Rethinking MIMO for Wireless Networks: Linear Throughput Increases with Multiple Receive Antennas

Nihar Jindal
ECE Department
University of Minnesota
nihar@umn.edu

Jeffrey G. Andrews
ECE Department
University of Texas at Austin
jandrews@ece.utexas.edu

Steven Weber
ECE Department
Drexel University
sweber@ece.drexel.edu

***Abstract*— The benefit of multiple antenna communication is investigated in wireless ad hoc networks, and the primary finding is that throughput can be made to scale linearly with the number of receive antennas even if each transmitting node uses only a single antenna. The linear throughput gain is achieved by (i) using the receive antennas to cancel the signals of nearby interferers as well as to increase signal power (i.e., for array gain), and (ii) maintaining a constant per-link rate and increasing the spatial density of simultaneous transmissions linearly with the number of antennas at each receiver. Numerical results show that even a few receive antennas provide substantial throughput gains, thereby illustrating that the asymptotic linear scaling result is also indicative of performance for reasonable numbers of antennas.**

I. Introduction

Multiple antenna communication is now a key component of virtually every contemporary high-rate wireless standard (e.g., LTE, 802.11n, and WiMax), and the theoretical result that sparked the intense academic and industrial investigation of MIMO communication was the finding by Foschini and Gans that the achievable throughput of a point-to-point MIMO channel scales *linearly* with the minimum of the number of transmit and receive antennas [1].

In this paper we are interested in the throughput gains that multiple antennas can provide in *ad hoc networks*, rather than in point-to-point channels. If multiple antennas are added at each node in the network and point-to-point MIMO techniques are used to increase the rate of every individual link (i.e., every hop in a multi-hop route) in the network according to the Foschini-Gans linear scaling results, then network-wide throughput increases linearly with the number of antennas per node.

Based upon this and our understanding of MIMO communication in links, it may seem that linear throughput scaling is only possible if multiple antennas are used at both transmitting and receiving nodes. However, the main finding of this paper is that network-wide throughput scales linearly with the number of receive antennas. In other words, *linear scaling can be achieved in wireless networks using only receive antenna arrays, without using multiple transmit antennas.* Linear scaling is achieved by:

1) Using each receive array to cancel the interference contributions of nearby transmissions and to increase the received power of the desired signal (i.e., for array gain)
2) Maintaining a constant per-link rate, and linearly increasing (with the number of receive antennas) the spatial density of simultaneous transmissions.

Although additional receive antennas can be used to increase the per-link SINR/rate if the density of transmissions is kept fixed, it has been shown that this provides a logarithmic increase in the per-link rate, and thus only a logarithmic increase in total system throughput/area spectral efficiency [2]. On the other hand, we show that the density of simultaneous transmissions can be increased linearly with the number of receive antennas if a constant per-link SINR/rate is maintained, thereby yielding a linear increase in system throughput (defined in terms of area spectral efficiency or transport capacity).

We consider a network in which transmitters are randomly located on an infinite plane according to a 2-D homogeneous Poisson point process, and each transmitter attempts communication with a receiver a fixed distance away from it. We study this network using the *transmission capacity* framework, in which each transmitter attempts communication at a fixed rate and the performance metric is the network-wide success probability, which is computed over the random transmitter locations and the random channels. In particular, we are interested in finding the maximum transmitter/interferer density such that a target success probability is maintained and wish to quantify the rate at which this density can be increased as the number of antennas at each receiving node is increased.

This paper is closely related to a number of prior works that have studied the use of multiple receive antennas in ad hoc networks with Poisson distributed transmitters/interferers. In [2] the performance of a single TX-RX link surrounded by a field of interferers with a *fixed density* is studied, and it is shown that the average per-link SINR increases with the number of receive antennas N as $N^{\frac{\alpha}{2}}$ where α is the path-loss exponent. As mentioned before, this corresponds to only a logarithmic increase in per-link rate and overall system throughput. References [3] and [4] consider precisely the same model as here but study the performance of slightly different receiver designs that turn out to yield very different performance: in [3] the receive filter is chosen according to the maximum-ratio criterion and thus provides only array gain (i.e., no interference cancellation is performed)[1], while [4] considers the other extreme where the N antennas are used to cancel the strongest $N-1$ interferers but no array gain is obtained. Both receiver designs achieve only a *sub-linear* density increase with N (at rates $N^{\frac{2}{\alpha}}$ and $N^{1-\frac{2}{\alpha}}$, respectively). On the other hand, we show that using a fraction of the receive degrees of freedom to provide array gain and the other fraction to perform interference cancellation allows for *linear* density scaling with N.

II. System Model and Metrics

We consider a network in which the set of transmitters are located according to a 2-D homogeneous Poisson point process (PPP) of density λ. Each transmitter communicates with a receiver a distance d meters away from it (i.e., each receiver is randomly located on a circle of radius d centered around its associated transmitter). Note that the receivers are not a part of the transmitter PPP. Each transmitter uses only one antenna, while each receiver has N antennas.

By the stationarity of the Poisson process we can consider the performance of an arbitrary TX-RX pair, which we refer to as TX0 and RX0. From the perspective of RX0, the set of interferers (which is the entire transmit process with the exception of TX0) also form a homogeneous PPP due to Slivnyaks Theorem; see [5] for additional discussion of this point and further explanation of the basic model.

As a result, network-wide performance is identical to the performance of a single TX-RX link (separated by d meters) surrounded by an infinite number of interferers located in space according to a homogeneous PPP with density λ interferers/m^2. Assuming a path-loss exponent of α ($\alpha > 2$) and a frequency-flat channel, the N-dimensional received signal $\mathbf{y}_0$ at RX0 is given by:

$$\mathbf{y}_0 = d^{-\alpha/2}\mathbf{h}_0 u_0 + \sum_{i\in\Pi(\lambda)} |X_i|^{-\alpha/2}\mathbf{h}_i u_i + \mathbf{z}$$

where $|X_i|$ is the distance to the i-th transmitter/interferer, $\mathbf{h}_i$ is the vector channel from the i-th transmitter to RX0, $\mathbf{z}$ is complex Gaussian noise with covariance $\eta\mathbf{I}$, and u_i is the power-ρ symbol transmitted by the i-th transmitter. We assume that each of the vector channels $\mathbf{h}_i$ have iid unit-variance complex Gaussian components (corresponding to a rich scattering environment), independent across transmitters and independent of the distances $|X_i|$.

Throughout the paper we make the assumption that RX0 knows all incoming channels. In practice this could be accomplished (with some degree of channel estimation error) through transmission of pilot symbols and/or listening to interfering communications, as discussed in [4].

[1]Although reference [3] also considers strategies involving multiple transmit antennas, here we have mentioned only the directly relevant single transmit/multiple-receive scenario results.

If unit norm receive filter $\mathbf{v}_0$ is used, the resulting signal-to-interference-and-noise ratio SINR is:

$$\text{SINR} = \frac{\rho d^{-\alpha}|\mathbf{v}_0^\dagger \mathbf{h}_0|^2}{\eta + \sum_{i\in\Pi(\lambda)} \rho |X_i|^{-\alpha}|\mathbf{v}_0^\dagger \mathbf{h}_i|^2}. \quad (1)$$

For simplicity of notation, we assume that the distances $|X_i|$ are in decreasing order. This allows us to take advantage of the property that the ordered squared-distances $|X_1|^2, |X_2|^2, \ldots$ follow a 1-D Poisson point process with intensity $\pi\lambda$ [6]. To further simplify notation, we define the constant $\text{SNR} \triangleq \frac{\rho d^{-\alpha}}{\eta}$ as the interference-free signal-to-noise ratio. This allows SINR to be rewritten in normalized form as:

$$\text{SINR} = \frac{|\mathbf{v}_0^\dagger \mathbf{h}_0|^2}{\frac{1}{\text{SNR}} + d^{\alpha}\sum_{i\in\Pi(\lambda)} |X_i|^{-\alpha}|\mathbf{v}_0^\dagger \mathbf{h}_i|^2}. \quad (2)$$

The received SINR depends on the interferer locations and the vector channels, both of which are random, and the performance metric of interest is the outage probability with respective to a pre-defined SINR threshold β:

$$P_{\text{out}}(\lambda) = P\left[\text{SINR} \le \beta\right]. \quad (3)$$

The outage probability is clearly an increasing function of the density λ, and we are most interested in the maximum density such that the outage probability does not exceed some constant $\epsilon > 0$; this quantity is denoted by λ_ϵ.

Because there are λ_ϵ transmitters per m^2, the *transmitted* area spectral efficiency (assuming transmission at a spectral efficiency of $\log_2(1+\beta)$ bps/Hz, corresponding to use of Gaussian inputs at all terminals) is $\lambda_\epsilon \log_2(1+\beta)$ bps/Hz/m^2, while the *successful* area spectral efficiency (ASE) is $\lambda_\epsilon(1-\epsilon)\log_2(1+\beta)$ bps/Hz/m^2. An additional metric of interest that is related to overall system throughput is the *transport capacity*, which is defined as the product of rate and distance summed over all transmissions [7]. Since each transmission occurs over a distance of d meters, the transport capacity (per m^2) is the successful ASE multiplied by d. Note that all three system-wide metrics (transmitted and successful ASE and transport capacity) are linear in λ_ϵ. As a result, showing that λ_ϵ scales linearly in N implies that each of these system-wide metrics also scale linearly in N.

III. Partial Zero Forcing

Having setup the basic system model, we can now move on to describing the receiver structure (i.e., the receive filter) analyzed throughout the paper. Generally speaking, the N-dimensional receive filter can either boost the power of the desired signal (by choosing $\mathbf{v}_0$ in, or close to, the direction of $\mathbf{h}_0$) or cancel the interference contributions of certain transmitters (by choosing $\mathbf{v}_0$ orthogonal to the channels of particular interferers), or some combination of the two. Maximum-ratio combining (MRC) corresponds to maximizing the desired signal power $\mathbf{v}_0 = \frac{\mathbf{h}_0}{||\mathbf{h}_0||}$ without any regard to the nature of the interference, whereas choosing $\mathbf{v}_0$ orthogonal to $N-1$ interferer channels corresponds to using all available degrees of freedom for interference cancellation.

We consider a *partial zero-forcing* (PZF) receiver that uses some of the receive degrees of freedom for interference cancellation while using the remaining degrees of freedom to boost desired signal power. More specifically, the filter $\mathbf{v}_0$ is chosen orthogonal to the channel vectors of the k *nearest* interferers $\mathbf{h}_1, \ldots, \mathbf{h}_k$ for some $k \leq N-1$. Amongst the filters satisfying the orthogonality requirement $|\mathbf{v}_0^\dagger \mathbf{h}_i|^2 = 0$ for $i = 1, \ldots, k$, we are interested in the one that maximizes the desired signal power $|\mathbf{v}_0^\dagger \mathbf{h}_0|^2$. By simple geometry, this corresponds to choosing $\mathbf{v}_0$ in the direction of the projection of vector $\mathbf{h}_0$ on the nullspace of vectors $(\mathbf{h}_1, \ldots, \mathbf{h}_k)$.

More precisely, if the columns of the $N \times (N-k)$ matrix $\mathbf{Q}$ form an orthonormal basis for the nullspace of $(\mathbf{h}_1, \ldots, \mathbf{h}_k)$ – such a matrix can be found by performing a full QR decomposition of matrix $[\mathbf{h}_1 \mathbf{h}_2 \cdots \mathbf{h}_k]$ – then the receive filter is chosen as:

$$\mathbf{v}_0 = \frac{\mathbf{Q}^\dagger \mathbf{h}_0}{||\mathbf{Q}^\dagger \mathbf{h}_0||}. \tag{4}$$

Note that $k = 0$ and $k = N-1$ correspond to the extremes of maximum-ratio combining (no interference cancellation) and full zero forcing/interference cancellation (no array gain).

At this point it is important to justify our consideration of this particular receiver. Although the MMSE receiver, which is written in unnormalized form as:

$$\mathbf{v}_0 = \left(\frac{1}{\text{SNR}} \mathbf{I} + d^\alpha \sum_{i \in \Pi(\lambda)} |X_i|^{-\alpha} \mathbf{h}_i \mathbf{h}_i^\dagger \right)^{-1} \mathbf{h}_0,$$

is well-known to be optimal (SINR-maximizing), we consider the sub-optimal PZF receiver primarily because we find it to be more amenable to analysis than the MMSE receiver[2]. Furthermore, by showing that linear scaling is achievable even with this sub-optimal receiver (this implies that linear scaling is also achievable with the optimal MMSE receiver), we are able to pinpoint the necessary requirements for linear scaling and also can easily explain why linear scaling is not achieved by either maximum-ratio combining or using all degrees of freedom for interference cancellation.

In the remainder of this section we analyze performance of the PZF-k receiver for arbitrary k, although later we will show that the PZF receiver achieves linear density scaling if k is increased with N according to $k = \theta N$ for some constant $0 < \theta < 1$.

A. Statistical Properties

We can now state some simple properties of the signal and interference coefficients that result from the partial ZF receiver. These properties result from the following basic result on the squared-norm of the projection of Gaussian vectors onto subspaces:

Lemma 1 ([8]): The squared-norm of the projection of a N-dimensional vector with iid unit-variance complex Gaussian components onto an independent s-dimensional subspace is χ^2_{2s}.

For simplicity, we denote the signal and interference coefficients as:

$$S \triangleq |\mathbf{v}_0^\dagger \mathbf{h}_0|^2 \tag{5}$$

$$H_i \triangleq |\mathbf{v}_0^\dagger \mathbf{h}_i|^2 \quad i = 1, 2, \ldots \tag{6}$$

and characterize the statistics of these coefficients in the following lemma:

Lemma 2: For PZF-k, the signal coefficient S is $\chi^2_{2(N-k)}$, the interference terms $I_1, \ldots, I_k$ are zero, and coefficients $I_{k+1}, I_{k+2}, \ldots$ are iid unit-mean exponential (i.e., χ^2_2). Furthermore, $S, I_{k+1}, I_{k+2}, \ldots$ are mutually independent.

For the sake of exposition, proofs of this and all subsequent lemmas are compiled in the appendix.

Using this statistical characterization and the definitions in (5)-(6) we can write the received SINR as:

$$\text{SINR} = \frac{S}{\frac{1}{\text{SNR}} + d^\alpha \sum_{i=k+1}^{\infty} \left(|X_i|^2\right)^{-\frac{\alpha}{2}} H_i} \tag{7}$$

where the S and H_i terms are characterized in Lemma 2, the quantities $|X_{k+1}|^2, |X_{k+2}|^2, \ldots$ are the $k+1, k+2, \ldots$ ordered points of a 1-D PPP with intensity $\pi\lambda$, and the distances $|X_i|^2$ are independent of the coefficients $S, H_{k+1}, H_{k+2}, \ldots$.

The aggregate interference power for PZF-k is denoted as:

$$I_k \triangleq d^\alpha \sum_{i=k+1}^{\infty} \left(|X_i|^2\right)^{-\frac{\alpha}{2}} H_i. \tag{8}$$

In the following lemma, we characterize the expected aggregate interference power:

Lemma 3: For $k > \frac{\alpha}{2} - 1$, the expected interference power is characterized as:

$$E[I_k] = \left(\pi d^2 \lambda\right)^{\frac{\alpha}{2}} \sum_{i=k+1}^{\infty} \frac{\Gamma\left(i - \frac{\alpha}{2}\right)}{\Gamma(i)} \tag{9}$$

$$< \left(\pi d^2 \lambda\right)^{\frac{\alpha}{2}} \mu^{-1} \left(k - \left\lceil \frac{\alpha}{2} \right\rceil\right)^{1-\frac{\alpha}{2}}, \tag{10}$$

where $\Gamma(\cdot)$ is the gamma function, $\lceil(\cdot)\rceil$ is the ceiling function, constant $\mu = \frac{\alpha}{2} - 1$, and the upper bound is valid for $k > \lceil \frac{\alpha}{2} \rceil$.

B. Bounds on Outage Probability and Maximum Density

Having upper bounded the expected interference power I_k, we can now use Markov's inequality to upper bound the outage probability.

Theorem 1: If partial ZF-k is used, the outage probability is upper bounded by:

$$P\left[\text{SINR} \leq \beta\right] \leq \frac{\beta\left(\left(\pi d^2 \lambda\right)^{\frac{\alpha}{2}} \mu^{-1} \left(k - \lceil \frac{\alpha}{2} \rceil\right)^{1-\frac{\alpha}{2}} + \frac{1}{\text{SNR}}\right)}{N-k-1}$$

and the maximum density λ_ϵ that achieves outage probability no larger than ϵ is lower bounded as:

$$\lambda_\epsilon \geq \left(\frac{\epsilon}{\beta}\right)^{\frac{2}{\alpha}} \frac{\mu^{\frac{2}{\alpha}}}{\pi d^2} \left(N - k - 1 - \frac{\beta}{\epsilon \text{ SNR}}\right)^{\frac{2}{\alpha}} \left(k - \left\lceil \frac{\alpha}{2} \right\rceil\right)^{1-\frac{2}{\alpha}}.$$

[2]In [2] the performance of the MMSE receiver is analyzed for fixed density and asymptotically large N; we are currently investigating whether this approach can be used in our setting where the density increases with N.

Proof 1: The outage probability upper bound is derived by rewriting the outage probability as the tail probability of random variable $\frac{1}{\text{SINR}}$ and then applying Markov's inequality:

$$\begin{aligned} P[\text{SINR} \le \beta] &= P\left[\frac{S}{I_k + \frac{1}{\text{SNR}}} \le \beta\right] \\ &= P\left[\frac{I_k + \frac{1}{\text{SNR}}}{S} \ge \frac{1}{\beta}\right] \\ &\overset{(a)}{\le} \beta E\left[\frac{I_k + \frac{1}{\text{SNR}}}{S}\right] \\ &\overset{(b)}{=} \beta E\left[I_k + \frac{1}{\text{SNR}}\right] E\left[\frac{1}{S}\right] \\ &\overset{(c)}{=} \beta E\left[I_k + \frac{1}{\text{SNR}}\right] (N-k-1)^{-1} \\ &\overset{(d)}{<} \frac{\beta\left(\left(\pi d^2 \lambda\right)^{\frac{\alpha}{2}} \mu^{-1} \left(k - \lceil \frac{\alpha}{2} \rceil\right)^{1-\frac{\alpha}{2}} + \frac{1}{\text{SNR}}\right)}{N-k-1}, \end{aligned}$$

where (a) holds due to Markov's inequality, (b) is due to the independence of random variables I_k and S, (c) holds because S is $\chi^2_{2(N-k)}$, and (d) follows from Lemma 3. Setting this bound equal to ϵ and then solving for λ yields the associated lower bound to λ_ϵ. □

It is worthwhile to note that the $\left(N - k - 1 - \frac{\beta}{\epsilon \text{ SNR}}\right)^{\frac{2}{\alpha}}$ term in the λ_ϵ lower bound is the density increase due to array gain (i.e., increased signal power), while the $\left(k - \lceil \frac{\alpha}{2} \rceil\right)^{1-\frac{2}{\alpha}}$ term is the density increase due to interference cancellation.

IV. Linear Density Scaling

Having established a lower bound to λ_ϵ as a function of k, the number of interferers cancelled by the partial-ZF receiver, we can now show that linear density scaling is achievable. If we choose the number of cancelled interferers $k = \theta N$ for some constant $0 < \theta < 1$, the density lower bound becomes:

$$\begin{aligned} \lambda_\epsilon \quad \ge \quad & \left(\frac{\epsilon}{\beta}\right)^{\frac{2}{\alpha}} \frac{\mu^{\frac{2}{\alpha}}}{\pi d^2} (1-\theta)^{\frac{2}{\alpha}} \theta^{1-\frac{2}{\alpha}} \times \\ & \left(N - \frac{1 + \frac{\beta}{\epsilon \text{ SNR}}}{1-\theta}\right)^{\frac{2}{\alpha}} \left(N - \theta^{-1} \left\lceil \frac{\alpha}{2} \right\rceil\right)^{1-\frac{2}{\alpha}} . \quad (11) \end{aligned}$$

This lower bound divided by N converges to a constant, thereby showing that linear density scaling is achieved for any θ satisfying $0 < \theta < 1$.[3]

This scaling result can be intuitively understood by examining how the signal and aggregate interference power increase with N. Choosing $\theta < 1$ ensures that the signal power, which is $\chi^2_{2(1-\theta)N}$, increases linearly with N. Based on the upper bound in Lemma 3 we can see that the condition $\theta > 0$ ensures that the interference power increases only linearly with N if λ is linear in N. These linear terms are offsetting, and thus allow an approximately constant SINR to be maintained as λ is increased linearly with N.

[3]This derivation proves that λ_ϵ scales *at least* linearly with N, but does not preclude super-linear scaling. We strongly believe that super-linear scaling is not achievable by either PZF or the optimal MMSE receiver, and a rigorous proof of this is under investigation.

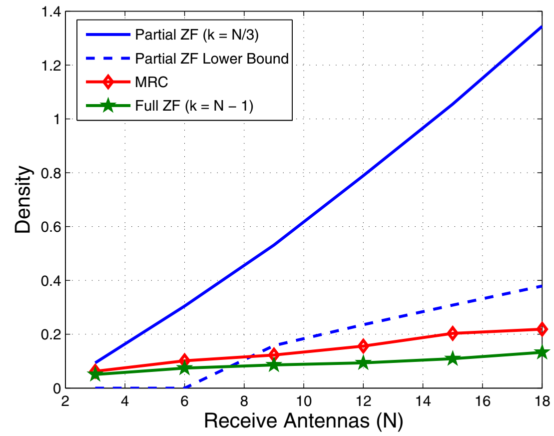

Fig. 1. Maximum density versus N for $\epsilon = .1$, $\beta = 1$, $\alpha = 3$, $\theta = \frac{1}{3}$, $d = 1$.

If an MRC receiver ($k = 0$) or PZF receiver with a constant value of k is used, the signal power increases linearly with N as desired but the interference power increases too quickly with the density (as $\lambda^{\frac{\alpha}{2}}$), thereby limiting the density growth to $N^{\frac{2}{\alpha}}$. At the other extreme, if full zero forcing ($k = N-1$) is used or $k = N - \kappa$ for any positive integer κ, the interference power scales appropriately with the density but the signal power is $\chi^2_{2\kappa}$ and thus does not increase with N, thereby limiting the density increase to $N^{1-\frac{2}{\alpha}}$.

A. Optimal Fraction of Cancelled Interferers

The parameter θ determines the fraction of the available receive degrees of freedom that are used to perform interference cancellation. Although any $\theta \in (0, 1)$ achieves linear scaling, we can find the optimal value of θ by maximizing the λ_ϵ lower bound in (11) in the limit of large N. This is clearly equivalent to maximizing the quantity $(1-\theta)^{\frac{2}{\alpha}} \theta^{1-\frac{2}{\alpha}}$. By simply setting the derivative (w.r.t. θ) to zero and solving, we find the optimal value of θ to be:

$$\theta^* = 1 - \frac{2}{\alpha}. \quad (12)$$

For $\alpha \approx 2$ the degrees of freedom should be used to boost signal power rather than to cancel interference (i.e., $\theta^* \to 0$, because far-away interference is not much weaker than nearby interference when the path loss exponent is small and thus there is little benefit in cancelling only the nearby interferers. On the other hand, as the path loss exponent increases $\theta^* \to 1$ the power from nearby interferers begins to dominate and thus it is preferable to use the antennas for interference cancellation rather than signal boosting.

V. Numerical Results

In Figures 1 and 2 the numerically computed maximum density λ_ϵ and the lower bound to λ_ϵ given in (11) are plotted versus N for $\epsilon = .1$, $\beta = 1$, $d = 1$ for $\alpha = 3$ (Fig. 1) and

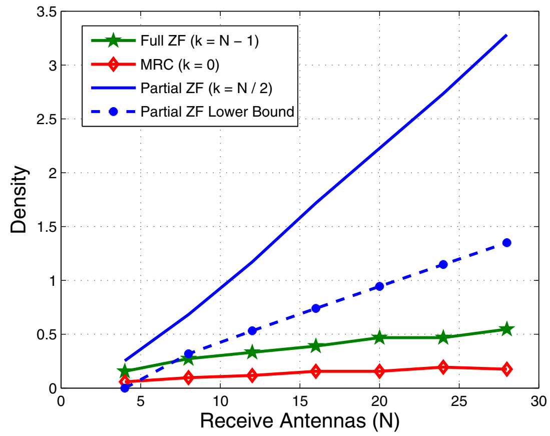

Fig. 2. Maximum density versus N for $\epsilon = .1$, $\beta = 1$, $\alpha = 4$, $\theta = \frac{1}{2}$, $d = 1$.

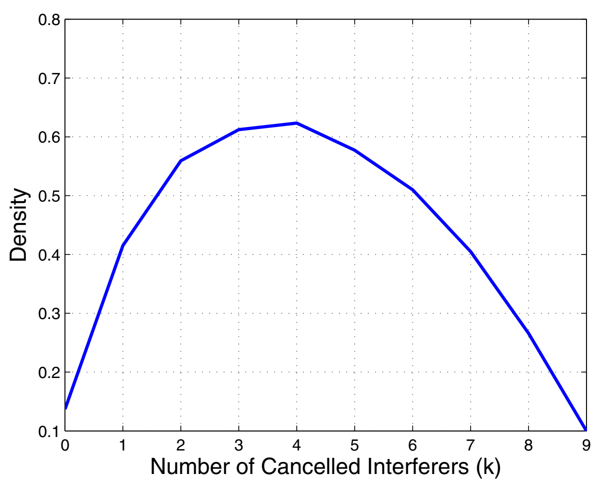

Fig. 3. Maximum density versus # of cancelled interferers k for $N = 10$, $\epsilon = .1$, $\beta = 1$, $\alpha = 3$, $d = 1$.

$\alpha = 4$ (Fig. 2). The optimizing value of θ is used in each plot ($\theta = \frac{1}{3}$ for $\alpha = 3$; $\theta = \frac{1}{2}$ for $\alpha = 4$). The maximum density achievable using MRC ($k = 0$) and full ZR ($k = N - 1$) are also shown in the figures, and both strategies are seen to be significantly inferior to partial ZF. Although the partial ZF lower bound is sufficient to prove linear scaling, the plots show the bound is quite loose, which is somewhat expected because it is based on the Markov inequality. It is also important to notice that the advantage of having multiple antennas kicks in immediately, as we see substantial gains even when only a few antennas are used. This indicates that the linear scaling result, which is technically asymptotic, is actually meaningful even for small values of N.

In Figure 3 the (numerically computed) maximum density λ_ϵ is plotted versus the number of cancelled interferers k for $N = 10$, $\alpha = 3$, $\epsilon = .1$, $\beta = 1$, and $d = 1$. The maximum density is achieved at $k = 4$, which essentially agrees with the fact that $\theta^* = \frac{1}{3}$ for $\alpha = 3$.

VI. Conclusion

We have showed that the throughput of an ad hoc network can be increased linearly with the number of receive antennas, without requiring multiple transmit antennas, which is in stark contrast to point-to-point channels for which a linear gain is achievable only with multiple transmit and multiple receive antennas. This linear gain is achieved by using the receive antennas to perform both interference cancellation and to increase the power of the desired signal, and by increasing the density of simultaneous transmissions while keeping the per-link rate fixed.

Although the density of simultaneous transmissions is, of course, bounded by the density of actual nodes, this density increase can be realized in practice by, for example, allowing for a larger contention probability if ALOHA is used or by precluding the need to use a transmission scheduling mechanism such as CSMA. More generally, the presence of multiple receive antennas allows for a considerably denser spatial packing of simultaneous transmissions than is possible without multiple antennas.

Furthermore, although it is likely that all or most nodes in the networks have the same number of antennas and that nodes function as both transmitters and receivers, the fact that multiple transmit antennas are not needed is important because it more precisely indicates what aspect of MIMO communication is needed to achieve large throughput gains. In addition, detection and decoding is significantly less complex for single-stream transmission as compared to multi-stream transmission (i.e., spatial multiplexing).

Appendix I
Proof of Lemma 2

By the definition of $\mathbf{v}_0$, the quantity $|\mathbf{v}_0^\dagger \mathbf{h}_0|^2$ is the squared-norm of the projection of vector $\mathbf{h}_0$ on Null$(\mathbf{h}_1, \ldots, \mathbf{h}_k)$. This nullspace is $N - k$ dimensional (with probability one) by basic properties of iid Gaussian vectors, and is independent of $\mathbf{h}_0$ by the independence of the channel vectors. By Lemma 1, $|\mathbf{v}_0^\dagger \mathbf{h}_0|^2$ is therefore $\chi^2_{2(N-k)}$. The second property holds by the definition of the PZF-k receiver. To prove the third property, note that $\mathbf{v}_0$ depends only on $\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_k$ and thus is independent of $\mathbf{h}_{k+1}, \mathbf{h}_{k+2}, \ldots$. Because the distribution of each channel vector is rotationally invariant (i.e., the distributions of $\mathbf{W}\mathbf{h}_i$ and $\mathbf{h}_i$ are the same for any unitary matrix $\mathbf{W}$), we can perform a change of basis such that $\mathbf{v}_0 = [1\ 0\ \cdots 0]^T$. After this change of basis, each $\mathbf{v}_0^\dagger \mathbf{h}_i$ (for $i \geq k+1$) is simply equal to the first component of $\mathbf{h}_i$. As a result $\mathbf{v}_0^\dagger \mathbf{h}_{k+1}, \mathbf{v}_0^\dagger \mathbf{h}_{k+2}, \ldots$ are iid complex Gaussians; thus the squared norms are iid exponentials, and furthermore these terms are independent of S.

APPENDIX II
PROOF OF LEMMA 3

First we have

$$E\left[\sum_{i=k+1}^{\infty}|X_i|^{-\alpha}H_i\right] = \sum_{i=k+1}^{\infty}E\left[|X_i|^{-\alpha}H_i\right] = \sum_{i=k+1}^{\infty}E\left[|X_i|^{-\alpha}\right]$$

where $E\left[|X_i|^{-\alpha}H_i\right] = E\left[|X_i|^{-\alpha}\right]E\left[H_i\right] = E\left[|X_i|^{-\alpha}\right]$ due to the independence of $|X_i|$ and H_i and the fact that $E[H_i] = 1$.

Because the ordered squared distances $|X_1|^2, |X_2|^2, \ldots$ are a 1-D PPP with intensity $\pi\lambda$, random variable $\pi\lambda|X_i|^2$ is χ^2_{2i} and thus has PDF $f(x) = \frac{x^{i-1}e^{-x}}{(i-1)!}$. Therefore

$$\begin{aligned} E\left[\left(|X_i|^2\right)^{-\alpha/2}\right] &= (\pi\lambda)^{\frac{\alpha}{2}}\int_0^{\infty}x^{-\alpha/2}\frac{x^{i-1}e^{-x}}{(i-1)!}dx \\ &= \frac{(\pi\lambda)^{\frac{\alpha}{2}}}{(i-1)!}\int_0^{\infty}x^{i-\frac{\alpha}{2}-1}e^{-x}dx \\ &= (\pi\lambda)^{\frac{\alpha}{2}}\frac{\Gamma\left(i-\frac{\alpha}{2}\right)}{\Gamma(i)}. \end{aligned}$$

This quantity is finite only for $i > \frac{\alpha}{2}$, and thus the expected power from the nearest uncancelled interferer is finite only if $k+1 > \frac{\alpha}{2}$.

To reach the upper bound, we use the inequality derived in [4] (which uses Kershaw's inequality to the gamma function):

$$\frac{\Gamma\left(i-\frac{\alpha}{2}\right)}{\Gamma(i)} < \left(i - \left\lceil\frac{\alpha}{2}\right\rceil\right)^{-\frac{\alpha}{2}}$$

where $\lceil\cdot\rceil$ is the ceiling function and we require $i > \left\lceil\frac{\alpha}{2}\right\rceil$. Therefore

$$\begin{aligned} \sum_{i=k+1}^{\infty}\frac{\Gamma\left(i-\frac{\alpha}{2}\right)}{\Gamma(i)} &< \sum_{i=k+1}^{\infty}\left(i-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{-\frac{\alpha}{2}} \\ &\le \int_k^{\infty}\left(x-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{-\frac{\alpha}{2}}dx \\ &= \left(\frac{\alpha}{2}-1\right)^{-1}\left(k-\left\lceil\frac{\alpha}{2}\right\rceil\right)^{1-\frac{\alpha}{2}}, \end{aligned}$$

where the inequality in the second line holds because $x^{-\frac{\alpha}{2}}$ is a decreasing function.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun. : Kluwer Academic Press*, no. 6, pp. 311–335, 1998.

[2] S. Govindasamy, D. Bliss, and D. H. Staelin, "Spectral efficiency in single-hop ad-hoc wireless networks with interference using adaptive antenna arrays," *IEEE Journal on Sel. Areas in Communications*, vol. 25, no. 7, pp. 1358–1369, Sept 2007.

[3] A. Hunter, J. G. Andrews, and S. Weber, "The transmission capacity of ad hoc networks with spatial diversity," *IEEE Trans. on Wireless Communications*, To Appear, 2008.

[4] K. Huang, J. G. Andrews, R. W. Heath, D. Guo, and R. Berry, "Spatial interference cancellation for multi-antenna mobile ad hoc networks," *IEEE Trans. on Info. Theory*, Submitted (arXiv:0807.1773).

[5] S. Weber, J. G. Andrews, and N. Jindal, "Transmission capacity: Applying stochastic geometry to uncoordinated ad hoc networks," *IEEE Journal on Sel. Areas in Communications*, Submitted (arXiv:0809.0016).

[6] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. on Info. Theory*, pp. 3584–6, Oct. 2005.

[7] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. on Info. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[8] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Wiley, 1982.