Interference and Noise Reduction by Beamforming in Cognitive Networks

Simon Yiu, *Member, IEEE,* Mai Vu, *Member, IEEE,* and Vahid Tarokh, *Senior Member, IEEE*

***Abstract*—We consider beamforming in a cognitive network with multiple primary users and a secondary user sharing the same spectrum. Each primary and secondary user consists of a transmitter and a receiver. In particular, we assume that the secondary transmitter has N_t antennas and transmits data to its single-antenna receiver using beamforming. The beamformer is designed to maximize the cognitive signal-to-interference ratio (CSIR)[1]. Using mathematical tools from random matrix theory, we derive both lower and upper bounds on the average interference created by the cognitive transmitter at the primary receivers and the average CSIR of the cognitive user. We further analyze and prove the convergence of these two performance measures asymptotically as the number of antennas N_t or primary users N_p increases. Specifically, we show that the average interference per primary receiver converges to $E[d^{-\alpha}]$, the expected value of the path loss in the network, whereas the average CSIR decays as $1/c$ when $c = N_p/N_t \to \infty$. In the special case of $N_t \geq N_p$, the lower bound of the average total interference approaches 0 and the upper bound of the average CSIR approaches $N_t E[d^{-\alpha}]/\sigma_{\mathrm{C}}^2$ where σ_{C}^2 is the noise variance at the cognitive receiver.**

***Index Terms*—Cognitive network, beamforming, fading channels, interference, random matrix theory.**

I. Introduction

THE Federal Communication Commissions (FCC) frequency allocation chart [1] indicates multiple allocations over all the frequency bands under 3 GHz. The intense competition for the use of spectrum at frequencies below 3 GHz creates the conception of spectrum shortage. However, studies by FCC show that the usage of the licensed spectrum is vastly under-utilized [2]. This motivates research in cognitive networks for the opportunistic use of the spectrum.

A cognitive network usually consists of the primary users who have the legacy priority access to the spectrum and the secondary users who use the spectrum only if communication does not create significant interference to the licensed primary users. Therefore, the unlicensed secondary users often employ cognitive radios for transmission to ensure non-interfering coexistence with the primary users [3]. This can be achieved in several ways as discussed in [4] and references therein. For example, the cognitive user can transmit concurrently with the primary users under an enforced spectral mask. Another strategy is to have the cognitive users monitor the spectrum and access it when an unused slot is detected.

Beamforming is a well-known spatial filtering technique which can be used for either directed transmission or reception of energy in the presence of noise and interference [5]. In multiple-antenna systems, beamforming exploits channel knowledge at the transmitter to maximize the signal-to-noise ratio (SNR) at the receiver by transmitting in the direction of the eigenvector corresponding to the largest eigenvalue of the channel [6]. Beamforming can also be used in the uplink or downlink of multiuser systems to maximize the signal-to-interference-plus-noise ratio (SINR) of a particular user [7].

In this paper, we study the effect of beamforming in cognitive networks, in which the primary users and the secondary receiver are uniformly distributed in a circular disc. The secondary transmitter located at the center of the disc is allowed to transmit concurrently with the primary transmitters (cf. Fig. 1). To minimize the interference caused to the primary receivers, the secondary transmitter is equipped with multiple antennas and employs beamforming for transmission. The beamforming vector the cognitive transmitter is designed such that it maximizes the desired signal power at its corresponding receiver while minimizing i) the noise at that receiver and ii) the total interference caused to all primary receivers. The ratio of the received signal power to the interference plus noise is referred to as the cognitive signal-to-interference ratio (CSIR). Since increasing the number of antennas improves the spatial directivity of signal energy, one would then expect a higher average CSIR and a lower average interference. On the other hand, with constant number of antennas, increasing the number of primary users in the network has the opposite effect. Therefore, there is an interesting trade-off between these parameters.

We investigate this trade-off by studying the average CSIR of the cognitive user and the average interference created at all primary receivers. In particular, by employing some known results in random matrix theory, we provide analytical bounds for these two performance measures. We prove that the average interference per primary receiver converges to $E[d^{-\alpha}]$, the average path loss of the network. The average CSIR of each cognitive user pair, on the other hand, decays as $1/c$ as $c = N_p/N_t \to \infty$, where N_p and N_t denote the number of primary receivers and the number of beamforming antennas at each cognitive transmitter, respectively. In the

Paper approved by A. Lozano, the Editor for Wireless Network Access and Performance of the IEEE Communications Society. Manuscript received September 27, 2008; revised February 9, 2009.

S. Yiu and V. Tarokh are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138 USA (e-mail: {simony, vahid}@seas.harvard.edu).

M. Vu was with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138 USA and is now with McGill University, Montreal, QC, H3A 2A7 Canada (e-mail: mai.h.vu@mcgill.ca).

Part of this paper was submitted to the IEEE Global Communications Conference (GLOBECOM) 2008.

Digital Object Identifier 10.1109/TCOMM.2009.10.080501

[1]The CSIR is defined as the ratio of the desired signal at the cognitive receiver to the total interference caused to all primary receivers plus the noise at the cognitive receiver. Please refer to (17) for a formal mathematical definition.

special case of $N_t \geq N_p$, the lower bound of the average total interference approaches 0 and the upper bound of the average CSIR approaches $N_t E[d^{-\alpha}]/\sigma_{\mathrm{C}}^2$ where σ_{C}^2 is the noise variance at the cognitive receiver. This implies that we can potentially create little to no interference to the primary users by employing more antennas in the cognitive transmitter than the number of primary receivers in the network.

Organization: This paper is organized as follows. In Section II, we introduce the system model of the proposed transmission scheme and formulate the beamforming optimization problem. We review some important results on random matrix theory in Section IV. In Section V, we study the average total interference and the average CSIR and derive bounds for these two terms. In Section VI, we provide some discussions on the average total interference and the average CSIR for two special cases where $N_p \gg N_t$ and $N_p \ll N_t$. Finally, we present simulation and numerical results in Section VII, and draw some conclusions in Section VIII.

Notation: In this paper, bold upper case and lower case letters denote matrices and vectors, respectively. $[\cdot]^H$, $E[\cdot]$, $\delta(\cdot)$, $j \triangleq \sqrt{-1}$, $|\cdot|$, $\ln(\cdot)$, $\mathrm{Im}(\cdot)$, $\mathbb{R}^+$, $\Leftrightarrow$, and $\mathrm{diag}\{\boldsymbol{x}\}$ denote Hermitian transposition, statistical expectation, the Dirac delta function, the imaginary unit, the absolute value of a scaler, the natural logarithm, the imaginary part of a complex number, the set of non-negative real number, mathematical equivalence, and a diagonal matrix with the elements of $\boldsymbol{x}$ in its main diagonal, respectively. In addition, $\boldsymbol{I}_N$, $[\boldsymbol{X}]_{i,j}$, $\lambda_{\min}(\boldsymbol{X})$, and $\lambda_{\max}(\boldsymbol{X})$ refer to the $N \times N$ identity matrix, the element in row i and column j of matrix $\boldsymbol{X}$, and the minimum and maximum eigenvalue of matrix $\boldsymbol{X}$, respectively.

II. Network and Channel Models

A. Network Model

Consider a cognitive network with N_p primary users and a single cognitive (secondary) user. Each primary user consists of a transmitter $\mathrm{P}_{\mathrm{T}}^i$ and a receiver $\mathrm{P}_{\mathrm{R}}^i$, $1 \leq i \leq N_p$. Likewise, the cognitive user has a transmitter $\mathrm{C_T}$ and a receiver $\mathrm{C_R}$. The primary users and the secondary receiver are uniformly distributed in a circular disc with radius R whereas the secondary transmitter is located at the center of the disc. Furthermore, we assume that each receiver (either primary or secondary) has a protected radius of $\epsilon > 0$ without any other interfering transmitter inside. This assumption inhibits infinite interference at any receiver. Fig. 1 shows an example of the network under consideration.

In this paper, we consider the interference at the primary receivers created by the cognitive transmitter. The average total interference created by the cognitive transmitter $\mathrm{C_T}$ is defined as

$$E[\mathrm{I}] = E\left[\sum_{i=1}^{N_p} \left|\text{Interference at } \mathrm{P}_{\mathrm{R}}^i\right|^2\right], \tag{1}$$

and the average CSIR of the cognitive user pair $\mathrm{C_T}$-$\mathrm{C_R}$ is defined as[2]

$$E[\mathrm{CSIR}] = E\left[\frac{\left|\text{Desired signal at } \mathrm{C_R}\right|^2}{\mathrm{I} + \sigma_{\mathrm{C}}^2}\right], \tag{2}$$

[2]The mathematical definition of I and CSIR will be formally given in (15) and (17), respectively.

where the expectation is taken over the spatial distribution of $\mathrm{C_R}$ and $\mathrm{P}_{\mathrm{R}}^i$, $\forall i$. In (2), σ_{C}^2 represents the noise variance at $\mathrm{C_R}$[3].

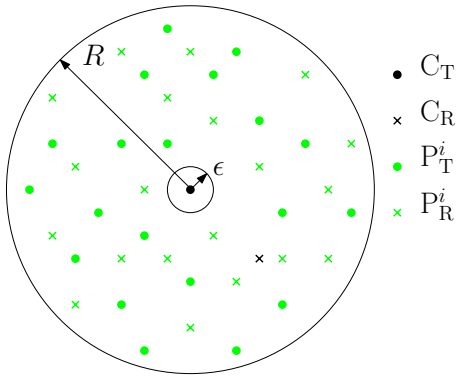

Fig. 1. A typical network with $N_p = 20$.

B. Channel and Signal Models

We assume that the cognitive transmitter $\mathrm{C_T}$ is equipped with N_t uncorrelated antennas whereas the cognitive receiver $\mathrm{C_R}$ and the primary receivers $\mathrm{P}_{\mathrm{R}}^i$, $1 \leq i \leq N_p$ are equipped with only a single antenna[4]. Denote the $N_t \times 1$ channel vector from $\mathrm{C_T}$ to $\mathrm{P}_{\mathrm{R}}^i$ as $\boldsymbol{h}_i$ and from $\mathrm{C_T}$ to $\mathrm{C_R}$ as $\boldsymbol{g}$. The elements of $\boldsymbol{h}_i$ and $\boldsymbol{g}$ are modeled as

$$h_i^n = \frac{1}{d_i^{\alpha/2}} \tilde{h}_i^n, \quad 1 \leq i \leq N_p, \quad 1 \leq n \leq N_t, \tag{3}$$

and

$$g^n = \frac{1}{d^{\alpha/2}} \tilde{g}^n, \quad 1 \leq n \leq N_t, \tag{4}$$

respectively, where $\alpha \geq 2$ is the path loss exponent and $\tilde{h}_i^n$ and $\tilde{g}^n$ are independent and identically distributed (i.i.d.) zero-mean complex Gaussian random variables with unit variance (Rayleigh fading). The distances d_i and d are i.i.d. random variables which represent the distance from $\mathrm{C_T}$ to $\mathrm{P}_{\mathrm{R}}^i$ and to $\mathrm{C_R}$, respectively. By the receiver-protected radius assumption, $d, d_i > \epsilon$. Finally, in order to provide theoretical bounds for the considered network, it is assumed that $\mathrm{C_T}$ has global channel state information (CSI) of the network, i.e., complete knowledge of $\boldsymbol{h}_i$ and $\boldsymbol{g}$.

The cognitive transmitter $\mathrm{C_T}$ employs a beamforming vector $\boldsymbol{w}$ with dimension $N_t \times 1$ for transmission of its data symbol x. The corresponding received signal at $\mathrm{C_R}$ and $\mathrm{P}_{\mathrm{R}}^i$ are given by

$$r_{\mathrm{C}} = \boldsymbol{w}^H \boldsymbol{g} x + n_{\mathrm{C}} \quad \text{and} \quad r_{\mathrm{P}}^i = \boldsymbol{w}^H \boldsymbol{h}_i x + n_i, \tag{5}$$

respectively. The noise terms n_{C} and n_i in (5) are modeled as i.i.d. zero-mean complex Gaussian random variables with variance σ_{C}^2 and σ_i^2, respectively. Finally, we impose an energy constraint on the beamforming vector,

$$\boldsymbol{w}^H \boldsymbol{w} = 1. \tag{6}$$

[3]We note that the primary and secondary users share the same frequency and therefore, the primary transmitters would cause interference to the secondary receiver when they transmit. This interference is a constant that is independent of the beamformer $\boldsymbol{w}$ and therefore, can be also absorbed into σ_{C}^2.

[4]Note that for the problem under consideration, the number of antennas at $\mathrm{P}_{\mathrm{T}}^i$ is not important, cf. (1) and (2).

C. Beamforming Formulation

We assume that the data symbols x in (5) are i.i.d. taken from an M-ary symbol alphabet. Therefore, the instantaneous total interference created by $\mathrm{C_T}$ is given by

$$\begin{aligned} \mathrm{I} &= \sum_{i=1}^{N_p} \left|\boldsymbol{w}^H \boldsymbol{h}_i x\right|^2 \\ &= \sum_{i=1}^{N_p} \boldsymbol{w}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w} = \boldsymbol{w}^H \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H \boldsymbol{w} = \boldsymbol{w}^H \boldsymbol{R}\boldsymbol{w}, \end{aligned} \quad (7)$$

where $[\boldsymbol{H}]_{n,i} = \tilde{h}_i^n$,

$$\boldsymbol{T} = \mathrm{diag}\{d_1^{-\alpha}, \ldots, d_{N_p}^{-\alpha}\}, \quad (8)$$

and $\boldsymbol{R} \triangleq \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$. We define the instantaneous CSIR of the cognitive user pair $\mathrm{C_T}$-$\mathrm{C_R}$ as

$$\begin{aligned} \mathrm{CSIR} &= \frac{|r_{\mathrm{C}}|^2}{\sum_{i=1}^{N_p} |r_{\mathrm{P}}^i|^2 + \sigma_{\mathrm{C}}^2} = \frac{\boldsymbol{w}^H \boldsymbol{G}\boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{R}\boldsymbol{w} + \sigma_{\mathrm{C}}^2} \\ &= \frac{\boldsymbol{w}^H \boldsymbol{G}\boldsymbol{w}}{\boldsymbol{w}^H (\boldsymbol{R} + \sigma_{\mathrm{C}}^2 \boldsymbol{I}_{N_t})\boldsymbol{w}}, \end{aligned} \quad (9)$$

where the definition $\boldsymbol{G} \triangleq \boldsymbol{g}\boldsymbol{g}^H$ is used and the third equality comes from the energy constraint of the beamforming vector, cf. (6). The maximum CSIR beamformer can now be obtained formally from the following optimization problem

$$\boldsymbol{w}_{\mathrm{opt}} = \underset{\boldsymbol{w}^H\boldsymbol{w}=1}{\mathrm{argmax}} \ \{\mathrm{CSIR}\}. \quad (10)$$

The optimal solution to the above optimization problem is the eigenvector corresponding to the maximum eigenvalue of the following generalized eigenvalue problem [7]

$$\boldsymbol{G}\boldsymbol{w} = \lambda(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})\boldsymbol{w} \Leftrightarrow (\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1}\boldsymbol{G}\boldsymbol{w} = \lambda\boldsymbol{w}. \quad (11)$$

For $N_p < N_t$, the matrix $\boldsymbol{R}$ is rank deficient and therefore, not invertible. However, $(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})$ is always invertible for $\sigma_{\mathrm{C}}^2 \neq 0$ regardless of the values of N_t and N_p. We also note that the above optimization problem (18) is closely related to the uplink and downlink beamforming problem considered in [7]. Finally, with the beamforming vector in (18), the instantaneous CSIR in (17) becomes

$$\mathrm{CSIR} = \lambda_{\max}\{(\boldsymbol{R} + \sigma_{\mathrm{C}}^2 \boldsymbol{I}_{N_t})^{-1}\boldsymbol{G}\}, \quad (12)$$

and the corresponding instantaneous total interference in (15) becomes

$$\mathrm{I} = \boldsymbol{w}_{\mathrm{opt}}^H \boldsymbol{R}\boldsymbol{w}_{\mathrm{opt}}. \quad (13)$$

Since $\boldsymbol{G}$ is a rank 1 matrix, $\boldsymbol{R}^{-1}\boldsymbol{G}$ has only one nonzero eigenvalue. Finally, by invoking the Rayleigh's principle [8], the interference can be bounded by

$$\lambda_{\min}(\boldsymbol{R}) \leq \frac{\boldsymbol{w}^H \boldsymbol{R}\boldsymbol{w}}{\boldsymbol{w}^H\boldsymbol{w}} \leq \lambda_{\max}(\boldsymbol{R}). \quad (14)$$

From (14), we can see that it is desirable to compute the largest and smallest eigenvalues of the matrix $\boldsymbol{R}$. Therefore, before we turn our attention to $E[\mathrm{CSIR}]$ and $E[\mathrm{I}]$ and provide bounds for these two average performance measures, in the next section we present some important results on random matrix theory which are useful in obtaining $\lambda_{\min}(\boldsymbol{R})$ and $\lambda_{\max}(\boldsymbol{R})$.

III. Beamforming Formulation

We assume that the data symbols x in (5) are i.i.d. taken from an M-ary symbol alphabet. Therefore, the instantaneous total interference created by $\mathrm{C_T}$ is given by

$$\begin{aligned} \mathrm{I} &= \sum_{i=1}^{N_p} \left|\boldsymbol{w}^H \boldsymbol{h}_i x\right|^2 = \sum_{i=1}^{N_p} \boldsymbol{w}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w} \\ &= \boldsymbol{w}^H \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H \boldsymbol{w} = \boldsymbol{w}^H \boldsymbol{R}\boldsymbol{w}, \end{aligned} \quad (15)$$

where $[\boldsymbol{H}]_{n,i} = \tilde{h}_i^n$,

$$\boldsymbol{T} = \mathrm{diag}\{d_1^{-\alpha}, \ldots, d_{N_p}^{-\alpha}\}, \quad (16)$$

and $\boldsymbol{R} \triangleq \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$. We define the instantaneous SINR of the cognitive user pair $\mathrm{C_T}$-$\mathrm{C_R}$ as

$$\begin{aligned} \mathrm{SINR} &= \frac{|r_{\mathrm{C}}|^2}{\sum_{i=1}^{N_p} |r_{\mathrm{P}}^i|^2 + \sigma_{\mathrm{C}}^2} = \frac{\boldsymbol{w}^H \boldsymbol{G}\boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{R}\boldsymbol{w} + \sigma_{\mathrm{C}}^2} \\ &= \frac{\boldsymbol{w}^H \boldsymbol{G}\boldsymbol{w}}{\boldsymbol{w}^H (\boldsymbol{R} + \sigma_{\mathrm{C}}^2 \boldsymbol{I}_{N_t})\boldsymbol{w}}, \end{aligned} \quad (17)$$

where the definition $\boldsymbol{G} \triangleq \boldsymbol{g}\boldsymbol{g}^H$ is used and the third equality comes from the energy constraint of the beamforming vector, cf. (6). The maximum SINR beamformer can now be obtained formally from the following optimization problem

$$\boldsymbol{w}_{\mathrm{opt}} = \underset{\boldsymbol{w}^H\boldsymbol{w}=1}{\mathrm{argmax}} \ \{\mathrm{SINR}\}. \quad (18)$$

The optimal solution to the above optimization problem is the eigenvector corresponding to the maximum eigenvalue of the following generalized eigenvalue problem [7]

$$\boldsymbol{G}\boldsymbol{w} = \lambda(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})\boldsymbol{w} \Leftrightarrow (\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1}\boldsymbol{G}\boldsymbol{w} = \lambda\boldsymbol{w}. \quad (19)$$

For $N_p < N_t$, the matrix $\boldsymbol{R}$ is rank deficient and therefore, not invertible. However, $(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})$ is always invertible for $\sigma_{\mathrm{C}}^2 \neq 0$ regardless of the values of N_t and N_p. We also note that the above optimization problem (18) is closely related to the uplink and downlink beamforming problem considered in [7]. Finally, with the beamforming vector in (18), the instantaneous SINR in (17) becomes

$$\mathrm{SINR} = \lambda_{\max}\{(\boldsymbol{R} + \sigma_{\mathrm{C}}^2 \boldsymbol{I}_{N_t})^{-1}\boldsymbol{G}\}, \quad (20)$$

and the corresponding instantaneous total interference in (15) becomes

$$\mathrm{I} = \boldsymbol{w}_{\mathrm{opt}}^H \boldsymbol{R}\boldsymbol{w}_{\mathrm{opt}}. \quad (21)$$

Since $\boldsymbol{G}$ is a rank 1 matrix, $\boldsymbol{R}^{-1}\boldsymbol{G}$ has only one nonzero eigenvalue. Next, we turn our attention to $E[\mathrm{SINR}]$ and $E[\mathrm{I}]$ and provide bounds for these two average performance measures.

IV. Methodologies for Computing $\lambda_{\min}(\boldsymbol{R})$ and $\lambda_{\max}(\boldsymbol{R})$

A. Special Case: $\boldsymbol{T} = \boldsymbol{I}_{N_p}$

Recall that the entries of $\boldsymbol{H}$ are i.i.d. Gaussian random variables with zero mean and unit variance. Therefore, if $\boldsymbol{T}$ is an identity matrix[5], the $N_t \times N_t$ matrix $\tilde{\boldsymbol{R}} \triangleq (1/N_t)\boldsymbol{H}\boldsymbol{H}^H$

[5]This corresponds to the case where all primary receivers have the same distance from the secondary transmitter.

is a complex Wishart matrix with N_p degrees of freedom and covariance matrix $\frac{N_p}{N_t}\boldsymbol{I}$. We note that it is a customary practice to consider the matrix $\tilde{\boldsymbol{R}}$ instead of $\boldsymbol{R} = \boldsymbol{H}\boldsymbol{H}^H$ in the literature. We shall apply the results obtained for $\tilde{\boldsymbol{R}}$ to $\boldsymbol{R}$ in Section V.

The Wishart matrix has been studied extensively in the literature and in particular, it is known that the empirical distribution function (e.d.f.) of its eigenvalues defined as

$$F_{\tilde{\boldsymbol{R}}}^{N_t}(x) = \frac{\text{Number of eigenvalues of } \tilde{\boldsymbol{R}} \leq x}{N_t} \tag{22}$$

converges almost surely, as $N_p/N_t \to c > 0$ as $N_t \to \infty$, to a nonrandom distribution function

$$F_{\tilde{\boldsymbol{R}}}(x) = \lim_{N_t\to\infty} E[F_{\tilde{\boldsymbol{R}}}^{N_t}(x)] \tag{23}$$

whose probability density function (p.d.f.) is the famous Marčenko-Pastur law [9]

$$\frac{dF_{\tilde{\boldsymbol{R}}}(x)}{dx} = f_c(x) = (1-c)^+\delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x}, \tag{24}$$

where $(z)^+ = \max(0, z)$, $a = (1-\sqrt{c})^2$, and $b = (1+\sqrt{c})^2$.

Clearly, the region of support associated with (24) is simply the region where $f_c(x) \neq 0$. By inspection, we can see that the support is $(\sqrt{c}-1)^2 \leq x \leq (\sqrt{c}+1)^2$ plus a mass point at $x = 0$ if $c < 1$. This mass point corresponds to the $N_t - N_p > 0$ zero eigenvalues when $N_t > N_p$. The bulk limit in (24) suggests $\lambda_{\min}(\tilde{\boldsymbol{R}}) \approx (\sqrt{c}-1)^2$ and $\lambda_{\max}(\tilde{\boldsymbol{R}}) \approx (\sqrt{c}+1)^2$. Indeed, if the entries in $\boldsymbol{H}$ has finite fourth moment, it has been proven in [10] that

$$\lim_{N_t\to\infty} \lambda_{\max}(\tilde{\boldsymbol{R}}) = (\sqrt{c}+1)^2, \tag{25}$$

whereas [11] has results on the smallest eigenvalue

$$\lim_{N_t\to\infty} \lambda_{\min}(\tilde{\boldsymbol{R}}) = (\sqrt{c}-1)^2. \tag{26}$$

B. $\boldsymbol{T}$ with Known Distribution

For $\tilde{\boldsymbol{R}} \triangleq \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$ where the distribution of $\boldsymbol{T}$ is known, an efficient tool to determine the limiting distribution (and therefore, also $\lambda_{\min}(\tilde{\boldsymbol{R}})$ and $\lambda_{\max}(\tilde{\boldsymbol{R}})$) is the so-called Stieltjes transform. The Stieltjes transform of a distribution function $F_{\tilde{\boldsymbol{R}}}(x)$ is given by

$$m_{\tilde{\boldsymbol{R}}}(z) = \int \frac{1}{x-z} \mathrm{d}F_{\tilde{\boldsymbol{R}}}(x),\ z \in D \equiv \{z \in \mathbb{C}, \text{Im } z > 0\}. \tag{27}$$

The above integral is over the support of $F_{\tilde{\boldsymbol{R}}}(x)$ which will be on $\mathbb{R}^+$ in our case because $\tilde{\boldsymbol{R}}$ is a positive semidefinite matrix with all of its eigenvalues being non-negative. The p.d.f. can be uniquely determined by the Stieltjes-Perron inversion formula [12]

$$\frac{\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x)}{\mathrm{d}x} = \frac{1}{\pi}\lim_{\eta\to 0} \text{Im } m_{\tilde{\boldsymbol{R}}}(\xi + j\eta). \tag{28}$$

It has been shown in [13] (see also [9]) that if the matrices $\boldsymbol{H}$ and $\boldsymbol{T}$ satisfy the following four conditions[6]:

1) $\boldsymbol{H}$ is a $N_t \times N_p$ matrix whose entries are i.i.d. complex random variables with zero mean and unit variance.
2) N_p is a function of N_t with $N_p/N_t \to c > 0$ as $N_t \to \infty$.
3) $\boldsymbol{T}$ is a diagonal matrix with real random entries and the e.d.f. of the entries $\{\tau_1, \ldots, \tau_{N_p}\}$ converges almost surely in distribution to a probability distribution function $F_{\boldsymbol{T}}(\tau)$ as $N_t \to \infty$.
4) $\boldsymbol{H}$ and $\boldsymbol{T}$ are independent.

Then, almost surely, the e.d.f. of $\tilde{\boldsymbol{R}} = (1/N_t)\boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$, namely $F_{\tilde{\boldsymbol{R}}}^{N_t}(x)$, converges in distribution to a nonrandom distribution function $F_{\tilde{\boldsymbol{R}}}(x)$ whose Stieltjes transform $m = m_{\tilde{\boldsymbol{R}}}(z)$ is the unique solution to the following equation

$$m = -\left(z - c\int \frac{\tau \mathrm{d}F_{\boldsymbol{T}}(\tau)}{1+\tau m}\right)^{-1}. \tag{29}$$

The above equation has a unique inverse, given by

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + c\int \frac{\tau \mathrm{d}F_{\boldsymbol{T}}(\tau)}{1+\tau m}, \quad m \in m_{\tilde{\boldsymbol{R}}}(D). \tag{30}$$

To determine the spectral density of $\tilde{\boldsymbol{R}}$ using (28), $m = m_{\tilde{\boldsymbol{R}}}(z)$ in (30) has to be solved explicitly. It is generally difficult, if not impossible, to obtain an analytical or even an easy numerical solution for the density of an arbitrary distribution. However, as shown in [14], much of the analytic behavior of $F_{\tilde{\boldsymbol{R}}}(x)$ can be inferred from (29)-(30) and in particular, the methodology presented in [14] can be used to find the support of $F_{\tilde{\boldsymbol{R}}}(x)$ and can be summarized in the following four steps:

1) Define $B \equiv \{m \in \mathbb{R} : m \neq 0, -m^{-1} \in S_{\boldsymbol{T}}^c\}$ where $S_{\boldsymbol{T}}^c$ denotes the complement of the support of $F_{\boldsymbol{T}}(\tau)$.
2) Plot (30) on B, i.e., $z_{\tilde{\boldsymbol{R}}}(m)$, $m \in B$.
3) Delete the range of values where the derivative $z'_{\tilde{\boldsymbol{R}}}(m) \geq 0$.
4) The remaining range of values is the support of $F_{\tilde{\boldsymbol{R}}}(x)$.

In general, convergence in distribution of $F_{\tilde{\boldsymbol{R}}}^{N_t}(x)$ does not imply that the extreme eigenvalues of $\tilde{\boldsymbol{R}}$, i.e., $\lambda_{\min}(\tilde{\boldsymbol{R}})$ and $\lambda_{\max}(\tilde{\boldsymbol{R}})$, converge to the minimum and maximum of the support of $F_{\tilde{\boldsymbol{R}}}(x)$. However, it has been shown that if the maximum (minimum) eigenvalue of $\boldsymbol{T}$ converges to the largest (smallest) number in the support of $F_{\boldsymbol{T}}(\tau)$, then the largest (smallest) eigenvalue of $\tilde{\boldsymbol{R}}$ converges almost surely to the largest (smallest) number in the support of $F_{\tilde{\boldsymbol{R}}}(x)$ [15]. Clearly, the eigenvalues of $\boldsymbol{T}$ are strictly bounded by the support of $F_{\boldsymbol{T}}(\tau)$, cf. (37). Therefore, $\lambda_{\min}(\tilde{\boldsymbol{R}})$ and $\lambda_{\max}(\tilde{\boldsymbol{R}})$, converge to the minimum and maximum of the support of $F_{\tilde{\boldsymbol{R}}}(x)$ in our case.

Finally, using the fact that

$$\frac{\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x)}{\mathrm{d}x} = \frac{1}{\pi}\lim_{\eta\to 0} \text{Im } m_{\tilde{\boldsymbol{R}}}(\xi + j\eta) = \frac{1}{\pi}\text{Im } m_{\tilde{\boldsymbol{R}}}(\xi), \tag{31}$$

where

$$x = \xi = z_{\tilde{\boldsymbol{R}}}(m) \in S_{\tilde{\boldsymbol{R}}} \in \mathbb{R}^+, \tag{32}$$

the density at x can be easily obtained by solving the following equation

$$z_{\tilde{\boldsymbol{R}}}(m) = x, \quad m = m_{\tilde{\boldsymbol{R}}}(x), \tag{33}$$

[6]We note that the original proof in [13], [9] considers matrix in the general form $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$ and there are 5 conditions with an additional condition concerning the requirement of the matrix $\boldsymbol{A}$.

where $z_{\tilde{\boldsymbol{R}}}(m)$ is given in (30). The root m can be obtained by applying the Newton's method to (33), i.e.,

$$m_{i+1} = m_i - \frac{f(m_i)}{f'(m_i)}, \tag{34}$$

where

$$f(m_i) = z_{\tilde{\boldsymbol{R}}}(m_i) - x, \tag{35}$$

and i is the iteration index.

V. Interference and CSIR Analysis

We derive the upper and lower bounds for the average interference and average CSIR in this section. These two measures depend on the path loss matrix $\boldsymbol{T}$ in (16). We first consider $E[\mathrm{I}]$ for $\boldsymbol{T} = \boldsymbol{I}_{N_p}$, then for $\boldsymbol{T}$ in (16), by using the results presented in the last section.

A. Interference Analysis: Special Case: $\boldsymbol{T} = \boldsymbol{I}_{N_p}$

By making use of (6), (14), (25), and (26) the average total interference $E[\mathrm{I}] = E[\boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}]$ can be bounded by [16]

$$N_t(\sqrt{c}-1)^2 \le E[\mathrm{I}] \le N_t(\sqrt{c}+1)^2, \quad 1 \le c \tag{36}$$

where the factor N_t comes from the fact that $\boldsymbol{R} = N_t \tilde{\boldsymbol{R}}$. For $c < 1$, there is a mass point at $x = 0$ and the lower limit of (36) is simply zero.

B. Interference Analysis: $\boldsymbol{T}$ with Known Distribution

Recall that $\boldsymbol{T}$ is a diagonal matrix whose diagonal entries are given by $[\boldsymbol{T}]_{i,i} = d_i^{-\alpha}$ and d_i ($d_i > \epsilon$) is the distance from $\mathrm{C_T}$ to $\mathrm{P_R^i}$, cf. (16). Therefore, for the problem at hand, the e.d.f. of the entries in $\boldsymbol{T}$ converges to a nonrandom distribution function, namely, the distribution of the random variable $\tau = d^{-\alpha}$ where d is the distance between the center of a disc with radius R and a random location in the disc. It is straightforward to show that the p.d.f. of τ is given by

$$\frac{\mathrm{d}F_{\boldsymbol{T}}(\tau)}{\mathrm{d}\tau} = \frac{2}{\alpha(R^2-\epsilon^2)\tau^{\frac{\alpha+2}{\alpha}}}, \quad \frac{1}{R^\alpha} \le \tau \le \frac{1}{\epsilon^\alpha}. \tag{37}$$

The limiting e.d.f. of $\tilde{\boldsymbol{R}}$ and the support of $F_{\tilde{\boldsymbol{R}}}(x)$ can be obtained by using the method presented in the last section. In particular, substituting (37) into (30) yields

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)\tau^{2/\alpha}}. \tag{38}$$

In general, the above definite integral has no closed-form solution for $\alpha > 2$ and one has to resort to numerical integration method such as the Trapezoid rule [17] for the evaluation of (38). For $\alpha = 2$, (38) can be computed in closed-form,

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + \frac{c}{R^2-\epsilon^2} \ln\left(\frac{m+R^2}{m+\epsilon^2}\right), \quad \alpha = 2. \tag{39}$$

Since $S_{\boldsymbol{T}} = \{1/R^\alpha \le m \le 1/\epsilon^\alpha\}$, we obtain the domain B as $B = \{m \in \mathbb{R} : m \ne 0, m < -R^\alpha, m > -\epsilon^\alpha\}$. As mentioned before, $\tilde{\boldsymbol{R}}$ is a positive semidefinite matrix with non-negative eigenvalues and therefore, we only need to consider the positive range of $z_{\tilde{\boldsymbol{R}}}(m)$, i.e., $z_{\tilde{\boldsymbol{R}}}(m) \ge 0$. It can be shown that $z_{\tilde{\boldsymbol{R}}}(m) < 0$ holds for $m \le -R^\alpha$ and $1 < c$ and the proof can be found in Proposition 1 of the Appendix. We note that for $c \le 1$, the smallest eigenvalue of $\boldsymbol{R}$ is simply zero. As a consequence, for our problem, we only have to plot (38) on $B \equiv \{m > -\epsilon^\alpha\}$. A typical $z_{\tilde{\boldsymbol{R}}}(m)$ plot is shown in Fig. 2. For this figure, we assume $R = 10$, $\epsilon = 3$, $\alpha = 2$, and $c = 50$. We choose a relatively large ϵ in this example for illustrative purpose only. In our simulations to be presented in Section VII, we pick $\epsilon = 0.05$. In Fig. 2, we can see that for $m > -\epsilon^\alpha$, there is a vertical asymptote at $m = 0$ and there is a local minimum and a local maximum on the left and on the right of the vertical asymptote, respectively. These critical points serve as the boundary points of the support of $F_{\tilde{\boldsymbol{R}}}(x)$ highlighted in bold line in the vertical axis given by $\lambda_{\max}(\tilde{\boldsymbol{R}}) = z_{\tilde{\boldsymbol{R}}}(m_1)$ and $\lambda_{\min}(\tilde{\boldsymbol{R}}) = z_{\tilde{\boldsymbol{R}}}(m_2)$ where m_1 and m_2 ($m_1 < m_2$) are the two zeros of the derivative of $z_{\tilde{\boldsymbol{R}}}(m)$,

$$z'_{\tilde{\boldsymbol{R}}}(m) = \frac{1}{m^2} - \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)^2 \tau^{(2-\alpha)/\alpha}} = 0. \tag{40}$$

Once again, for general α, one has to resort to numerical method for the evaluation of (40). For $\alpha = 2$, closed-form expression for (40) is possible and by removing the irrelevant terms in $z'_{\tilde{\boldsymbol{R}}}(m)$, the zeros can be obtained by solving the following second-order polynomial,

$$z'_{\tilde{\boldsymbol{R}}}(m) = 0 \Leftrightarrow (1-c)m^2 + (R^2+\epsilon^2)m + \epsilon^2 R^2 = 0, \quad \alpha = 2. \tag{41}$$

For Fig. 2, the resulting boundary points obtained are $\lambda_{\max}(\tilde{\boldsymbol{R}}) = z_{\tilde{\boldsymbol{R}}}(m_1) = 1.859$ and $\lambda_{\min}(\tilde{\boldsymbol{R}}) = z_{\tilde{\boldsymbol{R}}}(m_2) = 0.9086$. The corresponding limiting density $\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x)/\mathrm{d}x$ and the limiting e.d.f. $F_{\tilde{\boldsymbol{R}}}(x)$ of Fig. 2 are also shown in Fig. 3. Finally, by recalling $\boldsymbol{R} = N_t \tilde{\boldsymbol{R}}$, the average total interference $E[\mathrm{I}] = E[\boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}]$ may be bounded by

$$N_t z_{\tilde{\boldsymbol{R}}}(m_2) \le E[\mathrm{I}] \le N_t z_{\tilde{\boldsymbol{R}}}(m_1). \tag{42}$$

For performance comparison, it is more insightful to consider the average interference per primary receiver defined as $E_{\mathrm{a}}[\mathrm{I}] \triangleq E[\mathrm{I}]/N_p$,

$$\frac{z_{\tilde{\boldsymbol{R}}}(m_2)}{c} \le E_{\mathrm{a}}[\mathrm{I}] \triangleq \frac{E[\mathrm{I}]}{N_p} \le \frac{z_{\tilde{\boldsymbol{R}}}(m_1)}{c}. \tag{43}$$

Finally, for $c < 1$, there are $N_t - N_p > 0$ zero eigenvalues and $z_{\tilde{\boldsymbol{R}}}(m_2) = 0$.

C. CSIR Analysis

The upper and lower bounds for $E[\mathrm{CSIR}]$ are readily available by making use of the results obtained from the last subsection. In particular, the average CSIR given by $E[\mathrm{CSIR}] = E[\lambda_{\max}((\boldsymbol{R} + \sigma_{\mathrm{C}}^2 \boldsymbol{I}_{N_t})^{-1} \boldsymbol{G})]$ can be bounded by (44) given at the top of this page. We use $\lambda(\boldsymbol{G})$ to indicate the only nonzero eigenvalue of $\boldsymbol{G}$. In particular, $E[\lambda(\boldsymbol{G})]$ is given by $N_t E[d^{-\alpha}]$ where

$$\begin{aligned} E[d^{-\alpha}] &= \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{2\tau \mathrm{d}\tau}{\alpha(R^2-\epsilon^2)\tau^{\frac{\alpha+2}{\alpha}}} \\ &= \begin{cases} \dfrac{\ln(R^2/\epsilon^2)}{R^2-\epsilon^2}, & \alpha = 2 \\ \dfrac{2(\epsilon^{2-\alpha} - R^{2-\alpha})}{(\alpha-2)(R^2-\epsilon^2)}, & \alpha > 2 \end{cases} \end{aligned} \tag{45}$$

$$E[\lambda_{\min}((\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1})]E[\lambda(\boldsymbol{G})] \le E[\mathrm{CSIR}] \le E[\lambda_{\max}((\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1})]E[\lambda(\boldsymbol{G})]. \quad (44)$$

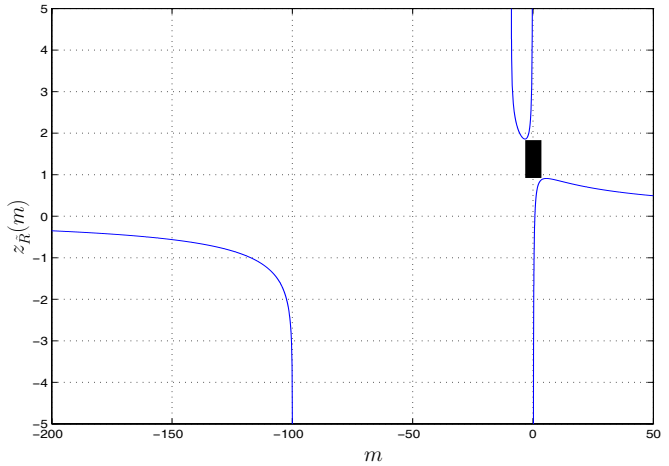

Fig. 2. $z_{\tilde{\boldsymbol{R}}}(m)$ vs. m for $R = 10$, $\epsilon = 3$, $\alpha = 2$, and $c = 50$. Support of $F_{\tilde{\boldsymbol{R}}}(x)$ is highlighted in bold line on the vertical axis where $z_{\tilde{\boldsymbol{R}}}(m_1) = 1.859$ and $z_{\tilde{\boldsymbol{R}}}(m_2) = 0.9086$.

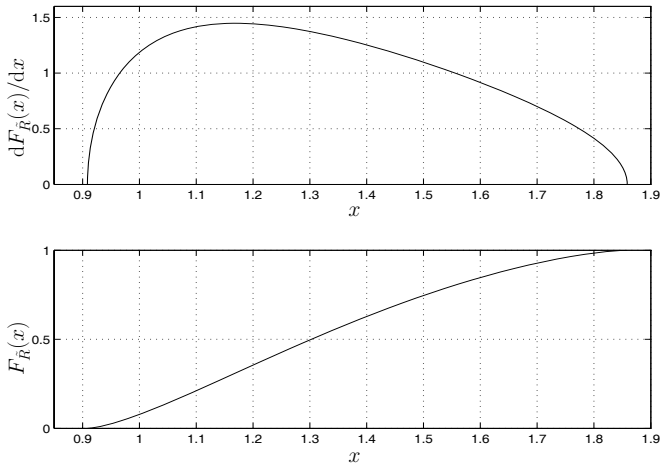

Fig. 3. The limiting density $\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x)/\mathrm{d}x$ and the limiting e.d.f. $F_{\tilde{\boldsymbol{R}}}(x)$ for $R = 10$, $\epsilon = 3$, $\alpha = 2$, and $c = 50$.

is the average path loss of the network. In addition, we note that

$$E[\lambda_{\min}((\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1})] = E[(\lambda_{\max}(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t}))^{-1}],$$
$$E[\lambda_{\max}((\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t})^{-1})] = E[(\lambda_{\min}(\boldsymbol{R}+\sigma_{\mathrm{C}}^2\boldsymbol{I}_{N_t}))^{-1}]. \quad (46)$$

As mentioned in Section IV-B, $\lambda_{\max}(\boldsymbol{R})$ and $\lambda_{\min}(\boldsymbol{R})$ converge almost surely to the maximum and the minimum support of $F_{\boldsymbol{R}}(x)$, respectively. Therefore, we arrive at the following bounds for $E[\mathrm{CSIR}]$

$$\frac{E[d^{-\alpha}]}{z_{\tilde{\boldsymbol{R}}}(m_1)+\sigma_{\mathrm{C}}^2/N_t} \le E[\mathrm{CSIR}] \le \frac{E[d^{-\alpha}]}{z_{\tilde{\boldsymbol{R}}}(m_2)+\sigma_{\mathrm{C}}^2/N_t}. \quad (47)$$

In the next section, we shall present results on $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ as $c \to 0$ and $c \to \infty$. They correspond respectively to the two extreme cases where $N_p \ll N_t$ and $N_p \gg N_t$.

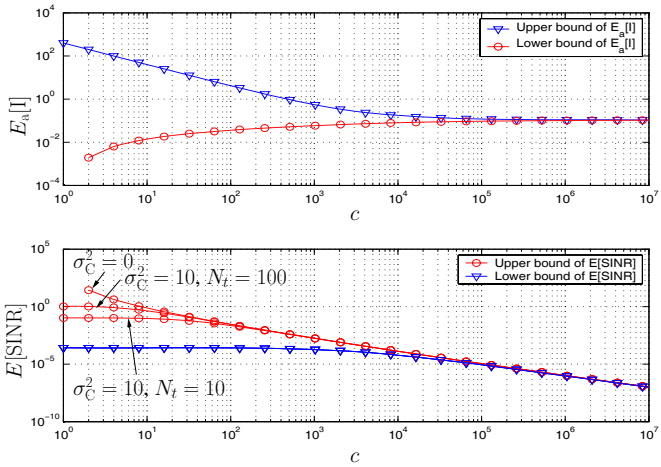

Fig. 4. Upper and lower bounds of $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ for $1 \le c \le 10^7$. $\alpha = 2$, $R = 10$ and $\epsilon = 0.05$.

VI. Discussion

Clearly, $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ are both functions of R, ϵ, and $c = N_p/N_t$. In this section, we provide some insights on how the two performance measures scale in the two extreme cases where i) $N_p/N_t \to c = 0$ as $N_t \to \infty$ and ii) $N_p/N_t \to c = \infty$ as $N_t \to \infty$.

A. $N_p/N_t \to c = 1$ as $N_t \to \infty$

For $c = 1$ (i.e., $N_t = N_p$), $z_{\tilde{\boldsymbol{R}}}(m) < 0$ holds also for $m > 0$ and therefore, the minimum support of $F_{\tilde{\boldsymbol{R}}}(x)$ is zero, i.e., $z_{\tilde{\boldsymbol{R}}}(m_2) = 0$. The proof is given in Proposition 2 of the Appendix. For the special case of $\alpha = 2$ and $c = 1$, the same result can be obtained. In particular, for $\alpha = 2$ and $c = 1$, (41) reduces to a linear function and has only one root. This root corresponds to the upper bound of (42) and $z_{\tilde{\boldsymbol{R}}}(m_2) = 0$. As mentioned in Section III, when $N_t > N_p$, $\tilde{\boldsymbol{R}}$ has $N_t - N_p > 0$ zero eigenvalues resulting in zero interference. For $N_t = N_p$, the lower bound of the interference is zero. Combining the two results, the corresponding upper bound of $E[\mathrm{CSIR}]$ becomes

$$E[\mathrm{CSIR}] \le \frac{N_t E[d^{-\alpha}]}{\sigma_{\mathrm{C}}^2}, \quad N_p \le N_t, \quad N_t \to \infty. \quad (48)$$

Clearly, for a noiseless system, i.e., $\sigma_{\mathrm{C}}^2 = 0$, it is enough to use $N_t = N_p$ antennas at the cognitive transmitter in order to achieve $E[\mathrm{CSIR}] = \infty$. On the other hand, when noise is presented at the cognitive receiver, it is desirable to use as many antennas at the cognitive transmitter as possible to achieve an infinite upper bound for $E[\mathrm{CSIR}]$.

Remark: When $c \to 0$, the limiting e.d.f. of $\tilde{\boldsymbol{R}}$ given by $F_{\tilde{\boldsymbol{R}}}(x)$ approaches the limiting e.d.f. of $\boldsymbol{T}$ given by $F_{\boldsymbol{T}}(\tau)$ [14] and the support of $F_{\tilde{\boldsymbol{R}}}(x)$ approaches the support of $F_{\boldsymbol{T}}(\tau)$, i.e., $1/R^{\alpha} \le x \le 1/\epsilon^{\alpha}$. However, it is important to note that for $c < 1$, there is a mass point of $(1-c)$ at $x = 0$ for $F_{\tilde{\boldsymbol{R}}}(x)$ due to the $N_t - N_p > 0$ zero eigenvalues of $\tilde{\boldsymbol{R}}$ [14].

B. $N_p/N_t \to c = \infty$ as $N_t \to \infty$

In the other extreme where the number of antennas at the cognitive transmitter is much less than the number of primary receivers, i.e., $c \to \infty$, we prove in the Appendix (Proposition 3) that the roots of (40) are approximately

$$m_1, m_2 \approx \pm\sqrt{\frac{(\epsilon^{2(\alpha-1)}R^{2(\alpha-1)})(R^2-\epsilon^2)(\alpha-1)}{c(R^{2(\alpha-1)}-\epsilon^{2(\alpha-1)})}}. \quad (49)$$

For $\alpha = 2$, (49) becomes $m_1, m_2 \approx \pm\sqrt{\epsilon^2 R^2/c}$. This is in accordance with the roots obtained by substituting $c = \infty$ into (41).

Substituting the resulting root into (38), we show in the Appendix (Proposition 4) that the maximum and minimum support of $F_{\tilde{\boldsymbol{R}}}(x)$ given by $z_{\tilde{\boldsymbol{R}}}(m_1)$ and $z_{\tilde{\boldsymbol{R}}}(m_2)$ converges to

$$z_{\tilde{\boldsymbol{R}}}(m_1) \approx z_{\tilde{\boldsymbol{R}}}(m_2) \approx cE[d^{-\alpha}]. \quad (50)$$

Applying the above result to (43) and (47), immediately, we see that for $c \to \infty$

$$E_{\mathrm{a}}[\mathrm{I}] \approx E[d^{-\alpha}] \quad (51)$$

and

$$E[\mathrm{CSIR}] \approx \frac{1}{c + \sigma_{\mathrm{C}}^2/(N_t E[d^{-\alpha}])} \approx \frac{1}{c}. \quad (52)$$

The above result can be also obtained by considering $E[\mathrm{I}]$ directly. By applying the law of large numbers for large N_p, $\boldsymbol{R} = \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$ is approximately a diagonal matrix with $[\boldsymbol{R}]_{n,n} = \sum_{i=1}^{N_p} |\tilde{h}_i^n|^2 d_i^{-\alpha}$. Consequently, $\frac{1}{N_p}\sum_{i=1}^{N_p} |\tilde{h}_i^n|^2 d_i^{-\alpha}$ approaches its expected value given by $E[d^{-\alpha}]$ as $N_p \to \infty$. (51) and (52) suggest that for very dense network, the average interference per primary receiver depends only on the average path loss from $\mathrm{C_T}$ to $\mathrm{P_R^i}$ and the average CSIR decreases exponentially with increasing N_p.

C. An Example

In Fig. (4), we plot the lower and upper bounds in (43) and (47) for $1 \le c \le 10^7$, $\alpha = 2$, $R = 10$ and $\epsilon = 0.05$. For the $E[\mathrm{CSIR}]$ vs. c figure, we consider three cases: i) σ_{C}^2=0, ii) $\sigma_{\mathrm{C}}^2 = 10$ with $N_t = 10$, and iii) $\sigma_{\mathrm{C}}^2 = 10$ with $N_t = 100$. As expected, as $c \to \infty$, the lower and upper bounds for both (43) and (47) for all cases converge according to (51) and (52), respectively. For small values of c ($c < 10^5$), the upper bound of $E[\mathrm{CSIR}]$ varies for the three considered cases. This is because for small c, $z_{\tilde{\boldsymbol{R}}}(m_2)$ is relatively small compared to $\sigma_{\mathrm{C}}^2/N_t$ and therefore, $E[\mathrm{CSIR}]$ is dominated by the effect of noise and N_t, cf. (47). It is also clear from the figure that increasing N_t helps improving $E[\mathrm{CSIR}]$. Finally, we note that for $c = 1$ and $\sigma_{\mathrm{C}}^2 = 0$, the lower bound of (43) and the upper bound of (47) are 0 and ∞, respectively.

VII. Numerical and Simulation Results

In this section, we present some numerical and simulation results. For all results shown we assume $\alpha = 2$, $R = 10$, and $\epsilon = 0.05$. Throughout our analysis, we assumed that $\mathrm{C_T}$ is located at the center of the disc. For comparison, we will show also the results where the location of $\mathrm{C_T}$ is random. In the Appendix (Proposition 5), we prove that the total interference caused to the primary receivers by $\mathrm{C_T}$ located at the center of a disc is always larger than the one if $\mathrm{C_T}$ were placed randomly in the disc.

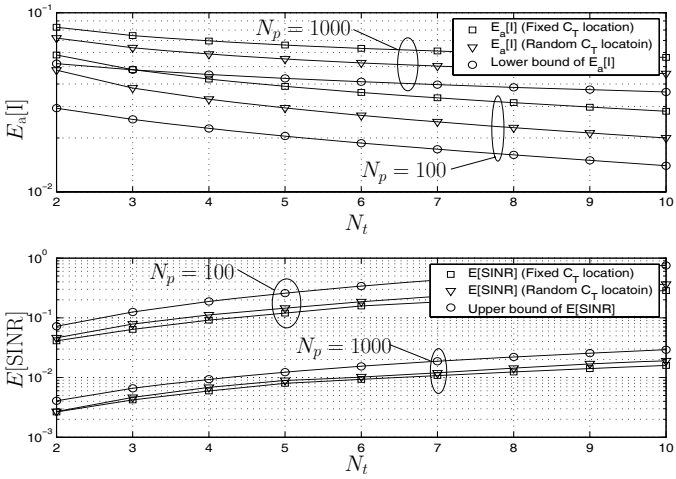

Fig. 5. Simulations of $E[\mathrm{CSIR}]$ and $E_{\mathrm{a}}[\mathrm{I}]$ as a function of N_t and different N_p. $\sigma_{\mathrm{C}}^2 = 0$. Fixed $\mathrm{C_T}$ (squares). Random $\mathrm{C_T}$ (triangles). The lower bound of $E_{\mathrm{a}}[\mathrm{I}]$ (43) and the upper bound of $E[\mathrm{CSIR}]$ (47) are also plotted (circles).

In Fig. 5, the simulated $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ are plotted as a function of N_t for $N_p = 100$ and 1000 and $\sigma_{\mathrm{C}}^2 = 0$. We have shown the results for both random $\mathrm{C_T}$ and fixed $\mathrm{C_T}$ (where $\mathrm{C_T}$ is placed at the center of the disc). For comparison, the lower bound of $E_{\mathrm{a}}[\mathrm{I}]$ (43) and the upper bound of $E[\mathrm{CSIR}]$ (47) are also depicted. Clearly, the average interference is smaller for random $\mathrm{C_T}$ which is in accordance with our discussion in Section V-B. Also as expected, increasing N_p increases $E_{\mathrm{a}}[\mathrm{I}]$ and results in a lower $E[\mathrm{CSIR}]$. On the other hand, increasing the number of antennas at $\mathrm{C_T}$ has the opposite effect. The simulation results are quite close to the theoretical limits given by the lower bound of $E_{\mathrm{a}}[\mathrm{I}]$ and the upper bound of $E[\mathrm{CSIR}]$. The results for the upper bound of $E_{\mathrm{a}}[\mathrm{I}]$ and the lower bound of $E[\mathrm{CSIR}]$ are not shown because the objective of the beamformer is to minimize the interference and maximize the CSIR, cf. (18). In fact, the upper bound of $E_{\mathrm{a}}[\mathrm{I}]$ and the lower bound of $E[\mathrm{CSIR}]$ are quite loose for the relatively small c considered in this figure. This is a good indication that the beamforming vectors are performing well in the small region of c. As we have seen in Fig. 4, as c increases, the system becomes saturated and the upper and lower bounds of $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ converge to the same value. In other words, choosing a random beamforming vector is as good as using the optimal one obtained from (18) for $c \to \infty$.

In general, the bounds obtained for $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ in (43) and (47) are asymptotic bounds for $N_p/N_t \to c > 0$ as $N_t \to \infty$. Therefore, a natural question to ask is how well these bounds perform in finite region of N_t. This question is answered in Fig. 6. In particular, the simulated $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{CSIR}]$ are plotted as a function of N_t for $c = 10$ and 100 (c is kept fixed by varying N_p for different N_t's). Again, we consider both random and fixed $\mathrm{C_T}$ and $\sigma_{\mathrm{C}}^2 = 0$. For reference, we have also plotted the lower bound of $E_{\mathrm{a}}[\mathrm{I}]$ (43) and the upper bound of $E[\mathrm{CSIR}]$ (47). Note that the bounds are constant for fixed value of c because they depend only on

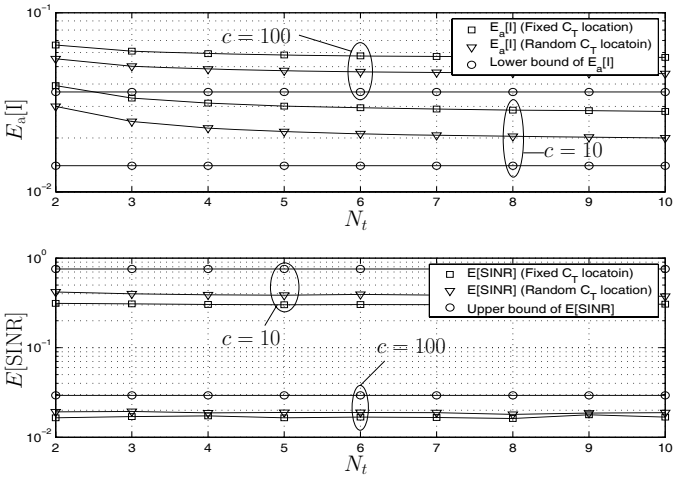

Fig. 6. Simulations of E[CSIR] and $E_{\rm a}$[I] as a function of N_t for different c. $\sigma_{\rm C}^2 = 0$. Fixed $\rm C_T$ (squares). Random $\rm C_T$ (triangles). The lower bound of $E_{\rm a}$[I] (43) and the upper bound of E[CSIR] (47) are also plotted (circles).

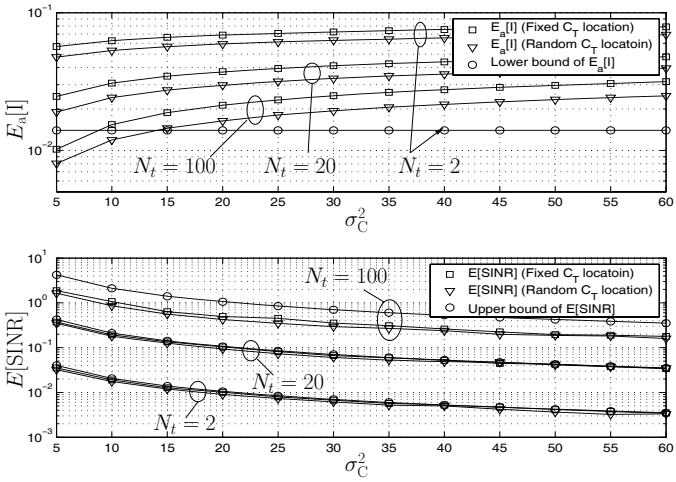

Fig. 7. Simulations of E[CSIR] and $E_{\rm a}$[I] as a function of $\sigma_{\rm C}^2$ for $N_p = 20$ and different N_t. Fixed $\rm C_T$ (squares). Random $\rm C_T$ (triangles). The lower bound of $E_{\rm a}$[I] (43) and the upper bound of E[CSIR] (47) are also plotted (circles).

the ratio $c = N_p/N_t$ and not the actual values of N_p and N_t, cf. (41), (43), and (47). The simulation results for E[CSIR] do not deviate much with fixed c and varying N_t and the bounds work well also for small values of N_t. On the other hand, for $c = 10$ and $N_t < 6$, the simulation results for $E_{\rm a}$[I] depend on the actual values of N_t and N_p because they deviate even if the ratio $c = N_p/N_t$ is fixed to a constant. This is not surprising, because the asymptotic bounds assume large N_t. Nevertheless, when $N_t \geq 6$, $E_{\rm a}$[I] becomes also dependent only on the ratio $c = N_p/N_t$ and not the actual values of N_p and N_t. In general, we find that when c is large enough, the simulation results for $E_{\rm a}$[I] depend also only on the ratio c even for small values of N_t, cf. the $E_{\rm a}$[I] curves for $c = 100$.

In our last example, we investigate the effect of noise on $E_{\rm a}$[I] and E[CSIR]. In Fig. 7, we plot $E_{\rm a}$[I] and E[CSIR] for both fixed and random $\rm C_T$ as a function of the noise variance ($\sigma_{\rm C}^2$) at the cognitive receiver. The lower bound of $E_{\rm a}$[I] and the upper bound of E[CSIR] are also shown. Note that for $N_p \leq N_t$, the lower bound of $E_{\rm a}$[I] is 0. It is clear from the figure that noise has a detrimental effect on both $E_{\rm a}$[I] and E[CSIR]. Although $\sigma_{\rm C}^2$ does not play a direct role in calculating the average interference, cf. (21), we see that $E_{\rm a}$[I] increases as $\sigma_{\rm C}^2$ increases. This is because as $\sigma_{\rm C}^2$ increases, it dominates the denominator of (17) and therefore, the resulting optimal beamforming vector emphasizes more on the maximization of the received signal at the cognitive receiver than on the minimization of the interference. As mentioned in Section VI, if noise is absent from the cognitive receiver, it is enough to use $N_t = N_p$ antennas at the cognitive transmitter in order to achieve $E_{\rm a}[{\rm I}] = 0$ and $E[{\rm CSIR}] = \infty$. However, this is certainly not true when noise is presented at the cognitive receiver and in that case, it is desirable to increase N_t beyond N_p. This is supported by our numerical and simulation results as we can see that for a given $\sigma_{\rm C}^2$, increasing the number of antennas at the cognitive transmitter decreases $E_{\rm a}$[I] while increasing E[CSIR]. On the other hand, if it is not affordable to use many antennas at the cognitive transmitter and minimizing the interference at the primary receivers is the key objective in the network, it may be desirable to drop the noise term and consider only the interference term in the denominator of (17) for the optimization of the beamforming vector. Finally, it is also interesting to see that E[CSIR] is slightly smaller for random $\rm C_T$ than for fixed $\rm C_T$. This is because when $\rm C_T$ is placed at the center, its average distance to $\rm C_R$ is closer than if $\rm C_T$ were placed randomly in the disc resulting in a higher average received power at $\rm C_R$.

VIII. Conclusion

In this paper, we consider a cognitive network which consists of multiple primary users and a single cognitive user. The secondary cognitive transmitter is allowed to transmit concurrently with the primary licensed transmitters. To mitigate interference, the secondary user transmits signals using multiple antennas with a beamforming vector. The beamforming vector is designed to maximize the CSIR. We derive bounds and provide asymptotic analyses for the average CSIR and the average interference caused to all primary receivers. In particular, we have shown that if the number of antennas at the secondary transmitter is greater than or equal to the number of primary receivers, the interference caused to the primary receivers can be made zero, creating an infinite CSIR at the cognitive user if noise is not presented at the cognitive receiver. On the other hand, if the number of primary receivers outgrows the number of antennas at the secondary transmitter, then both the average interference and the average CSIR approach fixed limits. These analyses can be useful in deciding the number of antennas to deploy in the secondary transmitters.

APPENDIX A

Proposition 1: For $m < -R^{\alpha}$ and $1 < c$, $z_{\tilde{\boldsymbol{R}}}(m) < 0$.

Proof: It is desirable to show that

$$\begin{aligned} z_{\tilde{\boldsymbol{R}}}(m) &= -\frac{1}{m} + \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)\tau^{2/\alpha}} \\ &< 0, \quad \text{for } m < -R^{\alpha}. \end{aligned} \tag{53}$$

Let $\tau^{2/\alpha} = x^{-1}$, $z_{\tilde{\boldsymbol{R}}}(m)$ becomes

$$\begin{aligned} z_{\tilde{\boldsymbol{R}}}(m) &= -\frac{1}{m} + \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^2}^{\epsilon^2} \frac{(-\alpha/2)\mathrm{d}x}{(1+mx^{-\alpha/2})x^{\alpha/2}} \\ &= -\frac{1}{m} + \frac{c}{(R^2-\epsilon^2)} \int_{\epsilon^2}^{R^2} \frac{\mathrm{d}x}{(x^{\alpha/2}+m)}. \end{aligned} \tag{54}$$

Therefore, the proof in (53) is equivalent to

$$\int_{\epsilon^2}^{R^2} \frac{-1}{m}\mathrm{d}x < c \int_{\epsilon^2}^{R^2} \frac{-\mathrm{d}x}{(x^{\alpha/2}+m)}. \tag{55}$$

Since $1 < c$, it is sufficient to prove (55) for the case of $c = 1$. Let $n = -m$, (55) becomes

$$\int_{\epsilon^2}^{R^2} \frac{1}{n}\mathrm{d}x < \int_{\epsilon^2}^{R^2} \frac{\mathrm{d}x}{(n-x^{\alpha/2})}. \tag{56}$$

If $n > x^{\alpha/2}$, (56) always holds. The maximum value x can take is R^2. Therefore, if

$$n > R^{\alpha} \Leftrightarrow -m > R^{\alpha} \Leftrightarrow m < -R^{\alpha}, \tag{57}$$

n is guaranteed to be greater than $x^{\alpha/2}$ and (56) always hold which completes this proof.

■

Proposition 2: For $c = 1$, $z_{\tilde{\boldsymbol{R}}}(m) < 0$ for $m > 0$.

Proof: It is desirable to show that

$$\begin{aligned} z_{\tilde{\boldsymbol{R}}}(m) &= -\frac{1}{m} + \frac{2}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)\tau^{2/\alpha}} \\ &< 0, \quad \text{for } m > 0. \end{aligned} \tag{58}$$

Making use of (54) derived in Proposition 2, (58) becomes

$$\frac{1}{(R^2-\epsilon^2)} \int_{\epsilon^2}^{R^2} \frac{\mathrm{d}x}{(x^{\alpha/2}+m)} < \frac{1}{m}. \tag{59}$$

The above equation can be written as

$$\int_{\epsilon^2}^{R^2} \frac{\mathrm{d}x}{(x^{\alpha/2}+m)} < \int_{\epsilon^2}^{R^2} \frac{\mathrm{d}x}{m}. \tag{60}$$

Since $\epsilon^2 \leq x \leq R^2$, (60) always holds and the proof is completed.

■

Proposition 3: For $c \to \infty$, the roots of $z'_{\tilde{\boldsymbol{R}}}(m)$ are

$$m_1, m_2 \approx \pm\sqrt{\frac{(\epsilon^{2(\alpha-1)}R^{2(\alpha-1)})(R^2-\epsilon^2)(\alpha-1)}{c(R^{2(\alpha-1)}-\epsilon^{2(\alpha-1)})}}. \tag{61}$$

Proof: It is desirable to solve

$$\begin{aligned} z'_{\tilde{\boldsymbol{R}}}(m) &= \frac{1}{m^2} - \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)\tau^{(2-\alpha)/\alpha}} \\ &= 0, \quad c \to \infty. \end{aligned} \tag{62}$$

We note that when $c \to \infty$, in order for (62) to hold, $m \propto \sqrt{c^{-1}} \approx 0$ where $\propto$ denotes proportional to. Therefore, $z'_{\tilde{\boldsymbol{R}}}(m)$ becomes

$$\begin{aligned} z'_{\tilde{\boldsymbol{R}}}(m) &= \frac{1}{m^2} - \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{\tau^{(2-\alpha)/\alpha}}, \quad \alpha \geq 2 \\ &= \frac{1}{m^2} - \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \tau^{(\alpha-2)/\alpha}\mathrm{d}\tau \\ &= \frac{1}{m^2} - \frac{2c}{\alpha(R^2-\epsilon^2)} \frac{\alpha}{2(\alpha-1)} \left(\epsilon^{2(1-\alpha)} - R^{2(1-\alpha)}\right) \\ &= \frac{1}{m^2} - \frac{c}{(R^2-\epsilon^2)(\alpha-1)} \left(\frac{R^{2(\alpha-1)}-\epsilon^{2(\alpha-1)}}{\epsilon^{2(\alpha-1)}R^{2(\alpha-1)}}\right). \end{aligned} \tag{63}$$

Setting (63) to zero, we obtain the roots in (61) and the proof is completed.

■

Proposition 4: For $c \to \infty$,

$$z_{\tilde{\boldsymbol{R}}}(m_1) \approx z_{\tilde{\boldsymbol{R}}}(m_2) \approx cE[d^{-\alpha}], \tag{64}$$

where m_1 and m_2 are the zeros of $z'_{\tilde{\boldsymbol{R}}}(m)$ given in (49) and (61) when $c \to \infty$.

Proof: The objective function $z_{\tilde{\boldsymbol{R}}}(m)$ is given by

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{(1+\tau m)\tau^{2/\alpha}}. \tag{65}$$

From (61), we can see that $m_1, m_2 \propto \pm\sqrt{c^{-1}}$. As $c \to \infty$, the first term in (65) is negligible when compared to the second term in (65) and $z_{\tilde{\boldsymbol{R}}}(m)$ is approximately given by

$$z_{\tilde{\boldsymbol{R}}}(m_1) \approx z_{\tilde{\boldsymbol{R}}}(m_2) \approx \frac{2c}{\alpha(R^2-\epsilon^2)} \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} \frac{\mathrm{d}\tau}{\tau^{2/\alpha}}, \quad c \to \infty. \tag{66}$$

Proceeding with the integration in (66) for the case where $\alpha = 2$ and $\alpha > 2$, we arrive with

$$\begin{aligned} z_{\tilde{\boldsymbol{R}}}(m_1) &\approx z_{\tilde{\boldsymbol{R}}}(m_2) \\ &\approx \begin{cases} c\dfrac{\ln(R^2/\epsilon^2)}{R^2-\epsilon^2} = cE[d^{-\alpha}], & \alpha = 2 \\ c\dfrac{2(\epsilon^{2-\alpha}-R^{2-\alpha})}{(R^2-\epsilon^2)(\alpha-2)} = cE[d^{-\alpha}], & \alpha > 2 \end{cases} \end{aligned} \tag{67}$$

where $E[d^{-\alpha}]$ is the average path loss given in (45).

■

Proposition 5: The total interference caused to the primary receivers by $\mathrm{C_T}$ located at the center of a disc is always larger than the one if $\mathrm{C_T}$ were placed randomly in the disc.

Proof: Let $t_c = d_c^{-\alpha}$ and $t_r = d_r^{-\alpha}$ where d_c is the distance from the center of a disc to a random location in the disc and d_r is the distance between two random locations in the disc. We denote the p.d.f. of t_c and t_r by $p_{t_c}(x)$ and $p_{t_r}(x)$, respectively. The domain of $p_{t_c}(x)$ and $p_{t_r}(x)$ are $R^{-\alpha} \leq t_c \leq \epsilon^{-\alpha}$ and $(2R)^{-\alpha} \leq t_r \leq \epsilon^{-\alpha}$, respectively. Note that the total interference $I(d_1^{-\alpha}, \ldots, d_{N_p}^{-\alpha})$ is an increasing function of $d_i^{-\alpha}$, $1 \leq i \leq N_p$ and we have to consider only one distance because d_i's are i.i.d.. Specifically, we have to prove that

$$\int_{(2R)^{-\alpha}}^{\epsilon^{-\alpha}} I(x)p_{t_r}(x)\mathrm{d}x \leq \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} I(x)p_{t_c}(x)\mathrm{d}x. \tag{68}$$

By breaking the integral on the left hand side of (68) into two parts, it can be seen that the prove is identical to

$$\int_{(2R)^{-\alpha}}^{R^{-\alpha}} I(x)p_{t_r}(x)\mathrm{d}x \leq \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} I(x)(p_{t_c}(x)-p_{t_r}(x))\mathrm{d}x. \tag{69}$$

Since $I(x)$ is an increasing function of x, it is sufficient to prove

$$I(R^{-\alpha})\int_{(2R)^{-\alpha}}^{R^{-\alpha}} p_{t_r}(x)\mathrm{d}x \leq \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} I(x)(p_{t_c}(x)-p_{t_r}(x))\mathrm{d}x. \tag{70}$$

Making use of the fact that the integral of any p.d.f. over its domain is 1, we have the following equality[7],

$$\int_{(2R)^{-\alpha}}^{R^{-\alpha}} p_{t_r}(x)\mathrm{d}x = \int_{R^{-\alpha}}^{\epsilon^{-\alpha}} (p_{t_c}(x)-p_{t_r}(x))\mathrm{d}x. \tag{71}$$

Therefore, we can replace the integral on the left hand side of (70) with the right hand side of (71). Finally, recall $I(x)$ is an increasing function of x, we have $I(R^{-\alpha}) \leq I(x)$, $R^{-\alpha} \leq x$ and therefore, the inequality in (70) holds and the proof is completed.

■

ACKNOWLEDGMENT

This research is supported in part by ARO MURI grant number W911NF-07-1-0376 and a NSERC PDF. The views expressed in this paper are those of the authors alone and not of the sponsors.

REFERENCES

[1] United States frequency allocations: The radio spectrum, U.S. Department of Commerce, National Telecommunications and Information Administration, Office of Spectrum Management, Oct. 2003.

[2] Spectrum policy task force report, Federal Communications Commision Tech. Rep. 02-155, Nov. 2002.

[3] J. Mitola, "Cognitive radio," Ph.D. thesis, Royal Institute of Technology (KTH), 2000.

[4] S. Srinivasa and S. A. Jafar, "The throughput potential of cognitive radio: a theoretical perspective," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 73–79, May 2007.

[5] B. D. V. Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag.*, pp. 4–24, 1988.

[6] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.

[7] M. Bengtsson and B. Ottersten, "Uplink and downlink beamforming for fading channels," in *Proc. Signal Processing Advances in Wireless Communications*, pp. 350–353, Annapolis, Maryland, USA, May 1999.

[8] K. Washizu, "On the bounds of eigenvalues," *Quarterly J. Mechanics & Applied Mathematics*, vol. 8, no. 3. pp. 311–325, 1955.

[9] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math USSR Sbornik*, vol. 1, pp. 457–483, 1967.

[10] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah, "On the limit of the largest eigenvalues of the large dimensional sample covariance matrix," *Probability Theory and Related Fields*, vol. 78, no. 4, pp. 509–521, Aug. 1988.

[11] Z. D. Bai and Y. Q. Yin, "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *The Annals of Probability*, vol. 21, no. 3, pp. 1275–1294, 1993.

[12] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*. New York: Hafner Pub. Co., 1965.

[7]Note also $p_{t_r}(x) \leq p_{t_c}(x)$ for $R^{-\alpha} \leq x \leq \epsilon^{-\alpha}$ and $p_{t_c}(x) = 0$ for $x \leq R^{-\alpha}$.

[13] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, Aug. 1995.

[14] J. W. Silverstein and S.-I. Choi, "Analysis of the limiting spectral distribution of large dimensional random matrices," *J. Multivariate Analysis*, vol. 54, no. 2, pp. 295–309, Aug. 1995.

[15] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Annals of Probability*, vol. 26, no. 1, pp. 316–345, Jan. 1998.

[16] A. M. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[17] R. A. Adams, *Single-Variable Calculus*. Addison-Wesley, 1995.

Simon Yiu (S'04, M'08) was born in Hong Kong in 1980. He received his Ph.D. degree in Electrical and Computer Engineering from the University of British Columbia in 2007. Since January 2008, he has been a Postdoctoral Research Fellow at the School of Engineering and Applied Sciences, Harvard University. His research interests span the broad area of wireless communication systems with a particular emphasis in cognitive networks, cooperative diversity systems, MIMO systems and space-time coding theory.

Mai Vu is an assistant professor in Electrical and Computer Engineering at McGill University. Before that, she was a lecturer and researcher at Harvard School of Engineering and Applied Sciences. She received a PhD degree in Electrical Engineering from Stanford University in 2006. Her research interests span the areas of wireless communications, cognitive networks, signal processing, information theory, and convex optimization. She has been working on precoding techniques for exploiting partial channel knowledge at the transmitter in MIMO wireless systems. Currently, she conducts research in cognitive wireless networks, studying the fundamental limits and designing distributed processing algorithms.

Vahid Tarokh worked at AT&T Labs-Research until August 2000, where he was the head of the Department of Wireless Communications and Signal Processing. He then joined the Department of Electrical Engineering and Computer Sciences (EECS) at MIT as an associate professor. In 2002, he joined Harvard University as a Gordon McKay Professor of Electrical Engineering. In July 2005, he was named Perkins Pofessor of Applied Mathematics and Hammond Vinton Hayes Senior Fellow of Electrical Engineering.

Dr. Tarokh's research interest is mainly focused in the areas of Signal processing, Communications (wireline and wireless) and Networking. He has received a number of awards, and holds two honorary degrees.