^2} = \max\left\{\mathsf{E}\left\{\frac{1}{N}\mathsf{tr}\left(\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\right)\right\},\right.$$
$$\left.\frac{1}{\theta_l}\mathsf{E}\left\{\mathbf{f}_l^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{f}_l\right\}, l=1,\ldots,L\right\}. \tag{11}$$

Here, the equality of (11) follows from (9). For Case II, we have

$$\nu = \max\left\{\mathsf{E}\left\{\frac{1}{N}\mathsf{tr}\left(\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\right)\right\},\right.$$
$$\left.\frac{1}{\theta_{\text{all}}}\mathsf{E}\left\{\mathsf{tr}\left(\mathbf{F}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{F}^H\right)\right\}\right\}. \tag{12}$$

Consequently, under the assumption of perfect CSI at both transmitter and receivers, the ergodic sum-rate of the CR network with Gaussian signaling can be defined as

$$R_{\text{sum}} \triangleq \sum_{k=1}^{K}\mathsf{E}\left\{\log\left(1+\gamma_k\right)\right\}. \tag{13}$$

Note that γ_k in the ergodic sum-rate is subject to the BS transmit power constraint in (3) and the interference power constraint (to the primary users) in (4).

B. Problem Formulation

The SINR γ_k in (10) is a function of the regularization parameter α and the projection control parameter β. In the literature, adopting incorrect regularization parameter would degrade performance significantly [18, 21, 27]. In light of the discussion in the previous subsection, one can realize that a proper projection control parameter can assist in decreasing the interference to the PUs. As a result, using the PP-RZF precoding effectively requires obtaining appropriate values of α and β to optimize certain performance metrics. In this paper, we are interested in finding (α, β), which maximizes the ergodic sum-rate (13). Formally, we have

$$\left\{\alpha^{\rm opt}, \beta^{\rm opt}\right\} = \underset{\alpha>0, 1\geq\beta\geq 0}{\arg\max}\ R_{\rm sum}. \tag{14}$$

The above problem does not admit a simple closed-form solution and the solution must be computed via a two-dimensional line search. Monte-Carlo averaging over the channels is required to evaluate the ergodic sum-rate (13) for each choice of α and β, which, unfortunately, makes the overall computational complexity prohibitive. This drawback hinders the development of the PP-RZF precoding. To address this problem, we resort to an asymptotic expression of (13) in the large-system regime in the next section.

III. Performance Analysis of Large Systems

This section presents the main results of the paper. First, we derive deterministic equivalents for the SINR γ_k and the ergodic sum-rate $R_{\rm sum}$ in a large-system regime. Then, we identify the asymptotically optimal regularization parameter and the asymptotically optimal projection control parameter to achieve the optimal deterministic equivalent for the ergodic sum-rate.

A. Deterministic Equivalents for the SINR and the Ergodic Sum-Rate

We present a deterministic equivalent for the SINR γ_k by considering the large-system regime, where N, K, and L approach infinity, whereas

$$c_1 = \frac{N}{K} \ \text{ and } \ c_2 = \frac{L}{N}$$

are fixed ratios, such that $0 < \liminf_N c_1 \leq \limsup_N c_1 < \infty, 0 < \liminf_N c_2 \leq \limsup_N c_2 \leq 1$. For brevity, we simply use $\mathcal{N} \to \infty$ to represent the quantity in such limit. In addition, we impose the assumptions below in our derivations.

Assumption 1: For the channel matrices $\mathbf{H}$ and $\mathbf{G}$ in (6), we have the following hypotheses:

1) $\tilde{\mathbf{H}} = [\frac{1}{\sqrt{N}}\tilde{h}_{ij}] \in \mathbb{C}^{K\times N}$, where $\tilde{h}_{ij}$'s are i.i.d. standard Gaussian.
2) $\tilde{\mathbf{F}} = [\frac{1}{\sqrt{N}}\tilde{f}_{ij}] \in \mathbb{C}^{L\times N}$, where $\tilde{f}_{ij}$'s have the same statistical properties as $\tilde{h}_{ij}$'s.

3) $\mathbf{R}_1 = \mathsf{diag}(r_{1,1}, \ldots, r_{1,K}) \in \mathbb{C}^{K\times K}$ and $\mathbf{R}_2 = \mathsf{diag}(r_{2,1}, \ldots, r_{2,L}) \in \mathbb{C}^{L\times L}$ are diagonal matrices with uniformly bounded spectral norm[2] with respect to K and L, respectively.

Based on the definition of $\mathbf{W}$ in (7), $\mathbf{W}^H\mathbf{W} = \mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1}\mathbf{F} = \tilde{\mathbf{F}}^H(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H)^{-1}\tilde{\mathbf{F}} = \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}$, where $\tilde{\mathbf{W}} \triangleq (\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H)^{-\frac{1}{2}}\tilde{\mathbf{F}}$. Therefore, $\tilde{\mathbf{W}}$ is $L \leq N$ rows of an $N \times N$ Haar-distributed unitary random matrix [29, Definition 4.6]. The partially-projected channel matrix $\check{\mathbf{H}}$ is clearly composed of the product of two different types of independent random matrices. Owing to recent advancements in large dimensional RMT [29], we arrive at the following theorem, and the details are given in Appendix A.

Theorem 1: Under Assumption 1, in Case I (per PU power constraint), as $\mathcal{N} \to \infty$, we have $\gamma_k - \overline{\gamma}_k \xrightarrow{a.s.} 0$, for $k = 1, \ldots, K$, where

$$\overline{\gamma}_k = \frac{\rho \overline{a}_k^2}{\rho \overline{b}_k + \overline{\nu}}, \tag{15}$$

with

$$\overline{a}_k = \frac{r_{1,k}(t_1 + t_2(1-\beta))}{\alpha + r_{1,k}(t_1 + t_2(1-\beta)^2)}, \tag{16a}$$

$$\overline{b}_k = r_{1,k}\left(\frac{(1-\overline{a}_k)^2\, t_1}{1+e} + \frac{(1-(1-\beta)\overline{a}_k)^2\,(1-\beta)^2 t_2}{1+(1-\beta)^2 e}\right)\frac{\partial e}{\partial \alpha}, \tag{16b}$$

$$\overline{\nu} = \max\left\{\left(\frac{t_1}{1+e} + \frac{(1-\beta)^2 t_2}{1+(1-\beta)^2 e}\right)\frac{\partial e}{\partial \alpha}, \frac{r_{2,l}}{\theta_l c_2}\frac{(1-\beta)^2 t_2}{1+(1-\beta)^2 e}\frac{\partial e}{\partial \alpha}, l = 1, \ldots, L\right\}, \tag{16c}$$

$$\frac{\partial e}{\partial \alpha} = \frac{\frac{1}{N}\mathsf{tr}\mathbf{R}_1\left(\alpha \mathbf{I}_K + (t_1 + t_2(1-\beta)^2)\,\mathbf{R}_1\right)^{-2}}{1 - \left(\frac{t_1}{1+e} + \frac{(1-\beta)^4 t_2}{1+(1-\beta)^2 e}\right)\frac{1}{N}\mathsf{tr}\left(\mathbf{R}_1\left(\alpha \mathbf{I}_K + (t_1 + t_2(1-\beta)^2)\,\mathbf{R}_1\right)^{-1}\right)^2}, \tag{16d}$$

$$t_1 = \frac{1-c_2}{1+e}, \qquad t_2 = \frac{c_2}{1+(1-\beta)^2 e}, \tag{16e}$$

and e is given as the unique solution to the fixed point equation

$$e = \frac{1}{N}\mathsf{tr}\mathbf{R}_1\left(\alpha \mathbf{I}_K + \left(t_1 + t_2(1-\beta)^2\right)\mathbf{R}_1\right)^{-1}. \tag{17}$$

Meanwhile, in Case II (sum power constraint), all asymptotic expressions remain, except for $\overline{\nu}$,

[2][35]: The spectral norm $\|\bullet\|_2$ is defined on $\mathbb{C}^{n\times n}$ by $\|\mathbf{A}\|_2 \equiv \max\{\sqrt{\lambda} : \lambda$ is an eigenvalue of $\mathbf{A}^*\mathbf{A}\}$.

which should be changed to

$$\overline{\nu} = \max\left\{ \left(\frac{t_1}{1+e} + \frac{(1-\beta)^2 t_2}{1+(1-\beta)^2 e}\right)\frac{\partial e}{\partial \alpha}, \frac{\mathrm{tr}\mathbf{R}_2}{\theta_{\mathrm{all}} c_2}\frac{(1-\beta)^2 t_2}{1+(1-\beta)^2 e}\frac{\partial e}{\partial \alpha}\right\}. \tag{18}$$

■

An intuitive application of Theorem 1 is that γ_k can be approximated by its deterministic equivalent[3] $\overline{\gamma}_k$, which can be determined based only on statistical channel knowledge, that is, $\mathbf{R}_1$, $\mathbf{R}_2$, and σ^2. Note that, according to the definition of the deterministic equivalent (see footnote 3), in the expression of the deterministic equivalent $\overline{\gamma}_k$, the parameters N, K, L, as well as the matrix dimensions of $\mathbf{R}_1$ and $\mathbf{R}_2$, are *finite*. Combining Theorem 1 with the continuous mapping theorem[4], we have $\log(1+\gamma_k) - \log(1+\overline{\gamma}_k) \xrightarrow{a.s.} 0$. An approximation $\overline{R}_{\mathrm{sum}}$ of the ergodic sum-rate R_{sum} in (13) is obtained by replacing the instantaneous SINR γ_k with its large system approximation $\overline{\gamma}_k$, that is,

$$\overline{R}_{\mathrm{sum}} = \sum_{k=1}^{K} \log(1+\overline{\gamma}_k). \tag{19}$$

Therefore, when $\mathcal{N} \to \infty$, $\frac{1}{K}\left(R_{\mathrm{sum}} - \overline{R}_{\mathrm{sum}}\right) \xrightarrow{a.s.} 0$ holds true almost surely.

To facilitate our understanding of Theorem 1, we look at it from the two special cases as follows:

1) In Theorem 1, we introduce the two variables t_1 and t_2 to reflect the effects of the projection control parameter β. If $\beta = 1$, from (16), then the deterministic equivalent $\overline{\gamma}_k$ does not depend on t_2. Substituting $\beta = 1$ into (15) and letting $\mathbf{R}_1 = \mathbf{I}_K$, we have

$$\overline{\gamma}_k = \frac{\rho\Big(c_1(1-c_2)(1+\zeta(\mu,\eta,\alpha))^2 - \zeta(\mu,\eta,\alpha)^2\Big)}{\rho + \Big(1+\zeta(\mu,\eta,\alpha)\Big)^2}, \tag{20}$$

[3][29, Definition 6.1] (also see [36]): Consider a series of Hermitian random matrices $\mathbf{B}_1$, $\mathbf{B}_2$, ..., with $\mathbf{B}_N \in \mathbb{C}^{N\times N}$ and a series $f_1, f_2, \ldots$ of functionals of $1\times 1, 2\times 2, \ldots$ matrices. A deterministic equivalent of $\mathbf{B}_N$ for functional f_N is a series $\mathbf{B}_1^\circ$, $\mathbf{B}_2^\circ$, ..., where $\mathbf{B}_N^\circ \in \mathbb{C}^{N\times N}$, of deterministic matrices, such that $\lim_{N\to\infty} f_N(\mathbf{B}_N) - f_N(\mathbf{B}_N^\circ) \to 0$. In this case, the convergence often be with probability one. Similarly, we term $g_N \triangleq f_N(\mathbf{B}_N^\circ)$ the deterministic equivalent of $f_N(\mathbf{B}_N)$, that is, the deterministic series $g_1, g_2, \ldots$, such that $f_N(\mathbf{B}_N) - g_N \to 0$ in some sense.
Note that the deterministic equivalent of the Hermitian random matrix $\mathbf{B}_N$ is a *deterministic* and a *finite dimensional* matrix $\mathbf{B}_N^\circ$. In addition, the deterministic equivalent of $f_N(\mathbf{B}_N)$ is $g_N \triangleq f_N(\mathbf{B}_N^\circ)$, which is a function of $\mathbf{B}_N^\circ$.

[4][37, Theorem 25.7-Corollary 2]: If $x_n \xrightarrow{a.s.} a$ and h is continuous function at a, then $h(x_n) \xrightarrow{a.s.} h(a)$.

where $\zeta(\mu,\eta,\alpha)\triangleq t_1/\alpha$, $\mu\triangleq 1-c_2$, and $\eta\triangleq 1/c_1$. Combining (16e) and (17), we obtain

$$\zeta(\mu,\eta,\alpha)\triangleq\frac{t_1}{\alpha}=\frac{1}{2}\left(\frac{\mu-\eta}{\alpha}-1+\sqrt{\frac{(\mu-\eta)^2}{\alpha^2}+\frac{2(\mu+\eta)}{\alpha}+1}\right). \tag{21}$$

Before providing an observation based on the above, we briefly review a well-known result from the large dimensional RMT. First, we consider the definition of $\mathbf{H}$ from (6). If $\mathbf{R}_1=\mathbf{I}_K$, the entries of the $K\times N$ matrix $\mathbf{H}$ are zero mean i.i.d. with variance $1/N$. Following [29, Chapter 3], we see that as $N,\,K\to\infty$ with $N/K\to c_1$, $\mathbf{h}_k^H\left(\mathbf{H}^H\mathbf{H}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{h}_k$ converges almost surely to

$$\int_a^b\frac{1}{\lambda+\alpha}f(\lambda)\,d\lambda, \tag{22}$$

where

$$f(\lambda)=(1-\eta)^+\,\delta(\lambda)+\frac{\sqrt{(\lambda-a)^+(b-\lambda)^+}}{2\pi\lambda} \tag{23}$$

with $(x)^+\triangleq\max\{x,\,0\}$, $a\triangleq(1-\sqrt{\eta})^2$, and $b\triangleq(1+\sqrt{\eta})^2$. In fact, $f(u)$ is the limiting empirical distribution of the eigenvalues of $\mathbf{H}^H\mathbf{H}$ and is known as the Marčcenko-Pastur law [38]. The integral of (22) can be evaluated in closed form

$$\frac{1}{2}\left(\frac{1-\eta}{\alpha}-1+\sqrt{\frac{(1-\eta)^2}{\alpha^2}+\frac{2(1+\eta)}{\alpha}+1}\right). \tag{24}$$

Note that (21) is equal to (24) when μ is replaced with 1, i.e., (24) is equal to $\zeta(1,\eta,\alpha)$. In fact, following the similar derivations of Theorem 1, we can show that (20) and the SINR of the conventional RZF precoding share the same formulation by replacing $\zeta(\mu,\eta,\alpha)$ in (20) with $\zeta(1,\eta,\alpha)$. Substituting the definitions of c_1,c_2 into μ and η, we have $\mu-\eta=1-c_2-1/c_1=(N-(L+K))/N$. Comparing this value with $1-\eta=(N-K)/N$ in (24), we thus conclude that if $\beta=1$, the SINR of the PP-RZF precoding is *similar*[5] to that of the conventional RZF precoding but with an increase in the number of active users from K to $K+L$. Hence, the degrees of freedom of the PP-RZF precoding is reduced to $N-(K+L)$ because the additional L degrees of freedom are used to suppress interference to the PUs.

[5]Notably, when $\beta=1$, the SINRs of the PP-RZF precoding and the conventional RZF precoding are similar but *not* identical because $\zeta(\mu,\eta,\alpha)$ is replaced with $\zeta(1,\eta,\alpha)$.

2) For another extreme case with $\beta = 0$ in Theorem 1, $t_1 + t_2 = \frac{1}{1+e}$. Letting $\mathbf{R}_1 = \mathbf{I}_K$, we obtain $\frac{1}{\alpha(1+e)} = \zeta(1, \eta, \alpha)$, such that

$$\overline{\gamma}_k = \frac{\rho\Big(c_1(1+\zeta(1,\eta,\alpha))^2 - \zeta(1,\eta,\alpha)^2\Big)}{\rho + \nu_0\Big(1+\zeta(1,\eta,\alpha)\Big)^2}, \tag{25}$$

where $\nu_0 = \max\{1, r_{2,l}/\theta_l, l = 1, \ldots, L\}$ is for Case I and $\nu_0 = \max\{1, \mathsf{tr}\mathbf{R}_2/\theta_{\text{all}}\}$ is for Case II. The received interference power constraint at the PUs (4) can be controlled only through ν_0, where β is not involved in ν_0. Therefore, the SINR $\overline{\gamma}_k$ is significantly degraded if the channel path gains between the BS and the PUs (that is, $r_{2,l}$'s) are strong. However, if the channel path gains between the BS and the PUs are weak, then $\nu_0 = 1$ and $\overline{\gamma}_k$ behave in a manner similar to but *not* identical to that of the conventional RZF precoding because c_1 is replaced with $c_1(1 - c_2)$.

Comparing (25) for $\beta = 0$ with (20) for $\beta = 1$ obtains notable results. First, we note that (20) and (25) share a similar formulation, except the additional ν_0 appears at the denominator of (25). When $\beta = 1$, the secondary BS yields zero interference on the PUs, such that the interference power constraint in (4) is always inactive. Therefore, no additional parameter ν_0 is required to reflect the received interference power constraint at the PUs. Although $\nu_0 \geq 1$, the SINR performance of the PP-RZF precoding with $\beta = 1$ is not implied to be always better than that with $\beta = 0$. An additional note should be given on $\zeta(\cdot, \eta, \alpha)$, where the argument $\cdot$ is μ for $\beta = 1$ and 1 for $\beta = 0$. The parameter $\mu = (N - L)/N$ for $\beta = 1$ implies that the additional L degrees of freedom is used to suppress interference to the PUs. Consequently, if the channel path gains between the BS and the PUs are weak, the SINR performance of the PP-RZF precoding with $\beta = 1$ shall not be better than that with $\beta = 0$. Thus, we infer that the projection control parameter should be decreased if the received interference power constraint at the PUs is relaxed.

Finally, we note that $\zeta(1, \eta, \alpha)$ agrees with $z(r, \alpha_0)$ in [10, 11, Theorem 1]. As a result, (25) is identical to the deterministic equivalent for the SINR obtained in [10, 11, Theorem 1], where the PP-RZF precoding with a *single* PU is considered. The deterministic equivalent for the SINR in [10, 11, Theorem 1] is clearly a special case of (15) with $\beta = 0$ even though the case of $\beta \neq 0$ is considered in [10, 11, Theorem 1] because a single PU results only in one-dimensional perturbation, and the effect of such perturbation *vanishes* in a large system. Even if the number of PUs L is finite and only N becomes large, the effect of β vanishes. The lack of a relation between β and the SINRs will result in a bias when the number of antennas at the BS is not so large. However, our analytical results show the effect of β by assuming that N, K, and L are large, whereas $c_1 = N/K$ and $c_2 = L/N$ are fixed ratios. Thus, our results are clearly more general than those in [10, 11].

Corollary 1: In addition to the assumptions of Theorem 1, we suppose further that $c_2 = 1$ (that is, $N = L$), $\mathbf{R}_1 = r_1 \mathbf{I}_K$, and $\beta \in [0, 1)$. Then, as $\mathcal{N} \to \infty$, we have $\gamma_k - \overline{\gamma} \xrightarrow{a.s.} 0$ for $k = 1, \dots, K$, where

$$\overline{\gamma} = \frac{\rho \left(c_1 r_1^2 - (c_1 \alpha e - r_1)^2\right)}{\rho (c_1 \alpha e)^2 + \nu_0}, \tag{26}$$

and e is given as an unique solution to the fixed point equation

$$e = \frac{r_1 (1 + e(1-\beta)^2)}{c_1 \alpha (1 + e(1-\beta)^2) + c_1 r_1 (1-\beta)^2},$$

and $\nu_0 = r_1 \max\{1, r_{2,l}/\theta_l, l = 1, \dots, L\}$ for Case I or $\nu_0 = r_1 \max\{1, \mathsf{tr}\mathbf{R}_2/\theta_{\rm all}\}$ for Case II.

Proof: By letting $c_2 = 1$ and $\mathbf{R}_1 = r_1 \mathbf{I}_K$, we immediately obtain the result from Theorems 1 and 2. ■

For a brief illustration, we consider only Case II of Corollary 1 because the same characteristics can be found in Case I. Given that $\theta_{\rm all} = P_{\rm all}/P_T = P_{\rm all}/(\sigma^2 \rho)$, (26) can be rewritten as

$$\overline{\gamma} = \begin{cases} \dfrac{c_1 r_1^2 - (c_1 \alpha e - r_1)^2}{(c_1 \alpha e)^2 + 1/\rho}, & 0 < \dfrac{\rho \sigma^2 \mathsf{tr}\mathbf{R}_2}{P_{\rm all}} \leq 1; \\ \dfrac{c_1 r_1^2 - (c_1 \alpha e - r_1)^2}{(c_1 \alpha e)^2 + \sigma^2 \mathsf{tr}\mathbf{R}_2 / P_{\rm all}}, & 1 < \dfrac{\rho \sigma^2 \mathsf{tr}\mathbf{R}_2}{P_{\rm all}}. \end{cases} \tag{27}$$

We can see that $\overline{\gamma}$ does not depend on the SNR ρ when $1 < \rho \sigma^2 \mathsf{tr}\mathbf{R}_2 / P_{\rm all}$. In this case, the system performance is interference-limited. Notably, the assumptions of $c_2 = 1$ and $\beta \neq 1$ are taken in Corollary 1. In the case of $c_2 = 1$ and $\beta = 1$, from (16a), we have $\overline{a}_k = 0$ and consequently $\overline{\gamma}_k = 0$, which implies a failure in the transmission. This result is reasonable because when $c_2 = 1$, the dimension of the null space of $\mathbf{F}$ is zero with probability one.[6] Therefore, the setting

[6]If $N = L$, we have $\mathsf{Rank}(\mathbf{I} - \mathbf{F}^H (\mathbf{F}\mathbf{F}^H)^{-1} \mathbf{F}) = 0$ with probability one because from [39, Theorem 1.1], $\mathbf{F}$ is a full rank square matrix with probability one.

of $\beta = 1$ results in transmission failure, *even* when the channel path gains between the BS and the PUs are weak. We thus show that a choice of appropriate β significantly affects the successful operation of the CR network, which serves as motivation for the remainder of this paper.

B. Asymptotically Optimal Parameters

Our numerical results confirm the high accuracy of the deterministic equivalent for the ergodic sum-rate $\overline{R}_{\rm sum}$ in the next section. Therefore, the deterministic equivalent for the ergodic sum-rate can be used to determine the regularization parameter α and the projection control parameter β. By replacing $R_{\rm sum}$ with $\overline{R}_{\rm sum}$ in (14), we focus on this particular optimization to maximize the deterministic equivalent for the ergodic sum-rate

$$\left\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\right\} = \mathop{\arg\max}_{\alpha>0, 1\geq\beta\geq 0} \overline{R}_{\rm sum}. \tag{28}$$

Similar to the problem in (14), the asymptotically optimal solutions $\overline{\alpha}^{\rm opt}$ and $\overline{\beta}^{\rm opt}$ do not permit closed-form solutions. However, the asymptotically optimal solution can be computed efficiently via the following methods without the need for Monte-Carlo averaging because $\overline{\gamma}_k$ is deterministic. First, given that β is fixed, the optimal $\overline{\alpha}^{\rm opt}(\beta) := \arg\max_{\alpha>0} \overline{R}_{\rm sum}(\beta)$ can be obtained efficiently via one-dimensional line search [21, 27], which performs the simple gradient method. The complexity in this part is linear. Then, we obtain the optimal $\overline{\beta}^{\rm opt} := \arg\max_{0\leq\beta\leq 1} \overline{R}_{\rm sum}(\overline{\alpha}^{\rm opt}(\beta), \beta)$ through the one-dimensional exhaustive search[7]. Finally, the optimal parameters are given by $\{\overline{\alpha}^{\rm opt}(\overline{\beta}^{\rm opt}), \overline{\beta}^{\rm opt}\}$. For a special case, we obtain a condition of the optimal solutions in the following proposition:

Proposition 1: Under the assumptions of Corollary 1, the asymptotically optimal parameters $\overline{\alpha}^{\rm opt}$ and $\overline{\beta}^{\rm opt}$ satisfy the equation

$$\overline{\alpha}^{\rm opt} = \frac{\nu_0 \left(1 - \overline{\beta}^{\rm opt}\right)^2}{\rho c_1 r_1}. \tag{29}$$

where $\overline{\beta}^{\rm opt} \in [0, 1)$.

[7]Although the one-dimensional exhaustive search seems burdensome, the case in question here is easy because the search is only over a closed set $0 \leq \beta \leq 1$.

Proof: By differentiating $\overline{R}_{\text{sum}}$ with respect to α and β, we immediately obtain the result from Corollary 1. ■

From Proposition 1, we note that the number of asymptotically optimal solutions is infinite. All α's and β's that satisfy (29) are optimal. This condition will be confirmed in the next section.

Similar to (27), we consider Case II for brief illustration. In this case, (29) can be rewritten as

$$\overline{\alpha}^{\text{opt}} = \begin{cases} \dfrac{(1-\overline{\beta}^{\text{opt}})^2}{\rho c_1 r_1}, & 0 < \dfrac{\rho\sigma^2 \mathrm{tr}\mathbf{R}_2}{P_{\text{all}}} \leq 1; \\ \dfrac{(1-\overline{\beta}^{\text{opt}})^2 \sigma^2 \mathrm{tr}\mathbf{R}_2}{c_1 r_1 P_{\text{all}}}, & 1 < \dfrac{\rho\sigma^2 \mathrm{tr}\mathbf{R}_2}{P_{\text{all}}}. \end{cases} \tag{30}$$

From (27), when $0 < \rho\sigma^2 \mathrm{tr}\mathbf{R}_2/P_{\text{all}} \leq 1$, the system performance is unaffected by the average received interference power constraint. In this case, $\overline{\beta}^{\text{opt}}$ is expected to be close to 0 because the weak interference at all the PUs is negligible. This condition is combined with the first term of (30) to reveal that $\overline{\alpha}^{\text{opt}}$ decreases with increasing ρ, where $\rho = P_T/\sigma^2$ is the same as previously defined. However, when $\rho\sigma^2 \mathrm{tr}\mathbf{R}_2/P_{\text{all}} > 1$, the system performance is limited by the average received interference power constraint. To decrease the interference, $\overline{\beta}^{\text{opt}}$ is expected to be close to 1. Therefore, the second term of (30) reveals that α decreases to 0 with an increase in $\overline{\beta}^{\text{opt}}$.

We end this section by observing two additional extreme cases in Theorem 1 for $\mathbf{R}_1 = \mathbf{I}_K$: If $\beta = 0$, by means of some algebraic manipulations, we obtain $\overline{\alpha}^{\text{opt}} = \nu_0/(c_1\rho)$. By contrast, if $\beta = 1$ and $c_2 \neq 1$, we obtain $\overline{\alpha}^{\text{opt}} = 1/(c_1\rho)$. We find that the optimal regularization parameter tends to decrease monotonically with increasing ρ, as expected. This characteristic is similar to that of the conventional RZF precoding in [18, 19], where $r_1 = 1$ is assumed and the asymptotically optimal regularization parameter $\overline{\alpha}^{\text{opt}} = 1/(c_1\rho)$ is derived.

IV. Simulations

In this section, we conduct simulations to confirm our analytical results. First, we compare the analytical results (19) in Theorem 1 and the Monte-Carlo simulation results (13) obtained from averaging over a large number of i.i.d. Rayleigh fading channels. In the simulations, we set channel path gains $r_{1,k} = 1$ and $r_{2,l} = 0.6$ for all k and l and assume that $P_l = P$ for all l in Case I and $P_{\text{all}} = LP$ in Case II. Several characteristics of Cases I and II are similar. Thus, without loss of generality, we provide the numerical results of Case I only.

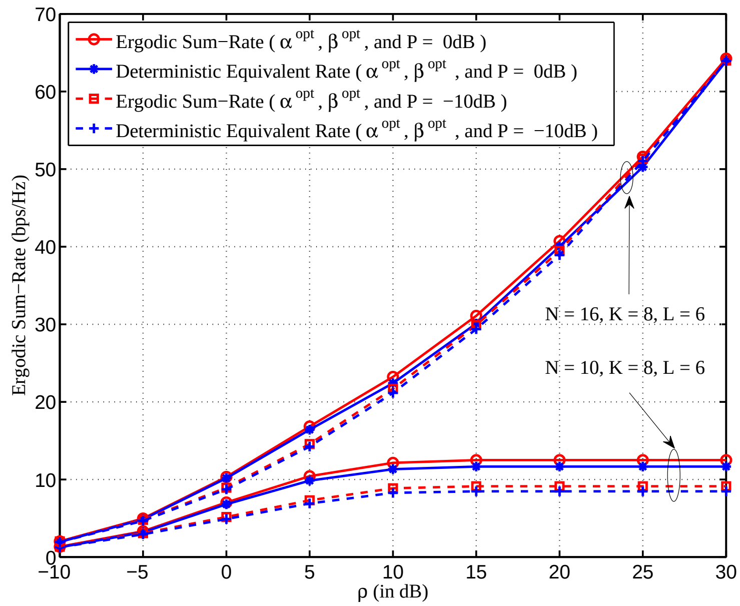

Fig. 2. Ergodic sum-rate and the deterministic equivalent results under different interference power threshold and and two different antenna configuration cases.

Fig. 2 compares the ergodic sum-rate and its deterministic equivalent result under different interference power thresholds $P \in \{-10\text{dB}, 0\text{dB}\}$ and two different antenna configuration cases: $\{N = 10,\ K = 8,\ L = 6\}$ and $\{N = 16,\ K = 8,\ L = 6\}$. In the simulation, $\{\alpha^{\text{opt}}, \beta^{\text{opt}}\}$ is obtained by using the two-dimensional line search in (14). We find that the deterministic equivalent is accurate under various settings even for systems with a not-so-large number of antennas. In addition, Fig. 2 illustrates that for the case with $\{N = 10,\ K = 8,\ L = 6\}$, the sum-rate of the SUs cannot increase linearly in SNR and becomes interference-limited because the sum-rate of the SUs is easily restricted by the average received interference power at each PU, particularly when the number of active users is larger than the number of antennas at the BS, that is, $L + K \geq N$.

In the above simulations, the best solutions of $\{\alpha^{\text{opt}}, \beta^{\text{opt}}\}$ are calculated by Monte-Carlo averaging over 10^4 independent trials; doing so which clearly results in a high computational cost. To confirm that the optimization based on the deterministic equivalent is not only more computationally efficient but also near-optimal, we compare the ergodic sum-rate of the PP-RZF precoding with $P = 0$dB and $\{N = 16,\ K = 8,\ L = 6\}$ in Fig. 3 for the following four cases: 1)

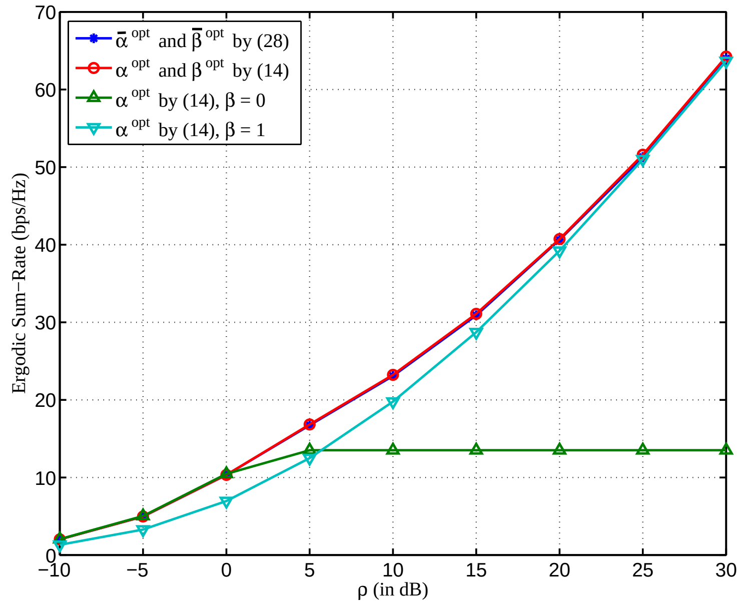

Fig. 3. Ergodic sum-rate results under various parameters with $P = 0$dB and $\{N = 16,\ K = 8,\ L = 6\}$.

$\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\}$, 2) $\{\alpha^{\rm opt}, \beta^{\rm opt}\}$, 3) $\{\alpha^{\rm opt}, \beta = 0\}$, and 4) $\{\alpha^{\rm opt}, \beta = 1\}$. The solution of $\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\}$ is obtained by using the two-dimensional line search in (28). $\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\}$ provides results that are indistinguishable from those achieved by $\{\alpha^{\rm opt}, \beta^{\rm opt}\}$, which demonstrates that the optimization based on the deterministic equivalent is promising. Moreover, the performance is significantly improved if the PP-RZF precoding with an appropriate choice of $\{\alpha, \beta\}$ is employed. In the low-SNR regime, the optimal transmission becomes the conventional RZF precoding, whereas the optimal transmission is the PP-RZF precoding with $\beta = 1$ in the high-SNR regime.

To provide further results on the optimal solutions of $\{\alpha, \beta\}$, Figs. 4 and 5 show the values of $\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\}$, $\{\alpha^{\rm opt}, \beta^{\rm opt}\}$ under various settings. We have observed that the optimal parameter $\{\overline{\alpha}^{\rm opt}, \overline{\beta}^{\rm opt}\}$ based on the deterministic equivalent result is almost consistent with $\{\alpha^{\rm opt}, \beta^{\rm opt}\}$ based on the ergodic sum-rate. Moreover, we have observed that with increasing ρ, $\alpha^{\rm opt}$ (or $\overline{\alpha}^{\rm opt}$) tends to monotonically decrease to 0, whereas $\beta^{\rm opt}$ (or $\overline{\beta}^{\rm opt}$) tends to monotonically increase from 0 to 1. These characteristics are expected based on the analysis in Section III.

Finally, we confirm the result in Proposition 1. Fig. 6 displays the ergodic sum-rate under various parameter settings with $P = 0$dB and $\{N = 10,\ K = 8,\ L = 10\}$. We find that when

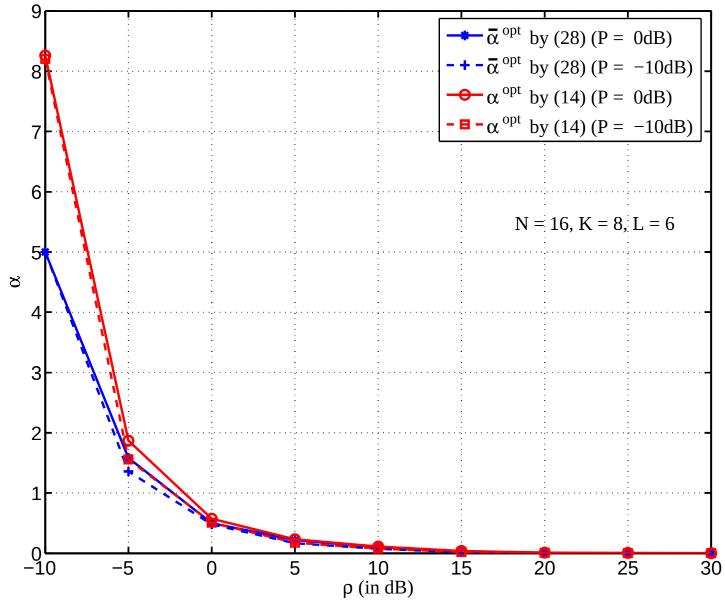

Fig. 4. Optimal α under different the interference power threshold for $\{N = 16,\ K = 8,\ L = 6\}$.

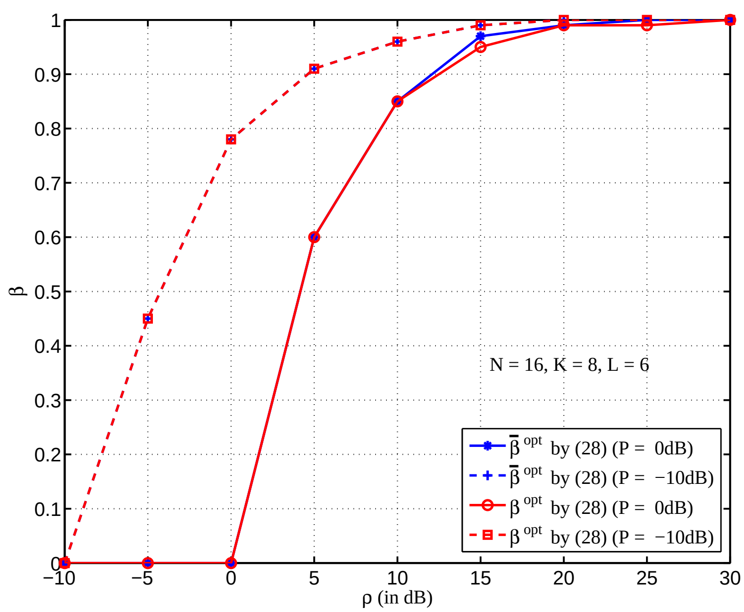

Fig. 5. Optimal β under different interference power threshold for $\{N = 16,\ K = 8,\ L = 6\}$.

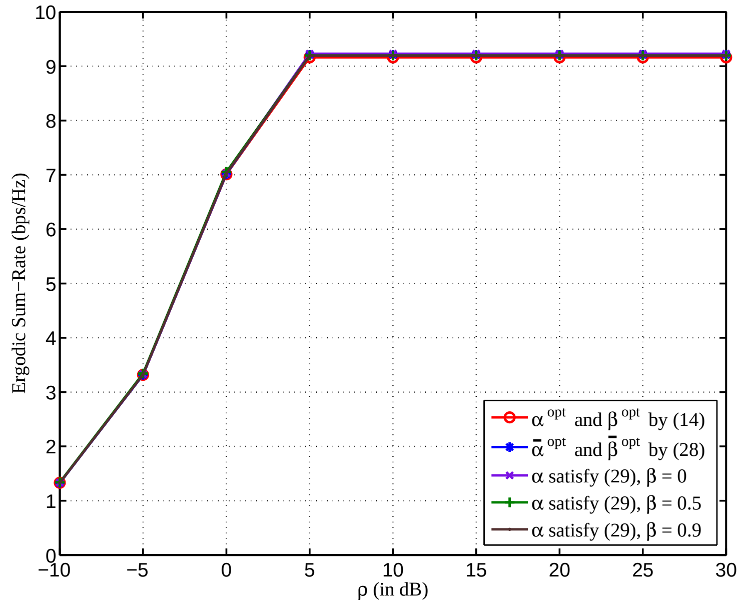

Fig. 6. Ergodic sum-rate results under various parameters for $P = 0\text{dB}$ and $\{N = 10,\ K = 8,\ L = 10\}$.

$c_2 = 1$, the parameters that satisfy (29) can achieve the asymptotically optimal sum-rate for any $\beta \in [0, 1)$, such that infinitely many asymptotically optimal solutions exist.

V. Conclusion

By exploiting the recent advancements in large dimensional RMT, we investigated downlink multiuser CR networks that consist of multiple SUs and multiple PUs. The deterministic equivalent of the ergodic sum-rate based on the PP-RZF precoding was derived. Numerical results revealed that the deterministic equivalent sum-rate provides reliable performance predictions even for systems with a not-so-large number of antennas. We thus used the deterministic equivalent result to identify the asymptotically optimal regularization parameter and the asymptotically optimal projection control parameter. In addition, we provided the condition that the regularization parameter and the projection control parameter are asymptotically optimal. Several insights have been gained into the optimal PP-RZF precoding design. A natural extension of this is to consider the PP-RZF precoding under various scenarios, such as spatial correlations and imperfect CSI at the transmitter. However, such development is still ongoing because of mathematical difficulties.

Appendix A: Proof of Theorem 1

To complete this proof, we first introduce the limiting distribution for a new class of random Hermitian matrix in Theorem 2. Such distribution serves as the mathematical basis for the latter derivation. We recall the definition of the Stieltjes transform (see, e.g., [40]). For a Hermitian matrix $\mathbf{B}_N \in \mathbb{C}^{N\times N}$, the Stieltjes transform of $\mathbf{B}_N$, is defined as

$$m_{\mathbf{B}_N}(\alpha) = \frac{1}{N}\mathsf{tr}\left(\mathbf{B}_N + \alpha\mathbf{I}_N\right)^{-1} \quad \text{for } \alpha \in \mathbb{R}^+.$$

For ease of explanation, we also define the matrix product Stieltjes transform of $\mathbf{B}_N$ as

$$m_{\mathbf{B}_N,\mathbf{Q}}(\alpha) = \frac{1}{N}\mathsf{tr}\mathbf{Q}\left(\mathbf{B}_N + \alpha\mathbf{I}_N\right)^{-1},$$

where $\mathbf{Q}$ is any matrix with bounded spectrum norm (with respect to N).

Notably, both $m_{\mathbf{B}_N}(\alpha)$ and $m_{\mathbf{B}_N,\mathbf{Q}}(\alpha)$ are functions of α, but for ease of notation, α is dropped. In addition, all the subsequent approximations will be performed under the limit $\mathcal{N} \to \infty$, and for ease of expression, $a \asymp b$ denotes that $a - b \xrightarrow{a.s.} 0$ as $\mathcal{N} \to \infty$.

Theorem 2: Consider an $N \times N$ matrix of the following form:

$$\mathbf{B}_N = \check{\mathbf{H}}^H\check{\mathbf{H}} = \big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)\tilde{\mathbf{H}}^H\mathbf{R}_1\tilde{\mathbf{H}}\big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big), \tag{31}$$

where $\tilde{\mathbf{W}}$, $\tilde{\mathbf{H}}$, and $\mathbf{R}_1$ follow the restrictions given by Assumption 1. Then, as $\mathcal{N} \to \infty$, we have

$$m_{\mathbf{B}_N,\mathbf{Q}} \asymp \frac{t_1 + t_2}{\alpha}\frac{1}{N}\mathsf{tr}\mathbf{Q}, \tag{32}$$

where $t_1 = \frac{1-c_2}{1+e}$ and $t_2 = \frac{c_2}{1+e(1-\beta)^2}$ with e being the unique solution to the fixed point equation

$$e = \frac{1}{N}\mathsf{tr}\mathbf{R}_1\left(\alpha\mathbf{I}_K + \left(t_1 + t_2(1-\beta)^2\right)\mathbf{R}_1\right)^{-1}. \tag{33}$$

Proof: If $c_2 = 1$ (i.e., $N = L$), $\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} = \mathbf{I}_L$, the result is directly obtained by Lemma 4 (see Appendix C).

We consider the case with $c_2 < 1$. Given that $m_{\mathbf{B}_N,\mathbf{Q}}$ is a function of two random matrices $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{H}}$, we aim to derive an iterative deterministic equivalent [41] of $m_{\mathbf{B}_N,\mathbf{Q}}$. In particular, we first find a function $\tilde{g}_N(\tilde{\mathbf{W}}, \alpha)$, such that $f_N((\tilde{\mathbf{H}}, \tilde{\mathbf{W}}), \alpha) \asymp \tilde{g}_N(\tilde{\mathbf{W}}, \alpha)$, where $f_N((\tilde{\mathbf{H}}, \tilde{\mathbf{W}}), \alpha) \triangleq m_{\mathbf{B}_N,\mathbf{Q}}$, and $\tilde{g}_N(\tilde{\mathbf{W}}, \alpha)$ is a function of $\tilde{\mathbf{W}}$ and is independent of $\{\tilde{\mathbf{H}}\}_{N\geq 1}$. Notably, $\tilde{g}_N(\tilde{\mathbf{W}}, \alpha)$ is a deterministic equivalent of $f_N((\tilde{\mathbf{H}}, \tilde{\mathbf{W}}), \alpha)$ with respect to random matrix sequences $\{\tilde{\mathbf{H}}\}_{N\geq 1}$. Second, we further find a function $g_N(\alpha)$, such that $\tilde{g}_N(\tilde{\mathbf{W}}, \alpha) \asymp g_N(\alpha)$. Thus, we obtain an iterative deterministic equivalent $g_N(\alpha)$ of $f_N((\tilde{\mathbf{H}}, \tilde{\mathbf{W}}), \alpha)$, i.e., $f_N((\tilde{\mathbf{H}}, \tilde{\mathbf{W}}), \alpha) \asymp g_N(\alpha)$.

When $\tilde{\mathbf{W}}$ is treated as a deterministic matrix, applying Lemma 4 (see Appendix C), we have

$$\frac{1}{N}\mathsf{tr}\mathbf{Q}\left(\mathbf{B}_N + \alpha\mathbf{I}_N\right)^{-1} \asymp \frac{1}{N}\mathsf{tr}\mathbf{Q}\left(\alpha\mathbf{I}_N + \alpha e\big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^2\right)^{-1}, \tag{34}$$

where

$$e = \frac{1}{N}\mathsf{tr}\mathbf{R}_1\left(\alpha\mathbf{I}_K + \tilde{e}\mathbf{R}_1\right)^{-1}, \tag{35}$$

$$\tilde{e} = \frac{1}{N}\mathsf{tr}\big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^2\left(\mathbf{I}_N + e\big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^2\right)^{-1}. \tag{36}$$

Notice the fact that $(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}})^2 = \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}$ so (34) and (36) can be written respectively as

$$\frac{1}{N}\mathsf{tr}\mathbf{Q}\left(\alpha\mathbf{I}_N + \alpha e\big(\mathbf{I}_N - \beta\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^2\right)^{-1} = \frac{1}{\alpha(\beta^2 - 2\beta)e}\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1}, \tag{37}$$

and

$$\tilde{e} = \frac{1}{(\beta^2 - 2\beta)e}\frac{1}{N}\mathsf{tr}\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1} + \frac{1}{e}\frac{1}{N}\sum_{l=1}^{L}\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1}\tilde{\mathbf{w}}_l, \tag{38}$$

where $\omega \triangleq \frac{1+e}{(\beta^2-2\beta)e}$ and $\tilde{\mathbf{w}}_l$ denotes the $l-$th row of $\tilde{\mathbf{W}}$.

Next, we aim to derive the deterministic equivalents of the terms $\frac{1}{N}\mathsf{tr}\mathbf{Q}(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}})^{-1}$ and $\tilde{\mathbf{w}}_l^H(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{w}}_l$. Applying a result of the Haar matrix in Lemma 5 (see Appendix C) to (37) and combing (34), we immediately get (32). Then, we deal with the deterministic equivalent of $\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1}\tilde{\mathbf{w}}_l$. According to the matrix inverse lemma (see, e.g., [42, Lemma 2.1][8]), we find

$$\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1}\tilde{\mathbf{w}}_l = \frac{\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)^{-1}\tilde{\mathbf{w}}_l}{1 + \mathbf{w}_l^H\big(\omega\mathbf{I}_N + \tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)^{-1}\tilde{\mathbf{w}}_l}, \tag{39}$$

where $\tilde{\mathbf{W}}_{[l]} \triangleq [\tilde{\mathbf{w}}_1, \ldots, \tilde{\mathbf{w}}_{l-1}, \tilde{\mathbf{w}}_{l+1}, \ldots, \tilde{\mathbf{w}}_L]^H \in \mathbb{C}^{(L-1)\times N}$. Then, the trace lemma for isometric

[8][42, Lemma 2.1]: For any $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{q} \in \mathbb{C}^n$ with $\mathbf{A}$ and $\mathbf{A} + \mathbf{q}\mathbf{q}^H$ invertible, we have

$$\mathbf{q}^H\left(\mathbf{A} + \mathbf{q}\mathbf{q}^H\right)^{-1} = \frac{1}{1 + \mathbf{q}^H\mathbf{A}^{-1}\mathbf{q}}\mathbf{q}^H\mathbf{A}^{-1}.$$

matrices [43, 44] gives us

$$\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N+\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)^{-1}\tilde{\mathbf{w}}_l \asymp \frac{1}{N-L}\mathsf{tr}\big(\mathbf{I}_N-\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)\big(\omega\mathbf{I}_N+\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)^{-1}$$
$$= \frac{1+\omega}{N-L}\mathsf{tr}\big(\omega\mathbf{I}_N+\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]}\big)^{-1}-\frac{N}{N-L}. \tag{40}$$

Now, applying [42, Lemma 2.2] and (72) to (40), we get

$$\tilde{\mathbf{w}}_l^H\big(\omega\mathbf{I}_N+\tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\big)^{-1}\tilde{\mathbf{w}}_l \asymp \frac{1}{\omega+1}. \tag{41}$$

Substituting (41) into (38) and using (72) and (35), we obtain (33). ■

Note that $m_{\mathbf{B}_N,\mathbf{Q}}$, e, t_1, and t_2 are all functions of α and β, but for ease of expression, α and β are dropped.

Theorem 2 indicates that $m_{\mathbf{B},\mathbf{Q}}$ can be approximated by its deterministic equivalent $\frac{t_1+t_2}{\alpha}\frac{1}{N}\mathsf{tr}\mathbf{Q}$ without knowing the actual realization of channel random components. The deterministic equivalent is analytical and is much easier to compute than $\mathsf{E}_{\mathbf{B}}\{m_{\mathbf{B},\mathbf{Q}}\}$, which requires time-consuming Monte-Carlo simulations. Motivated by this result in the large system limit, we aim to derive the deterministic equivalent of γ_k.

The SINR γ_k in (10) consists of three terms: (i) the signal power $|\mathbf{h}_k^H(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N)^{-1}\check{\mathbf{h}}_k|^2$, (ii) the interference power $\mathbf{h}_k^H(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N)^{-1}\check{\mathbf{H}}_{[k]}^H\mathbf{P}_{[k]}\check{\mathbf{H}}_{[k]}(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N)^{-1}\mathbf{h}_k$, and (iii) the noise power ν. Using Theorem 2, we establish the following three lemmas to derive the deterministic equivalent of each term, whose proofs are detailed in Appendices B-I, B-II, and B-III, successively.

Lemma 1: Under the assumption of Theorem 2, as $\mathcal{N}\to\infty$, we have

$$\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{h}}_k \asymp \overline{a}_k, \tag{42}$$

where $\overline{a}_k$ has been obtained by (16a).

Lemma 2: Under the assumption of Theorem 2, as $\mathcal{N}\to\infty$, we have

$$\mathbf{h}_k^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{h}_k \asymp \overline{b}_k, \tag{43}$$

where $\overline{b}_k$ has been obtained by (16b).

Lemma 3: Under the assumption of Theorem 2, as $\mathcal{N} \to \infty$, we have

$$\nu \asymp \overline{\nu}, \tag{44}$$

where $\overline{\nu}$ can be obtained by (16c) for Case I and by (18) for Case II.

According to Lemma 1, Lemma 2, and Lemma 3, we obtain the deterministic equivalent $\overline{\gamma}_k$ of γ_k in (15). The proof is then completed.

APPENDIX B: PROOFS OF LEMMA 1, LEMMA 2, AND LEMMA 3

B-I: Proof of Lemma 1

We start from an application of the matrix inverse lemma [42, Lemma 2.1] to the signal term, which results in

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k = \frac{\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]} + \alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{h}}_k}{1 + \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]} + \alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{h}}_k}. \tag{45}$$

Using [42, Lemma 2.3 and Lemma 2.2], we obtain

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]} + \alpha\mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \asymp r_{1,k}\frac{1}{N}\mathsf{tr}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1} - r_{1,k}\beta\frac{1}{N}\mathsf{tr}\mathbf{W}^H\mathbf{W}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}. \tag{46}$$

Similarly,

$$\begin{aligned} &\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H\check{\mathbf{H}}_{[k]} + \alpha\mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \\ \asymp\, & r_{1,k}\frac{1}{N}\mathsf{tr}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1} + r_{1,k}(\beta^2 - 2\beta)\frac{1}{N}\mathsf{tr}\mathbf{W}^H\mathbf{W}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1}. \end{aligned} \tag{47}$$

According to Theorem 2, we have

$$\frac{1}{N}\mathsf{tr}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1} \asymp \frac{t_1 + t_2}{\alpha}. \tag{48}$$

Noticing that $\mathbf{W}^H\mathbf{W} = \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}$ and by using the same approach as (38), we obtain

$$\frac{1}{N}\mathsf{tr}\mathbf{W}^H\mathbf{W}\left(\check{\mathbf{H}}^H\check{\mathbf{H}} + \alpha\mathbf{I}_N\right)^{-1} \asymp \frac{1}{\alpha(\beta^2 - 2\beta)e}\frac{1}{N}\sum_{l=1}^{L}\tilde{\mathbf{w}}_l^H\left(\omega\mathbf{I}_N + \tilde{\mathbf{W}}^H\tilde{\mathbf{W}}\right)^{-1}\tilde{\mathbf{w}}_l \asymp \frac{t_2}{\alpha}. \tag{49}$$

Substituting (48) and (49) into (46) and (47), we obtain

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \asymp \frac{r_{1,k}\left(t_1 + t_2(1-\beta)\right)}{\alpha}, \tag{50}$$

$$\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \asymp \frac{r_{1,k}\left(t_1 + t_2(1-\beta)^2\right)}{\alpha}. \tag{51}$$

Consequently, the expression of (45), together with (50) and (51), yields (42).

B-II: Proof of Lemma 2

Using the fact that $\mathbf{A}^{-1} - \mathbf{D}^{-1} = -\mathbf{A}^{-1}(\mathbf{A} - \mathbf{D})\mathbf{D}^{-1}$, we have

$$\begin{aligned}
&\mathbf{h}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
=&\mathbf{h}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k - \alpha \mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
&- \mathbf{h}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
&+ \alpha \mathbf{h}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k.
\end{aligned} \tag{52}$$

Applying the matrix inverse lemma, we obtain

$$\begin{aligned}
&\mathbf{h}_k^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
=&\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k - \frac{\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k}{1 + \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k}.
\end{aligned} \tag{53}$$

Similarly,

$$\begin{aligned}
&\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
=&\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{h}_k - \frac{\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \check{\mathbf{h}}_k \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k}{1 + \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k},
\end{aligned} \tag{54}$$

and

$$\begin{aligned}
&\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \\
=&\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{h}_k - \frac{\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \check{\mathbf{h}}_k \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k}{1 + \check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{h}}_k}.
\end{aligned} \tag{55}$$

According to Theorem 2, we have

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k \asymp \frac{r_{1,k}\left(t_1 + t_2\right)}{\alpha}. \tag{56}$$

Noticing that

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{h}_k = -\frac{\partial}{\partial \alpha} \mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{h}_k,$$

we thus obtain

$$\mathbf{h}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{h}_k \asymp -r_{1,k} \frac{\partial}{\partial \alpha} \left(\frac{t_1 + t_2}{\alpha}\right). \tag{57}$$

Similarly, combining (50) and (51) yields

$$\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{h}_k \asymp -r_{1,k} \frac{\partial}{\partial \alpha} \left(\frac{t_1 + (1-\beta) t_2}{\alpha}\right), \tag{58}$$

$$\check{\mathbf{h}}_k^H \left(\check{\mathbf{H}}_{[k]}^H \check{\mathbf{H}}_{[k]} + \alpha \mathbf{I}_N\right)^{-2} \check{\mathbf{h}}_k \asymp -r_{1,k} \frac{\partial}{\partial \alpha} \left(\frac{t_1 + (1-\beta)^2 t_2}{\alpha}\right). \tag{59}$$

Substituting (50), (51), (56), (57), (58), and (59) into (53), (54), and (55), and combining (42) and (52), we obtain (43).

B-III: Proof of Lemma 3

From (11), we first have

$$\begin{aligned} &\frac{1}{N} \mathsf{tr} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{H}}^H \check{\mathbf{H}} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \\ =&\frac{1}{N} \mathsf{tr} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} - \alpha \frac{1}{N} \mathsf{tr} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-2}, \end{aligned} \tag{60}$$

which, together with Theorem 2, yields

$$\frac{1}{N} \mathsf{tr} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{H}}^H \check{\mathbf{H}} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \asymp \frac{\partial t_1}{\partial \alpha} + \frac{\partial t_2}{\partial \alpha}. \tag{61}$$

For Case I, we have

$$\begin{aligned} &\mathbf{f}_l^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \check{\mathbf{H}}^H \check{\mathbf{H}} \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{f}_l \\ =&\mathbf{f}_l^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-1} \mathbf{f}_l - \alpha \mathbf{f}_l^H \left(\check{\mathbf{H}}^H \check{\mathbf{H}} + \alpha \mathbf{I}_N\right)^{-2} \mathbf{f}_l, \end{aligned} \tag{62}$$

where $l=1,\ldots,L$. From (34) and (37), we obtain

$$\mathbf{f}_l^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{f}_l \asymp \frac{1}{\alpha(\beta^2-2\beta)e}\mathsf{tr}\mathbf{f}_l\mathbf{f}_l^H\left(\omega\mathbf{I}_N+\mathbf{W}^H\mathbf{W}\right)^{-1}. \tag{63}$$

Noticing the fact that $\mathbf{W}=(\mathbf{F}\mathbf{F}^H)^{-\frac{1}{2}}\mathbf{F}$, by using the matrix inversion formula[9], we obtain

$$\begin{aligned}\mathsf{tr}\mathbf{f}_l\mathbf{f}_l^H\left(\omega\mathbf{I}_N+\mathbf{W}^H\mathbf{W}\right)^{-1} =&\mathsf{tr}\mathbf{f}_l\mathbf{f}_l^H\left(\omega^{-1}\mathbf{I}_N-\omega^{-1}\mathbf{F}^H\left(\mathbf{F}\mathbf{F}^H+\omega^{-1}\mathbf{F}\mathbf{F}^H\right)^{-1}\mathbf{F}\omega^{-1}\right)\\ =&\frac{1}{\omega+1}\mathsf{tr}\mathbf{f}_l\mathbf{f}_l^H \asymp \frac{r_{2,l}}{\omega+1}.\end{aligned} \tag{64}$$

As a result,

$$\mathbf{f}_l^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{f}_l \asymp \frac{r_{2,l}t_2}{c_2\alpha}. \tag{65}$$

Combining this with (62), we obtain

$$\mathbf{f}_l^H\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\left(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\right)^{-1}\mathbf{f}_l \asymp \frac{r_{2,l}}{c_2}\frac{\partial t_2}{\partial\alpha}. \tag{66}$$

From (61) and (66), the proof of (16c) can be accomplished using

$$\begin{aligned}&\frac{1}{N}\mathsf{tr}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\\ \asymp&\mathsf{E}\left\{\frac{1}{N}\mathsf{tr}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\right\},\end{aligned} \tag{67a}$$

$$\begin{aligned}&\mathbf{f}_l^H\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\mathbf{f}_l\\ \asymp&\mathsf{E}\left\{\mathbf{f}_l^H\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\mathbf{f}_l\right\}.\end{aligned} \tag{67b}$$

By using the martingale approach, we can prove (67) (See [45] for a similar application).

Similarly, for Case II, we have

$$\mathsf{E}\left\{\mathsf{tr}\mathbf{F}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\check{\mathbf{H}}^H\check{\mathbf{H}}\big(\check{\mathbf{H}}^H\check{\mathbf{H}}+\alpha\mathbf{I}_N\big)^{-1}\mathbf{F}^H\right\} \asymp \frac{\mathsf{tr}\mathbf{R}_2}{c_2}\frac{\partial t_2}{\partial\alpha}. \tag{68}$$

Therefore, we obtain (18).

Appendix C: Related Lemmas

In this appendix, we provide some lemmas needed in the proof of Appendix A.

[9]For invertible $\mathbf{A},\mathbf{B}$ and $\mathbf{R}$ matrices, suppose that $\mathbf{B}=\mathbf{A}+\mathbf{X}\mathbf{R}\mathbf{Y}$, then $\mathbf{B}^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{X}(\mathbf{R}^{-1}+\mathbf{Y}\mathbf{A}^{-1}\mathbf{X})^{-1}\mathbf{Y}\mathbf{A}^{-1}$.

Lemma 4: Let $\mathbf{X} \equiv [\frac{1}{\sqrt{N}} X_{ij}] \in \mathbb{C}^{N\times K}$, where X_{ij}'s are i.i.d. with zero mean, unit variance and finite 4-th order moment. In addition, let $\mathbf{Q} \in \mathbb{C}^{N\times N}$, $\mathbf{T} \in \mathbb{C}^{N\times N}$, and $\mathbf{R} \in \mathbb{C}^{K\times K}$ be nonnegative definite matrices with uniformly bounded spectral norm (with respect to N, N, and K, respectively). Consider an $N \times N$ matrix of the form $\mathbf{B}_N = \mathbf{T}^{\frac{1}{2}}\mathbf{X}\mathbf{R}\mathbf{X}^H\mathbf{T}^{\frac{1}{2}}$. Define $c_1 = N/K$. Then, as $K, N \to \infty$ such that $0 < \liminf_N c_1 \le \limsup_N c_1 < \infty$, the following holds for any $\omega \in \mathbb{R}^+$:

$$\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\mathbf{B}_N + \omega\mathbf{I}_N\big)^{-1} \asymp \frac{1}{N}\mathsf{tr}\mathbf{Q}\left(\omega\mathbf{I}_N + \omega e\mathbf{T}\right)^{-1}, \tag{69}$$

where e is given as the unique solution to the fixed-point equations

$$e = \frac{1}{N}\mathsf{tr}\mathbf{R}(\omega\mathbf{I}_K + \tilde{e}\mathbf{R})^{-1},$$
$$\tilde{e} = \frac{1}{N}\mathsf{tr}\mathbf{T}(\mathbf{I}_N + e\mathbf{T})^{-1}.$$

Proof: As a special case of [46, Theorem 1] or [21, Theorem 1], the result can be obtained immediately. ■

Lemma 5: Let $\mathbf{Q} \in \mathbb{C}^{N\times N}$ be a nonnegative definite matrix with uniformly bounded spectral norm (with respect to N) and $\tilde{\mathbf{W}} \in \mathbb{C}^{L\times N}$ be $L \le N$ rows of an $N \times N$ Haar-distributed unitary random matrix. Define $c_2 = L/N$. Then, as $L, N \to \infty$ such that $0 < \liminf_N c_2 \le \limsup_N c_2 \le 1$, the following holds for any $\omega \in \mathbb{R}^+$:

$$\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} \asymp \left(\frac{c_2}{\omega+1} + \frac{1-c_2}{\omega}\right)\frac{1}{N}\mathsf{tr}\mathbf{Q}. \tag{70}$$

Proof: Since $\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} = \mathbf{I}_L$ for $c_2 = 1$ (i.e., $N = L$), (70) evidently holds. We assume $c_2 < 1$ in the following proof. Firstly, we consider a special case with $\mathbf{Q} = \mathbf{I}$. Using the identity of the Stieltjes transform [29, Lemma 3.1] [10], we have

$$\frac{1}{N}\mathsf{tr}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} = \frac{c_2}{L}\mathsf{tr}\big(\tilde{\mathbf{W}}\tilde{\mathbf{W}}^H + \omega\mathbf{I}_L\big)^{-1} + \frac{1-c_2}{\omega}. \tag{71}$$

[10] [29, Lemma 3.1]: Let $\mathbf{A} \in \mathbb{C}^{N\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times N}$, such that $\mathbf{AB}$ is Hermitian. Then, for $z \in \mathbb{C}\backslash\mathbb{R}$

$$\frac{n}{N}m_{\mathbf{BA}}(z) = m_{\mathbf{AB}}(z) + \frac{N-n}{N}\frac{1}{z}.$$

Notice that the rows of $\tilde{\mathbf{W}}$ are orthogonal and hence $\tilde{\mathbf{W}}\tilde{\mathbf{W}}^H = \mathbf{I}_L$. Therefore,

$$\frac{1}{N}\mathsf{tr}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} \asymp \delta \triangleq \frac{c_2}{\omega+1} + \frac{1-c_2}{\omega}. \tag{72}$$

Next, for any nonnegative definite matrix with uniformly bounded spectral norm (with respect to N) $\mathbf{Q}$, we have

$$\begin{aligned}&\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} - \delta\frac{1}{N}\mathsf{tr}\mathbf{Q}\\ &= (1-\delta\omega)\,\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} - \delta\sum_{l=1}^{L}\mathbf{w}_l^H\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1}\mathbf{w}_l,\end{aligned} \tag{73}$$

where the first equality follows from the resolvent identity: $\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B}-\mathbf{A})\mathbf{B}^{-1}$ for invertible matrices $\mathbf{A}$ and $\mathbf{B}$. Using the matrix inverse lemma [42, Lemma 2.1], the trace lemma for isometric matrices [43, 44], and the fact that $\mathbf{Q}$ has uniformly bounded spectral norm (with respect to N), we obtain

$$\begin{aligned}\sum_{l=1}^{L}\mathbf{w}_l^H\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1}\mathbf{w}_l &= \sum_{l=1}^{L}\frac{\mathbf{w}_l^H\mathbf{Q}\big(\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]} + \omega\mathbf{I}_N\big)^{-1}\mathbf{w}_l}{1+\mathbf{w}_l^H\big(\tilde{\mathbf{W}}_{[l]}^H\tilde{\mathbf{W}}_{[l]} + \omega\mathbf{I}_N\big)^{-1}\mathbf{w}_l}\\ &\asymp c_2\frac{(1+\omega)\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} - \frac{1}{N}\mathsf{tr}\mathbf{Q}}{(1+\omega)\frac{1}{N}\mathsf{tr}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} - c_2}.\end{aligned} \tag{74}$$

Substituting (74) into (73), and combining (72), yields

$$\frac{(1+\omega)\delta\left(\frac{1}{N}\mathsf{tr}\mathbf{Q}\big(\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} + \omega\mathbf{I}_N\big)^{-1} - \delta\frac{1}{N}\mathsf{tr}\mathbf{Q}\right)}{(1+\omega)\delta - c_2} \asymp 0. \tag{75}$$

Therefore, we get (70). ■

REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radio more personal," *IEEE Personal Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.

[3] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[4] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sep. 2011.

[5] J. Kim, Y. Shin, T. W. Ban, and R. Schober, "Effect of spectrum sensing reliability on the capacity of multiuser uplink cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4349–4362, Nov. 2011.

[6] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Sig. Proc.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.

[7] S. Huang, X. Liu, and Z. Ding, "Decentralized cognitive radio control based on inference from primary link control information," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 394–406, Feb. 2011.

[8] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Sig. Proc.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[9] K. Hamdi, W. Zhang, and K. B. Letaief, "Opportunistic spectrum sharing in cognitive MIMO wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4098–4109, Aug. 2009.

[10] Y. Y. He and S. Dey, "Sum rate maximization for cognitive MISO broadcast channels: Beamforming design and large systems analysis," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2383–2401, May 2014.

[11] ——, "Weighted sum rate maximization for cognitive MISO broadcast channel: Large system analysis," in *Proc. ICASSP*, Vancouver, Canada, May 2013, pp. 4873–4877.

[12] X. Chen and C. Yuen, "Efficient resource allocation in rateless coded MU-MIMO cognitive radio network with QoS provisioning and limited feedback," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 395–399, Jan. 2013.

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

[14] D. Samardzija, H. Huang, T. Sizer, and R. Valenzuela, "Experimental downlink multiuser MIMO system with distributed and coherently coordinated transmit antennas," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Glasgow, Scotland, Jun. 2007, pp. 5365–5370.

[15] R. Irmer, H.-P. Mayer, A. Weber, V. Braun, M. Schmidt, M. Ohm, N. Ahr, A. Zoch, C. Jandura, P. Marsch, and G. Fettweis, "Multisite field trial for LTE and advanced concepts," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 92–98, Feb. 2009.

[16] H. A. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, and E. G. Larsson, "Multi-pair amplify-and-forward relaying with large antenna arrays," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 3228–3233.

[17] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter for the downlink of TDDDS-CDMA systems," in *Proc. IEEE 7th Int. Symp. Spread-Spectrum Tech. Appl. (ISSSTA)*, vol. 1, 2002, pp. 9–13.

[18] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication–Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[19] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, New Orleans, LA, Dec. 2008, pp. 1–4.

[20] R. Muharar and J. Evans, "Downlink beamforming with transmit-side channel correlation: A large system analysis," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[21] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[22] R. Muharar, R. Zakhour, and J. Evans, "Optimal power allocation and user loading for multiuser MISO channels with regularized channel inversion," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5030–5041, Dec. 2013.

[23] G. Geraci, R. Couillet, J. Yuan, M. Debbah, and I. B. Collings, "Large system analysis of linear precoding in MISO broadcast channels with confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1660–1671, Sep. 2013.

[24] J. Zhang, C. Yuen, C.-K. Wen, S. Jin, and X. Q. Gao, “Ergodic secrecy sum-rate for multiuser downlink transmission via regularized channel inversion: Large system analysis,” *IEEE Commun. Letters*, vol. 18, no. 9, pp. 1627–1630, Sep. 2014.

[25] R. Muharar, R. Zakhour, and J. Evans, “Base station cooperation with limited feedback: A large system analysis,” in *Proc. IEEE Int. Symp. on Inf. Theory Proceedings (ISIT)*, Cambridge, MA, Jul. 2012, pp. 1152–1156.

[26] Y. Huang, C. W. Tan, and B. D. Rao, “Joint beamforming and power control in coordinated multicell: Max-min duality, effective network and large system transition,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2730–2742, Jun. 2013.

[27] J. Zhang, C.-K. Wen, S. Jin, X. Q. Gao, and K.-K. Wong, “Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4801–4813, Oct. 2013.

[28] C.-K. Wen, J.-C. Chen, K.-K. Wong, and P. Ting, “Message passing algorithm for distributed downlink regularized zero-forcing beamforming with cooperative base stations,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2920–2930, May 2014.

[29] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge University Press, 2011.

[30] R. Zhang, “On peak versus average interference power constraints for protecting primary users in cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.

[31] C.-X. Wang, X. Hong, H.-H. Chen, and J. Thompson, “On capacity of cognitive radio networks with average interference power constraints,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1620–1625, Apr. 2009.

[32] L. Zhang, Y. Xin, and Y.-C. Liang, “Weighted sum rate optimization for cognitive radio MIMO broadcast channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2950–2957, Jun. 2009.

[33] L. Dai, Z. Wang, and Z. Yang, “Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 251–263, Feb. 2013.

[34] Z. Gao, L. Dai, Z. Lu, C. Yuen, Z. Wang, “Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations,” *IEEE Commun. Letters*, vol. 18, no. 7, pp. 1266–1269, Jul. 2014.

[35] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge Univ. Press, 1990.

[36] W. Hachem, P. Loubaton, and J. Najim, “Deterministic equivalents for certain functionals of large random matrices,” *Annals of Applied Probability*, 2007.

[37] P. Billingsley, *Probability and Measure*. Hoboken, NJ: Wiley, 2011.

[38] V. A. Marčenko and L. A. Pastur, “Distributions of eigenvalues for some sets of random matrices,” *Math. USSR-Sbornik*, vol. 1, pp. 457–483, Apr. 1967.

[39] M. Rudelson and R. Vershynin, “The smallest singular value of a random rectangular matrix,” *Commun. Pure Appl. Math.*, vol. 62, no. 12, pp. 1595–1739, 2009.

[40] J. W. Silverstein and Z. Bai, “On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” *J. Multiv. Anal.*, vol. 54, pp. 175–192, 1995.

[41] J. Hoydis, R. Couillet, and M. Debbah, “Iterative deterministic equivalents for the performance analysis of communication systems,” 2011. [Online]. Available: http://arxiv.org/abs/1112.4167.

[42] Z. Bai, Y. Chen, and Y.-C. Liang, *Random Matrix Theory and its Applications*. World Scientific Publishing Company, 2009.

[43] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville, “MMSE analysis of certain large isometric random precoded systems,” *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1293–1311, May 2003.

[44] R. Couillet, J. Hoydis, and M. Debbah, “Random beamforming over quasi-static and fading channels: A deterministic equivalent approach,” *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6392–6425, Oct. 2012.

[45] C.-K. Wen, G. Pan, K.-K. Wong, M. H. Guo, and J. C. Chen, “A deterministic equivalent for the analysis of non-Gaussian correlated MIMO multiple access channels,” *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 329–351, Jan. 2013.

[46] J. Zhang, C.-K. Wen, S. Jin, X. Q. Gao, and K.-K. Wong, “On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 133–148, Feb. 2013.