Interference Reduction by Beamforming in Cognitive Networks

Simon Yiu, Mai Vu, and Vahid Tarokh
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
Email: {simony, maivu, vahid}@seas.harvard.edu

Abstract— **We consider beamforming in a cognitive network with multiple primary users and secondary users sharing the same spectrum. In particular, we assume that each secondary transmitter has N_t antennas and transmits data to its single-antenna receiver using beamforming. The beamformer is designed to maximize the cognitive user's signal-to-interference ratio (SIR), defined as the ratio of the received signal power at the desired cognitive receiver to the total interference created at all the primary receivers. Using mathematical tools from random matrix theory, we derive both lower and upper bounds on the average interference at the primary receivers and the average SIR 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, the average interference per primary receiver converges to the expected value of the path loss in the network whereas the average SIR of the secondary user decays as $1/c$ when $c = N_p/N_t \to \infty$. In the special case of $N_t = N_p$, the average total interference approaches 0 and the average SIR approaches ∞.**

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 and secondary users are uniformly distributed in a circular disc. The secondary transmitters are allowed to transmit concurrently with the primary transmitters. To minimize the interference caused to the primary receivers, the secondary transmitters are equipped with multiple antennas and employ beamforming for transmission. The beamforming vector of each cognitive transmitter is designed such that it maximizes the desired signal power at its corresponding receiver while minimizing the total interference caused to all primary receivers. The ratio of the received signal power to the interference is referred to as the signal-to-interference ratio (SIR) of the cognitive user. Since increasing the number of antennas improves the spatial directivity of signal energy, one would then expect a higher average SIR 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 SIR of the cognitive users 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 the average path loss of the network. The average SIR 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 extreme when $N_t = N_p$, the lower bound of the average total interference approaches 0 and the upper bound of the average SIR approaches ∞. This implies that we can potentially create little to no interference to the primary users by employing as many antennas in the cognitive transmitter as the number of primary receivers.

Organization: This paper is organized as follows. In Section II, we introduce the system model of the proposed transmission scheme. We formulate the beamforming optimization problem in Section III. In Section IV, we study the average total interference and the average SIR and derive bounds for these two terms. We present simulation and numerical results in Section V, and draw some conclusions in Section VI.

978-1-4244-2324-8/08/$25.00 © 2008 IEEE.
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.

II. Network and Channel Models

A. Network Model

Consider a cognitive network in which N_p primary users and N_c cognitive (secondary) users are uniformly distributed in a circular disc with radius R. A primary user has a transmitter P_T^i and a receiver P_R^i, $1 \le i \le N_p$. Similarly, a cognitive user has a transmitter C_T^k and a receiver C_R^k, $1 \le k \le N_c$. 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 a network in which a C_T is located at the center of the disc.

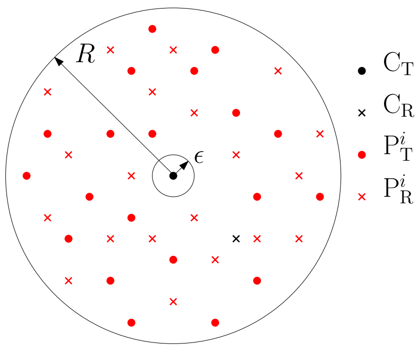

Fig. 1. Network model with $N_c = 1$ and $N_p = 20$. C_T is located at the center of the disc.

In such a multi-user network, there will be interference at both the primary and cognitive receivers. In this paper, we are mainly concerned with the interference at the primary receivers created by the cognitive transmitters. Define the average total interference created by the k^{th} cognitive transmitter C_T^k as[1]

$$E[\mathrm{I}_k] = E\left[\sum_{i=1}^{N_p} \left|\text{Interference at } \mathrm{P}_\mathrm{R}^i\right|^2 |\mathrm{C}_\mathrm{T}^k \text{ transmits}\right], \quad (1)$$

and the average SIR of the cognitive user pair C_T^k-C_R^k as[2]

$$E[\mathrm{SIR}_k] = E\left[\frac{\left|\text{Desired signal at } \mathrm{C}_\mathrm{R}^k\right|^2}{\mathrm{I}_k} |\mathrm{C}_\mathrm{T}^k \text{transmits}\right], \quad (2)$$

where the expectation is taken over the spatial distribution of C_T^k, C_R^k, and P_R^i, $\forall i$. We observe that $E[\mathrm{I}_k]$ and $E[\mathrm{SIR}_k]$ are independent of k because C_T^k and C_R^k are uniformly distributed in the disc and have the same statistical properties. Therefore, we can drop the dependent of k in (1) and (2) and consider only the simplified model of $N_c = 1$ and an arbitrary N_p. As a consequence, increasing the density of the cognitive users would only increase $E[\mathrm{I}]$ linearly by a factor of N_c.

[1] In this paper, bold upper case and lower case letters denote matrices and vectors, respectively. $[\cdot]^H$, $E[\cdot]$, $\delta(\cdot)$, $j \triangleq \sqrt{-1}$, $|\cdot|$, $\ln(\cdot)$, $\mathrm{Im}(\cdot)$, $\Leftrightarrow$, 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, 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.

[2] I_k and SIR_k will be formally defined in (6) and (8).

B. Channel and Signal Models

We assume that the cognitive transmitter C_T is equipped with N_t uncorrelated antennas whereas the cognitive receiver C_R and the primary receivers P_R^i, $1 \le i \le N_p$ are equipped with only a single antenna[3]. Denote the $N_t \times 1$ channel vector from C_T to P_R^i as $\boldsymbol{h}_i$ and from C_T to C_R as $\boldsymbol{g}$. The elements of $\boldsymbol{h}_i$ and $\boldsymbol{g}$ are modeled as

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

and

$$g^n = \frac{1}{d^{\alpha/2}} \tilde{g}^n, \quad 1 \le n \le N_t, \quad (4)$$

respectively, where α 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).

We assume that $\alpha = 2$ in this paper. The distances d_i and d are i.i.d. random variables which represent the distance from C_T to P_R^i and to 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 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 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 C_R and P_R^i are given by

$$r_\mathrm{C} = \boldsymbol{w}^H \boldsymbol{g} x \quad \text{and} \quad r_\mathrm{P}^i = \boldsymbol{w}^H \boldsymbol{h}_i x, \quad (5)$$

respectively.

III. Beamforming Formulation

As previously mentioned, in this paper, we focus on the interference created by the cognitive transmitter to the primary receivers and noise is not considered. However, it is straightforward to incorporate noise into our channel model and study the SINR instead.

We assume that the data symbols x in (5) are i.i.d. taken from an M-ary symbol alphabet with unit energy. Therefore, the instantaneous total interference is given by

$$\mathrm{I} = \sum_{i=1}^{N_p} \left|r_\mathrm{P}^i\right|^2 = \sum_{i=1}^{N_p} \boldsymbol{w}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w} = \boldsymbol{w}^H \boldsymbol{H} \boldsymbol{T} \boldsymbol{H}^H \boldsymbol{w} = \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}, \quad (6)$$

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

$$\boldsymbol{T} = \mathrm{diag}\{d_1^{-2}, \ldots, d_{N_p}^{-2}\}, \quad (7)$$

and $\boldsymbol{R} \triangleq \boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$. The instantaneous SIR of C_T-C_R is given by

$$\mathrm{SIR} = \frac{|r_\mathrm{C}|^2}{\sum_{i=1}^{N_p} \left|r_\mathrm{P}^i\right|^2} = \frac{\boldsymbol{w}^H \boldsymbol{G} \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}}, \quad (8)$$

[3] Note that for the problem under consideration, the number of antennas at P_T^i is not important, cf. (1) and (2).

where the definition $\boldsymbol{G} \triangleq \boldsymbol{g}\boldsymbol{g}^H$ is used. Therefore, the maximum SIR beamformer can be obtained formally from the following optimization problem

$$\boldsymbol{w}_{\rm opt} = \underset{\boldsymbol{w}^H\boldsymbol{w}=1}{\rm argmax} \ \{{\rm SIR}\}. \tag{9}$$

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

$$\boldsymbol{G}\boldsymbol{w} = \lambda\boldsymbol{R}\boldsymbol{w} \Leftrightarrow \boldsymbol{R}^{-1}\boldsymbol{G}\boldsymbol{w} = \lambda\boldsymbol{w}. \tag{10}$$

It is assumed that $N_t \leq N_p$ and therefore, $\boldsymbol{R}$ is invertible. If this condition is not imposed, there will be $N_t - N_p$ zero eigenvalues and it is theoretically possible for $\mathrm{C_T}$ to form a null in all the direction where $\mathrm{P}_{\mathrm{R}}^i$ are located resulting in $\mathrm{I} = 0$. We note that the above optimization problem (9) is closely related to the uplink and downlink beamforming problem considered in [7]. Finally, with the beamforming vector in (9), the instantaneous SIR in (8) becomes

$$\mathrm{SIR} = \lambda_{\max}\{\boldsymbol{R}^{-1}\boldsymbol{G}\}, \tag{11}$$

and the corresponding instantaneous total interference in (6) becomes

$$\mathrm{I} = \boldsymbol{w}_{\rm opt}^H\boldsymbol{R}\boldsymbol{w}_{\rm opt}. \tag{12}$$

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

IV. Interference and SIR analysis

We analyze the average interference and average SIR in this section and derive the upper and lower bounds. These two measures depend on the path loss matrix $\boldsymbol{T}$ in (7). We first study $E[\mathrm{I}]$ for a special case of $\boldsymbol{T}$, then for a general $\boldsymbol{T}$, by using some known random matrix results in the literature. At the end of this section, we provide some discussion on the implication of the results.

A. Interference Analysis: 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[4], the $N_t \times N_t$ matrix $\tilde{\boldsymbol{R}} \triangleq (1/N_t)\boldsymbol{H}\boldsymbol{H}^H$ 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 $\boldsymbol{R}$ at the end of this subsection.

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

$$F_{\tilde{\boldsymbol{R}}}^{N_t}(x) = \frac{\text{Number of eigenvalues of } \tilde{\boldsymbol{R}} \leq x}{N_t} \tag{13}$$

converges almost surely, as $N_p/N_t \to c > 0$ as $N_t \to \infty$, to a nonrandom distribution function[5]

$$F_{\tilde{\boldsymbol{R}}}(x) = \lim_{N_t\to\infty} E[F_{\tilde{\boldsymbol{R}}}^{N_t}(x)] \tag{14}$$

whose probability density function (p.d.f.) is the famous Marčenko-Pastur law [8]

$$\frac{dF_{\tilde{\boldsymbol{R}}}(x)}{dx} = f_c(x) = (1-c)^+\delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x}, \tag{15}$$

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

Clearly, the region of support associated with (15) is simply the region where $f_c(x) \neq 0$. By inspection, we can see that the support is $(\sqrt{c}-1)^2 \leq x \leq (\sqrt{c}+1)^2$. Invoking the Rayleigh's principle [9], we have

$$\lambda_{\min}(\tilde{\boldsymbol{R}}) \leq \frac{\boldsymbol{w}^H\tilde{\boldsymbol{R}}\boldsymbol{w}}{\boldsymbol{w}^H\boldsymbol{w}} \leq \lambda_{\max}(\tilde{\boldsymbol{R}}). \tag{16}$$

The bulk limit in (15) 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

$$\lim_{N_t\to\infty} \lambda_{\max}(\tilde{\boldsymbol{R}}) = (\sqrt{c}+1)^2, \tag{17}$$

whereas [11] has results on the smallest eigenvalue

$$\lim_{N_t\to\infty} \lambda_{\min}(\tilde{\boldsymbol{R}}) = (\sqrt{c}-1)^2. \tag{18}$$

Therefore, the average total interference $E[\mathrm{I}] = E[\boldsymbol{w}^H\boldsymbol{R}\boldsymbol{w}]$ can be bounded by [12]

$$N_t(\sqrt{c}-1)^2 \leq E[\mathrm{I}] \leq N_t(\sqrt{c}+1)^2, \tag{19}$$

where the factor N_t comes from the fact that $\boldsymbol{R} = N_t\tilde{\boldsymbol{R}}$.

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

Recall that $\boldsymbol{T}$ is a diagonal matrix whose diagonal entries are given by $[\boldsymbol{T}]_{i,i} = d_i^{-2}$ and d_i ($d_i > \epsilon$) is the distance from $\mathrm{C_T}$ to $\mathrm{P}_{\mathrm{R}}^i$. Since the location of $\mathrm{C_T}$ and $\mathrm{P}_{\mathrm{R}}^i$ are both uniformly distributed in a circular disc, the distribution of the distance d_i cannot be obtained trivially in closed-form expression. Here, we make a simplified assumption that $\mathrm{C_T}$ is always located at the center of the disc whereas $\mathrm{P}_{\mathrm{R}}^i$ are uniformly distributed in the circular disc. We claim that this always results in a larger average total interference $E[\mathrm{I}]$ and therefore, serves as an upper bound for $E[\mathrm{I}]$ where $\mathrm{C_T}$ is random. Due to space limitation, the proof is omitted and can be found in [13].

Similar to the last subsection, we bound the eigenvalues of $\tilde{\boldsymbol{R}} = (1/N_t)\boldsymbol{H}\boldsymbol{T}\boldsymbol{H}^H$ (and therefore, also the interference: $\boldsymbol{w}^H\boldsymbol{R}\boldsymbol{w}$, $\boldsymbol{R} = N_t\tilde{\boldsymbol{R}}$) by the support of its limiting e.d.f. (cf. (14)). An efficient tool to determine the limiting distribution is the so-called Stieltjes transform. Specifically, the Stieltjes transform of a distribution function $F_{\tilde{\boldsymbol{R}}}(x)$ is given by

$$m_{\tilde{\boldsymbol{R}}}(z) = \int \frac{1}{x-z}\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x),\ z \in D \equiv \{z \in \mathbb{C}, \mathrm{Im}\ z > 0\}. \tag{20}$$

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

[5]Note that $c > 0$ can be trivially satisfied because $1 \leq c$ is valid for our problem as we assume $N_t \leq N_p$.

The above integral is over the support of $F_{\tilde{\boldsymbol{R}}}(x)$ which will be on $x \geq 0$ 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 [14]

$$\frac{\mathrm{d}F_{\tilde{\boldsymbol{R}}}(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{\eta \to 0} \mathrm{Im}\ m_{\tilde{\boldsymbol{R}}}(\xi + j\eta). \tag{21}$$

It has been shown in [15] (see also [8]) 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 $\tilde{\boldsymbol{R}} = (1/N_t)\boldsymbol{HTH}^H$, 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

$$m = -\left(z - c\int \frac{\tau \mathrm{d}F_{\boldsymbol{T}}(\tau)}{1+\tau m}\right)^{-1}. \tag{22}$$

The above equation has a unique inverse, given by

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + c\int \frac{\tau \mathrm{d}F_{\boldsymbol{T}}(\tau)}{1+\tau m}, \quad m \in m_{\tilde{\boldsymbol{R}}}(D). \tag{23}$$

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 $\tau = d^{-2}$ 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

$$\frac{\mathrm{d}F_{\boldsymbol{T}}(\tau)}{\mathrm{d}\tau} = \frac{1}{(R^2-\epsilon^2)\tau^2}, \quad \frac{1}{R^2} \leq \tau \leq \frac{1}{\epsilon^2}. \tag{24}$$

Substituting (24) into (23) yields

$$z_{\tilde{\boldsymbol{R}}}(m) = -\frac{1}{m} + \frac{c}{R^2-\epsilon^2}\ln\left(\frac{m+R^2}{m+\epsilon^2}\right). \tag{25}$$

To determine the spectral density of $\tilde{\boldsymbol{R}}$ using (21), $m = m_{\tilde{\boldsymbol{R}}}(z)$ in (25) 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 [16], much of the analytic behavior of $F_{\tilde{\boldsymbol{R}}}(x)$ can be inferred from (22)-(23) and in particular, the methodology presented in [16] 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 (25) 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 $F_{\tilde{\boldsymbol{R}}}(x)$.

Since $S_{\boldsymbol{T}} = \{1/R^2 \leq m \leq 1/\epsilon^2\}$, $B = \{m \in \mathbb{R} : m \neq 0, m < -R^2, m > -\epsilon^2\}$. 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., $z_{\tilde{\boldsymbol{R}}}(m) \geq 0$. It can be shown that $z_{\tilde{\boldsymbol{R}}}(m) < 0$ holds for $m \leq -R^2$ and the proof can be found in [13]. As a consequence, for our problem, we only have to plot (25) on $B \equiv \{m > -\epsilon^2\}$. A typical $z_{\tilde{\boldsymbol{R}}}(m)$ plot is shown in Fig. 2. For this figure, we assume $R = 10$, $\epsilon = 3$, and $c = 50$. We choose a relatively large ϵ in this example for illustrative purpose only. In our simulations to be presented in Section V, we pick $\epsilon = 0.05$. In Fig. 2, we can see that for $m > -\epsilon^2$, 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 $z_{\tilde{\boldsymbol{R}}}(m_1)$ and $z_{\tilde{\boldsymbol{R}}}(m_2)$ ($z_{\tilde{\boldsymbol{R}}}(m_1) > z_{\tilde{\boldsymbol{R}}}(m_2)$, $m_1 < m_2$), where m_1 and m_2 are the two zeros of the derivative of $z_{\tilde{\boldsymbol{R}}}(m)$. By removing the irrelevant terms in $z'_{\tilde{\boldsymbol{R}}}(m)$, the zeros can be obtained by solving the following second-order polynomial,

$$z'_{\tilde{\boldsymbol{R}}}(m) = 0 \Leftrightarrow (1-c)m^2 + (R^2+\epsilon^2)m + \epsilon^2 R^2 = 0. \tag{26}$$

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 $F_{\tilde{\boldsymbol{R}}}(x)$ [17]. Clearly, the eigenvalues of $\boldsymbol{T}$ are bounded by the support of $F_{\boldsymbol{T}}(\tau)$ in our case, cf. (24). Consequently, 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 $N_t z_{\tilde{\boldsymbol{R}}}(m_2) \leq E[\mathrm{I}] \leq N_t z_{\tilde{\boldsymbol{R}}}(m_1)$. For performance comparison, it is more insightful to consider the average interference per primary receiver defined as $E_{\mathrm{a}}[\mathrm{I}] \triangleq E[\mathrm{I}]/N_p$,

$$\frac{z_{\tilde{\boldsymbol{R}}}(m_2)}{c} \leq E_{\mathrm{a}}[\mathrm{I}] \leq \frac{z_{\tilde{\boldsymbol{R}}}(m_1)}{c}. \tag{27}$$

C. SIR Analysis

The upper and lower bounds for $E[\mathrm{SIR}]$ are readily available by making use of the results obtained from the last subsection. In particular, the average SIR given by $E[\mathrm{SIR}] = E[\lambda_{\max}(\boldsymbol{R}^{-1}\boldsymbol{G})]$ can be bounded by

$$E[\lambda_{\min}(\boldsymbol{R}^{-1})]E[\lambda(\boldsymbol{G})] \leq E[\mathrm{SIR}] \leq E[\lambda_{\max}(\boldsymbol{R}^{-1})]E[\lambda(\boldsymbol{G})]. \tag{28}$$

We use $\lambda(\boldsymbol{G})$ to indicate the only nonzero eigenvalue of $\boldsymbol{G}$. In particular, $E[\lambda(\boldsymbol{G})]$ is equal to $N_t E[d^{-2}]$ where $E[d^{-2}]$ is given by

$$E[d^{-2}] = \int_{R^{-2}}^{\epsilon^{-2}} \frac{\tau}{(R^2-\epsilon^2)\tau^2}\mathrm{d}\tau = \frac{\ln(R^2/\epsilon^2)}{R^2-\epsilon^2}. \tag{29}$$

[6]We note that the original proof in [15], [8] considers matrix in the general form $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{HTH}^H$ and there are 5 conditions with an additional condition concerning the requirement of the matrix $\boldsymbol{A}$.

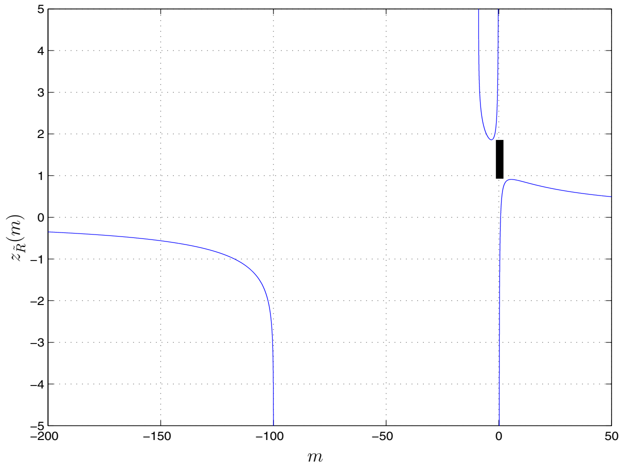

Fig. 2. $z_{\tilde{\boldsymbol{R}}}(m)$ vs. m for $R = 10$, $\epsilon = 3$, and $c = 50$. Support of $F_{\tilde{\boldsymbol{R}}}(x)$ is highlighted in bold line on the vertical axis where $z_{\tilde{\boldsymbol{R}}}(m_1) = 1.859$ and $z_{\tilde{\boldsymbol{R}}}(m_2) = 0.9086$.

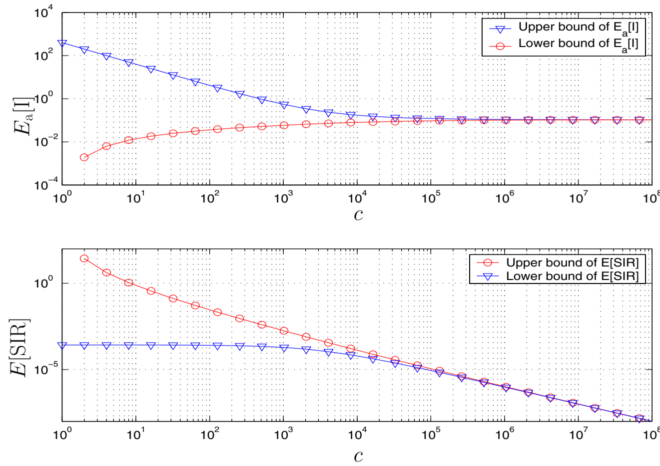

Fig. 3. Upper and lower bounds of $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ for $1 \leq c \leq 10^8$. $R = 10$ and $\epsilon = 0.05$.

We note that $E[\lambda_{\min}(\boldsymbol{R}^{-1})]$ and $E[\lambda_{\max}(\boldsymbol{R}^{-1})]$ are equivalent respectively to $E\left[(\lambda_{\max}(\boldsymbol{R}))^{-1}\right]$ and $E\left[(\lambda_{\min}(\boldsymbol{R}))^{-1}\right]$. As mentioned in the previous subsection, $\lambda_{\max}(\boldsymbol{R})$ 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 $E[\mathrm{SIR}]$

$$\frac{\ln(R^2/\epsilon^2)}{(R^2-\epsilon^2)z_{\tilde{\boldsymbol{R}}}(m_1)} \leq E[\mathrm{SIR}] \leq \frac{\ln(R^2/\epsilon^2)}{(R^2-\epsilon^2)z_{\tilde{\boldsymbol{R}}}(m_2)}. \quad (30)$$

In the next subsection, we shall present results on $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ as $c \to 1$ and $c \to \infty$. They correspond respectively to the two extreme cases where $N_p = N_t$ and $N_p \gg N_t$.

D. Discussion

Clearly, $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ are both functions of R, ϵ, and $c = N_p/N_t$. In this subsection, we provide some insights on how the two performance measures scale as $c \to 1$ and $c \to \infty$ for a given R and ϵ. For $c \to 1$, we note that (26) reduces to a linear function and has only one root. This root corresponds to the upper bound of (27). Therefore, for large N_t, it is theoretically possible to achieve $E_{\mathrm{a}}[\mathrm{I}] = 0$ and $E[\mathrm{SIR}] = \infty$ if there are as many antennas at the cognitive transmitter as the primary receivers, i.e., $N_t = N_p$.

In the other extreme where the number of antennas at the cognitive transmitter is much less than the number of primary receivers, i.e., $c \to \infty$, the roots of (26) are approximately $m_1, m_2 \approx \pm\epsilon R\sqrt{c^{-1}}$. Substituting the resulting root into (25), it can be shown 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

$$z_{\tilde{\boldsymbol{R}}}(m_1) \approx z_{\tilde{\boldsymbol{R}}}(m_2) \approx c\frac{\ln(\mathrm{R}^2/\epsilon^2)}{R^2-\epsilon^2} = cE[d^{-2}]. \quad (31)$$

Applying the above result to (27) and (30), immediately, we see that for $c \to \infty$

$$E_{\mathrm{a}}[\mathrm{I}] \approx E[d^{-2}] \quad \text{and} \quad E[\mathrm{SIR}] \approx \frac{1}{c}. \quad (32)$$

The above result can be also obtained by considering $E[\mathrm{I}]$ directly. By applying the law of large numbers for large N_p, $\boldsymbol{R} = \boldsymbol{HTH}^H$ is approximately a diagonal matrix with $[\boldsymbol{R}]_{n,n} = \sum_{i=1}^{N_p} |\tilde{h}_i^n|^2 d_i^{-2}$. Consequently, $\frac{1}{N_p}\sum_{i=1}^{N_p} |\tilde{h}_i^n|^2 d_i^{-2}$ approaches its expected value given by $E[d^{-2}]$ as $N_p \to \infty$. The above results 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 SIR decreases exponentially with increasing N_p. In Fig. (3), we plot the lower and upper bounds of (27) and (30) for $1 \leq c \leq 10^8$, $R = 10$ and $\epsilon = 0.05$. As expected, the lower and upper bounds for both (27) and (30) converge according to (32) as $c \to \infty$. Note also that for $c = 1$, the lower bound of (27) and the upper bound of (30) are 0 and ∞, respectively.

V. Numerical and Simulation Results

In this section, we present some numerical and simulation results. For all results shown we assume $R = 10$ and $\epsilon = 0.05$.

In Fig. 4, the simulated $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ are plotted as a function of N_t for $N_p = 100$ and 1000. 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 $E_{\mathrm{a}}[\mathrm{I}]$ (27) and the upper bound of $E[\mathrm{SIR}]$ (30) are also depicted. Clearly, the average interference is smaller for random $\mathrm{C_T}$ which is in accordance with our discussion in Section IV-B. Also as expected, increasing N_p increases $E_{\mathrm{a}}[\mathrm{I}]$ and results in a lower $E[\mathrm{SIR}]$. 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 $E[\mathrm{SIR}]$. The results for the upper bound of $E_{\mathrm{a}}[\mathrm{I}]$ and the lower bound of $E[\mathrm{SIR}]$ are not shown because the objective of the beamformer is to minimize the interference and maximize the SIR, cf. (9). In fact, the upper bound of $E_{\mathrm{a}}[\mathrm{I}]$ and the lower bound of $E[\mathrm{SIR}]$ 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. 3, as c increases, the system becomes saturated and the upper and lower bounds of $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ converge to the same value. In other words, choosing a random beamforming vector is as good as using the optimal one obtained from (9) for $c \to \infty$.

In general, the bounds obtained for $E_{\mathrm{a}}[\mathrm{I}]$ and $E[\mathrm{SIR}]$ in (27) and (30) are asymptotic bounds for $N_p/N_t \to c > 0$ as

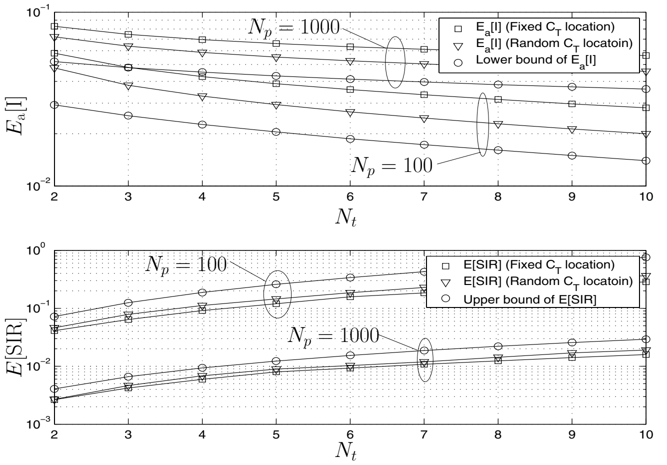

Fig. 4. Simulations of $E[\text{SIR}]$ and $E_{\text{a}}[\text{I}]$ as a function of N_t and different N_p. Fixed C_T (squares). Random C_T (triangles). The lower bound of $E_{\text{a}}[\text{I}]$ (27) and the upper bound of $E[\text{SIR}]$ (30) are also plotted (circles).

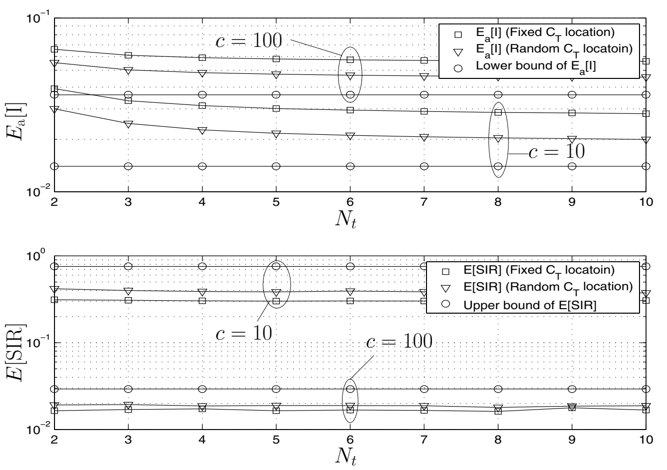

Fig. 5. Simulations of $E[\text{SIR}]$ and $E_{\text{a}}[\text{I}]$ as a function of N_t for different c. Fixed C_T (squares). Random C_T (triangles). The lower bound of $E_{\text{a}}[\text{I}]$ (27) and the upper bound of $E[\text{SIR}]$ (30) are also plotted (circles).

$N_t \to \infty$. Therefore, a natural question to ask is how well these bounds perform in finite region of N_t. This question is answered in Fig. 5. In particular, the simulated $E_{\text{a}}[\text{I}]$ and $E[\text{SIR}]$ 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 C_T. For reference, we have also plotted the lower bound of $E_{\text{a}}[\text{I}]$ (27) and the upper bound of $E[\text{SIR}]$ (30). Note that the bounds are constant for fixed value of c because they depend only on the ratio $c = N_p/N_t$ and not the actual values of N_p and N_t, cf. (26), (27), and (30). The simulation results for $E[\text{SIR}]$ 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 $E_{\text{a}}[\text{I}]$ 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 $N_t \geq 6$, $E_{\text{a}}[\text{I}]$ 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 $E_{\text{a}}[\text{I}]$ depend also only on the ratio c even for small values of N_t, cf. the $E_{\text{a}}[\text{I}]$ curves for $c = 100$.

VI. Conclusion

In this paper, we consider a cognitive network which consists of multiple primary users and multiple cognitive users. The secondary cognitive transmitters are allowed to transmit concurrently with the primary licensed transmitters. To mitigate interference, the secondary users transmit signals using multiple antennas with a beamforming vector. The beamforming vector is designed to maximize the SIR of the secondary user. We derive bounds and provide asymptotic analyses for the average SIR and the average interference caused to all primary receivers. In particular, we have shown that if the number of antennas at the secondary transmitters can be of the same as the number of primary receivers, the interference caused to the primary receivers can be made zero, creating an infinite SIR at the cognitive user. 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 SIR approach fixed limits. These analyses can be useful in deciding the number of antennas to deploy in the secondary transmitters.

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, October 2003.
[2] Spectrum policy task force report. Federal Communications Commision Tech. Rep. 02-155, November 2002.
[3] J. Mitola. *Cognitive Radio*. PhD thesis, Royal Institute of Technology (KTH), 2000.
[4] S. Srinivasa and S. A. Jafar. The throughput potential of cognitive radio: A theoretical perspective. *IEEE Communications Magazine*, 45(5):73–79, May 2007.
[5] B. D. V. Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, pages 4–24, 1988.
[6] G. Jongren, M. Skoglund, and B. Ottersten. Combining beamforming and orthogonal space-time block coding. *IEEE Trans. Inform. Theory*, 48(3):611–627, March 2002.
[7] M. Bengtsson and B. Ottersten. Uplink and downlink beamforming for fading channels. In *Proceedings of Signal Processing Advances in Wireless Communications*, pages 350–353, Annapolis, Maryland, USA, May 1999.
[8] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Math USSR Sbornik*, 1:457–483, 1967.
[9] K. Washizu. On the bounds of eigenvalues. *Quarterly Journal of Mechanics & Applied Mathematics*, 8(3):311–325, 1955.
[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*, 78(4):509–521, August 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*, 21(3):1275–1294, 1993.
[12] A. M. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.
[13] S. Yiu, M. Vu, and V. Tarokh. Interference and noise reduction by beamforming in cognitive networks. In preparation.
[14] N. I. Akhiezer. *The classical moment problem and some related questions in analysis*. Hafner Pub. Co., New York, 1965.
[15] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):175–192, August 1995.
[16] J. W. Silverstein and S.-I. Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):295–309, August 1995.
[17] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *The Annals of Probability*, 26(1):316–345, January 1998.