

Uplink Performance of Large Optimum-Combining Antenna Arrays in Poisson-Cell Networks

Siddhantan Govindasamy

F. W. Olin College of Engineering, Needham, MA, USA

Email: siddhantan.govindasamy@olin.edu

Abstract—The uplink of a wireless network with base stations distributed according to a Poisson Point Process (PPP) is analyzed. The base stations are assumed to have a large number of antennas and use linear minimum-mean-square-error (MMSE) spatial processing for multiple access. The number of active mobiles per cell is limited to permit channel estimation using pilot sequences that are orthogonal in each cell. The cumulative distribution function (CDF) of a randomly located link in a typical cell of such a system is derived when accurate channel estimation is available. A simple bound is provided for the spectral efficiency when channel estimates suffer from pilot contamination. The results provide insight into the performance of so-called massive Multiple-Input-Multiple-Output (MIMO) systems in spatially distributed cellular networks.

Index Terms—Massive MIMO, MMSE.

I. INTRODUCTION

Cellular systems with large numbers of base station antennas servicing a relatively small number of mobiles per cell has been proposed as a method to meet the increasing demand for wireless data communications. In such systems, mobiles transmit simultaneously in the same frequency band and the base-station separates the signals from the mobiles spatially [1], [2]. As the number of antennas at each base station grows large, the matched-filter (MF) receiver (and its transmit-side analog) are optimal [1]. However there is a significant range of parameters where the performance of the MMSE receiver greatly exceeds the performance of the simpler MF receiver [3] which makes analysis of the MMSE receiver in such networks interesting. The uplink performance of massive MIMO systems with MMSE processing has been analyzed before in [3] and [4], but in both those works, the spatial distribution of the network was not explicitly analyzed. Most works which analyze such systems have not explicitly modeled the spatial distribution of base stations and mobiles. Such analyses have the potential to provide valuable insight into the large-scale performance of cellular networks as noted in [5] which considered the downlink of single-antenna systems in Poisson-cell networks.

In this work, we analyze the performance of the uplink of a spatially distributed cellular system with multi-antenna, linear MMSE receivers at the base-stations in the interference-limited regime. The base stations are spatially distributed according to a homogenous PPP on the plane. The mobile nodes are also

assumed to be spatially distributed, but unlike in [6] where we assumed that the transmitting nodes are at independent spatial locations, in this work, we limit the number of *active* mobiles per cell to K , which results in correlation between the locations of the active mobiles in the network. Limiting the number of active mobiles per cell enables the use of orthogonal pilot sequences for channel estimation in each cell. This is the standard approach for channel estimation assumed in the literature [1]. Placing a limitation on the number of active mobiles per cell has other practical benefits as well, such as to meet quality of service requirements. However, limiting the number of active mobiles per cell causes the spatial positions of active mobiles in the network to become correlated which significantly complicates analysis. One approach to analyzing networks with spatially correlated users is by making the density of active users small, e.g., through the use of a medium-access-control protocol [7], [8]. This approach is not well suited to massive MIMO systems where a central assumption is that multiple mobiles transmit simultaneously in every cell. A second approach, which we proposed in [9], is to consider linear MMSE receivers with large numbers of antennas. In this work, we follow a similar approach, making use of the framework we introduced in [9] to derive asymptotic expressions for the spectral efficiency (assuming Gaussian codebooks) of a representative link with a large number of receiver antennas as a function of the number of antennas N , link length, mobile and base station density, path-loss exponent and maximum number of active mobiles per-cell, K . We also provide the CDF of the spectral efficiency when the representative link is randomly distributed in a typical cell of the network. In addition, we provide bounds to the spectral efficiency for systems where the channel estimate suffers from pilot contamination [10].

II. SYSTEM MODEL

Consider a cellular network with base stations distributed according to a PPP with density ρ_c base stations per unit area and suppose that the co-ordinates of the system are shifted such that the base-station closest to the origin is shifted to the origin. Assume that cells are formed by a Voronoi tessellation of the plane with the base stations as generator points. Such a tessellation is referred to as a Poisson-Voronoi Tessellation (PVT). Let the locations of the base stations be B_0, B_1, B_2, \dots , with B_0 denoting the base station at the origin

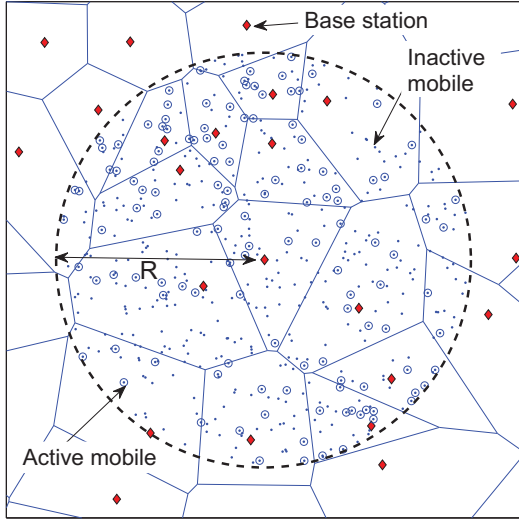


Fig. 1. Illustration of a Poisson-cell network with maximum number of active mobiles per cell, $K = 10$. Mobiles are represented by the dots and circles are used to highlight the active mobiles. Observe that smaller cells tend to have a higher density of active mobiles.

which we call the representative base station. Suppose that there is a mobile at X_0 which we call the representative transmitter which is transmitting to the representative base station. In the remainder of this work, we shall analyze the link between the representative transmitter and the representative base station which we shall also refer to as the representative link. We shall denote a realization of the base station point process by Π . Overlaid on the network of base stations is a circular network of radius R centered at the origin as shown in Figure 1, with n additional mobiles independent and identically distributed (i.i.d.) in the circular network. These mobiles, if active will be co-channel interferers to the representative link. Let the mobiles be located at X_1, X_2, \dots, X_n , and $r_i = |X_i|$ be the distance of the i -th mobile from the origin. The area density of mobiles ρ_m satisfies

$$n = \pi \rho_m R^2. \quad (1)$$

The average power (averaged over the fast-fading random variables) received at each antenna of the representative base station from a mobile at a distance r_i , transmitting with power P_i is $P_i r_i^{-\alpha}$, with $\alpha > 2$. The transmit power of the i -th mobile, P_i equals one or zero, depending on whether or not the mobile is active. Hence, we do not consider power control in this work. Up to $K \geq 1$ mobiles are active in each cell. If a cell has K or less mobiles they are all active, and if a cell has greater than K mobiles, K of the mobiles are selected randomly and with uniform probability to be active. The representative transmitter is always active and thus $P_0 = 1$. $\mathbf{y} \in \mathbb{C}^{N \times 1}$ contains the sampled signals at a given sampling time at the N antennas of the representative base

station and is given by

$$\mathbf{y} = r_0^{-\frac{\alpha}{2}} \mathbf{g}_0 x_0 + \sum_{i=1}^n r_i^{-\frac{\alpha}{2}} \mathbf{g}_i \sqrt{P_i} x_i, \quad (2)$$

where x_i is the transmitted symbol of the i -th mobile and $\mathbf{g}_i \in \mathbb{C}^{N \times 1}$ contains i.i.d., zero-mean, unit variance, circularly symmetric, complex Gaussian random variables denoted by $\mathcal{CN}(0, 1)$, which represent fast fading between the i -th mobile and the N antennas of the representative base station. Since we focus on the interference-limited regime where the noise power $\rightarrow 0$, (2) does not include noise. Thus, our results are applicable to networks with a high density of mobiles.

We shall consider the limit as N, n and $R \rightarrow \infty$, such that $n/N = c$ and ρ_m are constants, and (1) holds. Note that while our analysis provides insight into the scaling behavior of such systems, we use it primarily as a tool to analyze large networks of a fixed size. We additionally require $c > 1/\rho_c$ which ensures that as $R \rightarrow \infty$, with high probability, there would be a larger number of active mobiles in the *entire* network than antennas at the representative base station. This ensures that the matrix \mathbf{R} defined below is invertible with probability 1 if N, n, R is sufficiently large.

For the rest of this paper, whenever n, N or $R \rightarrow \infty$, it is assumed that the other two quantities go to infinity as well. The main results are given in terms of limiting values of a normalized version of the SIR, $\beta_N = N^{-\alpha/2} r_T^\alpha \text{SIR}$, at the output of a linear receiver which estimates x_0 using a weight vector \mathbf{w} . The estimate, $\hat{x}_0 = \mathbf{w}^\dagger \mathbf{y}$, and the weight vector,

$$\mathbf{w}^\dagger = \hat{\mathbf{h}}^\dagger \left(\sum_{j=1}^n r_j^{-\alpha} P_j \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1}, \quad (3)$$

where $\hat{\mathbf{h}}$ is an estimate of the channel vector \mathbf{g}_0 . Note that the weight vector above is the minimum-mean-square-error weight vector if $\hat{\mathbf{h}} = \mathbf{g}_0$ which is the assumption we use for the first part of this work. We also consider a pilot-contaminated estimate of \mathbf{g}_0 in Section III-C. We assume that the interference covariance matrix

$$\mathbf{R} = \sum_{j=1}^n r_j^{-\alpha} P_j \mathbf{g}_j \mathbf{g}_j^\dagger \quad (4)$$

is known perfectly at the representative base station. The SIR at the output of the receiver is given by

$$\text{SIR} = \frac{r_0^{-\alpha} |\mathbf{w}^\dagger \mathbf{g}_0|^2}{\sum_{i=1}^n r_i^{-\alpha} P_i |\mathbf{w}^\dagger \mathbf{g}_i|^2} \quad (5)$$

III. MAIN RESULTS

A. Density of Active Transmissions

The density of active mobiles in our system is the product of the density of all mobiles and the probability that a mobile is active in the limit as $R \rightarrow \infty$. This quantity is used in characterizing the spectral efficiency in the subsequent sections and is interesting in its own right as it determines the fraction of mobiles that can be active at any one time as a function

of K , ρ_c and ρ_m . The probability that a mobile is active is dependent the distribution of cell sizes, and is given in the following lemma.

Lemma 1:

$$\lim_{n \rightarrow \infty} \Pr(P_j = 1) = \rho_c E[|\mathcal{C}^T| h(|\mathcal{C}^T|)] . \quad (6)$$

where the expectation is with respect to the PDF of the area of the *typical cell* of a PVT with density ρ_c , and

$$h(a) \triangleq \frac{K - K e^{-\rho_m a} \sum_{k=0}^K \frac{(\rho_m a)^k}{k!}}{\rho_m a} + \sum_{m=0}^{K-1} \frac{(\rho_m a)^m}{m!} e^{-\rho_m a} . \quad (7)$$

Proof: Please see in Appendix B.

The typical cell (see e.g. [11] for its precise definition) is statistically equivalent to the cell containing the origin if the origin is added to the set of generator points of a PPP that underlies a PVT (see e.g. [11]). The exact PDF of $|\mathcal{C}^T|$ is given in [12], but it is expressed as an infinite series involving multiple integrals, which is challenging to compute numerically (e.g. the authors use monte-carlo integration to evaluate it in [12]). We do not include the explicit expressions here for the sake of brevity. Instead, we use an approximation to the PDF of $a = |\mathcal{C}^T|$, which is given in terms of a generalized gamma PDF as follows [13]. For $a > 0$,

$$f_A(a) \approx 15.225 \rho_c (\rho_c a)^{2.311} e^{-3.032 (\rho_c a)^{1.080}} \quad (8)$$

B. Spectral Efficiency with Perfect Channel Estimation

Assume that the representative base station has a perfect estimate of the channel vector, i.e. $\hat{\mathbf{h}} = \mathbf{g}_0$. The normalized SIR at the output of the MMSE receiver is then

$$\beta_N = N^{-\alpha/2} \hat{\mathbf{h}}^\dagger \mathbf{R}^{-1} \hat{\mathbf{h}} . \quad (9)$$

Conditioned on a specific realization of the base-station point process Π and the length of the representative link r_0 , which is assumed to be active, we have the following theorem.

Theorem 1: If the system model from the previous section holds, as $n, N, R \rightarrow \infty$ such that $n/N = c > 0$, and (1) hold, $\beta_N \rightarrow \beta$ in probability where β is the real, positive solution to

$$\frac{2\pi^2 \rho \beta^{\frac{2}{\alpha}}}{\alpha} \csc\left(\frac{2\pi}{\alpha}\right) = 1 + \frac{2(\pi\rho)^{2-\frac{2}{\alpha}} \beta}{(\alpha-2)(c + \pi\rho\beta)^{1-\frac{2}{\alpha}}} \times {}_2F_1\left(1 - \frac{2}{\alpha}, 1 - \frac{2}{\alpha}; 2 - \frac{2}{\alpha}; \frac{\pi\rho\beta}{\pi\rho\beta + c}\right) , \quad (10)$$

where $\rho = \rho_m \lim_{n \rightarrow \infty} \Pr(P_j = 1)$, and ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the Gauss hypergeometric function.

Proof: Given in Appendix A.

Note that ρ is the density of mobiles in the limit. Additionally, when the number of users in the *entire* network is much larger than the number of antennas at the representative base station, i.e. $n \gg N$, the second term on the RHS of (10) is small [9]. Approximating that term by zero, we can find an approximation for β which when combined with the Shannon

equation, yields the following expression for the spectral efficiency and its mean when Gaussian codebooks are used by each mobile [9].

$$\gamma \approx E[\gamma] \approx \log_2 \left(1 + \left[\frac{N\alpha}{2\pi^2 \rho r_0^2} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}} \right) . \quad (11)$$

If the representative transmitter is distributed with uniform probability in the cell at the origin, we can derive the CDF of the spectral efficiency from (11) using the nearest-neighbor distribution of a PPP (e.g. see [11]). This yields the following approximation to the CDF of the spectral efficiency of the representative link assuming a large number of base station antennas N and the number of users in the network greatly exceeding the number of antennas at the base station $n \gg N$.

$$\Pr(\gamma \leq \tau) \approx e^{-\frac{\rho_c N}{\rho} \frac{\alpha}{2\pi} \sin\left(\frac{2\pi}{\alpha}\right) \left(\frac{1}{2^{\tau-1}}\right)^{\frac{2}{\alpha}}} . \quad (12)$$

C. Pilot Contaminated Channel Estimation

We assume that the pilot signals used for channel estimation by the mobiles in a given cell are orthogonal, but that the same set of pilot signals is repeated in different cells. The estimated channel vector of the representative transmitter suffers from pilot contamination [10] from mobiles in other cells who shared the same pilot sequence. If we assume that the power used during the training sequence is high compared to the noise, we can neglect the effect of noise in the channel estimate. Thus the estimated channel between the representative base station and the representative transmitter $\hat{\mathbf{h}}$ is given by,

$$\hat{\mathbf{h}} = \sum_{i \in \mathcal{T}} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \mathbf{g}_i \quad (13)$$

where \mathcal{T} is the set of indices of the mobiles (including the representative transmitter) which shared the same pilot sequence with the representative transmitter during channel estimation. For the purposes of this section, we assume that $\alpha > 4$ which enables us to simplify the effect of the pilot contamination.

We assume that the receiver uses the weight vector in (3) with the pilot-contaminated channel estimate $\hat{\mathbf{h}}$ from (13). Note that we continue to assume that the interference covariance matrix is known at the receiver. We refer to this receiver as the PC-MMSE receiver. As in the proof of Theorem 1, we characterize the normalized SIR for the PC-MMSE receiver for a fixed r_0 . The normalized SIR for the PC-MMSE receiver is denoted by $\bar{\beta}_N = N^{-\alpha/2} r_0^\alpha \text{SIR}$. We can now state the following theorem.

Theorem 2: Conditioned Π and r_0 , $\bar{\beta}_N \rightarrow \bar{\beta}$ in probability, with $\bar{\beta}$ is bounded from below by the following random variable

$$\bar{\beta}_N \rightarrow \bar{\beta} \geq \frac{r_0^{-2\alpha}}{r_0^{-\alpha} + \sum_{j \in \mathcal{T}, j \neq 0} r_j^{-\alpha} P_j} \beta \quad (14)$$

Proof: Given in Appendix D.

We can further bound $\bar{\beta}$ by assuming that there is exactly one mobile in every cell that shares the pilot sequence with

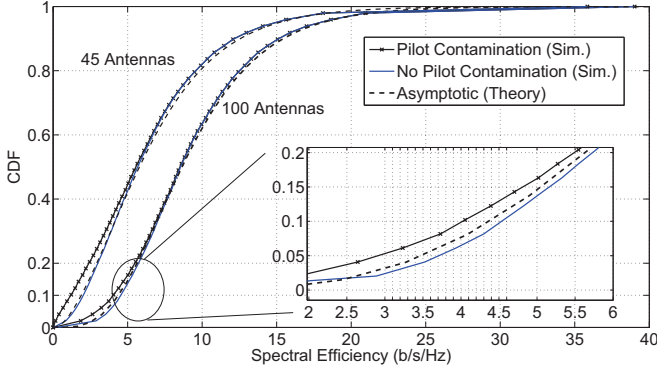


Fig. 2. Simulated CDF of the spectral efficiency for 45 and 100 antennas and with and without pilot contamination. The parameters used are $\rho_c = 2 \times 10^{-5}$, $\rho_m = 0.001$, $\alpha = 5$ and the representative transmitter was randomly placed in the cell at the origin. The maximum number of active mobiles per cell is $K = 10$.

the representative transmitter during channel estimation. Furthermore, for $j > 0$, observe that the closest point to the origin in the cell associated with the base station at B_j is bounded from above by $|B_j|/2$. Thus, all mobiles located in the cell associated with the base-station at B_j are at a distance $|B_j|/2$ or greater from the base station at the origin. This leads to the following bound.

$$\bar{\beta} \geq \frac{r_0^{-2\alpha}}{r_0^{-\alpha} + \sum_{j=1}^{\infty} \left| \frac{B_j}{2} \right|^{-\alpha}} \beta \quad (15)$$

Note here that while the bound given above is loose, to the best of our knowledge it is the only such bound which considers the effect of spatially distributed pilot contaminators.

IV. NUMERICAL SIMULATIONS AND RESULTS

We conducted Monte Carlo simulations to verify the accuracy of our asymptotic results. Figure 2 shows the CDFs of the spectral efficiency from simulations with and without pilot contamination as well as the asymptotic expression for the CDF from (12) for 45 and 100 antennas at the base station. The representative link was randomly distributed in the cell containing the origin. The remaining parameters are given in the figure caption. Observe that with 100 antennas and at an outage probability of 0.1, the theoretical prediction is within 0.25 b/s/Hz of the simulated values which validates (12). For 45 antennas, the asymptotic CDF is within 0.3 b/s/Hz of the simulated values. Additionally, note that the simulated pilot contaminated spectral efficiency for 100 antennas is within 0.5 b/s/Hz of the non-pilot contaminated spectral efficiency.

In Figure 3, we plotted the CDFs from (12) for 25 and 50 antennas per base station and $K = 1, 10$ and 20. This figure illustrates how the results in this paper can be used to analyze the tradeoff between increasing the density of transmissions by increasing K and the resulting reduction in per-link data rates due to increased interference. For a system with 50 antennas, at an outage probability of 0.1, there is approximately a five-fold increase in the spectral efficiency going from $K = 10$, to $K =$

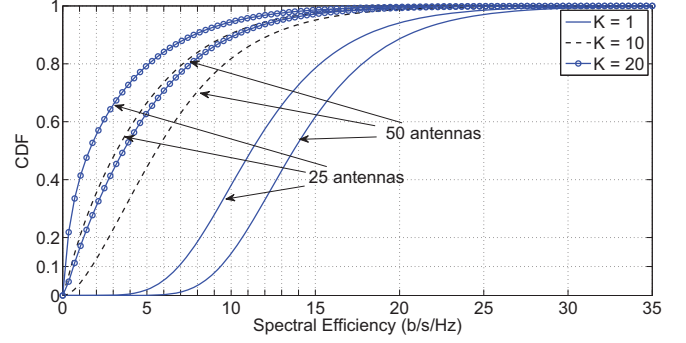


Fig. 3. Simulated CDF of spectral efficiency for 25 and 50 antennas. The parameters used are $\rho_c = 2 \times 10^{-5}$, $\rho_m = 0.001$, $\alpha = 5$ and the representative transmitter was randomly placed in the cell at the origin. The maximum number of active mobiles per cell is $K = 1, 10$ and 20.

1. The density of active mobiles with $K = 10$ is approximately 3.8×10^{-4} and the density of active mobiles with $K = 1$ is approximately 4×10^{-5} . Thus an approximately 10 fold increase in mobile density results in a reduction in the spectral efficiency of the representative link by approximately a factor of five, at an outage probability of 0.1. On the other hand, going from $K = 20$ to $K = 10$ results in nearly a doubling of the spectral efficiency of a representative link but with the density of active transmissions reduced by approximately a factor of less than two. Combined with models for channel coherence and training times (which depend on K), such an analysis could be used to optimize K .

Figure 4 illustrates the lower bound for the spectral efficiency for a fixed link length of $r_T = 100$ and $K = 5$, and the remaining parameters as given in the caption. The simulated markers represent the pilot contaminated mean spectral efficiency and the dashed line represents the lower bound from (15). For reference, the simulated asymptotic spectral efficiency with perfect channel estimation plotted using the solid line, and simulated spectral efficiencies for a system with non-zero noise such that the Signal-to-Noise-Ratio (SNR) is 20dB is plotted using the asterisk markers. Note that the lower bound is loose. However, it provides a guarantee on the worst case mean spectral efficiency under pilot contamination. Additionally, the simulations with 20dB SNR indicate that the interference-limited approximation is accurate even when the number of antennas is large.

V. SUMMARY AND CONCLUSIONS

In this paper, the spectral efficiency of the uplink in a spatially distributed, Poisson-cell network is analyzed. The base station is assumed to use a linear MMSE estimator and the number of active mobiles per cell is limited to K in order to permit channel estimation using intra-cell orthogonal pilots. The CDF of the spectral efficiency in the interference-limited regime is derived for a randomly distributed mobile in a typical cell of the network. The results can be used to statistically characterize achievable data rates in such networks as a function of tangible system parameters such as user and base-station density and number of antennas, and can help

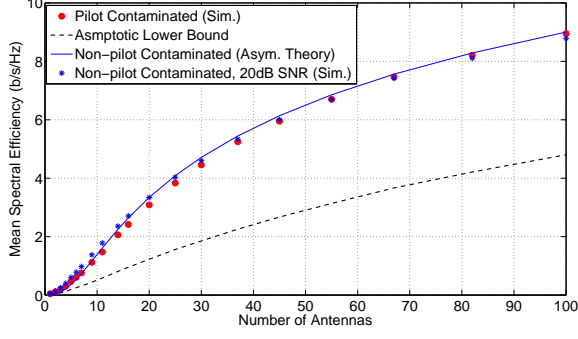


Fig. 4. Mean spectral efficiency vs. number of base station antennas with pilot contamination for a system with $\rho_c = 0.00002$, $\rho_m = 0.001$, $K = 10$, $\alpha = 5$, and a fixed link length $r_T = 100$. The dashed lines represent a lower bound for the mean spectral efficiency.

system designers optimize parameters such as the maximum number of active mobiles per cell.

APPENDIX

A. Proof of Main Result

The main result is proved using Theorem 1 of [9] in which the normalized SIR of a representative link with N antennas and a linear MMSE receiver in a network with interferers at correlated spatial positions is shown to converge in probability to a limit of the same form as in Theorem 1 in this paper. Theorem 1 in this paper follows directly from that result if the following two conditions hold.

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr(P_i r_i^{-\alpha} N^{\frac{\alpha}{2}} \leq x, P_j r_j^{-\alpha} N^{\frac{\alpha}{2}} \leq x) \\ = \lim_{n \rightarrow \infty} \Pr(P_i r_i^{-\alpha} N^{\frac{\alpha}{2}} \leq x) \Pr(P_j r_j^{-\alpha} N^{\frac{\alpha}{2}} \leq x). \end{aligned} \quad (16)$$

and

$$\lim_{n \rightarrow \infty} \Pr(r_i^{-\alpha} N^{\frac{\alpha}{2}} \leq x | P_i = 1) = \lim_{n \rightarrow \infty} \Pr(r_i^{-\alpha} N^{\frac{\alpha}{2}} \leq x). \quad (17)$$

Equation (16) is shown in Appendix C, and (17) follows directly from the following lemma

Lemma 2: For any $x > 0$,

$$\lim_{N \rightarrow \infty} \Pr(P_i = 0 | r_i \leq x^{-1/\alpha} \sqrt{N}) = \lim_{N \rightarrow \infty} \Pr(P_i = 0)$$

Proof: The proof follows similar steps to that used to prove Lemma 1 and is omitted for brevity.

B. Proof of Lemma 1

Let C_i denote the cell occupied by X_i . Let $A_i = |C_i \cap B(0, R)|$ be the area of intersection between C_i and $B(0, R)$. Additionally let $\#C_i$ denote the number of mobiles in C_i excluding X_i , and \mathcal{A} be the event that X_i and X_j are in different cells, i.e. $C_i \neq C_j$. Since the cells of a PVT are finite with probability 1 (e.g. see [11]), as $R \rightarrow \infty$, $\Pr(\mathcal{A}) \rightarrow 1$. Recall that all mobiles in a given cell are active if there are

K or fewer mobiles in the cell. If there are greater than K mobiles, K are selected randomly to be active. Thus,

$$\Pr(P_j = 1 | \#C(X_j) = \ell, \mathcal{A}) = \begin{cases} 1, & \text{if } \ell < K \\ \frac{K}{\ell+1}, & \text{otherwise.} \end{cases} \quad (18)$$

For $n - 2 \geq m$, we have

$$\begin{aligned} \Pr(\#C(X_j) = m | A_i = a_i, A_j = a_j, \mathcal{A}) \\ = \Pr(\#C(X_j) = m | A_j = a_j, \mathcal{A}) \\ = \binom{n-2}{m} \left(\frac{a_j}{\pi R^2} \right)^m \left(\frac{\pi R^2 - a_j}{\pi R^2} \right)^{n-m-2}. \end{aligned} \quad (19)$$

Let $\mathcal{C}_1, \mathcal{C}_2, \dots$ denote the cells in the network. Define the following set which consists of the portions of the cells wholly contained in $B(0, d)$.

$$\mathcal{C}_d := \{C_i \cap B(0, d) : i = 1, 2, \dots\}. \quad (20)$$

The j -th element of \mathcal{C}_d is denoted by \mathcal{C}_{dj} . Thus, A_j takes values in the set \mathcal{C}_R , where $\Pr(A_i = \mathcal{C}_{Rk}) = |\mathcal{C}_{Rk}|/(\pi R^2)$. Combining (1), (18) and (19), weighting by $\Pr(A_i = \mathcal{C}_{Rk})$ and summing over all possible values of A_j yields,

$$\begin{aligned} \Pr(P_j = 1 | \mathcal{A}) = \sum_{k=1}^{|\mathcal{C}_R|} \frac{|\mathcal{C}_{Rk}|}{\pi R^2} \left[\sum_{\ell=0}^{n-2} \min\left(1, \frac{K}{\ell+1}\right) \times \right. \\ \left. \binom{n-2}{\ell} \left(\frac{\rho_m |\mathcal{C}_{Rk}|}{n} \right)^\ell \left(1 - \frac{\rho_m |\mathcal{C}_{Rk}|}{n} \right)^{n-\ell-2} \right] \end{aligned} \quad (21)$$

The next step is to take the limit of (21) as $n, R \rightarrow \infty$. Since the term in the brackets is bounded for all n , we can its limit before taking the outer limit in (21) (see e.g. [14]). Writing $a = |\mathcal{C}_{Rk}|$, the limit of the term in the brackets in (21) is

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[\sum_{\ell=K}^{n-2} \frac{K}{\ell+1} \binom{n-2}{\ell} \left(\frac{\rho_m a}{n} \right)^\ell \left(1 - \frac{\rho_m a}{n} \right)^{n-\ell-2} a \right. \\ \left. + \sum_{m=0}^{K-1} \binom{n-2}{m} \left(\frac{\rho_m a}{n} \right)^m \left(1 - \frac{\rho_m a}{n} \right)^{n-m-2} \right] \\ = \frac{K - K e^{-\rho_m a} \sum_{k=0}^K \frac{(\rho_m a)^k}{k!}}{\rho_m a} + \sum_{m=0}^{K-1} \frac{1}{m!} (\rho_m a)^m e^{-\rho_m a} \\ = h(a). \end{aligned} \quad (22)$$

Thus, we can write the limit of (21) as

$$\lim_{R \rightarrow \infty} \Pr(P_j = 1 | \mathcal{A}) = \lim_{R \rightarrow \infty} \frac{|\mathcal{C}_R|}{\pi R^2} \sum_{k=1}^{|\mathcal{C}_R|} \frac{|\mathcal{C}_{Rk}|}{|\mathcal{C}_R|} h(|\mathcal{C}_{Rk}|). \quad (23)$$

We next state a result from stochastic geometry (Equation 5.2 of [11]), which relates the spatial average of a function over the cells of a single realization of a PVT to the ensemble average. For any bounded function f which maps convex sets in \mathbb{R}^2 to \mathbb{R} , the following holds with probability 1,

$$\lim_{R \rightarrow \infty} \frac{1}{|\mathcal{C}_R|} \sum_{k=1}^{|\mathcal{C}_R|} f(\mathcal{C}_{Rk}) = \frac{1}{E\left[\frac{1}{|\mathcal{C}^0|}\right]} E\left[\frac{f(\mathcal{C}^0)}{|\mathcal{C}^0|}\right], \quad (24)$$

where \mathcal{C}^0 is the *zero-cell* of a PVT, which is statistically equivalent to the cell that contains any given point in the plane. Additionally, we note that with probability 1, as $R \rightarrow \infty$, the number of cells per unit area in $B(0, R)$ approaches the density of the base stations with probability 1, i.e. $\lim_{R \rightarrow \infty} \frac{|\mathcal{C}_R|}{\pi R^2} = \rho_c$. Let $f(\mathcal{C}_{Rk}) = |\mathcal{C}_{Rk}|h(|\mathcal{C}_{Rk}|)$, which is only dependent on the area of \mathcal{C}_{Rk} . Thus, we can apply (24) cellular model we assume here, where the coordinates are shifted such that there is a base station at the origin. Combining (24), the fact that $\Pr(\mathcal{A}) \rightarrow 1$ and $E\left[\frac{1}{|\mathcal{C}^0|}\right] = \rho_c$ with (23) we have,

$$\lim_{n \rightarrow \infty} \Pr(P_j = 1) = E[h(|\mathcal{C}^0|)] . \quad (25)$$

Next we express the equation above in terms of the *typical cell* of a PVT which we denote by \mathcal{C}^T . The relationship between the mean of functions of the zero-cell and the typical cell is given by Equation 2.1 of [15]. Applying this result to the previous equation yields

$$\lim_{n \rightarrow \infty} \Pr(P_j = 1) = \frac{1}{E[|\mathcal{C}^T|]} E[|\mathcal{C}^T| h(|\mathcal{C}^T|)] . \quad (26)$$

Substituting the fact that $E[|\mathcal{C}^T|] = \frac{1}{\rho_c}$ with the above expression proves the lemma.

C. Proof of Equation (16)

Recalling the definitions used in the proof of Lemma 1, and following steps similar to that used to prove it, we can write

$$\begin{aligned} \Pr(P_i = 1, P_j = 1 | A_i = |\mathcal{C}_{Rs}|, A_j = |\mathcal{C}_{Rt}|, \mathcal{A}) = \\ \sum_{\ell=0}^{n-2} \min\left(1, \frac{K}{\ell+1}\right) \binom{n-2}{\ell} \left(\frac{\rho_m |\mathcal{C}_{Rt}|}{n}\right)^\ell \\ \times \left(1 - \frac{\rho_m |\mathcal{C}_{Rt}|}{n}\right)^{n-\ell-2} \times \\ \sum_{m=0}^{n-\ell-2} \min\left(1, \frac{K}{m+1}\right) \binom{n-\ell-2}{m} \left(\frac{\rho_m |\mathcal{C}_{Rs}|}{n-\rho_m |\mathcal{C}_{Rt}|}\right)^m \\ \times \left(1 - \frac{\rho_m |\mathcal{C}_{Rt}|}{n-\rho_m |\mathcal{C}_{Rt}|}\right)^{n-m-\ell-2} \end{aligned} \quad (27)$$

Applying steps similar to those used to prove Lemma 1 to (27), i.e. by weighting (27) by the probabilities that $A_i = |\mathcal{C}_{Rs}|, A_j = |\mathcal{C}_{Rt}|$, summing over all \mathcal{C}_{Rs} and \mathcal{C}_{Rt} , taking the limit as $n, N, R \rightarrow \infty$, and applying (24), we can show that

$$\begin{aligned} \lim_{R \rightarrow \infty} \Pr(P_i = 1, P_j = 1 | \mathcal{A}) = \\ \rho_c E[|\mathcal{C}^T| h(|\mathcal{C}^T|)] \rho_c E[|\mathcal{C}^T| h(|\mathcal{C}^T|)] . \\ = \lim_{R \rightarrow \infty} \Pr(P_i = 1) \lim_{R \rightarrow \infty} \Pr(P_j = 1) \end{aligned} \quad (28)$$

Since $x > 0$ and $\Pr(\mathcal{A}) \rightarrow 1$, we have

$$\begin{aligned} \lim_{R \rightarrow \infty} \Pr(P_i N^{\alpha/2} r_i^{-\alpha} > x, P_j N^{\alpha/2} r_j^{-\alpha} > x) \\ = \lim_{R \rightarrow \infty} \Pr(N^{\alpha/2} r_i^{-\alpha} > x, N^{\alpha/2} r_j^{-\alpha} > x | \\ P_i = 1, P_j = 1, \mathcal{A}) \Pr(P_i = 1, P_j = 1 | \mathcal{A}) \end{aligned}$$

Combining this with the fact that r_i and r_j are asymptotically independent when conditioned on \mathcal{A} and $\Pr(P_i = 1, P_j = 1)$,

$$\begin{aligned} \lim_{R \rightarrow \infty} \Pr(P_i N^{\alpha/2} r_i^{-\alpha} > x, P_j N^{\alpha/2} r_j^{-\alpha} > x) \\ = \lim_{R \rightarrow \infty} \Pr(P_i = 1) \lim_{R \rightarrow \infty} \Pr(N^{\alpha/2} r_i^{-\alpha} > x) \\ \times \lim_{R \rightarrow \infty} \Pr(P_j = 1) \lim_{R \rightarrow \infty} \Pr(N^{\alpha/2} r_j^{-\alpha} > x) \end{aligned} \quad (29)$$

Equation (16) follows from Lemma 2 and the fact that for $x > 0$, $\Pr(P_j N^{\alpha/2} r_j^{-\alpha} > x | P_j = 0) = 0$.

D. Proof of Theorem 2

Note that R, n , and N are related such that the CDF of $N^{\alpha/2} r_i^{-\alpha/2}$ does not vary with N, n, R and moreover, $p_i = N^{\alpha/2} r_i^{-\alpha}$ are bounded for all n, N, R . The signal power normalized by $N^{-\alpha}$ at the output of the PC-MMSE filter is

$$\begin{aligned} N^{-\alpha} S = r_0^{-\alpha} N^{-\alpha} |\mathbf{w}^\dagger \mathbf{g}_0|^2 \\ = r_0^{-\alpha} \left| \sum_{i \in \mathcal{T}} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \frac{1}{N} \mathbf{g}_i^\dagger \left(\sum_{j=1}^n \frac{1}{N} p_j \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right|^2 \\ = r_0^{-\alpha} \left| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \frac{1}{N} \mathbf{g}_i^\dagger \left(\sum_{j=1}^n \frac{1}{N} p_i \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right. \\ \left. + r_0^{-\alpha/2} \frac{1}{N} \mathbf{g}_0^\dagger \left(\sum_{j=1}^n \frac{1}{N} p_j \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right|^2 . \end{aligned} \quad (30)$$

Consider the expectation of the first term in the absolute value in (30).

$$\begin{aligned} E \left[\left| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \frac{1}{N} \mathbf{g}_i^\dagger \left(\sum_{j=1}^n \frac{1}{N} p_i \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right|^2 \right] \leq \\ E \left[\sum_{i \in \mathcal{T}, i \neq 0} \left| r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \right| \left| \frac{1}{N} \mathbf{g}_i^\dagger \left(\sum_{j=1, j \neq i}^n \frac{1}{N} p_i \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right|^2 \right] , \end{aligned}$$

where the inequality follows from applying the Sherman-Morrison-Woodbury matrix inversion lemma and the non-negative definiteness of the matrix

$$\left(\sum_{j=1, j \neq i}^n \frac{1}{N} p_i \mathbf{g}_j \mathbf{g}_j^\dagger \right) . \quad (31)$$

From Lemma 3 of [9] if (16) and (17) hold, $\exists n_0$, such that for all $n > n_0$, with probability 1, the minimum eigenvalue of the matrix above is bounded from below by $\lambda_{lb} > 0$. For $n > n_0$,

taking an eigen-decomposition of this matrix and simplifying:

$$\begin{aligned}
& E \left[\left\| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \frac{1}{N} \mathbf{g}_i^\dagger \left(\sum_{j=1, j \neq i}^n \frac{1}{N} p_i \mathbf{g}_j \mathbf{g}_j^\dagger \right)^{-1} \mathbf{g}_0 \right\|^2 \right] \\
& \leq E \left[\left\| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \sum_{j=1, j \neq i}^N \frac{u_j v_j}{N \lambda_{\ell b}} \right\|^2 \right] \\
& = E \left[\left\| \sum_{j=1}^{N-1} \frac{u_j v_j}{N \lambda_{\ell b}} \right\|^2 \right] E \left[\left\| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\frac{\alpha}{2}} \sqrt{P_i} \right\|^2 \right], \quad (32)
\end{aligned}$$

where u_i and v_i are i.i.d. $\mathcal{CN}(0, 1)$ random variables. Note that the expectation in (32) factors due to the isotropic nature of Gaussian random vectors. Note that one of the standard methods to prove the weak law of large numbers is to show that the sample mean of zero-mean i.i.d. random variables converges in the mean-square sense to zero. Since convergence in mean square implies convergence in mean, the first expectation on the RHS of (32) converges to zero. Since the mobiles which contribute towards the pilot contamination are located outside the cell at the origin, for $i \neq 0, i \in \mathcal{T}$, $r_i > D_{\min}$, where D_{\min} is the radius of the largest circle which is wholly contained in the cell at the origin. Since $\alpha > 4$, from Section III.A of [16], $E \left[\left\| \sum_{i \in \mathcal{T}, i \neq 0} r_i^{-\alpha/2} \right\|^2 \right]$ is bounded. Thus, we have (32) $\rightarrow 0$, which implies that the first term in the absolute value in (30) converges in probability to zero. From Theorem 1, the second term in the absolute value in (30) converges in probability to $r_0^{-\alpha} \beta$. Thus, the following holds in probability

$$N^{-\alpha} S \rightarrow r_0^{-2\alpha} \beta^2. \quad (33)$$

Next, consider the interference power normalized by $N^{-\alpha/2}$

$$\begin{aligned}
N^{-\frac{\alpha}{2}} I &= N^{-\frac{\alpha}{2}} \mathbf{w}^\dagger \sum_{i=1}^n r_i^{-\alpha} P_i \mathbf{g}_i \mathbf{g}_i^\dagger \mathbf{w} \\
&= \frac{1}{N} \hat{\mathbf{h}}^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \hat{\mathbf{h}}. \quad (34) \\
&= \frac{1}{N} \sum_{j \in \mathcal{T}} r_j^{-\frac{\alpha}{2}} \sqrt{P_j} \sum_{k \in \mathcal{T}} r_k^{-\frac{\alpha}{2}} \sqrt{P_k} \mathbf{g}_j^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_k. \\
&= \sum_{j \in \mathcal{T}} r_j^{-\frac{\alpha}{2}} \sqrt{P_j} \sum_{k \in \mathcal{T}, k \neq j} r_k^{-\frac{\alpha}{2}} \sqrt{P_k} \frac{1}{N} \mathbf{g}_j^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_k \\
&+ \sum_{j \in \mathcal{T}} r_j^{-\alpha} P_j \frac{1}{N} \mathbf{g}_j^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_j. \quad (35)
\end{aligned}$$

The first term on the RHS of (35), has a similar form to the first term in the absolute value in (30), which was shown to converge in probability to zero. Following a similar sequence of steps, we can show that the first term on the RHS of (35)

goes to zero as $n, N, R \rightarrow \infty$ which yields

$$\begin{aligned}
\lim_{N \rightarrow \infty} N^{-\frac{\alpha}{2}} I &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j \in \mathcal{T}} r_j^{-\alpha} P_j \mathbf{g}_j^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_j \\
&= \lim_{N \rightarrow \infty} \sum_{j \in \mathcal{T}, j \neq 0} \frac{\frac{1}{N} r_j^{-\alpha} P_j \mathbf{g}_j^\dagger \left(\sum_{i=1, i \neq j}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_j}{1 + \frac{1}{N} p_j \mathbf{g}_j^\dagger \left(\sum_{i=1, i \neq j}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_j} \\
&+ \frac{1}{N} r_0^{-\alpha} P_j \mathbf{g}_0^\dagger \left(\sum_{i=1}^n \frac{1}{N} p_i \mathbf{g}_i \mathbf{g}_i^\dagger \right)^{-1} \mathbf{g}_0 \\
&= r_0^{-\alpha} \beta + \sum_{j \in \mathcal{T}, j \neq 0} \frac{r_j^{-\alpha} P_j \beta}{1 + \bar{p}_j \beta} = \bar{\beta} \text{ in probability.} \quad (36) \\
&\leq r_0^{-\alpha} \beta + \sum_{j \in \mathcal{T}, j \neq 0} r_j^{-\alpha} P_j \beta, \quad (37)
\end{aligned}$$

where second equality is from applying the Sherman-Morrison-Woodbury matrix inversion lemma. The theorem is proved by dividing (33) by (37).

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [2] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *To appear in IEEE Commun. Mag.*, 2013. [Online]. Available: <http://arxiv.org/abs/1304.6690>
- [3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [4] N. Krishnan, R. D. Yates, and N. B. Mandayam, "Cellular systems with many antennas: Large system analysis under pilot contamination," *Proc. of the 50th Annual Allerton Conference on Communication, Control, and Computing.*, 2012.
- [5] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Apr. 2011.
- [6] S. Govindasamy, D. Bliss, and D. H. Staelin, "Asymptotic spectral efficiency of the uplink in spatially distributed wireless networks with multi-antenna base stations," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 100–112, Jun. 2013.
- [7] R. Ganti, J. Andrews, and M. Haenggi, "High-SIR transmission capacity of wireless networks with general fading and node distribution," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, May 2011.
- [8] R. Giacomelli, R. Ganti, and M. Haenggi, "Outage probability of general Ad Hoc networks in the high-reliability regime," *IEEE/ACM Trans. Netw.*, vol. 19, no. 4, pp. 1151–1163, Aug. 2011.
- [9] S. Govindasamy, "Asymptotic data rates of receive-diversity systems with MMSE estimation and spatially correlated interferers," *In review in IEEE Trans. Commun.*, vol. abs/1302.5002, 2013. [Online]. Available: <http://arxiv.org/abs/1302.5002>
- [10] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [11] W. S. Kendall and I. Molchanov, *New perspectives in stochastic geometry*. Oxford University Press, 2010.
- [12] P. Calka, "Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional poisson-voronoi tessellation and a poisson line process," *Advances in Applied Probability*, vol. 35, no. 3, pp. 551–562, 2003.
- [13] M. Tanemura, "Statistical distributions of poisson voronoi cells in two and three dimensions," *FORMA*, vol. 18, no. 4, pp. 221–247, 2003.
- [14] E. D. Habil, "Double sequences and double series," *Islamic University Journal-Series of Natural Studies and Engineering*, vol. 14, no. 1, pp. 223–254, 2006.

- [15] J. Mecke, "On the relationship between the 0-cell and the typical cell of a stationary random tessellation," *Pattern Recognition*, vol. 32, no. 9, pp. 1645–1648, 1999.
- [16] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.