Balancing Cooperation and Interference in Wireless Sensor Networks

Sam Vakil and Ben Liang

Department of Electrical and Computer Engineering

University of Toronto

Email:{vakil,liang}@comm.utoronto.ca

Abstract— We study the effect of cooperation in an interference limited, narrow-band wireless sensor network. Cooperation among available sensors can potentially lead to significant capacity increases. However, in an interference limited setting with asynchronous transmissions, exploiting more available sensors to help active sources will cause more interference to other sensors. Therefore, it is crucial to find the optimal trade-off between the amount of cooperation and the amount of interference introduced to the network. In this work we quantify the amount of cooperation using the notion of *relay zones* for each active sensor. The sensors that lie in such a zone are allowed to cooperate with the source. We then use the physical interference model to determine the probability that a relay node correctly decodes its corresponding source. Through numerical and simulation studies, we investigate the optimization of the relay-zone radius to maximize the network sum-rate based on relay availability and the sink reception capability. We show that the overall system capacity increases significantly under the proposed scheme, compared with cases where relay nodes are not exploited or where the relay zone radius is suboptimal.

I. Introduction

With the advent of a new generation of sensors with higher sensing and communication capabilities the challenge lies in forming a complex information gathering network to maximize the network capacity. In their landmark paper, Gupta and Kumar [1] showed that the per node capacity of an interference limited wireless network decreases as the number of nodes N increases and the achievable per node rate diminishes drastically.

However, relays can be exploited as a means to increase the capacity in a sensor network. The relay channel first introduced by van der Meulen in his PhD thesis leads to another communication scheme where instead of point to point communication between the source and destination, relays are exploited in a two-hop communication. The key capacity results for the case of a single relay were introduced by Cover and El Gamal in [2]. The capacity region for the relay channel with M relays is not known to date. However, Gastpar and Vetterli [3] have obtained upper and lower bounds on the capacity under Gaussian noise. These bounds are not tight for small values of M. However, as M increases the bounds tend to meet. In an information theoretic point of view the literature is rich on the subject of multi-user information theory, although the exact results are not known to date.

Cooperative diversity as a means to combat fading is another interesting issue in multi-user communication introduced mainly by Laneman *et al* [4] and Sendonaris *et al* [5]. In this setting, the relay channel is used to forward the data causing an increase in the capacity specifically for the cases where the source-destination channel experiences deep fades. The authors extend this work to benefit from co-operation of multiple nodes in [6]. Nabar *et al* [7] further evaluate the performance of co-operative schemes in the case of single source, single relay, and single destination and prove that full diversity can be obtained using this scheme. In [8] Sankaranarayanan, Kramer and Mandayam consider the case where multiple sources send their message to a relay, and the relay either simply forwards the data or first decodes and then forwards.

The studies above have shown that, cooperation among sensors can lead to significant capacity increases. However, if multiple source sensors and relaying nodes use the channel, their transmissions will interfere with each other. In the literature, it is usually assumed that there are a set of rules which lead to interference removal. For instance, different users might send their messages in a time slotted manner. However, this in general is not the optimum capacity achieving strategy. In a narrow-band sensor network with multiple sources the effect of interference greatly affects the network capacity and has to be considered.

Although the literature is rich in considering different relaying schemes in an information theoretic view point, the effect of relaying strategies and cooperation in a multiple source network, where the nodes have un-synchronized transmission has received much less attention. In a dense sensor network, a considerable fraction of sensors can be found idle during each transmission. The fundamental question is then "How should the idle sensors be optimally exploited to maximize the network capacity"? In this work we intend to answer this question and consider the optimal tradeoff between improved cooperation and increased interference when otherwise idle sensor node are used as relays.

The main contribution of this work includes

- A general framework for localized zone-based relaying in a large sensor network with multiple sources,
- A novel analytical framework to investigate the relation between the relay zone radius, the interference level, and the relay decoding probability,
- Derivation of the network sum-rate given multiple antennas at the data sink, as a main metric for relay zone optimization, in a MIMO multiple access setting, and
- Numerical and simulation studies to provide general

design guidelines for optimal relaying.

The rest of this paper is organized as follows. Section II explains the network model and presents a practical relaying algorithm. Section III presents details of our analytical framework to study the interaction between relaying and interference. Section IV computes the network sum-rate with relaying and multiple antennas at the data sink. Section V presents the numerical and experimental results. Finally, concluding remarks are given in Section VI

II. SENSOR NETWORK RELAYING ARCHITECTURE

In this section, we explain the sensor model under consideration and present a generic relaying architecture.

A. Data Dissemination

We consider a collection of N sensors placed randomly, uniformly and independently in the disk of unit area. It is assumed that the sensors can each measure a physical phenomenon. We assume the measurement of sensor i at time n to be represented by the random process $X_i(n)$. The measurements of two source sensors are assumed to be i.i.d processes justified by enough separation between them. Note that this is a result of the notion of relay zones in our model which causes the sensor within each zone to measure independently from a sensor in another zone. Each sensor measures the physical phenomenon in case it senses activity. The activity event in each slot is modeled using a Bernoulli random variable u_i for sensor i,

$$p(u_i) = \begin{cases} p_s, & \text{if} \quad u_i = 1 \\ 1 - p_s, & \text{if} \quad u_i = 0 \end{cases}. \tag{1}$$

Time is slotted to intervals of length L equal to the length of the data packet detected at an active sensor. The transmission is assumed to be half-duplex. Path-loss and channel variations are both considered in the model. When node i transmits with power $P_i[n]$, node j receives the transmission with power $P_i[n]\gamma_{ij}[n]$. The channel gain can be represented as

$$\gamma_{ij} = \frac{|h_{ij}|^2}{r_{ij}^\alpha}, \tag{2}$$

where r_{ij} is the distance between nodes i and j and h_{ij} models the fading channel from sensor i to j. Throughout the paper a block fading channel is considered for which the channel gain is constant over a block of length L.

We assume the sensor network has the same setting as SENMA [9]. The mobile access point (AP), i.e., data sink, is located at a variable height h, $0 < h < \infty$, above the center of the disk. The AP is equipped with sophisticated terminals and multiple receive antennas. In the following discussion we assume that AP has n_r receive antennas. However, each sensor only has one antenna.

B. Communication Scheme

The communication of a message at the active nodes is divided into two steps. The active sensor i first sends the message $X_i[n]$ to *potential relays* in its vicinity , forming a *relay zone*. In the second phase the potential relays which have

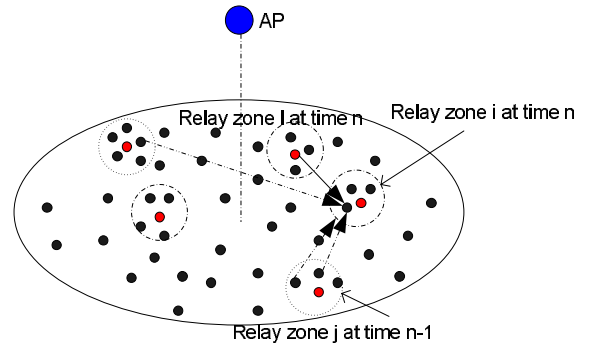Fig. 1. Sensor network layout. The relays that have decoded their message in time slot $n-1$ are interfering with the relays which are in their receive state in time n

successfully decoded the message forward it to the destination. The fraction of successful decoding nodes determines the optimum amount of cooperation among sensors. Since the successful decoding nodes are synchronized and deliver the same message $X_i[n]$, they are cooperating over the choice of the message sent towards the destination.

We assume that, the relays implement the *decode and forward* scheme. Then using the same physical model as the one introduced in [1], a relay node m is assumed to successfully decode the message sent from the source i if

$$\text{SINR}_m = \frac{P_i[k]\gamma_{im}[k]}{Z + \sum_{l \in \mathbb{S}[k], l \neq i} P_l[k]\gamma_{lm}[k] + \sum_{j \in \mathbb{R}[k-1]} P_j[k]\gamma_{jm}[k]} > \beta, \tag{3}$$

where the interference at potential relay nodes during the relay reception phase is either due to other sensors belonging to the active source set $\mathbb{S}[k]$, or the relays which have completed their reception in slot $k-1$ and are forwarding their corresponding message to AP in slot k. Such relays are represented as $\mathbb{R}[k-1]$. The noise is assumed to be Gaussian with variance Z. The parameter β is a design parameter and depends on the level of tolerated interference by the sensors. This is illustrated in Figure 1 with a simple example. Since we consider fading in our model the relays within each zone are still probable to go under deep fades. Therefore, some of the relays may not be capable of decoding successfully. However, the closer a relay sensor is to its corresponding source the higher the decoding probability is. Possibility of local communication between a source and its corresponding relays is another important benefit gained by introducing relay zones.

Since during each slot a different set of senders are activated, their corresponding relays are different. The relays from zone j in Figure 1 are one instant of the causes of interference at the potential relays in node i's relay zone. The source sensors which have started their transmission synchronously with source i are another cause of interference. Source l represents such an example.

Some transmission models allow the relays to transmit and receive simultaneously [2]. However, in the case of wireless sensor networks, physical constraints restrict the relay from simultaneous transmission and reception. Therefore, we define
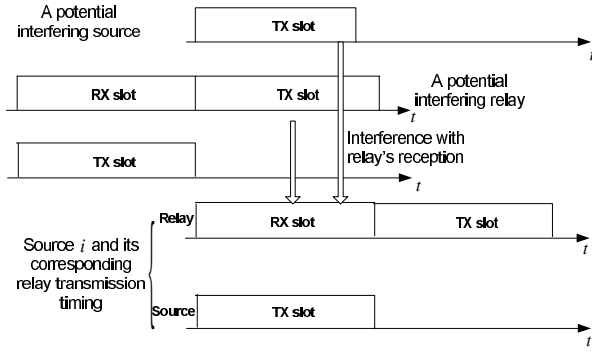
Fig. 2. Timing for a source and its corresponding relay in $\mathbb{REL}_i$ as well as possible interferer source and relays.

two states for each relay, *receive* and *transmit*. During the receive state the signal received at a relay is decoded correctly if (3) holds.

If a relay is in the relay zone of a source, and it has successfully decoded the message from the source, it transmits the message to the AP in the transmit state. During the relay transmission state we have the following expressions

$$Y_m[k] = 0, \quad \mathbf{Y}_d[k] = \sum_{i=1}^{N_p[k]} \mathbf{H}_i \mathbf{x}_i[k], \qquad (4)$$

where $Y_m[k]$ is the received message at a relay when it is in transmit phase, $\mathbf{Y}_d[k]$ is the received vector of dimension $n_r \times 1$ at the destination which is a superposition of the messages sent by all zones, $\mathbf{H}_i$ is the channel vector from the set of relays in zone i which have successfully decoded the message to the AP, $\mathbf{x}_i[k]$ represents the message vector sent out synchronously by the relays of zone i, and $N_p[k]$ represents the number of sources (relay zones) at time k.

C. Source Scheduling

To schedule a maximum number of active sources, we need to pack as many relay zones as possible in the disk area. As explained in Section II.A, each sensor generates a message at the beginning of slot k with probability p_s. Other sensors in the system at the beginning of each time slot can serve either as a relay or they can switch to the sleep mode and take no action. The timing of the network has been depicted in Fig. 2. The main challenge in the design of a relaying strategy is to decide whether an available node has to try to decode the source message or it has to remain in sleep mode over the next time slot.

We assume that the relay sensors can estimate their distance from the active sensors. The relaying decision will be made based on the relay's distance from the sources at the beginning of each transmission slot. Other active sources within the relay zone will not be allowed to send.

We define the ith relay-zone $\mathbb{REL}_i$ as a circle whose center is sensor i and has radius r_{rel}. The optimal value for this radius will be determined by the optimal relaying strategy. Then, in

slot $k-1$ the potential relay set is denoted by

$$\mathbb{R}[k-1] = \bigcup_{i=1}^{N_p[k-1]} \mathbb{REL}_i. \qquad (5)$$

During slot k the successful relay sensors forward their messages to the AP, and at the same time, a new set of sensors $\mathbb{S}[k]$ is activated with probability p_s. The source sensors $i \in \mathbb{S}[k]$ choose the potential relays based on the same criteria used during the prior slot (zone radius). The successful relays, $m[k-1] \in \mathbb{R}[k-1]$, from the last slot interfere with the reception of relays $m[k] \in \mathbb{R}[k]$ in slot k. Therefore, increasing the zone radius will result in more interference.

For a given optimal relay-zone area A (the details of derivation for zone radius are explained in Section III) the ultimate number of sources that can simultaneously transmit is $\frac{1}{A}$ since we considered the disk of unit area as the field over which the sensors are deployed. However, in practice this number strongly depends on the network topology. In the case of circular relay zones the source selection problem can be considered as a maximal circle packing problem which is NP-hard. Two sources i, j can simultaneously send their messages if they do not lie in the same relay zone. In other words their distance has to follow $d(i, j) \geq 2r_{rel}$. This problem can be translated into the maximal independent set (MIS) problem, by considering a graph and assuming that the active nodes represent the vertex set, and there is an edge between two vertices i and j if and only if their distance $d_{i,j} < 2r_{rel}$. Nodes within a relay zone can communicate locally to determine their relative distance. We have implemented the parallel algorithm presented in [10] by Luby to solve the MIS problem and found the maximal packing number in our experiments.

III. RELAY SET DECODING PROBABILITY AND INTERFERENCE ANALYSIS

In this section we quantify the interference and correct decoding probability at a relay, given the relay zone radius. Clearly, the amount of interference at a relay is a function of the number of interferers and their relative location to the relay. Therefore, we have to consider the permissible region (outside the relay zone) discussed in Section II-C as the possible interferer locations. During the source transmission phase a relay $m \in \mathbb{REL}_i$ decodes the corresponding source if and only if (3) holds. The interference at each relay can be due to either other sources $\{l | l \in \mathbb{S}[k], l \neq i\}$ sending simultaneously or the other relays which are in their transmission phase at time k and have started reception at time $k-1$, $\{j | j \in \mathbb{R}[k-1]\}$. The overall interference at m can then be expressed as

$$I_m[k] = \sum_{l \in \mathbb{S}[k], l \neq i} P\gamma_{lm}[k] + \sum_{j \in \mathbb{R}[k-1]} P\gamma_{jm}[k]$$
$$= \sum_{l=1}^{N_p} P\gamma_{lm}[k] + \sum_{j=1}^{D} P\gamma_{jm}[k], \qquad (6)$$

where D_i is the number of sensors which have been successful in decoding within $\mathbb{REL}_i$ and $D = \sum_{i=1}^{N_p} D_i$ represents the total number of interfering relays, and we have simplified the problem of relay selection by assuming that the sensors have

the same power P. We next formulate the relay-zone maximization problem and analytically find the expected number of successful decoding relays within each zone.

The goal of our design is to find the optimal area around each active sensor, the relay zone, in which the relays are allowed to *decode and forward*. We make the simplifying assumptions that this area is a circle and the radius of this circle is the same for all active nodes. The latter is justified by the fact that the amount of interference is not sensitive to relay zone location within the planar disk. However, the distance of the relay from its corresponding source is the determining factor. This fact will be elaborated in the Section III-C and depicted in Fig. 4. The metric of interest in this case is the network sum-rate $\sum_{i=1}^{N_p} R_i$, where R_i is the rate of information sent from the relay sensors within zone i towards the destination.

A. Expected Number of Decoding Relays

We denote the number of relays in zone i by the random variable N_r^i. Sensors are uniformly distributed over the unit disk area, therefore, the event $m \in \mathbb{REL}_i$ has a Bernoulli distribution with, $\Pr[m \in \mathbb{REL}_i] = A_i$, where $A_i = \pi r_{rel}^2$. N_r^i follows a binomial distribution as $\Pr(N_r^i = l) = \binom{N}{l} A_i^l (1 - A_i)^{N-l}$ with mean $E[N_r^i] = NA_i = N\pi r_{rel}^2$. In the following we quantify the expected number of successful decoding relays.

Theorem 1: In the given sensor network within region i, $E[D_i] = \int_0^{r_{rel}} N\Pr[SIR(r) > \beta] 2\pi r dr$, where $\Pr[SIR(r) > \beta]$ represents the successful decoding probability for a relay located at distance r relative to its source.

  Proof:

For the relay $m \in \mathbb{REL}_i$, the conditional distribution of $d(S(i), m)$, the distance between the source and the relay is

$$F_d(r|m \in \mathbb{REL}_i) = \Pr[d(S(i), m) < r|m \in \mathbb{REL}_i]$$
$$= \frac{\Pr[d(S(i), m) < r, m \in \mathbb{REL}_i]}{P[m \in \mathbb{REL}_i]} = \frac{\pi r^2}{\pi r_{rel}^2}, \quad (7)$$

since the sensors are uniformly distributed and the probability of lying in a specific region is proportional to the area of the region. Therefore, the conditional pdf of $d(S(i), m)$ can be written as

$$f_d(r|m \in \mathbb{REL}_i) = \frac{dF_d(r|m \in \mathbb{REL}_i)}{dr} = \frac{2\pi r}{\pi r_{rel}^2}. \quad (8)$$

We use a differential approach to find the expected number of decoding relays. Due to uniform sensor distribution, the number of sensors within each differential element of area inside $\mathbb{REL}_i$ is proportional to the element's area. Furthermore, the amount of SIR can be considered to be fixed within each such element. Thus, we can formulate the expected number of

successful decoding sensors as,

$$D_i = \int_0^{r_{rel}} N_r^i \Pr[SIR(r) > \beta] \frac{2\pi r}{\pi r_{rel}^2} dr \Rightarrow$$
$$E[D_i] = \int_0^{r_{rel}} E[N_r^i] \Pr[SIR(r) > \beta] \frac{2\pi r}{\pi r_{rel}^2} dr$$
$$= \int_0^{r_{rel}} \pi r_{rel}^2 \Pr[SIR(r) > \beta] \frac{2\pi r}{\pi r_{rel}^2} dr \quad (9)$$
$$= \int_0^{r_{opt}} N\Pr[SIR(r) > \beta] 2\pi r dr.$$

$\blacksquare$

We can further compute the expected number of total successful decoding relays within the network. We first present the following definition.

Definition 1: A *stopping rule* for a set of random variables Y_i is a positive, integer valued, random variable S such that for each $i > 1$ the indicator function $\delta[i]$, of $\{S \geq i\}$ is a function of $Y_1, \ldots Y_{i-1}$ [11].

Theorem 2: In the given sensor network the expected number of total successful decoding sensors satisfies $E[D] = E[N_p]E[D_i]$.

  Proof: Wald's equality ([11]) states that if $\{D_i\}$ are i.i.d random variables each with mean $E[D_i]$ and N_p is a stopping rule for $\{D_i\}$ and $D = D_1 + \ldots + D_{N_p}$, then $E[D] = E[D_i]E[N_p]$. We first show that N_p is a stopping rule for D_i. For any i the observation of D_i means that the scheduling algorithm permits zone i to be added to the set of relay zones. This is dependent on $1, \ldots, i-1$. In other words a new relay zone can be added if and only if it does not violate the requirements for source selection (its distance from other selected sources has to be more than $2r_{rel}$) and this will be dependent on the relative location of the new source with the previous ones. Therefore, N_p is a stopping rule and by Wald's equality $E[D] = E[N_p]E[D_i]$. $\blacksquare$

B. Expected Number of Interfering Sensors

Since the nodes are randomly distributed on the unit disk, the number of interferer nodes is a random variable. In time slot k the relays which have been in their reception phase in slot $k - 1$ interfere with the relays receiving in the current slot. The total number of interfering relays can therefore be formulated as

$$N_{I_{\text{relay}}}[k] = \sum_{j=1}^{N_p[k]} N_{I_{\text{relay}}}^j[k], \quad (10)$$

where $N_{I_{\text{relay}}}^j[k]$ represents the number of relays within relay zone j (which have been in receive mode during the previous slot $k - 1$ and are in transmit mode in the current slot) and $N_p[k]$ is the total number of permissible sources (relay zones) in slot k. The maximum number of relay zones that can be packed in the unit disk equals $\frac{1}{\pi r_{rel}^2}$ as explained in Section II-C. Furthermore, based on the network model assumption, the event of being an active source is Bernoulli with probability p_s. Therefore, the expected number of active sources equals $p_s N$. Since the number of sources can not exceed the above limit, we conclude $E[N_p] = \min(p_s N, \frac{1}{\pi r_{rel}^2})$.

The number of interferers $N_{I_{\text{relay}}}^j$ within each zone are i.i.d variables. Using a conditional expected value law we can write

$$E[N_{I_{\text{relay}}}] = E_{N_p}\big[E[N_{I_{\text{relay}}}|N_p = n_p]\big] = E_{N_p}\big[\sum_{j=1}^{N_p} E[N_{I_{\text{relay}}}^j]\big]$$
$$= E_{N_p}\big[N_p E[N_{I_{\text{relay}}}^j]\big] = E[N_p]E[N_{I_{\text{relay}}}^j]. \tag{11}$$

Finally, the total number $N_I[k]$ of interfering sensors is the sum of the above quantity and the number of sources in the current slot that are interfering with relay m's reception. This number is equal to $N_p[k] - 1$, since the source corresponding to the relay has to be removed from the interfering source set. Therefore, $E[N_I] = E[N_{I_{\text{relay}}}] + E[N_p] - 1$.

Thus far, for each relay sensor we have derived the expected number of interfering sensors.

C. Interference Analysis

In this section we explain the details for deriving an approximation to the interference value at each relay. The path loss roll-off factor is assumed to be equal to α. The channel from each interfering sensor l to the relay m is assumed to undergo Rayleigh fading. We further assume that the magnitude of fading is constant for each packet (quasi-static fading). The basis of our approach follows [12] in crux and is summarized below.

Interfering sensors to sensor m are constrained to be outside $\mathbb{REL}_i$. This is imposed by the scheduling discussed before. Figure 3 demonstrates a possible relay zone (the small circle). The region outside this circle is the potential interference zone. For each interfering node l the amount of interference to sensor m equals $I_{lm} = \frac{P}{r_{lm}^\alpha}|h_{lm}|^2$. We assume Rayleigh fading, therefore $|h_{lm}|^2$ has exponential distribution with parameter 1, and hence $E[I_{lm}] = \mu_l = \frac{P}{r_{lm}^\alpha}$.

We use a differential approach to evaluate the overall interference at m. Sensors are uniformly distributed, therefore, the number of sensors within each differential element of area is proportional to the element's area. Since there is a total of $N_I[k]$ such elements, in each differential element we have $N_I r dr d\theta$ sensors. A differential element $d\theta$ as depicted in Figure 3 causes the following amount of interference

$$dI = \frac{P}{r^\alpha} N_I r dr d\theta = \frac{P}{r^{\alpha-1}} N_I dr d\theta. \tag{12}$$

The overall expected interference at node m is, therefore, $E[I_m] = \int_{\mathfrak{S}} dI$, where $\mathfrak{S}$ is the potential interferer region.

For a differential element located at angle θ with respect to x' in Figure 3, the segment mh represents the distance of the maximum interferer, $d_{\max}(\theta)$, and the segment mj is the distance of m from the minimum possible interferer, $d_{\min}(\theta)$. The boundary equation for point h on the unit disk, with respect to axes x' and y' is

$$(x' + d(c,m))^2 + y'^2 = (\frac{1}{\sqrt{\pi}})^2, \tag{13}$$

where $d(c,m)$ represents the distance between disk center and the relay. Since $x' = d_{\max}(\theta)\cos\theta$ and $y' = d_{\max}(\theta)\sin\theta$, (13)
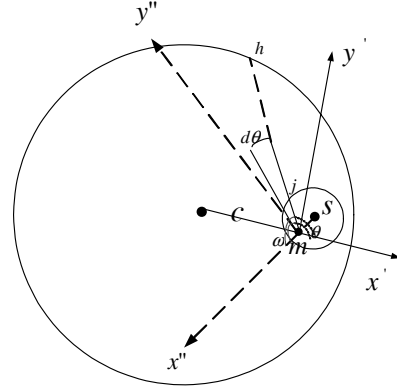

Fig. 3. Snapshot of the unit disk and a relaying zone. A source and its relay m have been considered. Differential elements of area have been considered to find the overall interference at m by integrating over this area.

results in $d_{\max}(\theta) = \sqrt{\frac{1}{\pi} - (d(c,m)\sin\theta)^2} - d(c,m)\cos\theta$. The same approach can be used to find the minimum distance $d_{\min}(\omega)$ by considering the small circle's equation relative to axis x'' and y'', which will result in $d_{\min}(\omega) = \sqrt{r_{\text{rel}}^2 - (d(s,m)\sin\omega)^2} - d(s,m)\cos\omega$, where ω is the angle of the differential element with respect to the axis x''. Note that $\omega + \theta = \angle x'mx''$ which is a constant. Therefore, the overall interference at m can be formulated as

$$E[I_m] = \int_0^{2\pi}\int_{d_{\min}(-\theta+\angle x'mx'')}^{d_{\max}(\theta)} N_I \frac{P}{r^{\alpha-1}} dr d\theta$$
$$= N_I P \int_0^{2\pi} \frac{r^{2-\alpha}}{2-\alpha}\Big|_{d_{\min}(-\theta+\angle x'mx'')}^{d_{\max}(\theta)} d\theta$$
$$= \frac{N_I P}{\alpha-2}\int_0^{2\pi}\big[\frac{1}{d_{\min}(-\theta+\angle x'mx'')^{\alpha-2}} - \frac{1}{d_{\max}(\theta)^{\alpha-2}}\big]d\theta. \tag{14}$$

In general, numerical integration is needed to be used to solve the above integral. However, in our analysis we have considered $\alpha = 4$. For this case it can be shown that

$$\int_0^\pi \frac{d\theta}{\left(\sqrt{a^2 - (b\sin\theta)^2} - b\cos\theta\right)^{\alpha-2}} = \frac{\pi a^2}{b^4 - 2b^2a^2 + a^4}. \tag{15}$$

Therefore, (14) can be further simplified as

$$E[I_m] = 2\frac{N_I P}{\alpha-2}\big(\frac{\pi^2}{\pi^2 d^4(c,m) - 2\pi d^2(c,m) + 1} - \frac{\pi r_{rel}^2}{d^4(s,m) - 2d^2(s,m)r_{rel}^2 + r_{rel}^4}\big). \tag{16}$$

Figure 4 represents the normalized value for the expected interference $\frac{E[I_m]}{N_I}$. The plot depicts the interference magnitude for different distances between the disk center c and the relay m. Also, the effect of the change of relay location relative to its corresponding source, within each zone has been considered. Note that interestingly the change of distance relative to the disk center does not cause substantial change in the expected interference value. However, within each zone, a possible relay which is closer to the zone boundaries undergoes a higher amount of interference as expected. The above observation suggests that the expected interference will not change dramatically as a function of the distance of the zone from
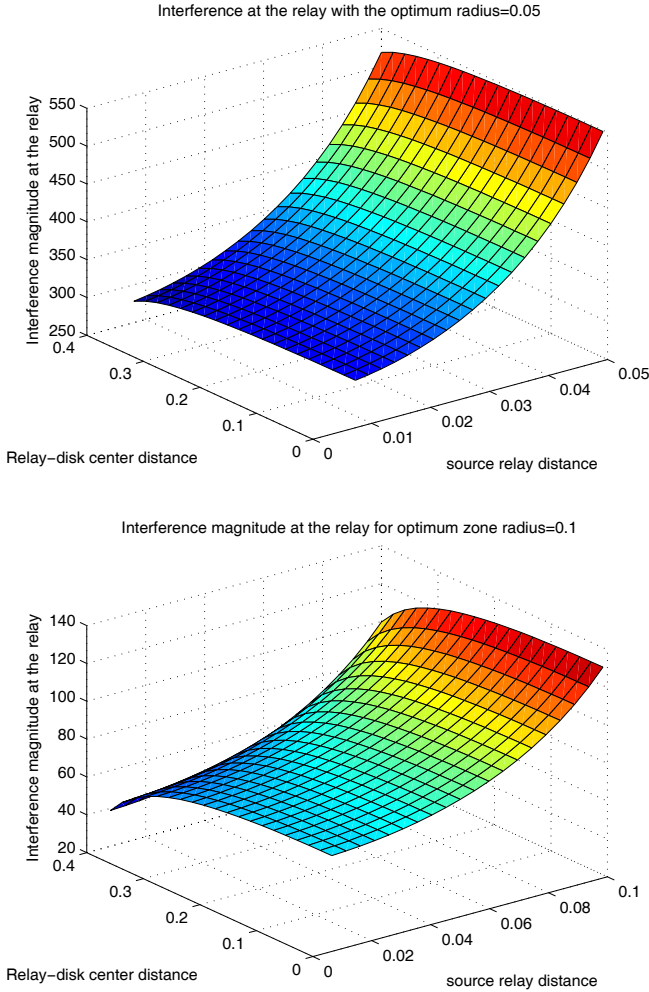
Fig. 4. The amount of interference (normalized by the number of interferers) at a relay versus relay's location in disk and its location relative to its source.

the disk center (neglecting the edge effects). However, the relative location of relay sensors within each zone is of more importance.

D. Decoding Probability

In this section we will give an approximation for $\Pr[SINR_m > \beta]$. Note that we consider equal power $P = 1$ for all nodes. We assumed that channel coefficients undergo Rayleigh fading, and therefore, the amount of signal and interference at each relay sensor are random variables. The interference at relay m can be formulated as $I_m = \sum_{j=1}^{N_I} I_m^j$, where each interference element has an exponential distribution with mean μ_l as we explained in section III-C. Computing this distribution in general becomes intractable, therefore, we approximate it by assuming that individual interference elements have equal means $\mu = \frac{E[I_m]}{E[N_I]}$. Using this approximation, it is clear that sum of $E[N_I]$ i.i.d exponential random variables with mean μ has Erlang distribution with parameter $E[N_I]$ and mean $\mu E[N_I] = E[I_m]$,

$$f_{I_m}(x) = \frac{x^{E[N_I]-1}}{(E[N_I]-1)!\mu^{E[N_I]}} e^{-\frac{x}{\mu}}, \quad \text{for} \quad x \geq 0. \quad (17)$$

The distribution of the signal power received at node m located at distance d from the corresponding source s can then be computed as follows. The received signal from s at m has the average power $\mu_s = \frac{P}{d^\alpha}$, and therefore, the distribution of the received power z in a Rayleigh environment follows, $f_Z(z) = \frac{1}{\mu_s} e^{-\frac{z}{\mu_s}}$. In a large network with interference, we assume that the amplitude of noise at each relay is small relative to the interference. Therefore, it suffices to find the distribution of SIR, $Y = \frac{Z}{I_m}$. Since Z and I_m are positive, the distribution of SIR is computed as,

$$
\begin{aligned}
F_Y(y) &= \int_{x=0}^{\infty} \int_{z=0}^{z=xy} f_Z(z) f_{I_m}(x) dz dx \\
&= \int_{x=0}^{\infty} \int_{z=0}^{z=xy} \frac{1}{\mu_s} e^{-\frac{z}{\mu_s}} \frac{x^{E[N_I]-1}}{(E[N_I]-1)!\mu^{E[N_I]}} e^{-\frac{x}{\mu}} dz dx \\
&= \int_{x=0}^{\infty} (1 - e^{-\frac{yx}{\mu_s}}) \frac{x^{E[N_I]-1}}{(E[N_I]-1)!\mu^{E[N_I]}} e^{-\frac{x}{\mu}} dx \\
&= 1 - \frac{1}{\left(1 + \frac{yE[I_m]}{\mu_s}\right)^{E[N_I]}},
\end{aligned}
$$
$$(18)$$

where we have used the table of integrals [13] to obtain the last equality. Thus, we can formulate $P[SIR > \beta] = 1 - F_Y(\beta) = \dfrac{1}{\left(1 + \frac{\mu\beta}{\mu_s}\right)^{E[N_I]}}$, where $\mu_s = \frac{P}{d^\alpha}$ is a function of the sensor's relative distance d with its corresponding source. The expected number of decoding nodes within each zone can therefore be formulated by the result of Theorem 1 and is a function of relay zone radius.

IV. NETWORK SUM-RATE

In Section III we introduced an analytical framework to quantify the number of successful decoding relays. To derive the network sum-rate, we consider the second phase during which the set of successful decoding relay sensors in slot k will forward the message to the AP. The successful decoding sensors $D_i[k+1]$ constitute a *cooperative MIMO* system, first analyzed in [14] in terms of power efficiency. The relay sensors within each zone $\mathbb{REL}_i$ serve as the multiple antennas sending a common message synchronously.

We formulate the problem of optimizing the relay-zone radius as maximization of the expected network sum-rate

$$r_{\text{opt}} = \arg\max_{r_{\text{rel}}} \sum_{i=1}^{N_p} \mathbb{E}[R_i(D_i)], \quad (19)$$

where D_i represents the number of nodes which have successfully decoded the message in zone i, and we use the notation $R_i(D_i)$ to clarify that the data rate corresponding to zone i is based on having D_i sensors in this zone. We upper bound the optimization problem given in (19) as

$$r_{\text{opt}} \simeq \arg\max_{r_{\text{rel}}} \sum_{i=1}^{N_p} R_i(\mathbb{E}[D_i]). \quad (20)$$

This approximation arises since the expected value of a concave function $f(x)$ obeys $\mathbb{E}[f(x)] \leq f[\mathbb{E}[x]]$ based on Jensen's

inequality. This makes intuitive sense, since the overall sum-rate is a monotonically increasing function of the number of transmitters. Therefore, the relay zone radius which results in the maximum expected number of decoding relays $\mathbb{E}[D_i]$ will lead to maximum the network sum-rate.

The capacity of a MIMO channel is derived in the landmark work of Telatar [15]. We further clarify the multiuser MIMO model and show that our setting follows the same scheme. We assume that the AP has access to channel state information for the capacity results to be true. Suppose the AP is equipped with n_r receive antennas. The number of transmit antennas in each zone i equals to $\mathbb{E}[D_i]$, the number of successful decoding sensors. Then, the uplink of a MIMO channel with multiple-users can be modeled as

$$\mathbf{Y} = [\mathbf{H}_1 \ldots \mathbf{H}_{N_p}] \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{N_p} \end{bmatrix} + \mathbf{Z}, \qquad (21)$$

where $\mathbf{Y}$ is the $n_r \times 1$ dimensional output vector, $\mathbf{H}_i$ is the $n_r \times \mathbb{E}[D_i]$ matrix representing the channel response from the cooperating sensors of zone i to the destination, and $\mathbf{x}_i$ represents the $\mathbb{E}[D_i] \times 1$ vector of the cooperative message sent from zone i. Note that since the sensors are located close to each other and at each instance we only consider the sensors which have successfully decoded the message, we can assume full cooperation and consider them as multiple-antennas sending the same message. Assuming the channel state information to be known at the receiver the formulation of capacity region with multiple receive antennas can be expressed as [16]

$$\sum_{i=1}^{M} R_i(\mathbb{E}(D_i)) \leq E_{\mathbf{H}}[\log \det(E_{n_r} + \frac{P}{Z} \sum_{i=1}^{M} \mathbf{H}_i \mathbf{H}_i^H)] \qquad (22)$$
$$\forall M, 1 \leq M \leq N_p,$$

where E_{n_r} is the $n_r \times n_r$ identity matrix. It is shown in [15] that for the case of sources with Gaussian distribution and channel matrices $\mathbf{H}_i$ with i.i.d complex Gaussian entries with mean zero, the above sum can be analytically expressed in terms of Laguerre polynomials. In our system model since it is assumed that the AP is located at a height h far enough from the sensors, the expected power received at the AP from all sensing sensors approximately equals $\frac{P}{h^\alpha}$. Assuming Rayleigh fading, the elements of each matrix $\mathbf{H}_i$ have a Gaussian distribution and are scaled by the above expected power factor. Hence, for our model $\mathbf{H}_i$ can be written as a scaled version of a matrix $\mathbf{H}_i'$ with zero mean complex Gaussian elements , $\mathbf{H}_i = \frac{1}{h^\alpha} \mathbf{H}_i'$, with P normalized to 1. We can now apply Theorem 2 in [15] to find an analytical expression for the network sum-rate. In Theorem 2 of [15] it is proved that the capacity of a single user MIMO channel with n_t transmit and n_r receive antennas with power constraint P_{total} on the transmit side and under Rayleigh fading equals

$$C(n_r, n_t, P_{\text{total}}) = \int_0^\infty \log(1 + \frac{P_{\text{total}}}{n_t}\lambda)$$
$$\sum_{s=0}^{f-1} \frac{s!}{(s+a-f)!}[L_k^{a-f}(\lambda)]^2 \lambda^{a-f} e^{-\lambda} d\lambda, \qquad (23)$$

where $f = \min(n_r, n_t)$, $a = \max(n_r, n_t)$, $L_s^{a-f}(x) = \frac{1}{s!}e^x x^{f-a} \frac{d^s}{dx^s}(e^{-x} x^{a-f+s})$ is the Associated Laguerre polynomial of order k [15]. The author further proves that under the multiuser setting with M senders each having power P_{total}, the sum-rate satisfies $\sum_{i=1}^{M} R_i(\mathbb{E}[D_i]) \leq C(n_r, Mn_t, MP_{\text{total}})$.

For the sensor network setting, we approximate the number of virtual antenna elements in each zone and the number of relay zones with their expected value, as we justified the use of Jensen's inequality. Therefore, the number of transmitter virtual antennas in each zone is $n_t = E[D_i]$, the number of receiver antennas is n_r and the power constraint for the transmitters within each zone is $P_{\text{total}} = n_t P$. Then, the achievable sum-rate satisfies $\sum_{i=1}^{E[N_p]} R_i \leq C(n_r, E[N_p]E[D_i], E[N_p]E[D_i]\frac{P}{h^\alpha})$. Since we assume all the $E[D_i]$ sensors within zone i have the same power P.

V. NUMERICAL ANALYSIS AND SIMULATION RESULTS

In this section we present numerical results based on the proposed analytical framework and compare them with simulation results. The capacity maximization problem has been solved numerically by changing the relaying zone radius and finding its optimum value. Clearly, the relay zone radius which causes the maximum number of *correctly decoding* relays within the network will lead to maximum network sum-rate. We have assumed the path loss roll-off factor to be $\alpha = 4$ within the planar disk as justified in [9] for sensor networks with low-lying antennas. The free space path loss factor between the sensors and the AP is however, considered to be $\alpha = 2$.

To avoid the possibility of very close sensors which will cause the amount of received signal to be unlimited in our model, a minimum distance ϵ is assumed between the sensors. For the unit disk with N sensors, $N\pi\epsilon^2 < A_{\text{disk}} = 1$ is needed to guarantee that all sensors can be located within the disk. We assumed $\epsilon = \frac{1}{\sqrt{5\pi N}}$. We also have considered the minimum threshold for decoding to be $\beta = 15$dB for the purpose of our simulation. For the simulation, the capacity results have been averaged over 25 different network topologies. In all cases, the AP is assumed to be located at height $h = 1$ above the network.

Figure 5 represents the network snapshot for one relay zone with the active source sensor for $r_{\text{rel}} = 0.05$ and its corresponding relays and the interferer sensors. Note that the idle sensors are not represented in this figure. In this example the average number of relays in a relaying zone with radius r_{rel} can be easily computed as $\mathbb{E}[N_r^i] = \pi r_{\text{rel}}^2 N = \pi (0.05)^2 1000 \simeq 8$.

Figure 6 presents log scaled plots of the sum rate for two different values of p_s. The effect of different numbers of receive antennas n_r based on the capacity results of the previous section has also been shown. As we expect, the total number of successful sensors in decoding determines the capacity. We observe that, the curves for different values of number of receive antennas have the same characteristic in terms of the point where the maximum sum-rate occurs. The determining factor for the network sum-rate is the total number of cooperative regions and the number of decoding sensors
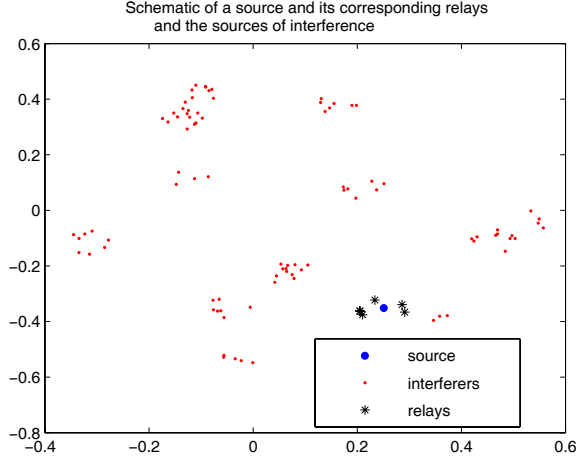
Fig. 5. Network schematic for a relay zone and the interferer sensors.
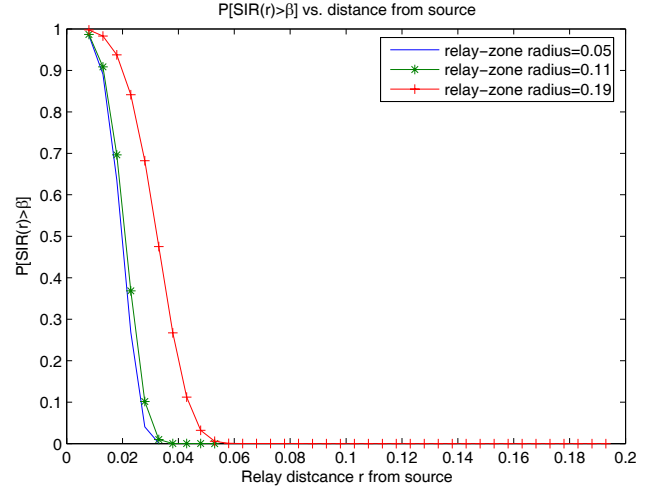

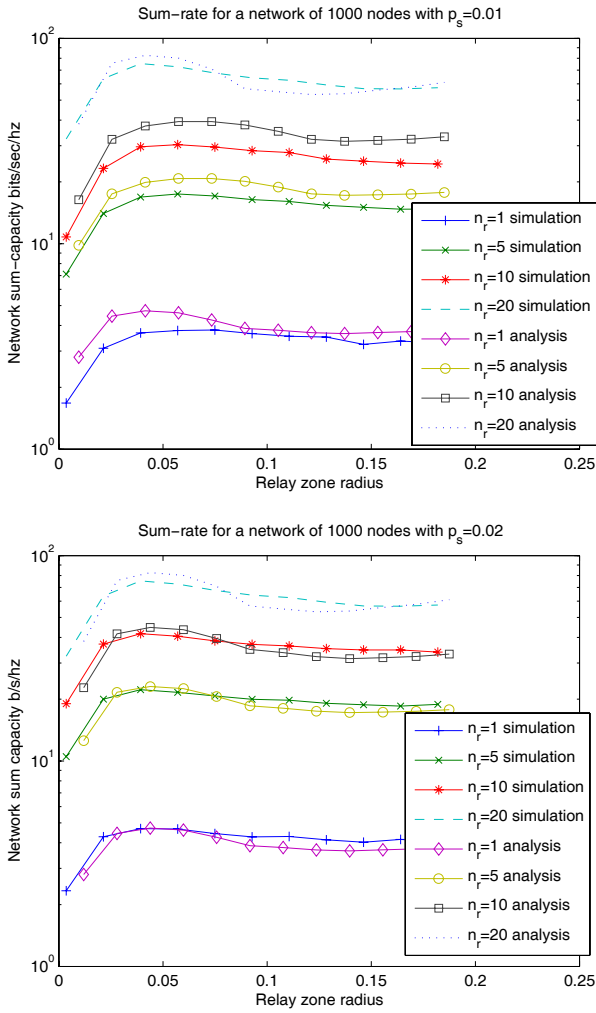Fig. 7. Probability of successful decoding versus relay distance from source.

Fig. 6. Network sum rate for different number of antennas at the destination.

within each region. Therefore, the optimum zone radius is the same for different values of n_r. However, the increase in the number of antennas will result in spatial multiplexing which causes the capacity increase shown in the curves. Furthermore, this figure suggests that the choice of the optimal zone radius is crucial for all values of n_r.

The main reasons for the difference between analysis and simulation are the edge effect, the approximations used, and the fact that it is not possible in general to quantify the number of active sources chosen by scheduling scheme analytically. Since sensors are randomly located, the actual number of sources chosen by the MIS algorithm is less than the number determined by the theoretical results.

In Figure 7 the probability of successful decoding at a relay vs. the relay distance from the source is given for three different values of $r_{\rm rel}$ with $r_{\rm opt} = 0.05$. As the figure suggests, increasing the zone-radius above the optimum value does not further improve the system performance.

From this figure it is clear that, the successful decoding probability is higher when a node is closer to its corresponding source. Also, as we increase the relay zone radius above its optimal value, the probability of successful decoding increases for the relays which are close to the source. This fact results from the scheduling algorithm, which imposes the interferers to be located further from the relays closer to the center of a relay zone for larger zones. However, as the figure demonstrates, the relays which are located far from the source decode successfully with a probability approaching zero. This is because the scheduling algorithm prevents the transmission of multiple sources within a relay zone, which results in more sources to be shut down in case of higher relay-zone radius. In this case the increase in the number of decoding relays (for the relays close to a specific source) can not compensate for the capacity loss due to the decrease in the number of sources. Therefore, the overall network capacity decreases.

As an example, consider a relay located at the distance 0.02 from its corresponding source in Figure 5. For the zone radius $r_{\rm opt} = 0.05$. the probability of successful decoding at this relay

equals 0.51. This value is almost the same for a zone radius $r_{opt} = 0.11$ and increases to 0.85 for $r_{opt} = 0.19$. However, in the latter case as we can see from the figure there is a very small probability of correctly decoding for the relays located further than $r_{opt} = 0.05$ from the source. In this case, by having a bigger zone the amount of interference in the network has increased while there is only little increase in the number of decoding relays which are close to the source. This results in the overall decrease of the sum-rate compared to the case of relaying with optimal zone radius.

VI. CONCLUSION

In this work we have studied the trade-off between exploiting sensors as relays and the increase in the amount of interference caused by asynchronous transmission of these relays in a dense sensor network. Based on this trade-off we have introduced the notion of relay zones whose radius can be optimized to maximize the overall network sum-rate. We give a closed form approximation, for the probability of correctly decoding at each potential relay node, and the average number of cooperating sensors within each zone is quantified. Assuming Rayleigh fading we have also computed the amount of interference within each zone and studied its variation at different locations within the disk and for different distances of the relay from its corresponding source. Numerical results based on the proposed analysis provide design guidelines for optimal relaying in interference limited sensor networks.

REFERENCES

[1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, March 2000.

[2] T. M. Cover and A. A. Elgamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 572–574, Sept. 1979.

[3] M. Gastpar and M. Vetterli, "The capacity of large gaussian relay networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 765–779, March 2005.

[4] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooprative diversity in wireless networks: Efficient protocols and outage behaviour," *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[5] A. Sendonaris, E. Erkip, and B. Azhang, "User cooperation diversitypart I: System description," *IEEE Transactions on Wireless Communications*, vol. 51, pp. 1927–1938, November 2003.

[6] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions Information Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[7] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: Performance limits and spacetime signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1099–1109, August 2004.

[8] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Hierarchical sensor networks: Capacity bounds and cooperative strategies using the multiple-access relay channel model," *IEEE SECON*, pp. 191–199, October 2004.

[9] P. Venkitasubramaniam, S. Adireddy, and L. Tong, "Sensor networks with mobile access: Optimal random access and coding," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1058–1068, August 2004.

[10] M. Luby, "A simple parallel algorithm for the maximal independent set problem," *Journal of the ACM*, vol. 15, pp. 1036–1055, November 1986.

[11] R. G. Gallager, *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1996.

[12] R. M. de Moraes, H. R. Sadjadpour, and J. Garcia-Luna-Aceves, "Throughput-delay analysis of mobile ad-hoc with a multi-copy relaying strategy," *IEEE SECON 2004*, pp. 200–209, October 2004.

[13] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York: Academic Press, 1980.

[14] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of mimo and cooperative mimo in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 2, pp. 1089–1098, August 2004.

[15] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunication*, vol. 10, pp. 585–595, November-December 1999.

[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, May 2005.