1 Introduction

In this chapter, we address an important question on the optimal role of multiple antennas in a wireless network. For a point-to-point channel with no interference, from Chapter **??**, we know that employing multiple antennas at both the transmitter and the receiver either linearly increases the capacity, or exponentially decreases the error rate with SNR. In contrast, in a wireless network, where interference is the performance limiter, finding how to best use the multiple antennas is a fairly complicated issue.

The problem is challenging because in the presence of interference, multiple antennas have dual roles at both the transmitter and the receiver side. On the transmitter side, multiple antennas can be used to beamform the signal towards the intended receiver or to suppress transmission (construed as interference) towards other receivers. Similarly, on the receiver side, each receiver can use its multiple antennas to improve the SNR from its intended transmitter or cancel the interference coming from other transmitters. To further compound the problem, the roles of multiple antennas at both the transmitter and the receiver side are inter-dependent on each other.

In this chapter, we derive results on the scaling of the transmission capacity with the number of antennas for two cases; i) CSIR case, where only the receivers have channel coefficient/state information (CSI), and ii) CSIT case, where in addition to CSIR, each transmitter also has CSI for its intended receiver. We derive upper and lower bounds on the transmission capacity with multiple antennas that do not match each other exactly, but have a negligible gap for path-loss exponent values close to 2.

We show that with linear decoders, e.g. zero-forcing or MMSE, the transmission capacity scales at least linearly with the number of antennas for both the CSIR and the CSIT case, and sending only data stream from each transmitter achieves the linear scaling of the transmission capacity in both cases. The derived upper and lower bounds are identical for both the CSIR and the CSIT case, thus, we conclude that the value of CSIT is limited in a wireless network. We obtain exact scaling results for transmission capacity with respect to the number of antennas for two important special cases: having only a single antenna at each transmitter/receiver, and a simplified receiver with no interference cancelation capability.

We close the chapter by characterizing the effect of the interference suppression capability of multiple antennas at the transmitter. For this end, we consider a cognitive/secondary wireless network, that is overlaid over a licensed/primary wireless network, that is allowed to operate under an outage probability constraint at each receiver node of the primary wireless network. The secondary nodes are equipped with multiple antennas, and use them at the transmitter side to suppress the interference they cause to any primary user, and at the receiver side to cancel dominant interferers. We obtain explicit results on the scaling of the transmission capacity of the secondary wireless network as a function of the number of transmit and receive antennas available at the secondary nodes.

2 Role of Multiple Antennas in Ad Hoc Networks

In a point-to-point channel with no interference, the only objective with multiple antennas at both the transmitter and the receiver is to improve the received signal strength. In an ad-hoc network, however, the role of multiple antennas is more diverse because of the presence of interference. For example, each transmitter could attempt to increase its own data rate by transmitting multiple data streams, or in the presence of CSI, could improve the signal strength by steering its beam towards the direction of the receiver, or suppress its interference towards other receivers by nulling its signal towards them. Similarly, each receiver could decode its signal of interest after mitigating interference using its multiple antennas. So inter-dependent questions like, how many data streams to transmit, how many interferers to cancel at receiver, are critical in finding the optimal role of multiple antennas in a wireless network.

One way to frame these questions more concretely is by defining the spatial transmit (receive) degrees of freedom (STDOF) (SRDOF). The STDOF refer to the signaling dimensions used at the transmitter for either transmitting to the intended receiver or to suppress interference towards other receivers. For example, with N transmit antennas there are total N STDOF, out of which possibly k can be used to send k independent streams, leaving the remaining $N-k$ STDOF for interference suppression in the presence of CSI at the transmitter, or not using the $N-k$ STDOF at all to decrease the overall interference at all other receivers. Similarly, the SRDOF refers to the number of spatial dimensions, that through linear processing (linear decoder/receiver), can be used to separate multiple source symbols at the receiver. For example, with N antennas at the receiver, the total SRDOF is equal to N, out of which m can be used for interference mitigation/cancelation and leaving the remainder of $N-m$ SRDOF for decoding the signal of interest.

When k STDOF are used by each transmitter to send k independent data streams, the number of interferers that can be canceled at any receiver using its m SRDOF is at most $\left\lfloor \frac{m}{k} \right\rfloor$. Larger STDOFs help in improving per-user transmission rates by sending more data streams but limit the interference suppression ability of any receiver. In this chapter, we find the optimal values of STDOF used for transmission k and SRDOF m used for interference cancelation that maximize the transmission capacity with linear decoders under different CSI assumptions. The choice of linear decoders is made for both their analytical tractability and low-complexity implementation.

3 Channel State Information Only at Receiver

We consider the fixed distance model of Section **??**, where each transmitter-receiver pair is at a fixed distance of d from each other. The transmitter locations $\{T_n\}$ are assumed to follow a PPP distribution with density λ_0. Each transmitter is assumed to transmit independently with probability p using an ALOHA protocol. Consequently, the active transmitter density is $\lambda = p\lambda_0$. We let $\Phi = \{T_n : T_n \text{ is active}\}$ to represent the active transmitter locations that is a PPP with density λ.

In this section, we consider the practically efficient model where each receiver has instantaneous channel state information (CSI), while no transmitter has any instanta-

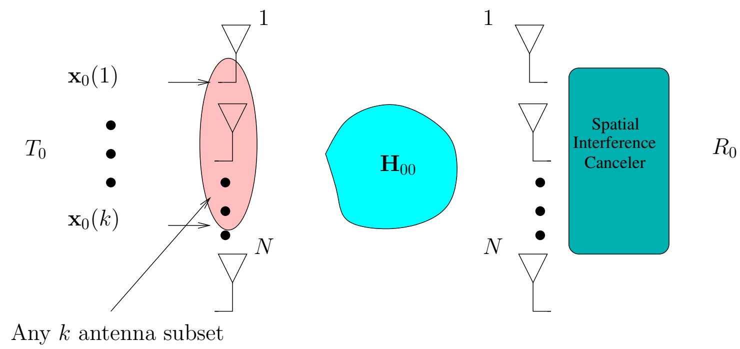

Figure 1: Transmit-receive strategy with no CSI at the transmitter

neous or delayed CSI. We refer to this scenario as CSIR (CSI at the receiver). The case with CSI at the transmitter (referred to as CSIT) is dealt in Section 4.

With no CSI at any transmitter, we assume that each transmitter uses any k, $k = 1, 2, \ldots, N$, of its N antennas to transmit k independent data streams to its receiver with equally distributing the power over all the k antennas. In terms of STDOF, this means that each transmitter uses k STDOF for transmission out of its total N STDOF.

Since CSI is not available, the choice of which antennas to use does not impact the performance. Each receiver is assumed to have CSI for the channel from its intended transmitter as well as from all the other interferers that are canceled/suppressed at that receiver.

Let $\mathbf{x}_n = [\mathbf{x}_n(1)\ \mathbf{x}_n(2) \ldots \mathbf{x}_n(k)]^T$ be the $k \times 1$ signal sent from transmitter T_n, where each element $\mathbf{x}_n(\ell),\ \ell = 1, 2, \ldots, k$ is independent and $\mathcal{CN}\left(0, \frac{1}{k}\right)$ distributed, so that the total power transmitted through $\mathbf{x}_n$ is unity. Then, the multiple antennas counterpart of received signal (**??**) at the typical receiver R_0 is

$$\mathbf{y}_0 = d^{-\alpha/2}\mathbf{H}_{00}\mathbf{x}_0 + \sum_{T_n \in \Phi \backslash \{T_0\}} d_n^{-\alpha/2}\mathbf{H}_{0n}\mathbf{x}_n, \tag{1}$$

where d_n is the distance between T_n and R_0, $\mathbf{H}_{0n} \in \mathbb{C}^{N \times k}$ is the channel coefficient matrix between T_n and R_0, such that the i, j^{th} entry $\mathbf{H}_{0n}(i, j)$ of $\mathbf{H}_{0n}$ is the channel coefficient between the i^{th} receive antenna of R_0 and j^{th} transmit antennas of T_n. Each entry of $\mathbf{H}_{0n}$ is assumed to be independent and Rayleigh distributed. We consider the interference limited regime and ignore the AWGN contribution. For analysis, we will consider the typical transmitter-receiver pair (T_0, R_0).

Interference cancelation: To cancel interference, each receiver multiplies its received signal with vector $\mathbf{q}^\dagger$ that lies in the null space of the channel matrices corresponding to the interferers that are chosen for cancelation. Thus, if $\mathcal{C} \in \Phi$ is the subset of interferers to be canceled, then $\mathbf{q} \in \mathsf{O}(\mathcal{H}_\mathcal{C})$, where $\mathsf{O}(\mathcal{H}_\mathcal{C})$ represents the null/orthogonal space of matrix $\mathcal{H}_\mathcal{C} = [\mathbf{H}_{0n}], n \in \mathcal{C}$.

Which interferers to cancel: Each receiver R_n with multiple antennas has to make a judicious choice of which interferers it should cancel before decoding its signal of interest $\mathbf{x}_n$. The most natural choice it to cancel those interferers that maximize the post-cancelation SIR, i.e. to find subset $\mathcal{C}$ that solves

$$\max_{\mathcal{C}} \mathsf{SIR} = \max_{\mathcal{C}} \frac{d^{-\alpha}\mathbf{q}\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}\mathbf{q}^{\dagger}}{\sum_{T_n\in\Phi\backslash\{T_0,\mathcal{C}\}} d_n^{-\alpha}\mathbf{q}\mathbf{H}_{0n}\mathbf{H}_{0n}^{\dagger}\mathbf{q}^{\dagger}}. \tag{2}$$

Solving (2), is however, complicated and also the performance analysis is difficult. As we have seen in Chapter **??**, typically, the closest interferers dominate the total interference seen at any receiver. This motivates the choice of canceling the nearest interferers in terms of distance from each receiver R_n. Canceling the nearest interferers is also efficient in terms of CSI requirement, since CSI from only the nearest interferers to be canceled is required, in comparison to the global CSI requirement for solving (2). Since any receiver with N antennas can cancel at most N interferers, CSI is only needed from at most N nearest interferers. Throughout this chapter, for analyzing the transmission capacity with multiple antennas, we assume that each receiver cancels its nearest interferers.

Another choice for interference cancelation is to cancel those interferers that have the largest interference power at the receiver. With this choice, some of the nearby interferers may not be canceled if their channel gains are very low. After multiplication by the cancelation vector $\mathbf{q}^{\dagger}$, however, the situation might change, and the post-cancelation channel gain values of the nearby uncanceled interferers could become moderately high, and they could start dominating the performance. We discuss this choice briefly in Remark 3.11 from the transmission capacity point of view.

Choice of Decoder: To decode $\mathbf{x}_0$ from (1), the optimal decoder is the ML decoder, that finds $\mathbf{x}_0$ that maximizes the likelihood $\mathbb{P}(\mathbf{y}_0|\mathbf{x}_0)$. As discussed in Chapter **??**, the complexity of the ML decoder is quite high since it finds the jointly optimal vector $\mathbf{x}_0$. Moreover, for the transmission capacity analysis with the ML decoder, we need to be able to analyze the outage probability $\mathbb{P}(I(\mathbf{x}_0;\mathbf{y}_0) < \beta)$, where $I(\mathbf{x}_0;\mathbf{y}_0)$ is the mutual information between input $\mathbf{x}_0$ and output $\mathbf{y}_0$. The exponent of outage probability $\mathbb{P}(I(\mathbf{x}_0;\mathbf{y}_0) < \beta)$ for the MIMO channel is only known for the high SNR regime [1] and that too in the absence of interference. Thus, in the presence of interference, meaningful analysis of transmission capacity is not possible with the optimal ML decoder. The obvious other choice is to consider linear decoders, such as ZF or minimum mean square error (MMSE) decoder. As discussed in Section **??**, with linear decoders, each element of the input signal vector $\mathbf{x}_0$ is decoded separately allowing the use of scalar outage probability expressions, while incurring linear decoding complexity in the size of vector $\mathbf{x}_0$. For detailed analysis purposes, we will consider the ZF decoder, and point out that identical results can be obtained for MMSE decoder as well in Remark 3.7. In particular, throughout this chapter, we consider a general ZF decoder called the *partial ZF decoder*, that allows the flexibility of choosing how many SRDOF to use for interference cancelation and leaving the remaining SRDOF for decoding the signal of interest.

3.1 Transmission Capacity With Partial ZF Decoder

With k data streams sent from each transmitter, and each receiver using m SRDOF for interference cancelation, let $N_{\mathsf{canc}} = \left\lfloor \frac{m}{k} \right\rfloor$ be the number of nearest canceled interferers. To cancel the nearest interferers, let the indices of the interferers be sorted in an increasing order in terms of their distance from the typical receiver R_0, i.e. $d_1 \le d_2 \le \ldots \le d_{N_{\mathsf{canc}}} \le d_{N_{\mathsf{canc}}+1} \le \ldots$. Then the received signal (1) is

$$\begin{aligned} \mathbf{y}_0 &= d^{-\alpha/2}\mathbf{H}_{00}(\ell)\mathbf{x}_0(\ell) + \sum_{j=1, j\neq \ell}^{k} d^{-\alpha/2}\mathbf{H}_{00}(j)\mathbf{x}_0(j) \\ &\quad + \sum_{n=1}^{\infty} d_n^{-\alpha/2} \sum_{j=1}^{k} \mathbf{H}_{0n}(j)\mathbf{x}_n(j), \qquad (3) \end{aligned}$$

where we have intentionally separated the data stream $\mathbf{x}_0(\ell)$ and the rest of the data streams $\mathbf{x}_0(1), \ldots, \mathbf{x}_0(\ell-1),\ \mathbf{x}_0(\ell+1), \ldots, \mathbf{x}_0(k)$, sent by the typical transmitter T_0.

To decode the $\mathbf{x}_0(\ell)^{th}$ data stream sent from transmitter T_0, $\ell = 1, 2, \ldots, k$, receiver R_0 uses partial ZF decoder to remove the inter-stream interference from all the other data streams

$$\mathbf{x}_0(1), \ldots, \mathbf{x}_0(\ell-1),\ \mathbf{x}_0(\ell+1), \ldots, \mathbf{x}_0(k)$$

sent by transmitter T_0, and all the k data streams transmitted by the N_{canc} nearest interferers $\mathbf{x}_n(j),\ n = 1, 2, \ldots, N_{\mathsf{canc}},\ j = 1, 2, \ldots, k$.

Let

$$\mathcal{H} = \left[\mathbf{H}_{00}(1) \ldots \mathbf{H}_{00}(\ell-1)\ \mathbf{H}_{00}(\ell+1) \ldots \mathbf{H}_{00}(k)\ \mathbf{H}_{01}\ \mathbf{H}_{02} \ldots \mathbf{H}_{0N_{\mathsf{canc}}}\right],$$

where $\mathcal{H} \in \mathbb{C}^{N\times m+k-1}$ be the channel matrix corresponding to the $k-1$ inter-stream interferers, and the N_{canc} nearest interferers in (3), where $\mathbf{H}_{0n} \in \mathbb{C}^{N\times k}$ is the channel matrix corresponding to the n^{th} nearest interferer, $N > m+k-1$. Since each channel coefficient is i.i.d. Rayleigh distributed, the rank of matrix $\mathcal{H}$ is $m+k-1$ with probability 1. Let S be the orthonormal basis of the null space $\mathsf{O}(\mathcal{H})$ of the matrix $\mathcal{H}$, where S has dimension $N-(m+k-1)$. To decode stream $\mathbf{x}_0(\ell)$, the receiver R_0 multiplies $\mathbf{q}_\ell^\dagger,\ \mathbf{q}_\ell^\dagger \in \mathsf{O}(\mathcal{H})$ to the received signal (3) to get

$$\mathbf{q}_\ell^\dagger \mathbf{y}_0 = d^{-\alpha/2}\mathbf{q}_\ell^\dagger \mathbf{H}_{00}(\ell)\mathbf{x}_0(\ell) + \sum_{n=N_{\mathsf{canc}}+1}^{\infty} d_n^{-\alpha/2} \sum_{j=1}^{k} \mathbf{q}_\ell^\dagger \mathbf{H}_{0n}(j)\mathbf{x}_n(j), \qquad (4)$$

$\ell = 1, 2, \ldots, k$. Similar to the choice of which interferers to cancel, there is choice for selecting the interference cancelation vector $\mathbf{q}_\ell^\dagger$. The obvious choice is the one that maximizes the SIR, however, that leads to analytic intractability.

So we consider the next best option of choosing $\mathbf{q}_\ell^\dagger \in \mathsf{O}(\mathcal{H})$ that maximizes the signal power $s = |\mathbf{q}_\ell^\dagger \mathbf{H}_{00}(\ell)|^2$. In Lemma 3.1, we show that the optimal $\mathbf{q}_\ell$ that maximizes the signal power s is given by

$$\mathbf{q}_\ell = \frac{\mathbf{H}_{00}(\ell)^\dagger \mathsf{S}\mathsf{S}^\dagger}{|\mathbf{H}_{00}(\ell)^\dagger \mathsf{S}\mathsf{S}^\dagger|},$$

and the signal power $s = |\mathbf{q}_\ell^\dagger \mathbf{H}_{00}(\ell)|^2$ is $\chi^2(2(N-m-k+1))$, since the dimension of S is $N-k-m+1$. Moreover, since $\mathbf{q}_\ell^\dagger$ is chosen to maximize the signal power

$s = |\mathbf{q}_\ell^\dagger \mathbf{H}_{00}(\ell)|^2$, it does not depend on the uncanceled interferer's channels $\mathbf{H}_{0n}$ for $n \ge N_{\mathsf{canc}} + 1$ in (3). Hence the power of the j^{th} stream of the n^{th} interferer, $n \ge N_{\mathsf{canc}} + 1$, $|\mathbf{q}_\ell^\dagger \mathbf{H}_{0n}(j)|^2$ is $\chi^2(2)$, since each entry of $\mathbf{H}_{0n}$ is independent and Rayleigh distributed. Adding the contribution from k independent data streams of each interferer, the total interference power of the n^{th} uncanceled interferer from its k data streams in (3) is $\mathsf{pow}_n = \sum_{j=1}^{k} |\mathbf{q}_\ell^\dagger \mathbf{H}_{0n}(j)|^2$ that is $\chi^2(2k)$ distributed.

Lemma 3.1 *Let $\mathbf{Q} \in \mathbb{C}^{N\times \ell}$, $\mathbf{Q}^\dagger\mathbf{Q} = \mathbf{I}$. Then*

$$\arg \max_{v \in \mathbf{Q}, |v|^2 = 1} |\mathbf{v}^\dagger \mathbf{h}_0|^2 = |\mathbf{Q}^\dagger \mathbf{h}_0|^2,$$

and $|\mathbf{Q}^\dagger \mathbf{h}_0|^2$ is $\chi^2(2\ell)$ if $\mathbf{h}_0 \in \mathbb{C}^{N\times 1}$ is complex Gaussian distributed with independent entries that have zero mean and unit variance.

Proof: Without loss of generality, let $\mathbf{v} = \frac{\mathbf{Q}x}{|\mathbf{Q}x|}$. From Cauchy-Schwarz inequality,

$$< \mathbf{h}_0^\dagger, \mathbf{Q}x > = < \mathbf{h}_0^\dagger \mathbf{Q}, x > \le |\mathbf{h}_0^\dagger \mathbf{Q}|^2,$$

and the maximum is achieved by $x = \mathbf{Q}^\dagger \mathbf{h}_0$. Thus, we get that

$$\max_{v \in \mathbf{Q}, |\mathbf{v}|^2 = 1} |\mathbf{v}^\dagger \mathbf{h}_0|^2 = |\mathbf{h}_0^\dagger \mathbf{Q}\mathbf{Q}^\dagger \mathbf{h}_0|,$$

which is the norm of vector $\mathbf{h}_0^\dagger \mathbf{Q}$.

Since the columns of $\mathbf{Q}$ are orthonormal, the covariance matrix of vector $\mathbf{h}_0^\dagger \mathbf{Q}$ of length ℓ is diagonal, where the expectation is with respect to entries of $\mathbf{h}_0$. Thus, the elements of $\mathbf{h}_0^\dagger \mathbf{Q}$ that are Gaussian distributed are uncorrelated, and hence are independent. Since the norm of a ℓ-length independent Gaussian vector is $\chi^2(2\ell)$ distributed, the result follows. More details can be found in [2]. □

Lemma 3.2 *Let $\mathsf{x}_1, \dots, \mathsf{x}_m$ be a set of m n-length Gaussian vectors, whose each element is independent with zero mean and unit variance. If vector y of length n is independent of $\mathsf{x}_1 \dots \mathsf{x}_m$, then $|\mathsf{y}^\dagger \mathsf{x}_i|^2 \sim \chi^2(2)$, $\forall\ i$ and $\sum_{i=1}^{m} |\mathsf{y}^\dagger \mathsf{x}_i|^2 \sim \chi^2(2m)$.*

Definition 3.3 *With partial ZF decoder, from (4), the SIR for the ℓ^{th} stream is given by*

$$\mathsf{SIR}_\ell = \frac{d^{-\alpha} s}{\sum_{n=N_{\mathsf{canc}}+1}^{\infty} d_n^{-\alpha} \mathsf{pow}_n}, \tag{5}$$

where from Lemma 3.1, signal power $s = |\mathbf{q}_\ell^\dagger \mathbf{H}_{00}(\ell)|^2 \sim \chi^2(2(N - k - m + 1))$, and from Lemma 3.2, interference power $\mathsf{pow}_n = \sum_{j=1}^{k} |\mathbf{q}_\ell^\dagger \mathbf{H}_{0n}(j)|^2 \sim \chi^2(2k)$.

Note that the same decoding strategy is used for each stream $\ell = 1, 2, \dots, k$ sent by any transmitter T_n, therefore the SIR for each stream ℓ is identically distributed. Henceforth we drop the subscript ℓ from SIR_ℓ, and represent it as SIR. Thus, for each stream ℓ, $\ell = 1, 2, \dots, k$, the outage probability at rate B bits/sec/Hz is given by

$$\begin{aligned} P_{out}(B) &= \mathbb{P}\left(\log(1 + \mathsf{SIR}) \le B\right), \\ &= \mathbb{P}\left(\mathsf{SIR} \le 2^B - 1\right). \end{aligned} \tag{6}$$

Let $2^B - 1 = \beta$. Since k streams are transmitted simultaneously, the transmission capacity is defined as

$$C = k\lambda(1-\epsilon)B \text{ bits/sec/Hz/m}^2, \tag{7}$$

where ϵ is the outage probability constraint for each data stream, and λ is the maximum density of nodes such that $P_{out}(B) \le \epsilon$ in (6). Here C represents the average successful rate of information transfer across the network when each transmitter sends k independent data streams.

From (5, 6),

$$P_{out}(B) \quad = \quad \mathbb{P}\left(\frac{d^{-\alpha}s}{I_{nc}} \le \beta\right), \tag{8}$$

where $I_{nc} = \sum_{n=N_{\mathsf{canc}}+1}^{\infty} d_n^{-\alpha}\mathsf{pow}_n$ is the total interference from uncanceled interferers, $d_n \le d_m,\ n < m$, where d_n's are ordered in increasing distance from the receiver R_0.

Remark 3.4 *For the case of $k = 1$ (single data stream transmission), the outage probability expression* (6) *and consequently the transmission capacity expression* (7) *is also valid for the ML decoder. Thus, the performance of ML decoder and ZF decoder is identical for $k = 1$. We will show in Theorem 3.5, that the optimal $k^\star = 1$ with the ZF decoder and the transmission capacity scales at least linearly with N. Thus, the same conclusion holds true for the ML decoder. What is left open is the fact whether $k^\star = 1$ for ML decoder or not? The simulation results (Fig. 3) point out that $k^\star = 1$ even for the ML decoder.*

To find the transmission capacity expression (7), and to maximize that with respect to the number of transmitted data streams k, and the number of SRDOF m used to cancel the nearest interferers, we need to find a closed form expression for the outage probability (8). Unfortunately, that is hard to find since the distribution of I_{nc} is unknown. We thus rely on deriving upper and lower bounds on the outage probability that allows us to find the optimal values of k and m that maximize the transmission capacity.

The main result of this section is as follows that is derived from [3] and [4].

Theorem 3.5 *The transmission capacity with multiple antennas and partial ZF decoder receiver scales as*

$$C = \Omega(N) \text{ and } C = \mathcal{O}(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}}),$$

with the number of antennas N. The optimal lower bound is achieved by sending a single data stream from each transmitter, $k^\star = 1$, and using $m^\star = \left(1 - \frac{2}{\alpha}\right) N$ SRDOF for interference cancelation to cancel the $\left(1 - \frac{2}{\alpha}\right) N$ nearest interferers.

Theorem 3.5 tells us that similar to point-to-point channels without interference, the transmission capacity of wireless network scales at least linearly with the number of antennas N. The derived upper bound does not match with the lower bound, however, the gap is negligible for path-loss exponents α close to 2 and the maximum gap is

$N^{1/4}$ at $\alpha = 4$ for any $2 < \alpha \leq 4$. The simulation results, however, suggest that this gap is only a manifest of the proof technique, and not the underlying principle and transmission capacity cannot scale faster than order N. Using simulations, we show that the transmission capacity can at best scale linearly with N by checking for many different combinations of k and m. Thus, a more finer analysis is required for exactly characterising the scaling of the transmission capacity with multiple antennas. For the special case, when each transmitter has a single antenna or chooses $k = 1$, we can obtain the exact results for the scaling of transmission capacity with respect to number of antennas N in Theorem 3.6.

A more important conclusion from Theorem 3.5 is that to maximize the lower bound, i.e. to achieve linear scaling with N, $k^\star = 1$, and $m^\star \propto N$, i.e. one should only transmit a single data stream from each transmitter, while linearly scaling the SRDOF dedicated for interference cancelation. To interpret this result, note that the number of transmitted data streams are directly related to the interference power, and the number of interferers that can be canceled by each receiver. Thus, in an interference limited network such as a wireless network, minimizing the number of data streams transmitted by each node keeps the interference power low and at the same time leaves enough room for canceling significant number of nearest interferers at the receiver. Moreover, by scaling the SRDOF used for interference cancelation linearly with N, the number of nearest interferers that can be canceled scales linearly with N, while leaving sufficient SRDOF (that also scales linearly in N) for decoding the signal of interest.

For the special case when each transmitter is equipped with a single transmit antenna, or chooses $k = 1$ irrespective of N, the proof of Theorem 3.5 can be used to obtain the exact scaling of transmission capacity with the number of receiver antennas, as follows.

Theorem 3.6 *With a single antenna at each transmitter or fixed $k = 1$, the transmission capacity with partial ZF decoder scales linearly with the number of receive antennas N, i.e.*

$$C = \Theta(N),$$

and the optimal SRDOF for interference cancelation is $m^\star = \left(1 - \frac{2}{\alpha}\right) N$.

Theorem (3.6) shows that linear scaling of transmission capacity is possible even if each transmitter has a single antenna. Thus, in a wireless network, the role of multiple antennas at the receiver side is more important than the transmit side. This result is in contrast to a point-to-point channel where the capacity scales linearly with the minimum of the transmit and the receive antennas.

Remark 3.7 *An alternate choice of linear decoder is the MMSE decoder, where to decode the data stream $\mathbf{x}_0(\ell)$ from received signal* (1)*, $\Sigma_\ell^{-1}\mathbf{H}_{00}(\ell)$ is multiplied to the received signal, where $\mathbf{H}_{00}(\ell)$ is the ℓ^{th} column of $\mathbf{H}_{00}$, and*

$$\Sigma_\ell = \sum_{i=1, i\neq \ell}^{k} d^{-\alpha}\mathbf{H}_{00}(i)\mathbf{H}_{00}(i)^\dagger + \sum_{T_n \in \Phi\backslash\{T_0\}} d_n^{-\alpha/2}\mathbf{H}_{0n}\mathbf{H}_{0n}^\dagger$$

is the spatial covariance matrix of the inter-stream interference and interference caused by other transmitters. The MMSE decoder is known to maximize the received SINR, *and hence the lower bound derived for the transmission capacity in Theorem 3.5 with partial ZF decoder also holds for the MMSE case as well.*

To upper bound the transmission capacity with MMSE decoder, we can let the signal power with the MMSE decoder to be distributed as $\chi^2(2N)$, which is clearly an idealization since after canceling interferers by multiplying $\Sigma_\ell^{-1}\mathbf{H}_{00}(\ell)$, the signal power is $\mathbf{H}_{00}^\dagger(\ell)\Sigma_\ell^{-1}\mathbf{H}_{00}(\ell)$ which is less than the norm of vector $\mathbf{H}_{00}(\ell)$ that is distributed as $\chi^2(2N)$. Moreover, by selecting $r = N$ (where r is the number of uncanceled nearest interferers used for lower bounding the outage probability in Theorem 3.5) an upper bound identical to Theorem 3.5 on the transmission capacity can also be found for the MMSE decoder as well [5]. Thus, results obtained for the ZF decoder also hold for the MMSE decoder as well.

Towards proving Theorem 3.5, we first upper and lower bound the outage probability (8) as follows.

Theorem 3.8 *The outage probability* (8) *when the transmitter sends k independent data streams, and the receiver cancels the N_{canc} nearest interferers using the partial ZF decoder, is lower bounded by*

$$P_{out}(B) \quad \geq \quad 1 - \frac{N-m-k+1}{(kr-1)d^\alpha\beta\left(\pi\lambda\right)^{\frac{\alpha}{2}}}\left(N_{\mathsf{canc}} + r + \frac{\alpha}{2}\right)^{\frac{\alpha}{2}},$$

for any $r \in \mathbb{N}^+$ such that $kr > 1$.

To derive this lower bound, we consider the interference contribution from only the r nearest uncanceled interferer (the $N_{\mathsf{canc}} + 1^{st}$ interferer to $N_{\mathsf{canc}} + r^{th}$ interferer) and consider their aggregate interference

$$I_{nc}^r = \sum_{j=1}^{r} d_{N_{\mathsf{canc}}+j}^{-\alpha}\mathsf{pow}_{N_{\mathsf{canc}}+j}. \tag{9}$$

Since $I_{nc}^r < I_{nc}$, from (8), we have

$$P_{out}(B) = \mathbb{P}\left(\frac{d^{-\alpha}s}{I_{nc}} \leq \beta\right) \geq \mathbb{P}\left(\frac{d^{-\alpha}s}{I_{nc}^r} \leq \beta\right).$$

For any r, we can efficiently bound the outage probability $\mathbb{P}\left(\frac{d^{-\alpha}s}{I_{nc}^r} \leq \beta\right)$ using the Markov's inequality as follows.

Proof: Consider the interference contribution from only the r nearest uncanceled interferers, I_{nc}^r. Fig. 2 illustrates this scenario, where the N_{canc} (squares) have been canceled, and only the interference coming from the r nearest uncanceled neighbors of receiver R_0 are considered towards computing the outage probability. To derive the lower bound, we use the Markov's inequality with s (signal power) as the random variable and compute the expectation with respect to the interference power I_{nc}^r. From (8),

$$\begin{aligned} 1 - P_{out}(B) &= \mathbb{P}\left(s > d^\alpha\beta I_{nc}\right) \\ &\leq \mathbb{P}\left(s > d^\alpha\beta I_{nc}^r\right), \text{ since } I_{nc}^r \leq I_{nc} \text{ from (9)}, \\ &\leq \mathbb{E}\left\{\frac{\mathbb{E}\left\{s\right\}}{d^\alpha\beta I_{nc}^r}\right\}, \quad \text{from Markov's inequality.} \end{aligned} \tag{10}$$

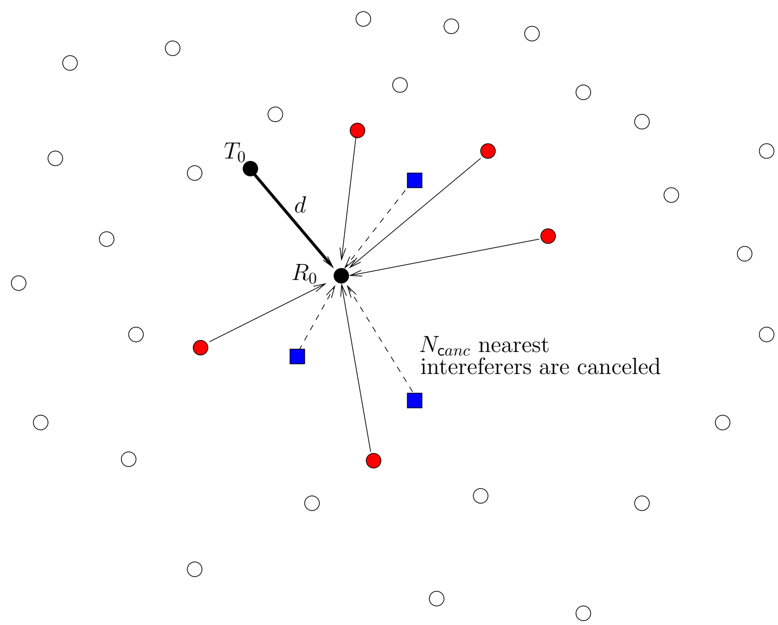

Figure 2: Squares represent the N_{canc} nearest canceled interferers with dashed lines, colored circles represent the r nearest uncanceled interferers whose interference contribution will be used to derive the lower bound on the outage probability, and uncolored circles are all the other uncanceled interferers.

Since interferers are ordered in increasing distance from the receiver R_0, $d_{N_{\mathsf{canc}}+j} \le d_{N_{\mathsf{canc}}+r},\ j = 1, \ldots, r-1$, we have $I_{nc}^f \ge I_{nc}^r$, where

$$I_{nc}^f = d_{N_{\mathsf{canc}}+r}^{-\alpha} \sum_{j=1}^{r} \mathsf{pow}_{N_{\mathsf{canc}}+j}, \tag{11}$$

is obtained by substituting for each of the path-loss term $d_{N_{\mathsf{canc}}+j}^{-\alpha}$ in I_{nc}^r by the path-loss term of the farthest interferer $d_{N_{\mathsf{canc}}+r}^{-\alpha}$. Hence, from (10),

$$1 - P_{out}(B) \le \mathbb{E}\left\{\frac{\mathbb{E}\{s\}}{d^\alpha \beta I_{nc}^f}\right\}. \tag{12}$$

There are three random variables involved in the analysis : signal power s, channel power of uncanceled interferers $\mathsf{pow}_{N_{\mathsf{canc}}+j}$'s, and the distance of the farthest uncanceled interferers $d_{N_{\mathsf{canc}}+r}$ considered for deriving the bound. From Definition 3.3, $s \sim \chi^2(2(N-m-k+1))$, and $\mathsf{pow}_{N_{\mathsf{canc}}+j}$'s are independent and $\chi^2(2k)$ distributed, hence their sum $\mathsf{pow} = \sum_{j=1}^r \mathsf{pow}_{N_{\mathsf{canc}}+j} \sim \chi^2(2kr)$. Moreover, from Lemma 3.9, we have $\pi\lambda d_{N_{\mathsf{canc}}+r}^2 \sim \chi^2(2(N_{\mathsf{canc}}+r))$.

Lemma 3.9 *Let d_n be the distance of the n^{th} nearest node of a PPP Φ with density λ from the origin. Then $\pi\lambda d_n^2 \sim \chi^2(2n)$.*

Proof: The result follows from the direct computation of the distribution of $\pi\lambda d_n^2$ by finding the distribution $\mathbb{P}(d_n > r)$ using the void probability of PPP Φ with density λ. □

Hence from (12), we have

$$\begin{aligned} 1 - P_{out}(B) &\stackrel{(a)}{\le} \frac{\mathbb{E}\{s\}}{d^\alpha\beta}\mathbb{E}\left\{\frac{1}{d_{N_{\mathsf{canc}}+r}^{-\alpha}}\right\}\mathbb{E}\left\{\frac{1}{\mathsf{pow}}\right\}, \\ &\stackrel{(b)}{=} \frac{N-k-m+1}{d^\alpha\beta\,(\pi\lambda)^{\frac{\alpha}{2}}} \int_0^\infty \frac{x^{N_{\mathsf{canc}}+r+\frac{\alpha}{2}}e^{-x}}{N_{\mathsf{canc}}+r!}dx \int_0^\infty \frac{\mathsf{pow}^{kr-2}e^{-\mathsf{pow}}}{kr-1!}d\mathsf{pow}, \\ &\le \frac{N-k-m+1}{d^\alpha\beta\,(\pi\lambda)^{\frac{\alpha}{2}}} \frac{\Gamma\left(N_{\mathsf{canc}}+r+1+\frac{\alpha}{2}\right)}{\Gamma\left(N_{\mathsf{canc}}+r+1\right)}\left(\frac{1}{kr-1}\right), \end{aligned} \tag{13}$$

where in (a) we have substituted for I_{nc}^f from (11), and (b) follows since $s \sim \chi^2(2(N-m-k+1))$, $\mathsf{pow} = \sum_{j=1}^r \mathsf{pow}_{N_{\mathsf{canc}}+j} \sim \chi^2(2kr)$, and $\pi\lambda d_{N_{\mathsf{canc}}+r}^2 \sim \chi^2(2r)$.

From Kershaw's inequality [6], that states that $\frac{\Gamma(x+1)}{\Gamma(x+s)} \le \left(x - \frac{1}{2} + \sqrt{s+\frac{1}{4}}\right)^{1-s}$ for $x > 0, 0 < s < 1$, since $\Gamma(x+1) = x\Gamma(x)$ we have

$$\frac{\Gamma\left(N_{\mathsf{canc}}+r+1+\frac{\alpha}{2}\right)}{\Gamma\left(N_{\mathsf{canc}}+r+1\right)} \le \left(N_{\mathsf{canc}}+r+1+\frac{\alpha}{2}\right)^{\frac{\alpha}{2}}.$$

Hence from (13), we have,

$$1 - P_{out}(B) \le \frac{N-k-m+1}{(kr-1)d^\alpha\beta\,(\pi\lambda)^{\frac{\alpha}{2}}}\left(N_{\mathsf{canc}}+r+\frac{\alpha}{2}\right)^{\frac{\alpha}{2}}.$$

□

Next, we derive an upper bound on the outage probability (8) using the Markov's inequality with I_{nc} as the random variable. For that purpose, we compute an upper bound on $\mathbb{E}\{I_{nc}\}$ using the Campbell's Theorem (Theorem **??**).

Theorem 3.10 *When each transmitter sends k independent data streams, and the receiver uses m SRDOF for canceling the N_{canc} nearest interferers using the partial ZF decoder, the outage probability* (8) *is upper bounded by*

$$P_{out}(B) \quad \leq \quad \begin{cases} \frac{2k\beta(\pi\lambda)^{\frac{\alpha}{2}}}{d^{-\alpha}(N-k-m)}\left(\frac{\alpha}{2}-1\right)^{-1}\left(N_{\mathsf{canc}}-\lceil\frac{\alpha}{2}\rceil\right)^{1-\frac{\alpha}{2}}, & k+m<N, \\ 1-\exp\left(\frac{-d^{\alpha}\beta 2k(\pi\lambda)^2\left(N_{\mathsf{canc}}-\lceil\frac{\alpha}{2}\rceil\right)^{1-\frac{\alpha}{2}}}{\left(\frac{\alpha}{2}-1\right)}\right), & k+m=N. \end{cases}$$

Proof: To upper bound the outage probability $\mathbb{P}\left(\frac{d^{-\alpha}s}{I_{nc}}\leq\beta\right)$ (8), we use Markov's inequality with $I_{nc}=\sum_{n=N_{\mathsf{canc}}+1}^{\infty} d_n^{-\alpha}\mathsf{pow}_n$, as the random variable. To apply Markov's inequality on $\mathbb{P}(I_{nc}\geq\frac{d^{-\alpha}s}{R})$, we need to bound the expected interference from uncanceled interferers $\mathbb{E}\{I_{nc}\}$ as follows. From (8),

$$\begin{aligned} \mathbb{E}\{I_{nc}\} &= \mathbb{E}\left\{\sum_{j\geq N_{\mathsf{canc}}+1} d_j^{-\alpha}\mathsf{pow}_j\right\}, \\ &\stackrel{(a)}{=} 2k\mathbb{E}\left\{\sum_{j\geq N_{\mathsf{canc}}+1} d_j^{-\alpha}\right\}, \\ &\stackrel{(b)}{=} 2k\sum_{j\geq N_{\mathsf{canc}}+1}(\pi\lambda)^{\frac{\alpha}{2}}\int_0^{\infty} x^{-\alpha/2}\frac{x^{j-1}\exp(-x)}{(j-1)!}dx, \\ &= 2k(\pi\lambda)^{\frac{\alpha}{2}}\sum_{j\geq N_{\mathsf{canc}}+1}\frac{\Gamma\left(j-\frac{\alpha}{2}\right)}{\Gamma(j)}, \end{aligned} \tag{14}$$

where (a) follows since the power of j^{th} interferer $\mathsf{pow}_j \sim \chi^2(2k)$, and (b) follows from Lemma 3.9 where $\pi\lambda d_j^2\sim\chi^2(2j)$. Note that $\Gamma\left(j-\frac{\alpha}{2}\right)$ is finite only for $j>\frac{\alpha}{2}$, thus we at least need to cancel at least $\frac{\alpha}{2}$ nearest interferers. Since, typically, $2<\alpha<4$, this is not much of a restriction. Using the Kershaw's inequality [6], $\frac{\Gamma(j-\frac{\alpha}{2})}{\Gamma(j)}\leq\left(j-\lceil\frac{\alpha}{2}\rceil\right)^{-\frac{\alpha}{2}}$, from (14), we get

$$\begin{aligned} \mathbb{E}\{I_{nc}\} &\leq 2k(\pi\lambda)^{\frac{\alpha}{2}}\sum_{j\geq N_{\mathsf{canc}}+1}\left(j-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{-\frac{\alpha}{2}}, \\ &\stackrel{(d)}{\leq} 2k(\pi\lambda)^{\frac{\alpha}{2}}\int_{N_{\mathsf{canc}}}\left(x-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{-\frac{\alpha}{2}}dx, \\ &= 2k(\pi\lambda)^{\frac{\alpha}{2}}\left(\frac{\alpha}{2}-1\right)^{-1}\left(N_{\mathsf{canc}}-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}, \end{aligned} \tag{15}$$

where (d) follows since $x^{-\alpha/2}$ is a decreasing function.

Using (15), we now derive the required upper bound. From (8), $P_{out}(B) =$

$$\begin{aligned}\mathbb{P}\left(I_{nc} \geq \frac{d^{-\alpha}s}{\beta}\right) &= \mathbb{E}_s\left\{\mathbb{P}\left(I_{nc} \geq \frac{d^{-\alpha}s}{\beta}\right)\right\}, \\ &\leq \frac{\mathbb{E}\{I_{nc}\}\beta}{d^{-\alpha}}\mathbb{E}\left\{\frac{1}{s}\right\}, \text{ from Markov's inequality.}\end{aligned}$$

From Definition 3.3, signal power $s \sim \chi^2(2(N-k-m+1))$. Thus, for $N > k+m$, we have $\mathbb{E}\left\{\frac{1}{s}\right\} = \frac{1}{N-k-m}$. Thus, substituting for the upper bound on the expected interference from (15),

$$\mathbb{P}\left(I_{nc} \geq \frac{d^{-\alpha}s}{\beta}\right) \leq \frac{2\beta k(\pi\lambda)^{\frac{\alpha}{2}}}{d^{-\alpha}(N-k-m)}\left(\frac{\alpha}{2}-1\right)^{-1}\left(N_{\mathsf{canc}} - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}.$$

Since $s \sim \chi^2(2(N-k-m+1))$, with $N = k+m$, s is an exponential random variable with parameter 1, and hence

$$\begin{aligned}P_{out}(B) &= \mathbb{P}(s \leq d^{\alpha}\beta I_{nc}), \\ &= \mathbb{E}\left\{1-\exp\left(-d^{\alpha}\beta I_{nc}\right)\right\}, \\ &= 1-\exp\left(-d^{\alpha}\beta\mathbb{E}\left\{I_{nc}\right\}\right), \qquad (16)\end{aligned}$$

since exp is a convex function. Thus, we get the required bound by plugging in the upper bound on $\mathbb{E}\left\{I_{nc}\right\}$ from (15).

□

Using Theorems 3.8 and 3.10, we prove Theorem 3.5 as follows.

Proof: (Theorem 3.5) Recall that the number of nearest canceled interferers are $N_{\mathsf{canc}} = \left\lfloor\frac{m}{k}\right\rfloor$. Using the definition of transmission capacity $C = (1-\epsilon)\lambda B$, and fixing $P_{out}(B) = \epsilon$, from the lower bound derived on outage probability in Theorem 3.8, for any $r \in \mathbb{N}$ such that $kr > 1$,

$$C \leq \frac{kR(1-\epsilon)^{1-\frac{2}{\alpha}}}{\pi}\left(\frac{N-m-k+1}{(kr-1)d^{\alpha}\beta}\right)^{\frac{2}{\alpha}}\left(\left\lfloor\frac{m}{k}\right\rfloor + r + \frac{\alpha}{2}\right). \qquad (17)$$

In terms of scaling with N, the upper bound is increasing in m, the SRDOF used for interference cancelation, as long as the total SRDOF do not exceed N, i.e., $k+m < N$. Thus we fix $m = \Theta(N)$, the largest scaling factor with respect to N. Let the number of data streams to transmit $k = \Theta(N^{\kappa})$. We will find the tightest upper bound as a function of κ. Recall that we can choose the parameter r, the number of uncanceled interferers whose aggregate interference we accounted for while deriving the upper bound in Theorem 3.10. We let $r = N^{2/\alpha}$.[1] Then, for $1-\kappa < \frac{\alpha}{2}$, the upper bound (17) is $\mathcal{O}\left(N^{\kappa\left(1-\frac{2}{\alpha}\right)+\frac{4}{\alpha}-\frac{4}{\alpha^2}}\right)$, for which the optimal $\kappa = 1$, and yields

$$C = \mathcal{O}\left(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}}\right).$$

[1]One can can take ceil or floor function if $N^{2/\alpha}$ is not an integer

For the other case of $1-\kappa \geq \frac{\alpha}{2}$, the upper bound (17) is $\mathcal{O}\left(N^{1-\frac{2}{\alpha}\left(1-\kappa-\frac{2}{\alpha}\right)}\right)$ and the optimal $\kappa = 0$ and upper bound is

$$C = \mathcal{O}\left(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}}\right).$$

Thus, for any κ, we get

$$C = \mathcal{O}\left(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}}\right).$$

Moving on to the lower bound, from Theorem 3.10, we consider the case of $k+m < N$ which provides a better lower bound than $k+m = N$. Equating $P_{out}(B) = \epsilon$, we have

$$C \quad \geq \quad \frac{kR(1-\epsilon)}{\pi}\left(\frac{\epsilon(N-k-m)}{k\beta d^{\alpha}}\right)^{\frac{2}{\alpha}}\left(\left\lfloor\frac{m}{k}\right\rfloor - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{2}{\alpha}}. \tag{18}$$

Clearly, $k = 1$, $m = \theta N, \theta \in (0,1]$, yields $C = \Omega(N)$. Finding the best constant θ, that maximizes the lower bound (18), is equivalent to solving,

$$\max_{\theta}(1-\theta)^{\frac{2}{\alpha}}\theta^{1-\frac{2}{\alpha}}.$$

By setting the derivative to zero, the optimal value of $\theta^\star = 1 - \frac{2}{\alpha}$. Note that the lower bound on the transmission capacity is concave in m. Thus, to enforce the integer constraint on m, m should be chosen as $\left\lfloor\left(1-\frac{2}{\alpha}\right)N\right\rfloor$ or $\left\lceil\left(1-\frac{2}{\alpha}\right)N\right\rceil$ depending on whichever value maximizes the lower bound. $\square$

Theorem 3.5 shows that with single data stream transmission, $k = 1$, and using a linearly increasing (with N) SRDOF for interference cancelation, $m = \theta N$, transmission capacity scales linearly with the number of antennas N. To interpret the optimality of $k = 1$ and $m = \theta N$, we need to look at (18), where the term $\left(\frac{(N-k-m)}{k}\right)^{\frac{2}{\alpha}}$ corresponds to the gain obtained by coherently combining the signal of interest using $N-k-m$ SRDOF, while the term $\left(\left\lfloor\frac{m}{k}\right\rfloor - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{2}{\alpha}}$ is attributed to the gain obtained by canceling the m nearest interferers. Thus, using $k = 1$ and $m = \theta N$, allows the two terms to balance out each other and allows a linear increase of the transmission capacity with number of antennas N.

To illustrate the scaling behavior of the transmission capacity with respect to the number of antennas N, we plot the simulated transmission capacity with different transmit-receive strategies, e.g. ($k = 1, m = N-1$), ($k = N/2, N-m = N/2$), and ($k = 1, m = (1-2/\alpha)N$) in Fig. 3 with increasing N. We plot the transmission capacity with both the partial ZF decoder as well as the ML decoder. Since for $k = 1$, both the ML and partial ZF decoder are identical, their transmission capacities are also the same. As expected from our derived results, sending a single data stream and using a constant fraction of SRDOF for interference cancelation, $k = 1, m = (1-2/\alpha)N$, achieves a linear increase of transmission capacity with increasing N, in contrast to sub-linear increase for the other two cases. More importantly, Fig. 3 shows that the upper bound derived in Theorem 3.5 where transmission capacity scales super-linearly with N is loose, and at best only linear increase in N is possible for transmission capacity. Fig. 3 also shows that performance of partial ZF decoder is very close to the

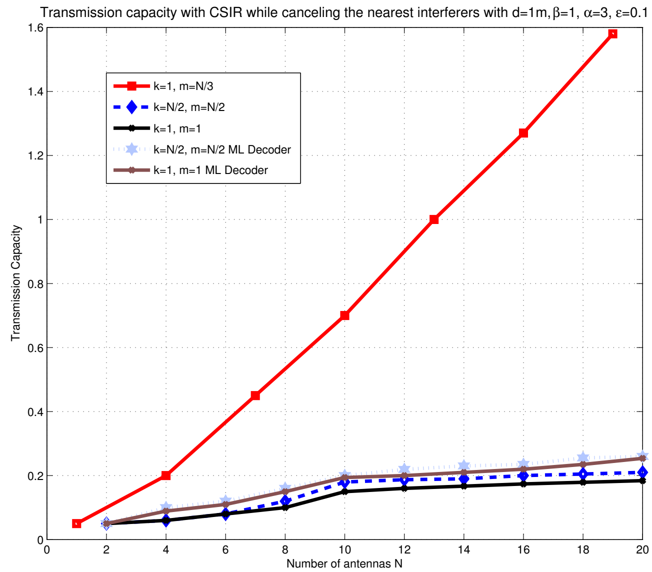

Figure 3: Transmission capacity versus N with CSIR while canceling the nearest interferers for $k = 1$, $d = 1$m, $\beta = 1$ bits, $\alpha = 3$, $\epsilon = 0.1$

ML decoder and all the conclusions we draw from Theorem 3.5 hold reasonably well for the ML decoder as well.

An important lesson from Theorem 3.5 is the utility of using Markov's inequality. Typically, bounds obtained by Markov's inequality are fairly loose, but for transmission capacity purposes this seems to be a handy tool for obtaining tight enough scaling bounds. Next, we indicate how to obtain the exact scaling results of Theorem 3.6, when each transmitter has a single antenna or chooses $k = 1$.

Proof: (Theorem 3.6) For a single transmit antenna or single stream transmission $k = 1$, we know from Theorem 3.5 that $C = \Omega(N)$ by using $m = \theta N$. Moreover, from (17), for $k = 1$, choosing $r = N$ with $m = \Theta(N)$, we get $C = \mathcal{O}(N)$, thus finishing the proof. □

Remark 3.11 *In this section, we have analyzed the case when each receiver cancels the nearest interferers using its multiple antennas. Another logical choice is to cancel those interferers that have the largest interference power at the receiver. With a single transmit antenna and N receive antennas, the scaling behavior of transmission capacity while canceling the $N-1$ strongest interferers has been analyzed in [7], and it is shown that the transmission capacity scales as $\epsilon^{1/N}$, where ϵ is the outage probability constraint. Thus, using multiple antennas for canceling the strongest interferers leads to diminished gains compared to canceling the nearest interferers, where the transmission capacity scales linearly with the number of antennas.*

This might appear counter intuitive, however, it can be explained by nothing that *while canceling the strongest interferers, some of the nearby interferers may not be canceled if their channel gains are very low. Post-cancelation, i.e. after multiplication by the cancelation vector, the situation might change, and the interference power from some of the nearest interferers could become moderate, in which case the nearby interferers dominate the performance.*

Next, using the results of this section, we find exact transmission capacity scaling result with respect to the number of antennas when no no CSI about other interferers' channels is available at any receiver, thereby precluding the possibility of interference cancelation.

3.2 No Interference Cancelation

In this section, we consider the case when no receiver employs any interference cancelation and uses all its SRDOF for decoding the data streams transmitted by its intended transmitter. This scenario is motivated for two important practical reasons. First, interference cancelation requires the knowledge of channel coefficients between the interferer and the receiver which is typically hard to get, especially in a wireless network. Secondly, the hardware complexity of the receiver is fairly low without the interference cancelation capability. We next show that there is no loss in terms of transmission capacity with or without interference cancelation in terms of scaling with respect to the number of antennas. Thus, the restricted receiver design has no effect on the transmission capacity performance, though the optimal transmit strategy used at each transmitter differs significantly with respect to the interference cancelation case. The advantage of CSI shows up in simplified encoding/decoding, since with CSI only 1 data stream needs to be transmitted and decoded to achieve linear scaling of the transmission capacity with multiple antennas, in comparison to a constant fraction of N data streams that are transmitted and decoded without CSI.

Theorem 3.12 *The transmission capacity with multiple antennas when receiver uses ZF decoder and does not employ any interference cancelation, scales as $C = \Theta(N)$, and the optimal number of data streams to transmit, k, scales linearly with N.*

Proof: When no interferers are canceled, the SRDOF used for interference cancelation is $m = 0$ or $N_{\mathsf{canc}} = 0$. Then from Theorem 3.10, for any r such that $kr > 1$, the transmission capacity is upper bounded by

$$C \quad \leq \quad \frac{kR(1-\epsilon)^{1-\frac{2}{\alpha}}}{\pi}\left(\frac{N-k+1}{(kr-1)d^{\alpha}\beta}\right)^{\frac{2}{\alpha}}\left(r+\frac{\alpha}{2}\right). \tag{19}$$

Since we have the freedom to choose r in the upper bound (Theorem 3.10), let r be a constant independent of N. Then for number of transmitted data streams $k = N^{\kappa}$, from (19), we have

$$C = \mathcal{O}\left(N^{\kappa}N^{\frac{2}{\alpha}(1-\kappa)}\right) = \mathcal{O}\left(N^{\frac{2}{\alpha}+\kappa\left(1-\frac{2}{\alpha}\right)}\right) = \mathcal{O}(N),$$

for the optimal value of $\kappa = 1$.

For the lower bound, similar to (18), for $k < N$ with no interference cancelation $m = 0$,

$$C \geq \frac{kR(1-\epsilon)}{\pi}\left(\frac{\epsilon(N-k)}{k\beta d^{\alpha}}\right)^{\frac{2}{\alpha}}. \tag{20}$$

Letting $k = \Theta(N)$, immediately from (20), we get $C = \Omega(N)$, finishing the proof. □

Remark 3.13 *If we fix single data stream transmission $k = 1$, then from* (19) *and* (20)*, we get that with no interference cancelation, the transmission capacity scales only sub-linearly as $\Theta\left(N^{\frac{2}{\alpha}}\right)$. This result was originally found in [8] using a direct outage probability computation. Thus to obtain linear scaling we have to scale the number of transmitted data streams with N.*

Remark 3.14 *Theorem 3.12 has been independently derived in [9] by explicitly computing the outage probability, rather than finding tight lower and upper bounds.*

Comparing Theorem 3.12 with Theorem 3.5, interestingly, we conclude that with or without interference cancelation, the transmission capacity scales linearly with N, and only the transmit-receive strategy changes. When no interference cancelation is employed, the number of data streams sent from each transmitter should scale with the number of antennas, as opposed to the case of interference cancelation where only single data stream should be transmitted. With no interference cancelation, if the number of data streams is not scaled with N, the transmission capacity scales only sub-linearly with N (Remark 3.13). Thus, with single data stream transmission $k = 1$, the maximal ratio combining (MRC) gain available at the receiver for decoding the only data stream scales as $N^{2/\alpha}$, and to achieve a linear growth of transmission capacity, we need to linearly scale the number of transmitted data streams with N.

From (20), one can easily show that using a $\left(1 - \frac{2}{\alpha}\right)$ fraction of the total transmit antennas N maximizes the lower bound on the transmission capacity with no interference cancelation. Thus, for small path loss exponents α i.e. when the interference is dominating, only a small number of data streams should be transmitted, while, for large path loss exponents which corresponds to the weak interference regime, almost all transmit antennas should be used to maximize the transmission capacity.

Remark 3.15 *For a cellular communication network, when each transmitter sends k independent data streams with equal power allocation, and no interference cancelation is employed at the receiver, using a single transmit antenna is shown to maximize the ergodic Shannon capacity in the presence of small number of strong co-channel interferers in [10–13]. Thus, the results obtained in this section with no interference cancelation for PPP distributed transmitter locations in a wireless network match with results on cellular networks only for small path loss exponents α.*

After discussing the case of having no CSI at any of the transmitters in this section, we move on to the more general (put practically challenging) scenario in the next section, when each transmitter is also assumed to have CSI for its corresponding receiver, and find the impact of CSI availability at each transmitter on the transmission capacity.

4 Channel State Information at Both Transmitter and Receiver

In this section, we consider the case when in addition to each receiver having the CSI for all the channels (Section 3), the transmitter also has CSI for the channel between itself and its intended receiver. We refer to this as the CSIT case. From a transmission capacity perspective, the CSIT case is fundamentally different than the CSIR case, since with CSI, each transmitter can increase the signal power at its intended receiver by steering the beam towards it, and possibly the role of multiple receiver antennas and consequently the transmission capacity scaling is different from the CSIR case.

Remark 4.1 *With global CSIT, each transmitter could also use its multiple antennas for interference suppression by nulling out its signal towards unintended receivers. We consider the interference suppression role of multiple transmit antennas in Section 5 together with a cognitive radio network.*

We begin with a brief background on using CSIT for a point-to-point multiple antenna channel without interference. Lets consider a point-to-point multiple antenna channel, where $\mathbf{H} \in \mathbb{C}^{N \times N}$ is the channel matrix between the transmitter-receiver pair with N antennas each. If $\tilde{\mathbf{x}}$ is the transmit signal, then the received signal is given by

$$\mathbf{y} = \mathbf{H}\tilde{\mathbf{x}} + \mathbf{w}, \tag{21}$$

where $\mathbf{w}$ is the AWGN vector with independent entries that have zero mean and unit variance. Let $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^{\dagger}$ be the singularvalue decomposition of $\mathbf{H}$. To maximize the mutual information, the transmitter sends its signal over the strongest singular values of the channel $\mathbf{H}$ [**?**]. Let $\mathbf{V}^k$ be the matrix consisting of the first k columns of $\mathbf{V}$ corresponding to the k strongest eigenvalues of $\mathbf{H}\mathbf{H}^{\dagger}$. Thus, if $\mathbf{x} \in \mathbb{C}^{k \times 1}$ is the input signal, then the transmitter sends $\tilde{\mathbf{x}} = \mathbf{V}^k\mathbf{x}$ through its N antennas, and the received signal is

$$\mathbf{y} = \mathbf{U}\mathbf{D}\mathbf{V}^{\dagger}\mathbf{V}^k\mathbf{x} + \mathbf{w}. \tag{22}$$

The case of $k = 1$ is referred to as beamforming, while the case of $k > 1$ is called multi-mode beamforming. With multi-mode beamforming, if the receiver multiplies the received signal (22) with $\mathbf{U}^{\dagger}$, the equivalent received signal is given by

$$\mathbf{y}(\ell) = \mathbf{D}(\ell, \ell)\mathbf{x}(\ell) + \hat{\mathbf{w}}(\ell), \tag{23}$$

for $\ell = 1, \ldots, k$, where noise contributions $\hat{\mathbf{w}}(\ell)$ are Gaussian and independent $\forall\ \ell$, since $\mathbf{U}^{\dagger}$ is unitary. Thus, with CSIT, the received signal gets decouples into k independent signals, where the signal power of the ℓ^{th} channel is equal to the ℓ^{th} eigenvalue of $\mathbf{H}\mathbf{H}^{\dagger}$. Thus, the knowledge of $\mathbf{H}$ not only helps in increasing the received SNR, but also simplifies the decoding since each element of the input signal $\mathbf{x}$ can be decoded independently.

Now we look at our model with interference. We assume that each transmitter uses multi-mode beamforming even in the presence of interferers. As before,

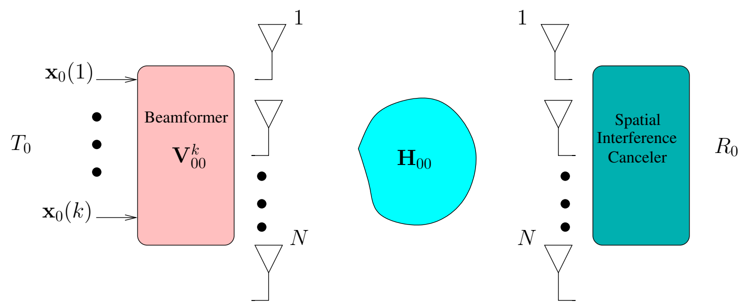

Figure 4: Transmit-receive strategy with beamforming at the transmitter.

$\mathbf{H}_{nm} \in \mathbb{C}^{N\times N}$ represents the channel matrix between transmitter T_m and receiver R_n. Transmitter T_n is assumed to only know $\mathbf{H}_{nn}$, the channel between itself and its corresponding receiver.

Consider the typical transmitter-receiver pair (T_0, R_0). Let the singularvalue decomposition of $\mathbf{H}_{00} \in \mathbb{C}^{N\times N}$ (channel between T_0 and R_0) be $\mathbf{U}_{00}\mathbf{D}_{00}\mathbf{V}_{00}^{\dagger}$. Let $k,\ k \in [1, 2, \dots, N]$ denote the number of independent data streams (STDOF) sent by each transmitter to its receiver. Then with multi-mode beamforming, transmitter T_0 sends $\mathbf{V}_{00}^k\mathbf{x}_0$, where $\mathbf{V}_{00}^k$ be the matrix consisting of first k columns of $\mathbf{V}_{00}$ corresponding to the k strongest eigenvalues of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$, and $\mathbf{x}_0 \in \mathbb{C}^{k\times 1}$ is the data vector consisting of k independent streams, where each stream is $\mathcal{CN}\left(0, \frac{1}{k}\right)$ distributed. Note that in contrast to the CSIR case, with CSIT, the k data streams are transmitted by all the N transmit antennas via processing through $\mathbf{V}_{00}^k$. For keeping the analysis tractable, we consider equal power allocation among the k transmitted streams.

Similar to the case of point-to-point channel with no interference (23), we will show in (27) that even in a wireless network, with multi-mode beamforming, the channel between each transmitter-receiver pair is equivalent to k scalar parallel channels with no inter-stream interference in contrast to the CSIR case. Thus, we assume that each receiver uses k SRDOF to receive the intended signal, while the remaining $N - k$ SRDOF are used for canceling the $c(k) = \left\lfloor \frac{N}{k} \right\rfloor - 1$ nearest interferers.

To cancel the $c(k)$ interferers, the receiver projects the received signal on to the null space of the $c(k)$ interferers. A block diagram depicting the transmit-receive strategy is illustrated in Fig. 4.

Using multi-mode beamforming at each transmitter, the received signal $\mathbf{y}_0 \in \mathbb{C}^{N\times 1}$ at receiver R_0 is

$$\mathbf{y}_0 = d^{-\alpha/2}\mathbf{H}_{00}\mathbf{V}_{00}^k\,\mathbf{x}_0 + \sum_{T_n:\Phi\setminus\{T_0\}} d_n^{-\alpha/2}\mathbf{H}_{0n}\mathbf{V}_{nn}^k\mathbf{x}_n, \tag{24}$$

where $\mathbf{V}_{nn}$ is the matrix of the right singular vectors of the channel between transmitter n and receiver n, $\mathbf{V}_{nn}^k$ are the first k columns of $\mathbf{V}_{nn}$, $\mathbf{H}_{nn} = \mathbf{U}_{nn}\mathbf{D}_{nn}\mathbf{V}_{nn}^{\dagger}$. Let the indices of the interferers be sorted in an increasing order in terms of their distance from R_0, i.e. $d_1 \le d_2 \le \dots \le d_{c(k)} \le d_{c(k)+1} \le \dots$. Let S be the basis of the null space of the channel matrices $[\mathbf{H}_{01} \dots \mathbf{H}_{0c(k)}]$ corresponding to the $c(k)$ nearest interferers to be canceled. Since $N-k$ SRDOF are used for interference cancelation, $\mathsf{S} \in \mathbb{C}^{k\times N}$. Multiplying S to the received signal (24),

$$\begin{aligned} \mathsf{S}\mathbf{y}_0 &= d^{-\alpha/2}\mathsf{S}\mathbf{U}_{00}\mathbf{D}_{00}\mathbf{V}_{00}^{\dagger}\mathbf{V}_{00}^{k}\,\mathbf{x}_0 + \sum_{n=c(k)+1}^{\infty} d_n^{-\alpha/2}\mathsf{S}\mathbf{H}_{0n}\mathbf{V}_{nn}^{k}\mathbf{x}_n, \\ &= d^{-\alpha/2}\mathsf{S}\mathbf{U}_{00}^{k}\mathbf{D}_{00}^{k}\mathbf{x}_0 + \sum_{n=c(k)+1}^{\infty} d_n^{-\alpha/2}\mathsf{S}\mathbf{H}_{0n}\mathbf{V}_{nn}^{k}\mathbf{x}_n, \end{aligned} \tag{25}$$

where $\mathbf{U}_{00}^{k}$ is the $N\times k$ matrix consisting of the first k columns of $\mathbf{U}_{00}$, and $\mathbf{D}_{00}^{k} \in \mathbb{C}^{k\times k}$ is the diagonal matrix consisting of the first k entries of $\mathbf{D}_{00}$. Since S, and $\mathbf{U}_{00}^{k}$ are both of rank k, and are independent of each other with each entry drawn from a continuous distribution, $\mathsf{S}\mathbf{U}_{00}^{k}$ is full rank with probability 1. Multiplying $\left(\mathsf{S}\mathbf{U}_{00}^{k}\right)^{-1}$ to the received signal (25)

$$\hat{\mathbf{y}}_0 = d^{-\alpha/2}\mathbf{D}_{00}^{k}\mathbf{x}_0 + \sum_{n=c(k)+1}^{\infty} d_n^{-\alpha/2}\left(\mathsf{S}\mathbf{U}_{00}^{k}\right)^{-1}\mathsf{S}\mathbf{H}_{0n}\mathbf{V}_{nn}^{k}\mathbf{x}_n. \tag{26}$$

Note that $\mathbf{D}_{00}$ is the diagonal matrix of the eigenvalues of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$. Denoting the ℓ^{th} eigenvalue of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$ by $\sigma_\ell(\mathbf{H}_{00})$, the received signal (26) can be decomposed into k parallel channels as

$$\hat{\mathbf{y}}_0(\ell) = d^{-\alpha/2}\sqrt{\sigma_\ell(\mathbf{H}_{00})}\mathbf{x}_0(\ell) + \sum_{n=c(k)+1}^{\infty} d_n^{-\alpha/2}\sum_{j=1}^{k} g_n(\ell,j)\mathbf{x}_n(j),\ \ell = 1,2,\dots,k, \tag{27}$$

where $\mathbf{x}_n(j)$ is the j^{th} element of the transmitted vector $\mathbf{x}_n$, $g_n(\ell,j)$ is the $(\ell,j)^{th}$ element of $\left(\mathsf{S}\mathbf{U}_{00}^{k}\right)^{-1}\mathsf{S}\mathbf{H}_{0n}\mathbf{V}_{nn}^{k}$.

Thus, with multi-mode beamforming, as shown in (27), the received signal can be decomposed into k parallel channels, with the ℓ^{th} channel corresponding to the data stream $\mathbf{x}_0(\ell)$ having no contribution from data streams $\mathbf{x}_0(1),\dots,\mathbf{x}_0(\ell-1),\mathbf{x}_0(\ell+1),\dots\mathbf{x}_0(k)$. Therefore, with multi-mode beamforming, there is no inter-stream interference from the other $k-1$ data streams sent by the same transmitter. Thus, using $N-k$ SRDOF for interference cancelation, $c(k) = \left\lfloor \frac{N-k}{k} \right\rfloor$ nearest interferers can be canceled at each receiver.

Since S, $\mathbf{U}_{00}^{k}$, and $\mathbf{V}_{nn}$, are independent of $\mathbf{H}_{0n}$, $g_n(\ell,j)$'s in (27) are independent for $j=1,\dots,k$, and each $g_n(\ell,j) \sim \chi^2(2)$ from Lemma 3.2. Thus, the interference power of the ℓ^{th} data stream of the n^{th} interferer

$$\mathsf{pow}_n(\ell) = \mathbb{E}\left\{\left|\sum_{j=1}^{k} g_n(\ell,j)\mathbf{x}_n(j)\right|^2\right\} \sim \chi^2(2k).$$

Let

$$I_n(\ell) = d_n^{-\alpha}\mathsf{pow}_n(\ell)$$

be the interference power of the n^{th} interferer for the ℓ^{th} channel in (27). Since S and $\mathbf{U}_{00}^k$ are independent of $\mathbf{H}_{0n}$, $g_n(\ell, j)$'s and consequently $\mathsf{pow}_n(\ell)$'s are identically distributed for all ℓ, and it follows that $I_n(\ell)$ is identically distributed for all ℓ. Then the total interference power seen at receiver R_0 for the ℓ^{th} channel corresponding to signal $\mathbf{x}_0(\ell)$ in (27) is $I_{nc}(\ell) = \sum_{n=c(k)+1}^{\infty} I_n(\ell)$. Since $I_n(\ell)$ is identically distributed for all ℓ, it follows that $I_{nc}(\ell)$ is also identically distributed for all $\ell = 1, \dots, k$ channels.

We assume an uniform data rate of B bits/sec/Hz on each of the k transmitted data streams.[2] By combining the k streams, the total rate of transmission between a source and destination is kB bits/sec/Hz. To define outage probability, we consider the outage event of the data stream with the worst channel gain, which in this case is the k^{th} data stream, since the eigenvalues of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$ are indexed in the decreasing order. Thus, the outage probability for any channel in (27) is at most

$$\begin{aligned} P_{out}(B) &= \mathbb{P}\left(\log\left(1 + \frac{d^{-\alpha}\sigma_k(\mathbf{H}_{00})}{I_{nc}(k)}\right) \le B\right), \\ &= \mathbb{P}\left(\frac{d^{-\alpha}\sigma_k(\mathbf{H}_{00})}{I_{nc}(k)} \le 2^B - 1\right), \end{aligned} \tag{28}$$

where $\sigma_k(\mathbf{H}_{00})$ is the k^{th} eigenvalue of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$. Since, $I_{nc}(\ell)$ is identically distributed for each $\ell = 1, 2, \dots, k$, from here on we drop the index ℓ and represent $I_{nc}(\ell)$ as I_{nc} for each $\ell = 1, 2, \dots, k$. Thus, with $2^B - 1 = \beta$

$$P_{out}(B) = \mathbb{P}\left(\frac{d^{-\alpha}\sigma_k(\mathbf{H}_{00})}{I_{nc}} \le \beta\right), \tag{29}$$

where

$$I_{nc} = \sum_{n=c(k)+1}^{\infty} d_n^{-\alpha}\mathsf{pow}_n, d_i \le d_j,\ i < j,$$

and pow_n are i.i.d. with $\chi^2(2k)$ distribution.

This definition of outage probability (29) implies that if $P_{out}(B) = \epsilon$, then all the k streams can at least support a data rate of B bits/sec/Hz with probability $1 - \epsilon$, and the transmission capacity is defined as

$$C = k\lambda(1-\epsilon)B \text{ bits/sec/Hz/m}^2,$$

by combining the contribution from all the k transmitted data streams. Deriving a closed form expression for the outage probability requires the distribution of I_{nc}, and $\sigma_k(\mathbf{H}_{00})$, the k^{th} maximum eigenvalue of the Wishart matrix $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$. Unfortunately, both these distributions are unknown, and hence finding an exact expression for the outage probability is difficult. To facilitate analysis, we use upper and lower bounds on the outage probability derived in Theorems 3.8 and 3.10, and then find the optimal number of data streams k that maximize the transmission capacity.

[2] In general with multi-mode beamforming, data rates can be a function of the magnitude of the eigenvalues, however, that requires finding the optimal rate allocation that minimizes the maximum of the outage probability on k different streams, which is an unsolved problem.

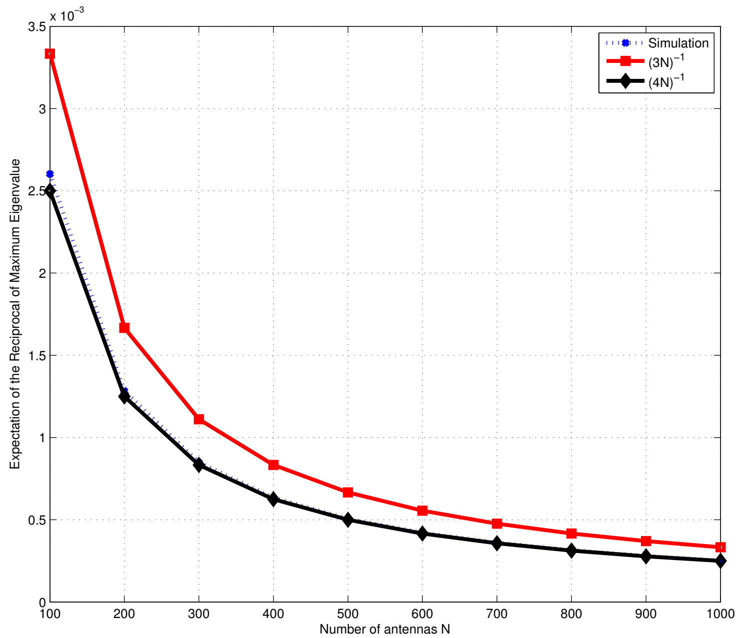

Figure 5: Empirical expected value of the reciprocal of the largest eigenvalue of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$.

Use of Theorems 3.8 and 3.10 also allows us to circumvent the problem of requiring a simple closed form expression for the probability density function (PDF) of $\sigma_k(\mathbf{H}_{00})$. For our analysis, it will suffice to know the expected value of the maximum eigenvalue of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$, $\sigma_1(\mathbf{H}_{00})$, and the expected value of the reciprocal of $\sigma_1(\mathbf{H}_{00})$. For large N, the maximum eigenvalue of $\mathbf{H}_{00}\mathbf{H}_{00}^{\dagger}$, $\sigma_1(\mathbf{H}_{00})$, converges to $4N$ [14], and $\mathbb{E}\{\sigma_1(\mathbf{H}_{00})\} \approx 4N$. With extensive simulation results, Fig. 5, we observe that $\mathbb{E}\{\frac{1}{\sigma_1(\mathbf{H}_{00})}\} \approx \frac{1}{3.5N}$, however, an analytical proof for this result cannot be found readily in literature. Note that the constant $1/3.5$ is immaterial for us, we are only interested in the scaling of the mean of the reciprocal of the maximum eigenvalue of Wishart matrix with N and our simulations show that mean of the reciprocal of the maximum eigenvalue of Wishart matrix does not decrease faster than N^{-1}. We will use both these large N approximations on $\mathbb{E}\{\sigma_1(\mathbf{H}_{00})\}$, and $\mathbb{E}\{\frac{1}{\sigma_1(\mathbf{H}_{00})}\}$, for our analysis.

The main result of this section is as follows that characterizes the scaling of transmission capacity with multiple antennas using multi-mode beamforming.

Theorem 4.2 *With multi-mode beamforming and ZF decoder, the transmission capacity scales as*

$$C = \Omega(N), \text{ and } C = \mathcal{O}(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}})$$

with the number of antennas N. The optimal lower bound is achieved by $k^\star = 1$ and $c(k) = N - 1$, i.e. sending only one data stream on the strongest eigenvector, and canceling the maximum number of interferers $N-1$, is optimal.

Proof: With the number of nearest canceled interferers to be $c(k) = \left\lfloor \frac{N}{k} \right\rfloor - 1$, from Theorem 3.10, for any r such that $kr > 1$,

$$C \leq \frac{kR(1-\epsilon)^{1-\frac{2}{\alpha}}}{\pi} \left(\frac{\mathbb{E}\{\sigma_k(\mathbf{H}_{00})\}}{(kr-1)d^\alpha \beta} \right)^{\frac{2}{\alpha}} \left(\left\lfloor \frac{N}{k} \right\rfloor - 1 + r + \frac{\alpha}{2} \right), \tag{30}$$

where we have replaced the expected signal power $\mathbb{E}\{s\} = N - m - k + 1$ of the CSIR case with $\mathbb{E}\{\sigma_k(\mathbf{H}_{00})\}$, the expected signal power of the k^{th} data stream with multi-mode beamforming (26). Recall that we have ordered the eigenvalues $\sigma_k(\mathbf{H}_{00})$ in decreasing order, $\sigma_k(\mathbf{H}_{00}) \geq \sigma_m$ if $k > m$. From [14], $\mathbb{E}\{\sigma_1(\mathbf{H}_{00})\} = 4N$, hence $\mathbb{E}\{\sigma_k(\mathbf{H}_{00})\} < 4N$ for $k > 1$. Hence, from (30), similar to the proof of Theorem 3.5, we can show that

$$C = \mathcal{O}\left(N^{1+\frac{2}{\alpha}-\frac{4}{\alpha^2}}\right)$$

with $r = N^{2/\alpha}$ by parametrizing $k = N^\kappa$ and finding the best κ.

For the lower bound, by substituting $\mathbb{E}\left\{\frac{1}{s}\right\} = \mathbb{E}\left\{\frac{1}{\sigma_k(\mathbf{H}_{00})}\right\}$ in Theorem 3.8 for $k < N$, we have

$$C \geq \frac{kR(1-\epsilon)}{\pi} \left(\frac{\epsilon}{k\mathbb{E}\left\{\frac{1}{\sigma_k(\mathbf{H}_{00})}\right\} \beta d^\alpha} \right)^{\frac{2}{\alpha}} \left(\left\lfloor \frac{N}{k} \right\rfloor - 1 - \left\lceil \frac{\alpha}{2} \right\rceil \right)^{1-\frac{2}{\alpha}}. \tag{31}$$

As pointed out earlier, $\mathbb{E}\{\frac{1}{\sigma_1(\mathbf{H}_{00})}\} = \frac{1}{3.5N}$, $\mathbb{E}\{\frac{1}{\sigma_k(\mathbf{H}_{00})}\} > \frac{1}{3.5N}$ for $k > 1$. Thus, evaluating the lower bound (31) at $k = 1$, we get $C = \Omega(N)$.

□

With multi-mode beamforming, the lower bound on the transmission capacity is maximized by using single stream beamforming ($k = 1$) together with canceling the $N-1$ nearest interferers, and the lower bound scales linearly with N. Thus, comparing the CSIT and CSIR cases, the transmission strategy remains identical, but the reception strategy is completely different (with the CSIR case $m = \Theta(N)$ nearest interferers are canceled). This difference is because in the CSIT case, the average signal power (strongest eigenvalue) scales linearly with N without any processing at the receiver, while in the CSIR case, it is independent of N if signals received at multiple receive antennas are not combined at the receiver. Thus, in the CSIR case, to boost the signal power so that it scales with N, $\Theta(N)$ SRDOF are required for decoding the signal of interest allowing only $m = \Theta(N)$ nearest interferers to be canceled.

The derived bounds on the transmission capacity in both the CSIR and CSIT cases are identical, implying that the value of channel feedback (which is generally costly) is fairly limited. There is, however, a constant multiplicative gain of 4 in terms of signal power with CSIT, since with the optimal mode of $k = 1$, the signal power $\mathbb{E}\{\sigma_1(\mathbf{H}_{00})\} = 4N$ in comparison to order N for the CSIR case. The real advantage of CSIT is the simplified encoding and decoding, since with CSIT, the multiple data

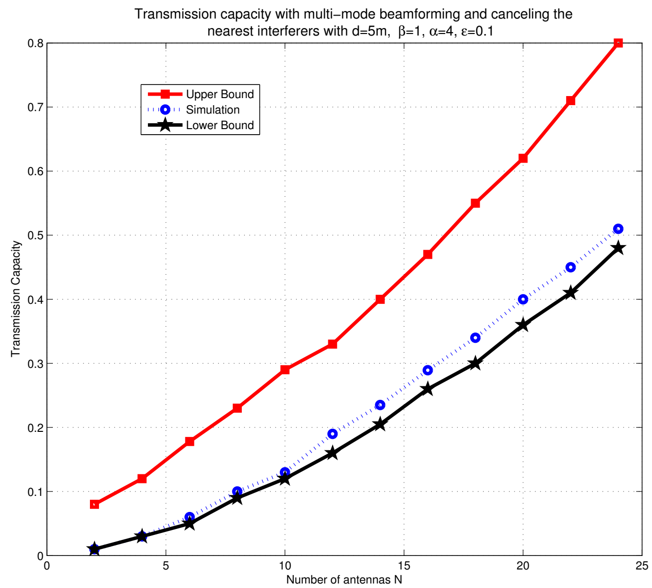

Figure 6: Transmission capacity versus the number of antennas N with multi-mode beamforming and canceling the nearest interferers with single stream data transmission $k = 1$, $d = 5$m, $\beta = 1$ ($B = 1$ bits/sec/Hz), $\alpha = 4$, $\epsilon = 0.1$

streams sent by the transmitter can be resolved as parallel channels at the receiver resulting in independent decoding.

To illustrate the scaling behavior of the transmission capacity with CSIT as a function of the number of antennas N, we plot the derived lower and upper bound, and the simulated transmission capacity in Fig. 6, for $k = 1$, $d = 5$m, path-loss exponent $\alpha = 4$ and outage probability constraint of $\epsilon = 0.1$ with increasing N. We see that the transmission capacity grows linearly even for $\alpha = 4$, for which the upper bound suggests super-linear scaling of $N^{1+1/4}$. Thus, the derived lower bound that scales linearly N is tight, however, some more analytical work is required to tighten the upper bound to make it scale linearly with N. To show the optimality of using a single data stream from each transmitter k, in Fig. 7, we plot the transmission capacity as a function of k for total $N = 8$ antennas. Fig. 7, clearly shows that the transmission capacity with multi-mode beamforming is a decreasing function of k, and sending a single data stream is optimal in a wireless network.

Similar to the CSIR case, for the CSIT case, we can get the exact result for the special case when each receiver employs no interference cancelation. We show that with no interference cancelation, the transmission capacity scales as $\Theta(N)$ and the optimal number of data streams to transmit is $k = \Theta(N)$. Further, if the number of receive antennas is 1, then we show that the transmission capacity is $\Theta(N^{\frac{2}{\alpha}})$, i.e. scales sub-linearly with N.

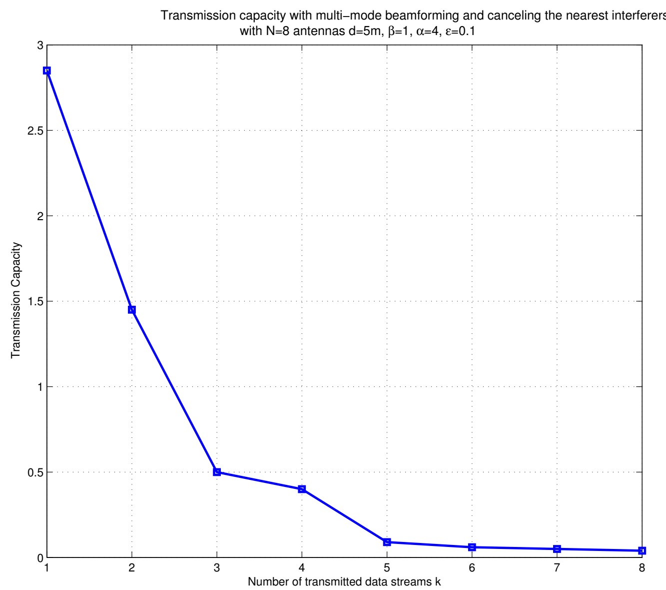

Figure 7: Transmission capacity versus the number of transmitted data streams k with multi-mode beamforming and canceling the nearest interferers with $d = 5$m, $\beta = 1$, $\alpha = 4$, $\epsilon = 0.1$, total number of antennas $N = 8$.

Theorem 4.3 *Without interference cancelation at any receiver, with multi-mode beamforming, the transmission capacity is $C = \Theta(N)$, and the optimal number of data streams to transmit is $k = \theta N, \theta \in (0, 1]$. If the number of receive antennas is 1, then the transmission capacity is $C = \Theta(N^{\frac{2}{\alpha}})$, where N is the number of transmit antennas.*

Proof: Follows similarly to the proof of Theorem 4.2 □

Remark 4.4 *In this section, even though we have assumed the availability of CSI at each transmitter, we have not accounted for resources required for feeding back CSI from each receiver. In general, it is a hard problem to quantify the effects of feedback. In Chapter* **??**, *we present some results in that direction.*

5 Spectrum-Sharing/Cognitive Radios

After considering the dual role of multiple antennas in previous sections, sending multiple data streams from the transmitter and canceling interference at the receiver, in this section, we look at the third possible role of multiple antennas in a wireless network: using them at the transmitter for suppressing interference towards other receivers. To highlight this feature, we consider a cognitive/secondary wireless network that is overlaid over a pre-existing/primary wireless network that consists of licensed/primary nodes.

In particular, we consider two co-existing networks, one primary and other secondary, where primary network is oblivious to the presence of the secondary network, while the secondary network is aware of the primary network. For the primary network, we assume the same model as in Section 3, where each primary transmitter has a primary receiver associated with it at a fixed of fixed distance d_p, with SIR threshold β_p, and under an outage probability constraint of ϵ_p at each receiver, except that each transmitter and receiver has a single antenna. Thus, from Theorem **??**, the maximum density of primary network is

$$\lambda_p^\star = \frac{\ln(1-\epsilon_p)}{c\beta_p^{\frac{2}{\alpha}} d_p^2},$$

for a constant c.

The secondary network is overlaid on top of the primary network, where each secondary transmitter has a secondary receiver associated with it at a fixed of fixed distance d_s, with density λ_s and SIR threshold β_s, under an outage probability constraint of ϵ_s at each secondary receiver. Clearly, the presence of secondary transmitters increases the interference seen at any primary receiver, thus, if $\lambda_s \neq 0$, the primary outage probability constraint of ϵ_p cannot be met if the primary network is operating with density $\lambda_p^\star$. Therefore, $\lambda_s = 0$, if primary network density is $\lambda_p^\star$ and primary outage probability constraint is ϵ_p.

To make the problem non-degenerate, we relax the primary outage probability constraint of ϵ_p to $\epsilon_p + \Delta_p$ while keeping the primary network density to be $\lambda_p^\star$, and find the maximum value of λ_s such that the relaxed primary outage probability constraint of $\epsilon_p + \Delta_p$, and the secondary outage probability constraint of ϵ_s is satisfied simultaneously.

We assume that the secondary nodes are equipped with multiple transmit and receive antennas. Multiple antennas at each secondary transmitter node are used for interference suppression towards primary receivers, while multiple receive antennas are interference cancelation at each secondary receiver Since the secondary network has to operate under an outage probability constraint at each primary node, it is important to control the interference that each secondary node creates, and this is where the interference suppression feature of multiple transmit antennas comes to the fore.

We let the locations of primary and secondary transmitters to be distributed as two independent homogenous PPPs with density λ_1, and λ_2, respectively. We consider an ALOHA random access protocol for both the primary and secondary transmitters, with access probability p. Consequently, the active primary and secondary transmitter processes are also homogenous PPPs on a two-dimensional plane with density $\lambda_p = p\lambda_1$, and $\lambda_s = p\lambda_2$, respectively.

Let the location of the n^{th} active primary transmitter be T_{pn}, and the n^{th} active secondary transmitter be T_{sn}. The set of all active primary and secondary transmitters is denoted by $\Phi_p = \{T_{pn},\ n \in \mathbb{N}\}$ and $\Phi_s = \{T_{sn},\ n \in \mathbb{N}\}$, respectively.

We assume that each secondary transmitter has N_t antennas, while each secondary receiver has N_r antennas. We also assume that each secondary transmitter has CSI for its corresponding receiver, as well as for its N_t nearest primary receivers that is used to suppress interference towards them. Each secondary receiver is assumed to have CSI for its intended transmitter as well as for its N_r nearest interferers (from the union $\Phi_s \cup \Phi_p$). The system model of overlaid wireless networks under consideration is illustrated in Fig. 8, where the squares represent the primary transmitters and receivers, while the dots represent the secondary transmitters and receivers, and a dashed line indicates a suppressed interferer. We restrict ourselves to the case when each secondary transmitter sends only one data stream using its multiple antennas to its intended secondary receiver.

Let the beamformer used by the n^{th} secondary transmitter for interference suppression towards primary receivers is denoted by $\mathbf{b}_n \in \mathbb{C}^{N\times 1}$. Then, the received signal at the primary receiver R_{p0} is

$$\begin{aligned} y_0 &= \sqrt{P_p} d_p^{-\alpha/2} h_{00} x_{p0} + \sum_{n:T_{pn}\in\Phi_p\setminus\{T_{p0}\}} \sqrt{P_p} d_{pp,n}^{-\alpha/2} h_{0n} x_{pn} \\ &\quad + \sum_{n:T_{sn}\in\Phi_s} \sqrt{\frac{P_s}{N}} d_{sp,n}^{-\alpha/2} \mathbf{g}_{0n} \mathbf{b}_n x_{sn}, \end{aligned} \tag{32}$$

where P_p and P_s is the transmit power of each primary and secondary transmitter, respectively, $h_{0n} \in \mathbb{C}$ is the channel between the n^{th} primary transmitter T_{pn} and a primary receiver R_{p0}, $\mathbf{g}_{0n} \in \mathbb{C}^{1\times N}$ is the channel vector between the n^{th} secondary transmitter T_{sn} with N_t antennas and R_{p0}, $d_{pp,n}$ and $d_{sp,n}$ is the distance between T_{pn} and R_{p0}, and T_{sn} and R_{p0}, respectively, x_{pn} and x_{sn} are data signals transmitted from T_{pn} and T_{sn}, respectively, with $x_{pn}, x_{sn} \sim \mathcal{C}N(0,1)$.

The second term of (32) corresponds to the interference received from primary transmitters at the primary receiver R_{p0}, while the third term corresponds to the interference received from secondary transmitters at the primary receiver R_{p0}.

We consider the interference limited regime, i.e. noise power is negligible compared to the interference power, and drop the AWGN contribution. We assume that each h_{0n}, and each entry of $\mathbf{g}_{0n}$ is i.i.d. Rayleigh distributed.

Similarly, the $\mathbb{C}^{N_r \times 1}$ received signal $\mathbf{v}_0$ at the secondary receiver R_{s0} is

$$\begin{aligned}\mathbf{v}_0 &= \sqrt{\frac{P_s}{N}} d_s^{-\alpha/2} \mathbf{Q}_{00} \mathbf{b}_0 x_{s0} + \sum_{n: T_{sn} \in \Phi_s \setminus \{T_{s0}\}} \sqrt{\frac{P_s}{N}} d_{ss,n}^{-\alpha/2} \mathbf{Q}_{0n} \mathbf{b}_n x_{sn} \\ &\quad + \sum_{n: T_{pn} \in \Phi_p} \sqrt{P_p} d_{ps,n}^{-\alpha/2} \mathbf{f}_{0n} x_{pn}, \end{aligned} \tag{33}$$

where $d_{ss,n}$ and $d_{ps,n}$ is the distance between T_{sn} and R_{s0}, and T_{pn} and R_{s0}, respectively, $\mathbf{Q}_{0n} \in \mathbb{C}^{N_r \times N}$ is the multiple antenna channel between the secondary transmitter T_{sn} and secondary receiver R_{s0}, $\mathbf{f}_{0n} \in \mathbb{C}^{N_r \times 1}$ is the channel vector between T_{pn} and R_{s0}. Each of the channel coefficients are assumed to be Rayleigh distributed.

Secondary Transmitter Interference Suppression: To minimize the interference caused at primary receivers, the N_t transmit antennas at each secondary transmitter are used to suppress interference towards its $N_t - 1$ nearest primary receivers. Thus, the beamformer (suppressing vector) employed by the n^{th} secondary transmitter $\mathbf{b}_n$ lies in the null space of the channel vectors of the N_t nearest primary receivers, i.e., $[\mathbf{g}_{1n}^\dagger \cdots \mathbf{g}_{N_t-1n}^\dagger]$, to suppress the interference towards its $N_t - 1$ nearest primary receivers.

Remark 5.1 *Note that each secondary transmitter nulls/suppresses its signal towards its $N_t - 1$ nearest primary receivers. However, from a primary receiver's perspective this does not translate into not receiving any interference from its $N_t - 1$ nearest secondary transmitters.*

Let N_{supp} be the random variable denoting the number of consecutive nearest secondary interferers that appear suppressed at the typical primary receiver R_{p0}. For example, as shown in Fig. 9, each secondary transmitter tries to suppress interference towards its 3 nearest primary receivers. A dashed line indicates suppressed interferer while a solid line indicates non-suppressed interferer. In Fig. 9, we can see that the primary receiver R_{p0} can still receive interference from its second nearest secondary transmitter T_{s1}, in which case $N_{\mathsf{supp}} = 1$.

With $N_{\mathsf{supp}} = \mathsf{c}$ nearest secondary interferers suppressed at primary receiver R_{p0}, the interference received from both the primary and secondary transmitters at the primary receiver R_{p0} in (32) is

$$I_{mimo}(\mathsf{c}) = \underbrace{\sum_{n: T_{pn} \in \Phi_p \setminus \{T_{p0}\}} P_p d_{pp,n}^{-\alpha} |h_{0n}|^2}_{I_{pp}} + \underbrace{\sum_{n: n > \mathsf{c},\ T_{sn} \in \Phi_s} P_s d_{sp,n}^{-\alpha} |\mathbf{g}_{0n} \mathbf{b}_n|^2}_{I_{sp}^{\mathsf{c}}}. \tag{34}$$

Thus, with signal model (32), the SIR at R_{p0} is

$$\mathsf{SIR}_p = \frac{P_p d_p^{-\alpha} |h_{00}|^2}{I_{mimo}(\mathsf{c})}. \tag{35}$$

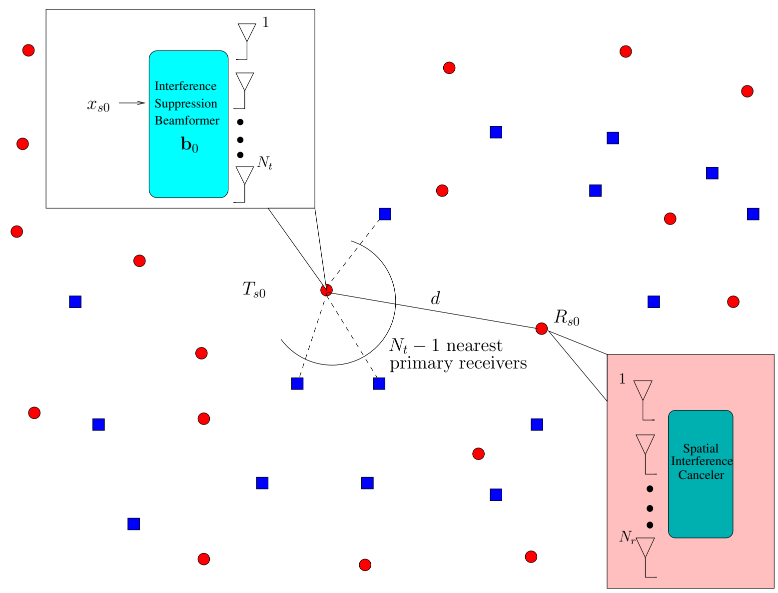

Figure 8: Transmit-receive strategy of secondary transmitters and receivers (dots) and primary transmitters and receivers (squares), where each secondary transmitter suppresses its interference towards its $N_t - 1$ nearest primary receivers.

Secondary Receiver Interference Cancelation: Similar to Section 3, we consider the use of a partial ZF decoder at each secondary receiver, that uses its m SRDOF for canceling the nearest interferers from the union of the primary and the secondary interferers, and the remaining $N - m$ SRDOF are used for decoding the signal of interest. Since each primary and secondary transmitter sends a single data stream, the number of interferers that can be canceled at each secondary receiver using m SRDOF is m.

For interference cancelation, let the n^{th} secondary receiver multiply $\mathbf{t}_n^\dagger$ to the received signal (33). Then $\mathbf{t}_n^\dagger$ lies in the null space of channel vectors corresponding to its m nearest interferers from $\{\Phi_p \cup \Phi_s\} \backslash \{T_{sn}\}$ chosen such that it maximizes the signal power $|\mathbf{t}_n^\dagger \mathbf{Q}_{nn} \mathbf{b}_n|^2$ in (33).

Thus, from (33), the SIR at the secondary receiver R_{s0} is

$$\mathsf{SIR}_s = \frac{P_s d_s^{-\alpha} |\mathbf{t}_0^\dagger \mathbf{Q}_{00} \mathbf{b}_0|^2}{\sum_{n: T_{sn} \in \Phi_s \backslash \{T_{s0}\}} P_s d_{ss,n}^{-\alpha} |\mathbf{t}_n^\dagger \mathbf{Q}_{0n} \mathbf{b}_n|^2 + \sum_{n: T_{pn} \in \Phi_p} P_p d_{ps,n}^{-\alpha} |\mathbf{t}_0^\dagger \mathbf{f}_{0n}|^2}, \quad (36)$$

where the beamforming vector $\mathbf{b}_n$ used by secondary transmitter T_{sn} lies in the null space of $[\mathbf{g}_{1n}^\dagger \ldots \mathbf{g}_{N-1n}^\dagger]$ to suppress the interference from T_{sn} towards its $N_t - 1$ nearest primary receivers, and $\mathbf{t}_n$ lies in the null space of channel vectors corresponding to the m nearest interferers of R_{s0} from $\{\Phi_p \cup \Phi_s\} \backslash \{T_{sn}\}$ chosen such that it maximizes

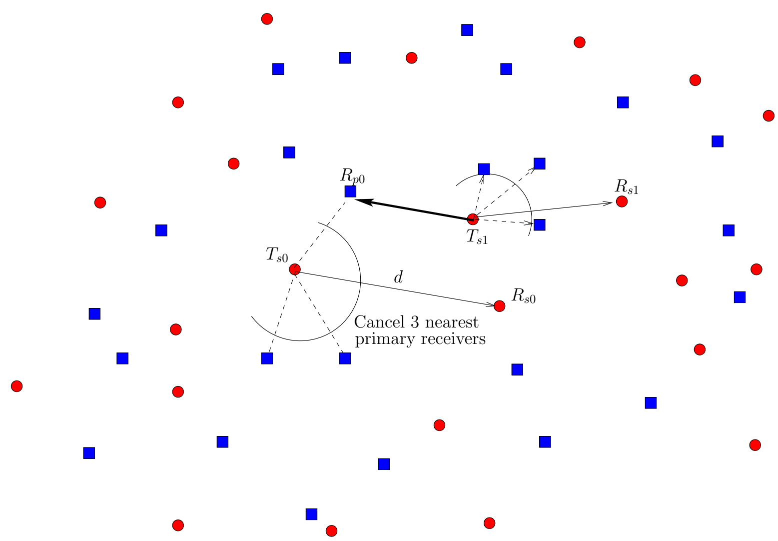

Figure 9: Each dot (secondary transmitter) suppresses its interference towards its 3 nearest squares (primary receivers) denoted by dashed lines, but still a blue node can receive interference from one of its 3 nearest red nodes.

the signal power $|\mathbf{t}^\dagger \mathbf{Q}_{nn}\mathbf{b}_n|^2$. From Lemma 5.2, optimal

$$\mathbf{t}_n = \frac{(\mathbf{Q}_{nn}\mathbf{u})^\dagger \mathsf{S}\mathsf{S}^\dagger}{|(\mathbf{Q}_{nn}\mathbf{u})^\dagger \mathsf{S}\mathsf{S}^\dagger|},$$

where $\mathsf{S} \in \mathbb{C}^{N_r \times N_r - m}$ is the orthonormal basis of the null space of channel vectors corresponding to the m nearest interferers of R_{s0} from $\Phi_p \cup \Phi_s \backslash \{T_{sn}\}$.

Lemma 5.2 *The signal power $s = |\mathbf{t}_0^\dagger \mathbf{Q}_{00}\mathbf{b}_0|^2$ in (36) at the secondary receiver with $\mathbf{t}_n = \frac{(\mathbf{Q}_{nn}\mathbf{u})^\dagger \mathsf{S}\mathsf{S}^\dagger}{|(\mathbf{Q}_{nn}\mathbf{u})^\dagger \mathsf{S}\mathsf{S}^\dagger|}$ is $\sim \chi^2(2(N_r - m))$. The interference power at secondary receiver from the secondary transmitter n in (36), $\mathsf{pow}_{ss}^{0n} = |\mathbf{t}_0^\dagger \mathbf{q}_{0n}\mathbf{b}_n|^2$, and the interference power at secondary receiver from the primary transmitter n in (36), $\mathsf{pow}_{ps}^{0n} = |\mathbf{t}_0^\dagger \mathbf{f}_{0n}|^2$ is $\sim \chi^2(2)$.*

Proof: The first statement follows from Lemma 3.1. The second and third statement follows since $\mathbf{t}_0^\dagger$, $\mathbf{b}_n$, and $\mathbf{q}_{0n}$ are independent, and since each entry of $\mathbf{q}_{0n}, \mathbf{f}_{0n} \sim \mathcal{CN}(0,1)$. $\square$

We next present an alternate way of representing the interference term in (36), that allows both easy analysis and simpler notation.

Lemma 5.3 *The interference term in (36)*

$$I_s = \sum_{n:T_{sn} \in \Phi_s \backslash \{T_{s0}\}} P_s d_{ss,n}^{-\alpha} \mathsf{pow}_{ss}^{0n} + \sum_{n:T_{pn} \in \Phi_p} P_p d_{ps,n}^{-\alpha} \mathsf{pow}_{ps}^{0n}$$

received at the typical secondary receiver R_{s0} can also be written as

$$\sum_{n:T\in\Phi\setminus\{T_{s0}\}} P_n d_n^{-\alpha} \mathsf{pow}^{0n},$$

where pow^{0n} *is* $\sim \chi^2(2)$, $\Phi = \{\Phi_s \cup \Phi_p\}$, *and* P_n *is a binary random variable which takes value* P_p *with probability* $\frac{\lambda_p}{\lambda_p+\lambda_s}$, *and value* P_s *with probability* $\frac{\lambda_s}{\lambda_p+\lambda_s}$.

Essentially, the Lemma says that the aggregate interference seen at a secondary receiver can be thought of as interference coming for a single PPP Φ that is a union of the primary and the secondary transmitter's PPP, and where each node of PPP Φ transmits with either power P_s or P_p with probability $\frac{\lambda_s}{\lambda_p+\lambda_s}$ and $\frac{\lambda_p}{\lambda_p+\lambda_s}$, respectively.

Proof: Since the superposition of two independent PPP's is a PPP, consider the union of Φ_p and Φ_s that are independent as a single PPP $\Phi = \{\Phi_s \cup \Phi_p\}$. Thus, the interference received at the typical secondary receiver R_{s0} is derived from the transmitters corresponding to Φ with channel gains pow_{ss}^{0n} or pow_{ps}^{0n}, where both pow_{ss}^{0n} and pow_{ps}^{0n} are $\sim \chi^2(2)$, and is denoted as pow. Note that the primary transmitters use power P_p, and the secondary transmitters use power P_s. The probability that any randomly chosen node of Φ belongs to Φ_p is $\frac{\lambda_p}{\lambda_p+\lambda_s}$, hence the power transmitted by any node of Φ is P_p with probability $\frac{\lambda_p}{\lambda_p+\lambda_s}$, and P_s with probability $\frac{\lambda_s}{\lambda_p+\lambda_s}$. □

Thus, we can write the SIR expression (36) at the typical secondary receiver R_{s0} after canceling the m nearest interferers from $\Phi = \Phi_p \cup \Phi_s$ at secondary receiver R_{s0}, more compactly as

$$\mathsf{SIR}_s = \frac{P_s d_s^{-\alpha} |\mathbf{t}_0^\dagger \mathbf{Q}_{00} \mathbf{b}_0|^2}{\sum_{n>m,\ T_n\in\Phi\setminus\{T_{s0}\}} P_n d_n^{-\alpha} \mathsf{pow}^{0n}}. \tag{37}$$

We assume that the rate of transmission for each primary (secondary) transmitter is $R_p = \log(1+\beta_p)$ ($R_s = \log(1+\beta_s)$) bits/sec/Hz. Therefore, a packet transmitted by T_{p0} (T_{s0}) can be successfully decoded at R_{p0} (R_{s0}), if $\mathsf{SIR}_p \ge \beta_p$ ($\mathsf{SIR}_s \ge \beta_s$). Without the presence of secondary network, the SIR at the primary receiver R_{p0} is

$$\mathsf{SIR}_p^{nc} = \frac{P_p d_p^{-\alpha} |h_{00}|^2}{\sum_{n:T\in\Phi_p\setminus\{T_{s0}\}} P_n d_n^{-\alpha} \mathsf{pow}^{0n}}. \tag{38}$$

Primary Network Outage Model: For a given rate R_p bits/sec/Hz and primary outage probability constraint ϵ_p, let $\lambda_p^\star$ be the maximum density for which the outage probability of the primary network

$$P_{p,out}^{nc} = \mathbb{P}\left(\mathsf{SIR}_p^{nc} \le \beta_p\right) \le \epsilon_p. \tag{39}$$

From Theorem **??**, $\lambda_p^\star = \frac{\ln(1-\epsilon_p)}{c\beta_p^{\frac{2}{\alpha}} d_p^2}$. We assume that the primary network operates at the largest permissible density $\lambda_p^\star$.

Allowing secondary transmissions to co-exist with the primary transmissions, increases the interference received at R_{p0} as quantified in SIR_p (35) compared to SIR_p^{nc} (38), and thereby increases the outage probability from $P_{p,out}(\beta_p)^{nc}$ (39) to

$$P_{p,out}(\beta_p) = \mathbb{P}\left(\mathsf{SIR}_p \le \beta_p\right). \tag{40}$$

Thus, if the primary outage probability constraint is fixed at ϵ_p, and the primary network density is $\lambda_p^\star$, the density of the secondary transmitters cannot be non-zero.

To make the problem non-trivial, we consider an increased outage probability tolerance at the primary receivers of $\epsilon_p + \Delta_p$.

Secondary Network Outage Model: For the secondary network we consider the usual outage probability constraint of $P_{s,out}(\beta_s) = \mathbb{P}(\mathsf{SIR}_s \le \beta_s) \le \epsilon_s$. Thus, we want to find the maximum density of secondary transmitters λ_s satisfying both the outage constraint at primary receivers $P_{p,out}(\beta_p) \le \epsilon_p + \Delta_p$, and secondary receivers $P_{s,out}(\beta_s) = \mathbb{P}(\mathsf{SIR}_s \le \beta_s) \le \epsilon_s$ for primary nodes' density $\lambda_p^\star$. Thus, the maximum density of the secondary network is

$$\lambda_s^\star = \max_{P_{p,out}(\beta_p) \le \epsilon_p + \Delta_p,\ P_{s,out}(\beta_s) \le \epsilon_s} \lambda.$$

Consequently, the transmission capacity of the secondary network is defined as

$$C_s = \lambda_s^\star (1 - \epsilon_s) R_s \text{ bits/sec/Hz/m}^2.$$

In the following, we derive $\lambda_s^\star$ as a function of secondary transmit (N_t) and receive (N_r) antennas via computing the outage probabilities. To compute the outage probability $P_{p,out}$ and $P_{s,out}(\beta_s)$, we once again consider a typical transmitter receiver pair (T_{p0}, R_{p0}) and (T_{s0}, R_{s0}), respectively.

We next state the main Theorem of this section, on the scaling of transmission capacity of secondary nodes with multiple antennas under a primary and secondary outage probability constraint.

Theorem 5.4 *When each secondary transmitter uses $N_t - 1$ STDOF for suppressing interference towards its $N_t - 1$ nearest primary receivers, and each secondary receiver uses m SRDOF for canceling the m nearest interferers from $\{\Phi_s \cup \Phi_p\} \backslash \{T_{s0}\}$, then*

$$C_s = \Omega\left(\min\{N_r, {N_t}^{1-\frac{2}{\alpha}}\}\right), \text{ and } C_s = \mathcal{O}\left(\min\{N_t, N_r\}\right),$$

and $m = \theta N_r,\ \theta \in (0,1]$ maximizes the lower bound on the transmission capacity of the secondary wireless network.

Theorem 5.4 highlights the dependence of transmission capacity of the secondary network on the number of transmit and receive antennas, when multiple antennas are allowed to exploit their full capability; perform interference suppression at the transmit side and interference cancelation at the receive side. It also identifies that increasing only the transmit or receiver antennas is futile and to get non-vanishing gain, both the transmit and receive antennas have to be increased simultaneously, which is expected since there are two outage probability constraints.

Theorem 5.4 shows that if the number of transmit antennas is much larger than the number of receive antennas $N_t >> N_r$, then the transmission capacity increases linearly with N_r, the number of receive antennas. With large number of transmit antennas at the secondary nodes $N_t >> N_r$, each secondary transmitter can suppress its interference towards a very large number of primary receivers and hence the outage probability constraint at each primary receiver is always met. Thus, with $N_t >> N_r$,

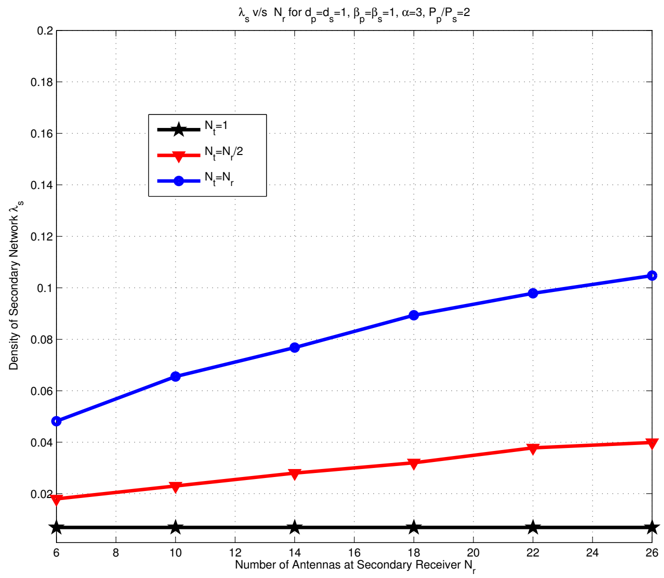

Figure 10: Density of the secondary network with respect to number of transmit and receive antennas N_t, N_r at the secondary nodes.

only the outage probability constraint at the secondary nodes is active, and the situation is identical to that of Theorem 3.5, where each transmitter has a single antenna and each receiver has N_r antennas with a single outage probability constraint, and hence the result is identical to that of Theorem 3.5.

When $N_r >> N_t$, the transmission capacity is limited by the interference suppression capability of secondary transmitters, and Theorem 5.4 shows that the transmission capacity scales at least as ${N_t}^{1-2/\alpha}$. This result is intuitive since larger the path-loss exponent α, less is the interference caused by each secondary transmitter at any primary receiver.

In Fig. 10, we plot the density of the secondary network with respect to the number of secondary transmit and receive antennas N_t and N_r for outage probabilities $\epsilon_p = \epsilon_s = .1$. We see that for $N_t = N_r$, the density of the secondary network scales sublinearly with N_t, however, for $N_t = 1$ the density of the secondary network is constant as expected.

Proof: Similar to the proof of Theorem 3.5, we will first find an upper and a lower bound on the outage probability, but in this case we have two outage probabilities to bound, one at the primary receiver, and the other at the secondary receiver. The outage probability bounds for the secondary receiver follow from Theorems 3.8 and 3.10, since the secondary receiver employs partial ZF decoder for interference cancelation, similar to Section 3. Thus, we only need to derive the bounds for the outage probability expression at the primary receiver in Theorem 5.5, when each secondary transmitter uses its $N_t - 1$ STDOF for interference suppression.

Considering the relaxed outage probability constraint of $\epsilon_p + \Delta_p$ at any primary receiver when each secondary transmitter uses $k = N_t - 1$ STDOF for interference suppression, from Theorem 5.5,

$$\lambda_s^\star = \Omega\left({N_t}^{1-\frac{2}{\alpha}}\right), \text{and } \lambda_s^\star = \mathcal{O}\left(N_t\right). \tag{41}$$

Next, we consider the outage probability constraint of ϵ_s on each secondary receiver. From (37), the outage probability at the secondary receiver R_{s0} is

$$P_{s,out}(\beta_s) \quad = \quad \mathbb{P}\left(\frac{P_s d_s^{-\alpha}|\mathbf{t}_0^\dagger \mathbf{Q}_{00}\mathbf{b}_0|^2}{\sum_{n>m,\ T_n\in\Phi\backslash\{T_{s0}\}} P_n d_n^{-\alpha}\mathsf{pow}^{0n}} \le \beta_s\right),$$

where $\Phi = \{\Phi_s \cup \Phi_p\}$, and interference power $\mathsf{pow}^{0n} \sim \chi^2(2)$ (Lemma 5.2) and signal power $|\mathbf{t}_0^\dagger \mathbf{Q}_{00}\mathbf{b}_0|^2 \sim \chi^2(2(N_r - m))$ (Lemma 5.2). Thus, with m SRDOF used for interference cancelation at each secondary receiver, from Theorems 3.8 and 3.10, for $\Phi = \Phi_p \cup \Phi_s$ with density $\lambda_p + \lambda_s$, we get for any $r > 1$,[3] with a single data stream transmission $k = 1$,

$$P_{s,out}(\beta_s) \ge 1 - \frac{(N_r - m)(m + r + \frac{\alpha}{2})^{\frac{2}{\alpha}}}{(r-1)\frac{d_p^\alpha \beta_s}{P_s}(\pi(\lambda_s + \lambda_p))^{\frac{\alpha}{2}}}\left(\frac{\lambda_p}{P_p(\lambda_p + \lambda_s)} + \frac{\lambda_s}{P_s(\lambda_p + \lambda_s)}\right), \tag{42}$$

and

$$P_{s,out}(\beta_s) \le \frac{(\pi(\lambda_p + \lambda_s))^{\frac{\alpha}{2}}\beta_p d_p^\alpha\left(\frac{\alpha}{2} - 1\right)^{-1}\left(m - \lceil\frac{\alpha}{2}\rceil\right)^{1-\frac{\alpha}{2}}}{N_r - m - 1}\left(\frac{\lambda_p P_p}{\lambda_p + \lambda_s} + \frac{\lambda_s P_s}{\lambda_p + \lambda_s}\right), \tag{43}$$

respectively, where we have taken the expectation with respect to power transmitted P_n by any node of $\Phi = \Phi_p \cup \Phi_s$, which is a binary random variable taking values P_p and P_s, with probability $\frac{\lambda_p}{\lambda_p+\lambda_s}$ and $\frac{\lambda_s}{\lambda_p+\lambda_s}$, respectively.

From the lower bound on the outage probability (42), we get that

$$\lambda_s^\star = \mathcal{O}(N_r), \tag{44}$$

by choosing $r = {N_r}^{2/\alpha}$, similar to Theorem 3.5, by fixing outage probability $P_{s,out}(\beta_s) = \epsilon_s$. Moreover, with $m = \theta N_r,\ \theta \in (0, 1]$, using the upper bound on the outage probability (43), we get

$$\lambda_s^\star = \Omega(N_r) \tag{45}$$

Hence considering both the outage probability constraints together, from (41), (44) and (45), we get

$$\lambda_s^\star = \Omega\left(\min\{N_r, {N_t}^{1-\frac{2}{\alpha}}\}\right), \text{ and } \lambda_s^\star = \mathcal{O}\left(\min\{N_r, N_t\}\right).$$

□

[3] r represents the number of nearest uncanceled interferers considered for bounding the outage probability.

Finally, we prove a more general result than required by Theorem 5.4, where we derive bounds on the outage probability at any primary receiver when each secondary transmitter uses k STDOF for interference suppression towards its k nearest primary receivers. The proof of Theorem 5.4 is slightly long and complicated. For Theorem 5.4, we fix $k = N_t - 1$ to get (41).

Theorem 5.5 *If k STDOF are used for interference suppression at each secondary transmitter, then*

$$\lambda_s^\star = \Omega\left(k^{1-\frac{2}{\alpha}}\right), \text{ and } \lambda_s^\star = \mathcal{O}\left(k\right).$$

Proof: Since we are interested in establishing the scaling behavior of the density of the secondary network with respect to N_t, we consider the case when both N_t and k, the number of STDOF used for interference suppression, are large enough. We bound the outage probability (40) at a typical primary receiver, and find the density of secondary network λ_s that satisfies the primary outage constraint of $P_{p,out} = \epsilon_p + \Delta_p$.

Lower Bound: Recall that N_{supp} is the random variable representing the number of consecutive nearest secondary interferers suppressed at the typical primary receiver R_{p0}. Let $N_{\mathsf{supp}} = \mathsf{c}$, and recall the definition of interference received at R_{p0}, $I_{mimo}(\mathsf{c}) = I_{pp} + I_{sp}^{\mathsf{c}}$ from (34), where I_{pp} is the interference contribution from primary transmitters other than T_{p0}, and I_{sp}^{c} is the secondary transmitters other than the c consecutive nearest secondary interferers, at the primary receiver R_{p0}.

From (40) and (35), the outage probability at primary receiver R_{p0} is $P_{p,out}$

$$\begin{aligned}
&= \mathbb{E}_{N_{\mathsf{supp}}}\left\{\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(\mathsf{c})} \le \beta_p\right)\right\},\\
&\stackrel{(a)}{=} \mathbb{E}_{N_{\mathsf{supp}}}\left\{\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(\mathsf{c})} \le \beta_p\right)\middle| N_{\mathsf{supp}} < \lfloor k/\eta\rfloor\right\} \times \mathbb{P}\left(N_{\mathsf{supp}} < \lfloor k/\eta\rfloor\right)\\
&\quad + \mathbb{E}_{N_{\mathsf{supp}}}\left\{\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(\mathsf{c})} \le \beta_p\right)\middle| N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right\} \times \mathbb{P}\left(N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right),\\
&\stackrel{(b)}{\le} \delta + \mathbb{E}_{N_{\mathsf{supp}}}\left\{\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(\mathsf{c})} \le \beta_p\right)\middle| N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right\} \times \mathbb{P}\left(N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right),\\
&\stackrel{(c)}{\le} \delta + \mathbb{E}_{N_{\mathsf{supp}}}\left\{1 - \mathbb{E}\left\{\exp\left(-\frac{\beta_p I_{mimo}(\mathsf{c}) d_p^\alpha}{P_p}\right)\right\}\middle| N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right\},\\
&\stackrel{(d)}{=} \delta + \mathbb{E}_{N_{\mathsf{supp}}}\left\{1 - \int_0^\infty \exp\left(-\beta_p (I_{pp}) d_p^\alpha\right) f_{I_{pp}}(s) ds\right.\\
&\quad \left.\int_0^\infty \exp\left(-\frac{\beta_p (I_{sp}^{\mathsf{c}}) P_s d^\alpha}{P_p}\right) f_{I_{sp}^{\mathsf{c}}}(t) dt \middle| N_{\mathsf{supp}} \ge \lfloor k/\eta\rfloor\right\},\\
&\stackrel{(e)}{=} \delta + \mathbb{E}_{N_{\mathsf{supp}}}\left\{1 - \mathcal{L}_{I_{pp}}\left(\beta_p d_p^\alpha\right)\right.\\
&\quad \left.\left(1 - \mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{\sum_{n:\, n>\mathsf{c}, T_{sn}\in\Phi_s} P_s d_{sp,n}^{-\alpha}|g_{0n}|^2} \le \beta_p\right)\right)\middle| C \ge \lfloor k/\eta\rfloor\right\},
\end{aligned} \tag{46}$$

where (a) follows by splitting the expectation over conditioning the event $N_{\mathsf{supp}} < \lfloor k/\eta \rfloor$ where η is a constant, (b) follows by letting $\eta \in \mathbb{N}$ such that $\mathbb{P}\left(N_{\mathsf{supp}} < \lfloor k/\eta \rfloor\right) \le \delta,\ \delta \le \Delta_p$, where Δ_p is the additional tolerance of outage probability at the primary receivers, and η is independent of k. Existence of $\eta \in \mathbb{N}$ such that $\mathbb{P}\left(N_{\mathsf{supp}} < \lfloor k/\eta \rfloor\right) \le \Delta_p$ is guaranteed, since for large values of k, canceling only a few nearest secondary interferers has a very small probability. Inequality (c) follows by taking expectation with respect to $|h_{00}|^2 \sim \exp(1)$ and using $\mathbb{P}\left(N_{\mathsf{supp}} \ge \lfloor k/\eta \rfloor\right) \le 1$. Equality (d) follows since $I_{mimo}(\mathsf{c}) = I_{pp} + I_{sp}^{\mathsf{c}}$, and I_{pp} and I_{sp}^{c} are independent. Equality (e) follows by defining $\mathcal{L}_I(.)$ as the Laplace transform of I, and noting that

$$\int_0^{\infty} \exp\left(-\frac{\beta_p(I_{sp}^{\mathsf{c}})P_s d^{\alpha}}{P_p}\right) f_{I_{sp}^{\mathsf{c}}}(t)dt = \mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{sp}^{\mathsf{c}}} > \beta_p\right), \tag{47}$$

since $|h_{00}|^2 \sim \exp(1)$, and where the expectation in the R.H.S. is taken only with respect to $|h_{00}|^2$. Thus, from (46)

$$\begin{aligned}
P_{p,out} &\stackrel{(f)}{\le} \delta + \mathbb{E}_{N_{\mathsf{supp}}}\left\{1 - \exp\left(-\lambda_p c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right)\right. \\
&\qquad \left.\left(1 - (\pi\lambda_s)^{\frac{\alpha}{2}}\beta_p\left(\frac{P_s}{P_p}\right)d_p^{\alpha}\left(\frac{\alpha}{2}-1\right)^{-1}\left(\mathsf{c} - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}\right)\middle| N_{\mathsf{supp}} \ge \lfloor k/\eta \rfloor\right\}, \\
&= \delta + 1 - \exp\left(-\lambda_p c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right) + \exp\left(-\lambda_p c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right)(\pi\lambda_s)^{\frac{\alpha}{2}}\beta_p \\
&\qquad \left(\frac{P_s}{P_p}\right)d_p^{\alpha}\left(\left(\frac{\alpha}{2}-1\right)^{-1}\mathbb{E}_{N_{\mathsf{supp}}}\left\{\left(\mathsf{c} - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}\middle|\ N_{\mathsf{supp}} \ge \lfloor k/\eta \rfloor\right\}\right), \\
&\stackrel{(g)}{\le} \delta + \epsilon_p + \exp\left(-\lambda_p c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right)(\pi\lambda_s)^{\frac{\alpha}{2}}\beta_p\left(\frac{P_s}{P_p}\right)d_p^{\alpha} \\
&\qquad \left(\left(\frac{\alpha}{2}-1\right)^{-1}\left(\lfloor k/\eta \rfloor + \left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}\right),
\end{aligned} \tag{48}$$

where (f) follows by using the lower bound on the success probability

$$\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{\sum_{n:\ n>\mathsf{c}, T_{sn}\in\Phi_s} P_s d_{sp,n}^{-\alpha}|g_{0n}|^2} > \beta_p\right)$$

from Theorem 3.10, by substituting $k = 1$ data stream, and $N_r - k - m = 1$, since the signal strength $|h_{00}|^2 \sim \chi^2(2)$ in this case. Finally (g) follows since for $N_{\mathsf{supp}} \ge \lfloor k/\eta \rfloor$, $\mathbb{E}_{N_{\mathsf{supp}}}\left\{(\mathsf{c} - \lceil\frac{\alpha}{2}\rceil)^{1-\frac{\alpha}{2}}|N_{\mathsf{supp}} \ge \lfloor k/\eta \rfloor\right\} \le (\lfloor k/\eta \rfloor - \lceil\frac{\alpha}{2}\rceil)^{1-\frac{\alpha}{2}}$ for $\alpha > 2$.

From the primary outage probability constraint in the absence of a secondary wireless network (Theorem **??**)

$$\epsilon_p = 1 - \exp\left(-\lambda_p^{\star} c \beta_p^{\frac{2}{\alpha}} d_p^2\right),$$

where $\lambda_p^{\star}$ is the largest density of primary nodes satisfying the outage constraint of ϵ_p.

Hence, equating (48) with the relaxed outage probability of $P_{p,out} = \epsilon_p + \Delta_p$ at each primary receiver, and substituting for ϵ_p, we get

$$\lambda_s \geq \frac{1}{\pi}\left(\frac{\Delta_p - \delta}{\exp\left(-\lambda_p^\star c_1 \beta_p^{\frac{2}{\alpha}}\right) d_p^2 \beta_p \left(\frac{P_s}{P_p}\right) d_p^\alpha \left(\left(\frac{\alpha}{2}-1\right)^{-1} \left(\lfloor k/\eta \rfloor + 1\right)^{1-\frac{\alpha}{2}} + c_3\right)^{\frac{2}{\alpha}}}\right),$$

and

$$\lambda_s = \Omega\left(k^{1-\frac{2}{\alpha}}\right). \tag{49}$$

Upper bound: To find an upper bound on λ_s, we consider the case when exactly k consecutive nearest secondary interferers are suppressed at each primary receiver. Clearly, when each secondary transmitter uses k STDOF for interference suppression towards the primary receivers, at best k consecutive nearest secondary interferers are suppressed at each primary receiver, thus yielding the upper bound. This can also be seen from Fig. 9, where each secondary transmitter tries to suppress interference towards its 3 nearest primary receivers.

Thus, from (40) and (35),

$$\begin{aligned}
P_{p,out}(\beta_p) &= \mathbb{E}_{N_{\mathsf{supp}}}\left\{\mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(\mathsf{c})} \leq \beta_p\right)\right\}, \\
&\stackrel{(a)}{\geq} \mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{I_{mimo}(k)} \leq \beta_p\right), \\
&\stackrel{(b)}{=} 1 - \mathbb{E}\left\{\exp\left(-\frac{\beta_p(P_p I_{pp} + P_s I_{sp}^k)d_p^\alpha}{P_p}\right)\right\}, \\
&= 1 - \int_0^\infty \exp\left(-\beta_p(I_{pp})d_p^\alpha\right) f_{I_{pp}}(s)ds \\
&\quad \int_0^\infty \exp\left(-\frac{\beta_p(I_{sp}^k)P_s d^\alpha}{P_p}\right) f_{I_{sp}^x}(t)dt, \\
&\stackrel{(c)}{=} 1 - \mathcal{L}_{I_{pp}}\left(\beta_p d_p^\alpha\right) \int_0^\infty \exp\left(-\frac{\beta_p(I_{sp}^k)P_s d^\alpha}{P_p}\right) f_{I_{sp}^x}(t)dt, \\
&\stackrel{(d)}{=} 1 - \mathcal{L}_{I_{pp}}\left(\beta_p d_p^\alpha\right)\left(1 - \mathbb{P}\left(\frac{P_p d_p^{-\alpha}|h_{00}|^2}{\sum_{n>k, T_{sn}\in\Phi_s} P_s d_{sp,n}^{-\alpha}|g_{0n}|^2} \leq \beta_p\right)\right), \\
&\stackrel{(e)}{\geq} 1 - \exp\left(-\lambda_p^\star c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right) \frac{\left(k + r + \frac{\alpha}{2}\right)^{\frac{\alpha}{2}}}{d^\alpha \beta (\pi\lambda_s)^{\frac{2}{\alpha}}},
\end{aligned}$$

where (a) follows from the fact that using k STDOF for interference suppression by each secondary transmitter, at best k consecutive nearest secondary interferers are suppressed at each primary receiver, (b) follows by definition of $I_{mimo}(k)$ (34). Equality (c) follows since the Laplace transform of I_{pp}, the interference contribution from PPP

Φ_p with density $\lambda_p^\star$, evaluated at $\left(\beta_p d_p^\alpha\right)$ is

$$\mathcal{L}_{I_{pp}}\left(\beta_p d_p^\alpha\right) = \exp\left(-\beta_p(I_{pp})d_p^\alpha\right) f_{I_{pp}}(s)ds = \exp\left(-\lambda_p^\star c_1 \beta_p^{\frac{2}{\alpha}} d_p^2\right).$$

Equality (d) follows similar to (47), since $|h_{00}|^2 \sim \exp(1)$, and finally (e) follows from the lower bound on outage probability (Theorem 3.8), since the signal power is $\sim \chi^2(2)$ instead of $N-m-k+1$ for Theorem 3.8 and $N_{\mathsf{canc}} = k$ nearest interferers are canceled at each primary receiver.

Thus, we get the upper bound

$$\lambda_s = \mathcal{O}(k). \tag{50}$$

Combining (49), and (50),

$$\lambda_s^\star = \Omega\left(k^{1-\frac{2}{\alpha}}\right), \text{ and } \lambda_s^\star = \mathcal{O}\left(k\right).$$

□

6 Reference Notes

The results presented in Section 3 and 4 can be found in [3]. The study of transmission capacity with multiple antennas was initiated in [8], followed up in [4, 5, 7, 9], and mostly settled in [3]. Results on multiple antennas in cellular networks can be found in [10–13]. Results on space-division multiple access with multiple antennas can be found in [15], and impact of multiple antennas with scheduling can be found in [16]. The results of Section 5 with multiple antennas in overlaid networks are presented from [17]. Transmission capacity result for single antenna equipped secondary nodes can be found in [18–23, 23–26].

References

[1] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] R. J. Muirhead, *Aspects of multivariate statistical theory*. Wiley. com, 2009, vol. 197.

[3] R. Vaze and R. Heath, "Transmission capacity of ad-hoc networks with multiple antennas using transmit stream adaptation and interference cancellation," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 780 –792, Feb. 2012.

[4] N. Jindal, J. Andrews, and S. Weber, "Rethinking MIMO for wireless networks: Linear throughput increases with multiple receive antennas," in *IEEE International Conference on Communications, 2009. ICC '09.*, June 2009, pp. 1–6.

[5] ——, "Multi-antenna communication in ad hoc networks: Achieving MIMO gains with SIMO transmission," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 529–540, 2011.

[6] D. Kershaw, "Some extensions of W. Gautschis inequalities for the gamma function," *Mathematics of Computation*, vol. 41, no. 164, pp. 607–611, Oct. 1983.

[7] K. Huang, J. Andrews, D. Guo, R. Heath, and R. Berry, "Spatial interference cancellation for multiantenna mobile ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1660–1676, 2012.

[8] A. M. Hunter, J. G. Andrews, and S. P. Weber, "Capacity scaling of ad hoc networks with spatial diversity," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5058–71, Dec. 2008.

[9] R. Louie, M. McKay, and I. Collings, "Spatial multiplexing with MRC and ZF receivers in ad hoc networks," in *IEEE International Conference on Communications, 2009. ICC '09.*, June 2009, pp. 1–5.

[10] M. Demirkol and M. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proc. IEEE Vehicular Technology Conference (VTC 2001)*, 2001.

[11] R. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, June 2003.

[12] A. Lozano and A. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec 2002.

[13] A. Moustakas, S. Simon, and A. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large n analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[14] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*. PhD. Dissertation, MIT, 1989.

[15] M. Kountouris and J. G. Andrews, "Transmission capacity scaling of sdma in wireless ad hoc networks," in *Information Theory Workshop, 2009. ITW 2009. IEEE*, 2009, pp. 534–538.

[16] A. M. Hunter, R. K. Ganti, and J. G. Andrews, "Transmission capacity of multi-antenna ad hoc networks with csma," in *Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2010*, 2010, pp. 1577–1581.

[17] R. Vaze, "Transmission capacity of spectrum sharing ad hoc networks with multiple antennas," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2334 –2340, July 2011.

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

[19] M. Vu and V. Tarokh, "Scaling laws of single-hop cognitive networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4089 –4097, august 2009.

[20] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," in *7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009.*, 23-27 2009, pp. 1 –5.

[21] C. Yin, C. Chen, and S. Cui, "Stable distribution based analysis of transmission capacities for overlaid wireless networks," in *International Conference on Wireless Communications Signal Processing, WCSP 2009.*, Nov. 2009, pp. 1 –5.

[22] C. Yin, C. Chen, T. Liu, and S. Cui, “Generalized results of transmission capacities for overlaid wireless networks,” in *IEEE International Symposium on Information Theory, ISIT 2009.*, June 28-July 3 2009, pp. 1774 –1778.

[23] K. Huang, V. Lau, and Y. Chen, “Spectrum sharing betwurln cellular and mobile ad hoc networks: transmission-capacity trade-off,” *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1256 –1267, september 2009.

[24] C. Yin, L. Gao, T. Liu, and S. Cui, “Transmission capacities for overlaid wireless ad hoc networks with outage constraints,” in *IEEE International Conference on Communications, 2009. ICC ’09.*, 14-18 2009, pp. 1 –5.

[25] O. Bakr, M. Johnson, R. Mudumbai, and K. Ramchandran, “Multi-antenna interference cancellation techniques for cognitive radio applications,” in *IEEE WCNC 2009.*, 5-8 2009, pp. 1 –6.

[26] C. Yin, L. Gao, and S. Cui, “Scaling laws for overlaid wireless networks: A cognitive radio network versus a primary network,” vol. PP, no. 99, pp. 1 –1, 2010.

Index