The ergodic rate density of ALOHA wireless ad-hoc networks

Yaniv George, Itsik Bergel ,*Senior Member, IEEE,* and Ephraim Zehavi, *Fellow, IEEE*

***Abstract*—In recent years, much attention has been paid to the analysis of random wireless ad hoc networks (WANETs) that combine the effect of the physical layer and the medium access layer. However, most works have concentrated on an outage rate model which does not accurately describe the performance of modern communication systems. In this work we consider the ergodic rate density (ERD) of a random ALOHA WANET with a homogenous Poisson point process node distribution. We present two novel lower bounds on the ERD, one for general transmission and reception strategies and the other for receivers with a spatial interference cancellation capability. The bounds are simpler than previously published results for ALOHA WANETs (that have primarily considered the outage rate density). In addition, the bounds and the bounding technique are quite general and enable the derivation of closed form expressions of the ERD for various network models. The efficiency and simplicity of the bounds are demonstrated through several applications, and insights are drawn on the behavior of the network performance as function of the path-loss factor, transmission strategies and number of antennas. Simulation results demonstrate that these simple lower bounds predict the performance of ALOHA WANETs with high accuracy.**

I. Introduction

Wireless ad-hoc networks (WANETs) offer simplicity and flexibility, which make them suitable for many practical applications. These networks do not depend on infrastructure such as base stations, and are typically coordinated by decentralized multiple access protocols. The ALOHA protocol, [2], is the simplest packet-based access mechanism. Together with its modified versions (e.g., [3], [4]) the ALOHA access protocol has attracted much attention both in theory and in practice.

While some interesting works on the capacity of WANETs have considered specific network structures (e.g., [5]), more general insights have been obtained from the analysis of random networks. The most popular model for the positions of active users in random WANETs is the homogeneous Poisson Point Process (PPP), [3]. In this model, the users' locations are assumed to be uniformly distributed over an infinite plane. The PPP model enables the analysis of WANET performance, and formulates the performance dependence on user density, without having to take the specific users' locations into account. In this work we use a PPP model to analyze the WANET performance. We focus on the access protocol and the physical layer processing, and use simplifying assumptions on the operation of higher communication layers.

Most previous analyses of WANETs have focused on the concept of outage capacity. In this approach, the transmitters encode each message using an error correction code, and transmit it (once or several times). The analysis is based on the calculation of the probability that the receiver will be able to decode the message correctly. If the receiver cannot decode the message, it is said to be in 'outage'. For example in the context of PPP, the popular Transmission Capacity (TC) measure, [6], is defined as the maximal area spectral efficiency subject to a fixed outage probability.

The effect of transmission strategies on the TC of ALOHA WANETS under fading channels was studied in [7] for fixed transmission power, channel inversion and threshold scheduling. Using upper and lower bounds on the outage probabilities they showed that the channel inversion scheme is inferior to the fixed transmission power scheme.

Upper and lower bounds on the TC of an ad-hoc network when each node is equipped with multiple antennas were derived in [8]. Although they considered the transmission of multiple streams by each transmitter, the optimal number of transmitting streams was shown to be one. Assuming channel state information in the transmitter, the best scaling of the transmission capacity with the number of antennas was achieved by transmit beamforming combined with interference cancellation in the receiver. In this case, the TC was shown to scale linearly with the number of antennas per node.

However, generally speaking, the analysis of the network TC involves quite complicated expressions. In order to make the TC analysis tractable, simplifying assumptions are commonly used; e.g., small outage [9], specific fading models [3] or specific path-loss factors [10].

An alternative approach, adopted here, evaluates the ergodic rate density (ERD) using the maximum mutual information between the transmitted signal and the received signal given the interferers activity, [1]. Basic results in information theory guarantee that such a rate, which is higher than the outage rate, is indeed achievable although it may incur significant delays (e.g., [11]).

An example of the use of the ergodic rate as a performance metric can be found in [5] where the authors developed a mathematical framework for the ergodic capacity region of specific topology WANETs under several transmission protocols. Haenggi, [12], introduced a lower bound on the ERD of users in ALOHA WANET. This bound was developed for the special case in which both the desired and interference channels fading follow the Rayleigh distribution. Stamatiou et al, [13], introduced upper and lower bounds on the performance of WANETs applying frequency hopping scheme. Their framework included Rayleigh fading channels and the ERD was evaluated under the assumption of no channel

Y. George, I. Bergel and E. Zehavi are with the Faculty of Engineering, Bar-Ilan University, Ramat-Gan.

Part of this work was published in SPAWC 2013, [1].

state information (CSI) on the interfering transmitters[1]. The resulting lower bound was shown to be tight for low user densities.

An asymptotic upper bound on the ERD of WANETs equipped with N receive antennas was introduced in [14]. In their model they assumed that the interferers keep a minimum distance from receivers. The SINR was shown to grow as $N^{\alpha/2}$ where α is the path-loss factor and N is the number of antennas.

Although the analysis of the ERD is simpler than the analysis of the Outage-Rate Density (ORD), to date there is no simple expression for ERD in the general case.

In this paper we analyze a random PPP WANET utilizing the ALOHA protocol, and present novel lower bounds on its ERD. The first bound holds for any reception strategy, while the second bound holds for receivers that apply spatial interference cancellation. Both bounds are much simpler than equivalent results obtained for the outage rate metric (e.g., TC). Nevertheless, these bounds are very close to the actual ERD, and can even predict the behavior of ORD as function of the network parameters; e.g.,path-loss factor, network density and number of antennas per node. The usefulness of the bounds is illustrated by several common applications: For the single antenna case, the applications include fixed transmission power, channel inversion and threshold scheduling. For multiple antenna WANETs, we present cases of transmit beamforming with or without interference cancellation in the receiver. The tightness of the derived bounds is also evaluated and shown by simulation.

The ERD analysis have three distinct advantages, which will be emphasized in what follows. First, the ergodic rate is achievable and higher than the outage rate. Second, this analysis results in a much simpler performance bounds, and third, its behavior is very similar to the outage rate, so that the derived results can also predict the general behavior in outage models.

The rest of this paper is organized as follows: The following subsection discusses presents various schemes that can achieve the ERD. Section II describes the system model. Section III introduces the novel lower bounds on the ERD. Section IV demonstrates the application of these bounds to five different scenarios and introduce further insights on the bounds. Section V presents our concluding remarks.

A. Achieving the Ergodic capacity

As mentioned above, most analyses of PPP WANETs have been based on outage rates. This approach is suitable for more traditional systems where a message is encoded and transmitted, and then either received successfully or discarded (retransmissions are also possible, but are decoded in the same way as new transmissions). However, modern communication schemes allow networks to approach the achievable ergodic rate. In this subsection we briefly describe some of these schemes to demonstrate the achievability of the ERD.

[1]This CSI assumption describes well system that use fast frequency hopping. In this paper we consider the complete CSI case, which can characterize slow/medium frequency hopping systems.

The most straightforward schemes which approach the ergodic rate utilizes time diversity. In this scheme, the data is encoded into code words which are spread over multiple ALOHA packets. If the number of packets per code word is large enough, the achievable rate will approach the ergodic rate. Note, however, that this method requires a significant delay, and hence it is not suitable for all applications.

A frequency diversity scheme can be implemented for example by random frequency hopping in which each code word is spread over different frequency bands [13]. Assuming that the amount of frequency bands is large, this scheme will effectively approach the ergodic rate without significant delay.

An alternative scheme is based on feedback from the receiver to the transmitter and is implemented by an incremental-redundancy hybrid automatic repeat request (IR-HARQ), [15]–[17], also known as HARQ type III. In this scheme the transmitter basically continues to generate new parity bits from the channel encoder until it receives an acknowledge (ACK) message from the receiver [18]. The IR-HARQ scheme is implemented in several wireless standards as UMTS HSDPA/HSUPA [19], mobile WiMAX [20], and the 3GPP Long Term Evolution (LTE) [21]. The optimal HARQ coding rate for ALOHA WANETs was studied in [22].

The incremental redundancy scheme can be integrated with the ALOHA access protocol using two possible strategies. In the first strategy all incremental redundancy bits are transmitted by ALOHA packets of fixed length. The transmitter continues to send packets until it receives an ACK message, and then moves on to the next codeword. This strategy will approach the ergodic rate when the average number of packets per codeword is large.

The second strategy uses variable packet sizes. The transmitter expands the length of the ALOHA packet and continues to transmit new parity bits until it receives the ACK message. In this scheme the mean packet length is the inverse of the ergodic rate, and using the renewal reward theorem [23] we conclude that the mean of the achievable rate approaches the ergodic rate. This strategy requires a much shorter delay than the previous strategy and is very useful in practical systems.

It is important to note that due to the network homogeneity, the same ergodic rate is achievable by each of the nodes, but only if we consider long enough averaging so that nodes mobility is significant. On the other hand, the HARQ schemes achieve the network ERD by allowing each user to transmit data at a different rate (that matches its instantaneous conditions). Thus, the ERD can be achieved in the network with very short delays, and longer delays are required only to achieve fairness between users.

II. System Model

We assume a decentralized wireless ad-hoc network utilizing an ALOHA protocol (e.g., [3]). Assuming the operation of a routing mechanism, some of the nodes have data that needs to be transmitted to specific destinations. For simplicity we assume that the next destination for each message is located at fixed distance, d, from the transmission source[2]. Nodes that

[2]Note that as the network density increases, this assumption become more realistic.

have data to transmit randomly decide on an access time to the network. The active transmitter distribution is modeled by a two dimensional PPP with density of λ.

The desired power, received at receiver i from its paired (i-th) transmitter is given by:

$$S_i = \rho(Y_i) d^{-\alpha} Y_i \tag{1}$$

where Y_i is the effective power fading between the i-th transmitter and its desired i-th receiver. This effective power fading represents any random change in the power of the received signal, for example due to the transmitter preprocessing, channel fading and receiver postprocessing.

We assume that the transmission power can be adapted according to the power gain of the desired channel, following some predetermined transmission policy. For ease of notation, we mark the transmitted power by $\rho(Y_i)$, where the function ρ describes the transmission strategy. The path-loss factor is denoted by $\alpha > 2$.

The power received at receiver i from transmitter j for $i \neq j$ is:

$$W_{i,j} = \rho(Y_j) X_{i,j}^{-\alpha} V_{i,j} \tag{2}$$

where $V_{i,j}$ is the effective power fading and $X_{i,j}$ is the distance between the i-th transmitter and the j-th receiver respectively. The fading variables $V_{i,j}$ are independent and identically distributed (i.i.d) and statistically independent of all distance variables. In the following we also use the notation V and Y when we discuss the statistical nature of one of the Random Variables (RVs) $V_{i,j}$ and Y_j, bearing in mind that these are single representatives of families of iid RVs. Note that in some cases the distribution of Y may be identical to the distribution of V, whereas in other cases these two RVs may have different distributions (as will be shown in section IV).

We use the shift invariant property of the system, [24], to analyze the performance of the network using user 0 as a probe receiver. Without loss of generality, we assume that the probe receiver is located at the origin.

For notational simplicity, in the following we drop the probe receiver index and define the set $\bar{\mathcal{A}}$ as the set of all active pairs, excluding the probe pair. The aggregate interference, measured at the probe receiver, can be written as:

$$I = \sum_{j \in \bar{\mathcal{A}}} W_j = \sum_{j \in \bar{\mathcal{A}}} \rho(Y_j) X_j^{-\alpha} V_j \tag{3}$$

and the power of the desired signal, measured at the probe receiver, can be written as: $S = \rho Y d^{-\alpha}$.

Without loss of generality we assume that $d = 1$. Denoting by σ^2 the contribution of the thermal-noise, the signal-to-interference-and-noise-ratio (SINR) measured at the probe receiver is given by $\dfrac{S}{\sigma^2 + I}$. The ergodic rate density (ERD) of a network with an active user density of λ is given by:

$$R(\lambda) = \lambda \cdot E\left[\log_2\left(1 + \frac{S}{\sigma^2 + I}\right)\right]. \tag{4}$$

III. Lower Bounds On the ERD

In this section we present two lower bounds on the ERD of ALOHA WANETs. Theorem 1 presents a bound that holds in general, whereas Theorem 2 presents a tighter bound for the specific case of WANETs with spatial interference cancellation capabilities.

Theorem 1 (General case): A lower bound on the ERD of a network with an active user density of λ is:

$$R(\lambda) \geq R_{\text{LB}}(\lambda) \tag{5}$$

where

$$R_{\text{LB}}(\lambda) = \lambda e^{\frac{2}{\alpha}-1} \cdot E\left[\log_2\left(1 + \frac{\rho(Y) \cdot Y}{\sigma^2 + C_\alpha \cdot \lambda^{\frac{\alpha}{2}}}\right)\right] \tag{6}$$

and

$$C_\alpha \triangleq \frac{2}{\alpha(\alpha-2)^{\frac{\alpha}{2}}} \left(\pi \alpha E\left[V^{\frac{2}{\alpha}}\right] E\left[\rho^{\frac{2}{\alpha}}(Y)\right]\right)^{\frac{\alpha}{2}}. \tag{7}$$

The proof of the lower bound is given in appendix A below and is based on Jensen's inequality. However, since the expectation on the aggregate interference in ALOHA WANETs is infinite, the derivation is split into two cases. The cases are defined by the power received from the strongest interferer, which is above or below a predefined threshold parameter. In the first case, the rate is lower bounded by 0, and in the second, the rate is lower bounded using Jensen's inequality. The power threshold parameter is then optimized to achieve a tighter bound.

Theorem 2 (Spatial interference cancellation): A lower bound on the ERD of a network with an active user density of λ, when each receiver cancels its M closest transmitters and $M \geqslant \left\lfloor \dfrac{\alpha}{2} \right\rfloor$ is:

$$R^M(\lambda) \geq R_{\text{LB}}^M(\lambda) \tag{8}$$

where

$$R_{\text{LB}}^M(\lambda) = \lambda \cdot E\left[\log_2\left(1 + \frac{\rho(Y) \cdot Y}{\sigma^2 + C_{\alpha,M} \cdot \lambda^{\frac{\alpha}{2}}}\right)\right] \tag{9}$$

and for $2 < \alpha < 4$

$$C_{\alpha,M} \triangleq \frac{2\pi^{\frac{\alpha}{2}}}{\alpha - 2} E[V] E[\rho(Y)] \left(M - \frac{\alpha}{4}\right)^{1-\frac{\alpha}{2}} \tag{10}$$

while for $4 \leq \alpha < 6$

$$C_{\alpha,M} \triangleq \frac{2\pi^{\frac{\alpha}{2}}}{\alpha - 2} E[V] E[\rho(Y)] \frac{\left(M - \frac{1}{2} - \frac{\alpha}{4}\right)^{2-\frac{\alpha}{2}}}{M-1}. \tag{11}$$

The proof of Theorem 2 is given in appendix B below. The proof is again based on Jensen's inequality, but this time using the fact that in this case the expectation of the aggregate interference is bounded. The next corollary shows that theorem 2 is tight for large number of antennas (for small number of antennas, a tighter bound can be obtained by combining the bounding techniques of Theorem 1 and Theorem 2).

Corollary 1: The lower bound on the ERD from Theorem 2 converges to the network ERD when the number of antennas grows to infinity:

$$\lim_{M \to \infty} \frac{R_{\text{LB}}^M(\lambda)}{R^M(\lambda)} = 1. \tag{12}$$

The proof of Corollary 1 is given in appendix C below.

IV. Applications of the Lower Bounds

In this section we apply the theorems of Section III to various scenarios. The obtained bounds for the different applications are also discussed, evaluated and compared to simulation results. At the end of the section we present several insights on the impact of the system parameters and the network policies on the ERD.

A. Single Antenna WANETs

In the single antenna case we have no pre/post-processing. Thus, the distribution of Y is identical to the distribution of V, and we assume that both follow an exponential distribution ($\sqrt{V}$ and $\sqrt{Y}$ have a Rayleigh distribution). Due to power normalization we set the fading power to $E[V] = E[Y] = 1$ (i.e., $V, Y \sim \text{Exp}(1)$). In this case, the expectation over V in (7) is:

$$E\left[V^{\frac{2}{\alpha}}\right] = \Gamma\left(1+\frac{2}{\alpha}\right) \tag{13}$$

where $\Gamma(\cdot)$ represents the Gamma function.

1) Single Antenna - Fixed Transmission Power: Consider a single antenna WANET that applies a fixed transmission power strategy. The fixed transmission power is applied by $\rho(Y) = \rho_0$. Substituting (13) into Theorem 1 results in:

$$\begin{aligned} R_{\text{LB}}^{\text{C}}(\lambda) &= \lambda e^{\frac{2}{\alpha}-1} \cdot E\left[\log_2\left(1+\frac{Y}{\frac{1}{\rho_0}\left(\sigma^2+C_\alpha^{\text{C}}\cdot\lambda^{\frac{\alpha}{2}}\right)}\right)\right] \\ &= \frac{\lambda}{\ln(2)} \cdot \exp\left(\frac{2}{\alpha}-1+\frac{1}{\rho_0}\left(\sigma^2+C_\alpha^{\text{C}}\cdot\lambda^{\frac{\alpha}{2}}\right)\right) \\ &\quad \cdot E_i\left(\frac{1}{\rho_0}\left(\sigma^2+C_\alpha^{\text{C}}\cdot\lambda^{\frac{\alpha}{2}}\right)\right) \end{aligned} \tag{14}$$

where

$$C_\alpha^{\text{C}} \triangleq \frac{2}{\alpha\left(\alpha-2\right)^{\frac{\alpha}{2}}}\left(\pi\alpha\Gamma\left(1+\frac{2}{\alpha}\right)\rho_0^{\frac{2}{\alpha}}\right)^{\frac{\alpha}{2}} \tag{15}$$

and $E_i(z)$ is the exponential integral function, which is the solution to the integral $\int_z^\infty \frac{e^{-t}}{t}dt$.

The ERD and the bound of (14) are depicted in Fig. 1 as a function of the active user density, λ. As can be seen, in this case the bound is about 30% below the ERD. Note also that the bound exhibits the same behavior as the ERD; hence it is useful to draw insights from its behavior. As a reference, Fig. 1 also depicts the lower bound introduced by Stamatiou et al., [13, eq. 26]. This bound was derived for the case of partial CSI. Yet, it is also a lower bound in the complete CSI case described herein. The Stamatiou bound is even tighter then the bound presented herein for very low user densities. Yet, for medium to high user densities, and in particular near the optimal user density, the bound derived herein is more useful.

Note on simulations:

The results are plotted using Monte-Carlo simulations and were averaged over 50,000 Network realizations. All simulations were performed in the interference limited regime in which the contribution of the thermal noise can be neglected. The channel was characterized by Raleigh fading and for most of the simulations the path-loss factor was $\alpha = 3$.

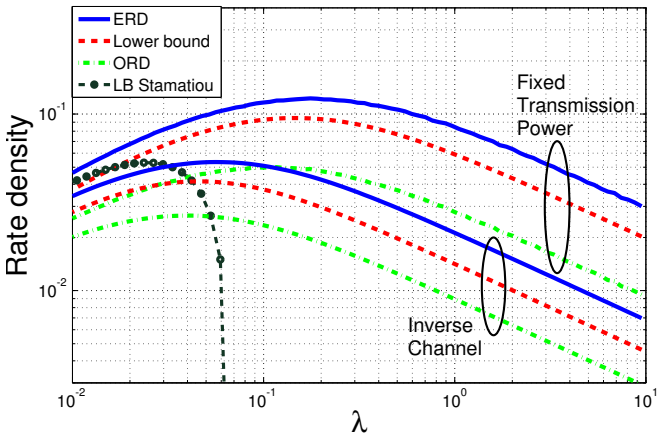

Fig. 1. Rate density as a function of the active user density for the fixed transmission power and inverse channel cases.

In most figures we also evaluated (as a reference) the performance of an outage scheme, in which each packet contained a code word of fixed length. In this case the performance metric was the ORD, given by:

$$R_{\text{out}}(\lambda) = \max_\beta \lambda \cdot P_{\text{out}}(\lambda,\beta)\cdot\log_2(1+\beta) \tag{16}$$

where, [25, eq. 5.83],

$$P_{\text{out}}(\lambda,\beta) \triangleq \Pr\left(\sum_{k=1}^{N_s}\log_2\left(1+\frac{S}{\sigma^2+I_k}\right)\geq N_s\cdot\log_2(1+\beta)\right), \tag{17}$$

N_s is the number of symbols in a packet, S and σ^2 are the signal and noise powers, and I_k is the interference power, experienced at the k-th symbol. Several works, [26]–[28], showed that (17) describes the behavior of modern error-correction codes[3].

Fig. 2 depicts the maximal rate densities (i.e., using the optimal user density) as function of the path-loss factor, α, for the fixed transmission power strategy. The figure depicts the ERD, the lower bound introduced by Haenggi [12], the lower bound described in (14), the ORD and a lower bound on the ORD[4]. As can be seen, both the ERD and its lower bound ,(14), share a similar slope; i.e., the lower bound can be used to anticipate the behavior of the ERD as function of the path-loss factor. The lower bound introduced by Haenggi [12] is tighter for large α and looser for smaller α. Note that the bound in [12] is valid only for Rayleigh fading channels and WANETs applying fixed transmission power.

2) Single antenna - Threshold Scheduling Strategy : Consider a WANET with single antenna nodes, that apply threshold scheduling strategy with a channel threshold of Y_t. The transmission strategy is described by:

$$\rho(Y) = \begin{cases} \rho_0 & \text{if } Y \geq Y_t \\ 0 & \text{Otherwise} \end{cases}. \tag{18}$$

[3]Unlike the ERD, the ORD is different for slotted and unslotted ALOHA. In slotted ALOHA the interference is constant within each slot, and the sum in (17) can refer to slots instead of symbols. In the following simulations we assume slotted ALOHA with one slot per packet; i.e., we used $N_s = 1$.

[4]For the last curve we used the upper bound expression on the outage probability as function of the SIR threshold (β) given in [7]. Then we optimized the SIR threshold parameter as described in (16).

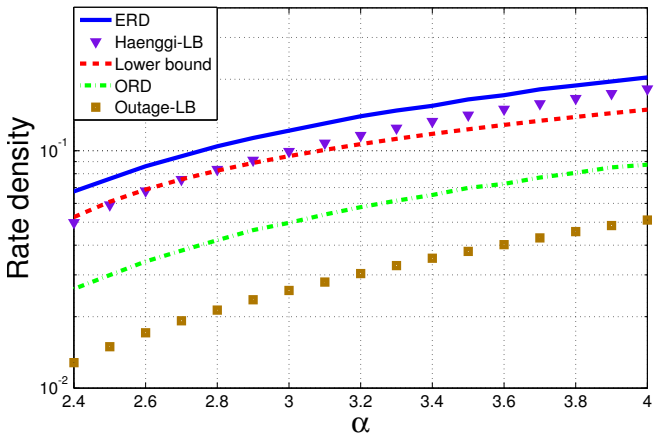

Fig. 2. Maximal rate density as function of the path-loss factor applying fixed transmission power.

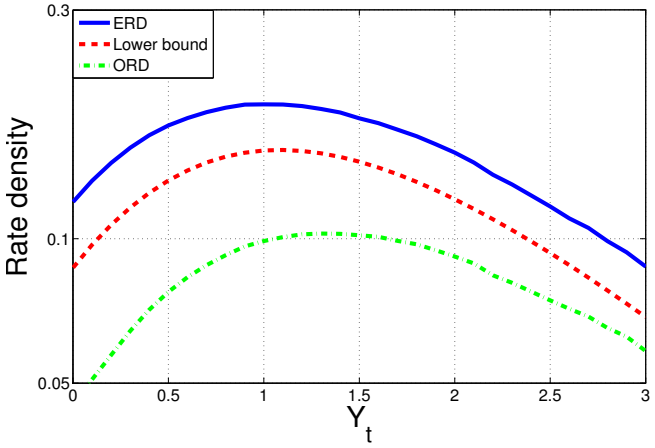

Fig. 3. Rate density as a function of the channel threshold for threshold scheduling with $\lambda = 0.26$.

For the Rayleigh fading case, the probability of transmission is $\Pr(Y > Y_t) = e^{-Y_t}$ and the probability density function of the desired channel is $f_Y(y|Y > Y_t) = e^{-y+Y_t}$. Using Theorem 1, the lower bound can be written as:

$$\begin{aligned} R_{\text{LB}}^{\text{T}}(\lambda) &= \lambda e^{\frac{2}{\alpha}-1} e^{-Y_t} \cdot E\left[\log_2\left(1+\frac{Y}{K_{\alpha,Y_t}}\right)\middle| Y > Y_t\right] \\ &= \lambda e^{\frac{2}{\alpha}-1-Y_t}\left(\log_2\left(1+\frac{Y_t}{K_{\alpha,Y_t}}\right)\right. \\ &\quad \left. + \frac{e^{K_{\alpha,Y_t}+Y_t}}{\ln(2)} \cdot E_i(K_{\alpha,Y_t}+Y_t)\right) \end{aligned} \tag{19}$$

where

$$K_{\alpha,Y_t}^{\text{T}} \triangleq \frac{1}{\rho_0}\left(\sigma^2 + C_{\alpha,Y_t}^{\text{T}} \cdot \lambda^{\frac{\alpha}{2}}\right) \tag{20}$$

and

$$C_{\alpha,Y_t}^{\text{T}} \triangleq \frac{2}{\alpha(\alpha-2)^{\frac{\alpha}{2}}}\left(\pi\alpha\Gamma\left(1+\frac{2}{\alpha}\right)\left(\rho_0^{\frac{2}{\alpha}} e^{-Y_t}\right)\right)^{\frac{\alpha}{2}}. \tag{21}$$

Using $E_i(z) > \frac{1}{2}e^{-z}\ln\left(1+\frac{2}{z}\right)$, [29], the lower bound in (19) can be simplified to:

$$\begin{aligned} R_{\text{LB}}^{\text{T}}(\lambda) &= \frac{1}{2}\lambda e^{\frac{2}{\alpha}-1-Y_t}\left(\log_2\left(K_{\alpha,Y_t}^{\text{T}}+Y_t\right)\right. \\ &\quad \left. + \log_2\left(K_{\alpha,Y_t}^{\text{T}}+Y_t+2\right) - 2\log_2\left(K_{\alpha,Y_t}^{\text{T}}\right)\right) \end{aligned} \tag{22}$$

which is very close to (19) for large values of Y_t.

Fig. 3 depicts three rate densities as function of the channel threshold parameter, Y_t, for an active user density of $\lambda = 0.26$. The rate densities are the ERD, the lower bound described in (22) and the ORD. In this case, the difference between the ERD and its bound is around 25%, and again both curves show the same behavior. In particular, both the ERD and its lower bound have a maxima around $Y_t = 1$, which implies that (22) can also be used to find the optimum threshold for WANET when applying threshold scheduling mechanism. The ORD is somewhat lower from the lower bound (on the ERD). The ORD of the threshold scheduling scheme was discussed in [7], but without any closed form expression.

3) Single antenna - Channel Inversion Strategy: We next analyze the case of a single antenna with a channel inversion power control strategy. Each transmitter adapts the transmission power to achieve a fixed received power in its destination receiver; i.e. $\rho(Y) = \rho_0 Y^{-1}$ where ρ_0 is a constant. This strategy leads to:

$$E\left[\rho^{\frac{2}{\alpha}}(Y)\right] = \rho_0^{\frac{2}{\alpha}} \cdot E\left[Y^{-\frac{2}{\alpha}}\right]. \tag{23}$$

Substituting (23) into Theorem 1 leads to the following lower bound:

$$R_{\text{LB}}^{\text{I}}(\lambda) = \lambda e^{\frac{2}{\alpha}-1} \cdot \log_2\left(1+\frac{1}{\frac{1}{\rho_0}\left(\sigma^2 + C_\alpha^{\text{I}} \cdot \lambda^{\frac{\alpha}{2}}\right)}\right) \tag{24}$$

where

$$C_\alpha^{\text{I}} \triangleq \frac{2}{\alpha}\left(\frac{\pi\alpha E\left[Y^{\frac{2}{\alpha}}\right] E\left[Y^{-\frac{2}{\alpha}}\right] \rho_0^{\frac{2}{\alpha}}}{(\alpha-2)}\right)^{\frac{\alpha}{2}} \tag{25}$$

and for Rayleigh fading

$$E\left[Y^{\frac{2}{\alpha}}\right] = \Gamma\left(1+\frac{2}{\alpha}\right),\ E\left[Y^{-\frac{2}{\alpha}}\right] = \Gamma\left(1-\frac{2}{\alpha}\right). \tag{26}$$

In the interference limited regime, in which the thermal noise can be neglected, we can also write a closed form expression for the optimal user density:

$$\begin{aligned} \lambda_\star^{\text{I}} &\triangleq \arg\max_\lambda R_{\text{LB}}(\lambda) \\ &= \frac{\left(\frac{\alpha-2}{\pi\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)^{-\frac{2}{\alpha}}}{E\left[Y^{\frac{2}{\alpha}}\right] E\left[Y^{-\frac{2}{\alpha}}\right]} \end{aligned} \tag{27}$$

where $W(\cdot)$ is the product-log function also known as the Lambert W function, which is defined as the inverse function of $f(w) = we^w$. The lower bound in this case can be written as:

$$\begin{aligned} R_{\text{LB}}^{\text{I}}\left(\lambda_\star^{\text{I}}\right) &= \frac{e^{\frac{2}{\alpha}-1}\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)}{\ln(2)\cdot E\left[Y^{\frac{2}{\alpha}}\right] E\left[Y^{-\frac{2}{\alpha}}\right]} \\ &\quad \cdot \frac{\left(\frac{\alpha-2}{\pi\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}}{\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)^{\frac{2}{\alpha}}}. \end{aligned} \tag{28}$$

Note that the equivalent closed form expression for the ORD metric is unknown, and in [7] only bounds on its outage probability were derived. Fig. 1 depicts the ERD, the lower bound, (24), and the ORD for this case. An analytical comparison between the performance of the fixed transmission power and the channel inversion strategies is presented in IV-C2.

B. Multi antenna WANETs

1) Transmit Beamforming: In this subsubsection we analyze the case of transmit beamforming with N_T antennas and a single receive antenna. We also assume the use of the simple channel inversion power control strategy (i.e., $\rho(Y)=\rho_0 Y^{-1}$ where ρ_0 is a constant). The preprocessing of beamforming over N_T antennas results in a Chi-square desired channel distribution, with $2N_T$ degrees of freedom, $2Y \sim \chi^2_{2N_T}$, which leads to:

$$E\left[\rho^{\frac{2}{\alpha}}(Y)\right]=\rho_0^{\frac{2}{\alpha}}\cdot E\left[Y^{-\frac{2}{\alpha}}\right]=\rho_0^{\frac{2}{\alpha}}\frac{\Gamma\left(N_T-\frac{2}{\alpha}\right)}{\Gamma(N_T)}. \quad (29)$$

This preprocessing does not change the interference channel statistics, V. Substituting (29) into Theorem 1 leads to:

$$R^{\mathrm{B}}_{\mathrm{LB}}(\lambda)=\lambda e^{\frac{2}{\alpha}-1}\cdot\log_2\left(1+\frac{1}{\frac{1}{\rho_0}\left(\sigma^2+C^{\mathrm{B}}_{\alpha}\cdot\lambda^{\frac{\alpha}{2}}\right)}\right) \quad (30)$$

where

$$C^{\mathrm{B}}_{\alpha}\triangleq\frac{2}{\alpha}\left(\frac{\pi\alpha\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(N_T-\frac{2}{\alpha}\right)\rho_0^{\frac{2}{\alpha}}}{(\alpha-2)\Gamma(N_T)}\right)^{\frac{\alpha}{2}}. \quad (31)$$

In the interference limited regime we can also write a closed form expression for the optimal user density:

$$\lambda^{\mathrm{B}}_{\star}=\frac{\left(\frac{\alpha-2}{\pi\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)^{-\frac{2}{\alpha}}\Gamma(N_T)}{\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(N_T-\frac{2}{\alpha}\right)} \quad (32)$$

where $W(\cdot)$ is the product-log function as defined in subsubsection IV-A3. The lower bound in this case can be written as:

$$R^{\mathrm{B}}_{\mathrm{LB}}\left(\lambda^{\mathrm{B}}_{\star}\right)=\frac{e^{\frac{2}{\alpha}-1}\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)}{\ln(2)\cdot\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(N_T-\frac{2}{\alpha}\right)} \cdot\frac{\left(\frac{\alpha-2}{\pi\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}\Gamma(N_T)}{\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)^{\frac{2}{\alpha}}}. \quad (33)$$

Using Kershaw's inequality, [30], [31],

$$\frac{\Gamma(N_T)}{\Gamma(N_T-\frac{2}{\alpha})}\geq\left(N_T-\frac{1}{2}-\frac{1}{\alpha}\right)^{\frac{2}{\alpha}} \quad (34)$$

and substituting into (33) simplifies the lower bound to:

$$R^{\mathrm{B}}_{\mathrm{LB}}\left(\lambda^{\mathrm{B}}_{\star}\right)\geq\frac{e^{\frac{2}{\alpha}-1}\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)}{\ln(2)\cdot\Gamma\left(1+\frac{2}{\alpha}\right)} \cdot\frac{\left(\frac{\alpha-2}{\pi\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}\left(N_T-\frac{1}{2}-\frac{1}{\alpha}\right)^{\frac{2}{\alpha}}}{\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)^{\frac{2}{\alpha}}}. \quad (35)$$

Fig. 4 depicts the ERD, the lower bound described in (30) and the ORD as function of the active user density, λ, for the case of transmit beamforming with inverse channel and 1, 3 and 9 transmit antennas. As the figure illustrates, the lower bounds for these three cases describe the network performance very well as function of both the density and the number of transmit antennas. As expected, the optimal density of users increases with the number of antennas.

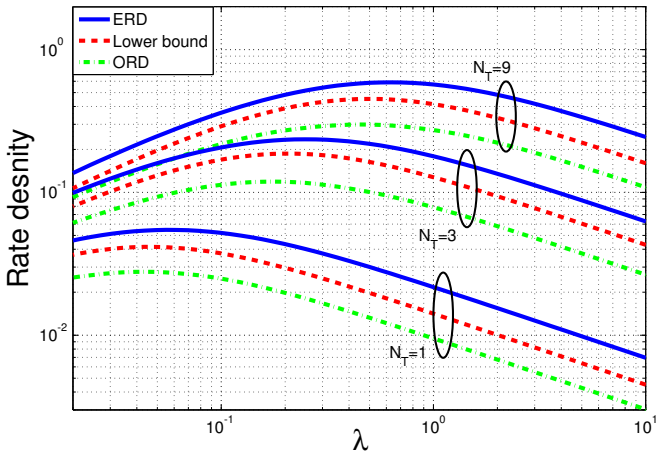

Fig. 4. Rate density as a function of the active user density for transmit-beamforming with 1, 3 and 9 antennas.

Note that Theorem 1 can also be applied to the case of receive diversity or combined receive and transmit diversity, by using the relevant probability function of the desired channel RV, Y (in this cases, the interference fading distribution does not change). In particular, the case of single transmit antenna and N_R receive antennas, results in the same ERD , (30), substituting only N_T by N_R.

2) Transmit Beamforming and Interference Cancellation: We assume $N_T\geq 2$ transmit antennas and $N_R\geq\lfloor\alpha/2\rfloor$ receive antennas for each node. The transmitter performs beamforming and channel inversion as described in subsubsection IV-B1. Each receiver uses its antennas to cancel its N_R-1 closest interferers.

The transmission power can be written as $\rho(Y)=\rho_0 Y^{-1}$. For the Rayleigh fading channel, $2Y$ again has a chi-square distribution with $2N_T$ degrees of freedom and

$$E\left[\rho(Y)\right]=\rho_0\cdot E\left[Y^{-1}\right]=\frac{\rho_0}{N_T-1}. \quad (36)$$

Defining $C^{\mathrm{S}}_{\alpha,N_R,N_T}\triangleq\rho_0\cdot K^{\mathrm{S}}_{\alpha,N_R,N_T}$ and substituting (36) into Theorem 2 leads to:

$$R^{\mathrm{S}}_{\mathrm{LB}}(\lambda)=\lambda\cdot\log_2\left(1+\frac{1}{\frac{1}{\rho_0}\left(\sigma^2+C^{\mathrm{S}}_{\alpha,N_R,N_T}\cdot\lambda^{\frac{\alpha}{2}}\right)}\right) \quad (37)$$

where for $2<\alpha\leq 4$ and $N_R\geq 2$:

$$K^{\mathrm{S}}_{\alpha,N_R,N_T}\triangleq\frac{2\pi^{\frac{\alpha}{2}}\left(N_R-1-\frac{\alpha}{4}\right)^{1-\frac{\alpha}{2}}}{(\alpha-2)(N_T-1)} \quad (38)$$

and for $4\leq\alpha<6$ and $N_R\geq 3$:

$$K^{\mathrm{S}}_{\alpha,N_R,N_T}\triangleq\frac{2\pi^{\frac{\alpha}{2}}\left(N_R-\frac{3}{2}-\frac{\alpha}{4}\right)^{2-\frac{\alpha}{2}}}{(\alpha-2)(N_T-1)(N_R-2)}. \quad (39)$$

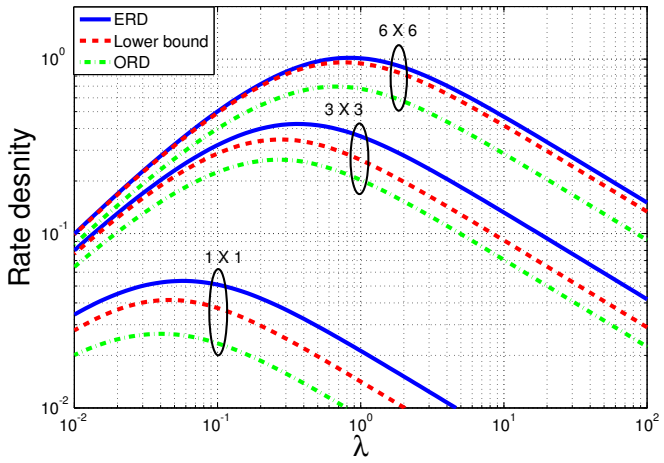

Fig. 5. Rate density as a function of the active user density for beamforming with interference cancellation for 1X1 , 3X3 and 6X6 antennas.

In the interference limited regime we can also optimize (37) and obtain a closed form expression for the optimal user density:

$$\lambda^{\mathrm{S}}_{\star}=\left(K^{\mathrm{S}}_{\alpha,N_R,N_T}\cdot\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)\right)^{-\frac{2}{\alpha}}. \quad (40)$$

Substituting (40) into (37) leads to:

$$R^{\mathrm{S}}_{\mathrm{LB}}\left(\lambda^{\mathrm{S}}_{\star}\right)=\frac{\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)}{\ln(2)} \cdot\left(K^{\mathrm{S}}_{\alpha,N_R,N_T}\left(\exp\left(\frac{\alpha}{2}+W\left(-\frac{\alpha}{2}e^{-\frac{\alpha}{2}}\right)\right)-1\right)\right)^{-\frac{2}{\alpha}}. \quad (41)$$

Fig. 5 depicts the rate densities for transmit beamforming and interference cancellation at the receiver, as function of the active user density, λ. For this simulation we used $N = N_T = N_R$. The curves of the ERD, the lower bound, (37), and the ORD, are plotted for the cases of $N = 3, 6$ antennas. As a reference we added the sum-rate curve of the single antenna inverse channel, (24). Note that the accuracy of the lower bound, ,(37), is indeed improving with the number of antennas as shown in Corollary 1.

C. Further Insights from the Lower Bounds

1) Impact of the path-loss factor on the ERD: As shown below, the lower bound on the maximal ERD for the inverse channel scheme in the interference limited regime, (28), can be approximated for large values of α by:

$$R^{\mathrm{I}}_{\mathrm{LB}}\left(\lambda^{\mathrm{I}}_{\star}\right)\cong\frac{\frac{\alpha}{2}+\log\left(\frac{\alpha}{2}\right)+\frac{1}{\alpha}\left(\log\left(\frac{\alpha}{2}\right)\right)^2}{\pi e^2\cdot\ln(2)\cdot E\left[V^{\frac{2}{\alpha}}\right]E\left[V^{-\frac{2}{\alpha}}\right]}. \quad (42)$$

This approximation shows the dependence of the maximal ERD in the channel model when the density of active users is optimized. In particular, it shows that the effect of the fading is summarized by a single constant (the product of the expectations in the denominator). The dependence of the ERD on the path-loss factor is more complicated, but, for large values of α the maximal ERD grows linearly with α.

The accuracy of (42) is depicted in Fig. 6, where the two upper curves depict the exact ERD, (28), and its approximation, (42), for the case of a single antenna and Rayleigh fading. As can be observed, this simple approximation captures the effect of α on the maximal ERD with high accuracy.

The approximated expression, (42), is derived from (28) through the following steps:

$$R^{\mathrm{I}}_{\mathrm{LB}}\left(\lambda^{\mathrm{I}}_{\star}\right)\overset{(a)}{\cong}\frac{\frac{\alpha}{2}e^{\frac{2}{\alpha}}\left(1-\frac{2}{\alpha}\right)\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}}{\pi e\cdot\ln(2)\cdot\left(e^{\frac{\alpha}{2}}-1\right)^{\frac{2}{\alpha}}\cdot E\left[V^{\frac{2}{\alpha}}\right]E\left[V^{-\frac{2}{\alpha}}\right]} \overset{(b)}{\cong}\frac{\left(\frac{\alpha}{2}\right)^{1+\frac{2}{\alpha}}}{\pi e^2\cdot\ln(2)\cdot E\left[V^{\frac{2}{\alpha}}\right]E\left[V^{-\frac{2}{\alpha}}\right]} \quad (43)$$

where we assumed large values of α and (a) used $\left|W\left(-\alpha/2e^{-\frac{\alpha}{2}}\right)\right|\ll\alpha/2$ (and hence we omitted the Lambert W terms) and (b) used $1-2/\alpha\approx e^{-\frac{2}{\alpha}}$ and $e^{\frac{\alpha}{2}}-1\approx e^{\frac{\alpha}{2}}$. Using the three dominant terms of the Maclaurin series of x^{-x} for $x=2/\alpha$ results in

$$\left(\frac{\alpha}{2}\right)^{\frac{2}{\alpha}}\approx 1-\frac{2}{\alpha}\log\left(\frac{2}{\alpha}\right)+\frac{2}{\alpha^2}\log^2\left(\frac{2}{\alpha}\right) \quad (44)$$

which together with (43) led to (42).

Note that this growth of the ERD in α is guaranteed only for the maximal ERD, i.e., where the density of active users is optimized. This conclusion does not hold for an arbitrary selection of active user density. For some user densities the ERD even decreases with the rise in α as shown by the following Proposition:

Proposition 1: For any $\alpha > 2$ there exists λ_α such that $\left.\frac{\partial R_{\mathrm{LB}}(\lambda)}{\partial\alpha}\right|_{\lambda=\lambda_\alpha}<0$.

Proof of Proposition 1: The lower bound (6) is the multiplication of two terms. The left expression, $\lambda e^{\frac{2}{\alpha}-1}$, clearly decreases with α. The right expression, which can be written as

$$\left.E\left[\log_2\left(1+\frac{\rho(Y)\cdot Y}{\sigma^2+x}\right)\right]\right|_{x=C_\alpha\cdot\lambda^{\frac{\alpha}{2}}} \quad (45)$$

decreases in x. Selecting λ that will guarantee that $\frac{\partial\left(C_\alpha\cdot\lambda^{\frac{\alpha}{2}}\right)}{\partial\alpha}\geq 0$ will ensure that (45) is decreasing with α. This can be obtained by selecting the active user density to be $\lambda_\alpha\geq\exp\left(-\frac{\partial C_\alpha}{\partial\alpha}\cdot\frac{2}{C_\alpha}\right)$ (which is positive and finite for any $\alpha > 2$). □

The lower curve in Fig. 6 depicts the lower bound (24) for Rayleigh fading and $\lambda = 3$ as function of the path-loss factor. As can be noticed this curve is not monotonic, and actually decreases for $\alpha > 3.1$.

2) Fixed transmission Power versus Inverse Channel: In the following subsubsection we compare the lower bound of the fixed transmission power, (14), to the lower bound of the inverse channel strategy, (24), for any active user density.

Lemma 1: For a Rayleigh fading channel, using fixed transmission power gives a higher ERD than channel inversion, for any active user density:

$$R^{\mathrm{C}}_{\mathrm{LB}}(\lambda) > R^{\mathrm{I}}_{\mathrm{LB}}(\lambda) \quad (46)$$

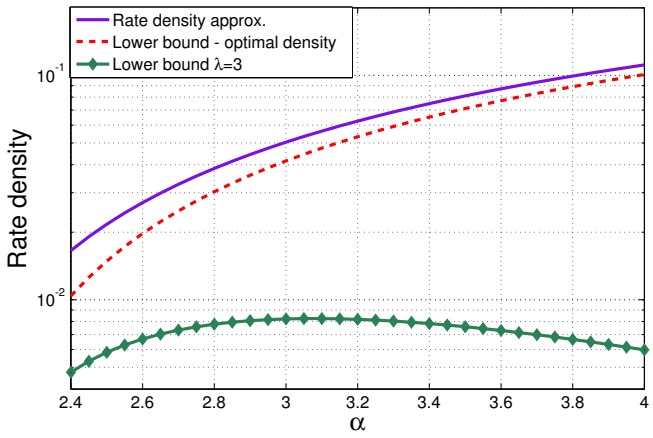

Fig. 6. Rate density as function of the path-loss factor for the rate density approximation (42), the lower bound on the maximal ERD (28) and the lower bound (24) with $\lambda = 3$.

Proof of Lemma 1: Define:

$$g(x) \triangleq \frac{R^{\mathrm{C}}_{\mathrm{LB}}((x/C^{\mathrm{a}}_{\alpha})^{2/\alpha}) - R^{\mathrm{I}}_{\mathrm{LB}}((x/C^{\mathrm{a}}_{\alpha})^{2/\alpha})}{\frac{\lambda}{\ln(2)} e^{\frac{2}{\alpha}-1}}$$
$$= e^{x} \cdot E_i\left(x\right) - \ln\left(1 + x^{-1}\left(\Gamma\left(1-\frac{2}{\alpha}\right)\right)^{-\frac{\alpha}{2}}\right). \quad (47)$$

We next prove that $g(x) > 0$ for any $x > 0$. We start by noting that $g(x)$ decreases monotonically with α, and hence we can lower bound it by

$$g(x) \geq \lim_{\alpha\to\infty} g(x)$$
$$= \lim_{\epsilon\to 0} e^{x} \cdot E_i\left(x\right) - \ln\left(1 + x^{-1}\left(\Gamma\left(1-\epsilon\right)\right)^{-\frac{1}{\epsilon}}\right)$$
$$= e^{x} \cdot E_i\left(x\right) - \ln\left(1 + x^{-1}e^{-\gamma}\right) \quad (48)$$

where the last equality used the power series of $\Gamma(1+z)$, [32], and γ is the Euler-Mascheroni constant ($\gamma = 0.57721...$). It is easy to show that $g(x) > 0 \ \forall \ x > 0$. □

This superiority of the fixed transmission power policy over the channel inversion policy was also shown in [7] for the ORD metric. Fig. 1 compares the channel inversion and fixed power transmission strategies. The lower bounds of the fixed transmission power scheme is given by (14), and of the channel inversion scheme is given by (24). As was proven above, the ERD bound for the fixed power scheme is superior to the ERD bound for the channel inversion scheme. The figure shows that this conclusion also holds for the actual ERD and ORD.

3) The Expected ERD Gain from Utilization of Beamforming: From the ratio between (35) and (28) we can find the gain from utilization of N_T transmit antennas:

$$\frac{R^{\mathrm{B}}_{\mathrm{LB}}\left(\lambda^{\mathrm{B}}_{\star}\right)}{R^{\mathrm{I}}_{\mathrm{LB}}\left(\lambda^{\mathrm{I}}_{\star}\right)} = \Gamma\left(1-\frac{2}{\alpha}\right) \cdot \left(N_T - \frac{1}{2} - \frac{1}{\alpha}\right)^{\frac{2}{\alpha}}. \quad (49)$$

Note that for $N_T \to \infty$, the lower bound (35) scales as $N_T^{\frac{2}{\alpha}}$. The same scaling was also observed for the ORD [9] for large enough N_T.

Equation (49) anticipates that the expected gain of utilizing 3 and 9 antennas compared to a single antenna is 4.5 and 10.9 respectively. The measured ratio between the ERD utilizing 3 and 9 antennas compared to a single antenna in Figure 4 is 4.3 and 10.8 respectively. Hence, equation (49) is useful for estimating the expected gain as function of the number of transmit antennas.

4) The ERD Gain of Beamforming and Interference Cancellation: Defining $J_{\alpha} \triangleq \left(\frac{1-\frac{2}{\alpha}}{e}\right)^{\frac{2}{\alpha}-1} \cdot \Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)$ and finding the ratio between (41) and (28), we can obtain the ERD gain from the utilization of $N_T \times N_R$ antennas.

The gain ratio for $2 < \alpha \leq 4$ and $N_R \geq 2$ is :

$$\frac{R^{\mathrm{S}}_{\mathrm{LB}}\left(\lambda^{\mathrm{S}}_{\star}\right)}{R^{\mathrm{I}}_{\mathrm{LB}}\left(\lambda^{\mathrm{I}}_{\star}\right)} = J_{\alpha} \cdot \left(N_T - 1\right)^{\frac{2}{\alpha}} \left(N_R - 1 - \frac{\alpha}{4}\right)^{1-\frac{2}{\alpha}} \quad (50)$$

and for $4 \leq \alpha < 6$ and $N_R \geq 3$ we get:

$$\frac{R^{\mathrm{S}}_{\mathrm{LB}}\left(\lambda^{\mathrm{S}}_{\star}\right)}{R^{\mathrm{I}}_{\mathrm{LB}}\left(\lambda^{\mathrm{I}}_{\star}\right)} = J_{\alpha} \cdot \left(N_T - 1\right)^{\frac{2}{\alpha}} \left(N_R - 2\right)^{\frac{2}{\alpha}} \left(N_R - \frac{3}{2} - \frac{\alpha}{4}\right)^{1-\frac{4}{\alpha}}. \quad (51)$$

Note that in the special case where $N_T = N_R = N$ the ERD scales linearly with N for large enough N. This scaling was again observed for the ORD metric [8]. Note, however, that [8] only presented asymptotic results for an infinite number of antennas.

Moreover, from (50) and (51) one can observe that for $\alpha < 4$ the role of the beamforming and the transmit antennas is more important than the interference cancellation and the receive antennas. For $\alpha > 4$ the roles change and the receive antennas has more effect than the transmit antennas.

Fig. 7 depicts the maximal rate density as function of the number of transmit antennas for transmit beamforming with and without interference cancellation at the receiver. In the case of interference cancellation, we assumed the same number of antennas in the receiver and transmitter; i.e., $N = N_T = N_R$. Equations (41) and (35) were used to plot the lower bound curves for the cases with and without interference cancellation respectively. As shown in subsubsection IV-C3, the curve of the ERD scales as $N^{\frac{2}{\alpha}}$ for without interference cancellation and linearly with N for the interference cancellation case. Moreover, as stated in Corollary 1 the lower bound for the interference cancellation case converges to the ERD for large N.

V. Conclusions

In this work we analyzed the ergodic rate density (ERD) of ALOHA WANETs assuming a homogenous PPP distribution of nodes. The ERD is achievable in modern communication systems and was shown to result in a simpler analysis. We presented two novel lower bounds on the ERD. The first bound holds for a general reception strategy and the second bound holds for receivers with spatial interference cancellation.

The usefulness of the two lower bounds was demonstrated by five applications. The first three are single antenna

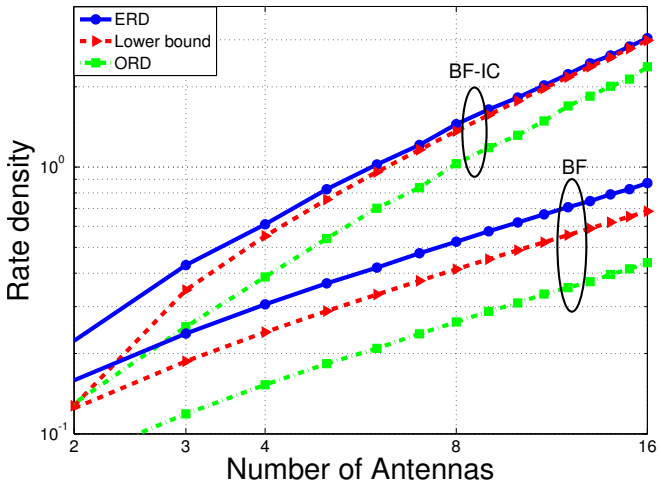

Fig. 7. Rate density as a function of number of antennas for transmit beamforming with and without interference cancellation in the receiver.

WANETs with different transmission strategies (fixed transmission power, threshold scheduling and inverse channel). The fourth and fifth applications are multiple antenna WANETs applying transmit beamforming, with or without interference cancellation in the receivers. For each application we present closed from expressions of the bounds as function of the system parameters.

These closed form expressions are further used to draw insights on the impact of the system parameters on the ERD. The maximal ERD is shown to grow linearly with the path-loss factor for large values of the path-loss factor. However, for fixed active user density we showed that the ERD is not necessarily a monotonic function of the path-loss factor. We also presented accurate and simple expressions for the maximal-ERD gain as function of the number of antennas with and without interference cancellation. These expressions can be used to study the scaling of the ERD for large number of antennas. In addition, these expressions can also provide the accurate gain for small number of antennas.

The bounds are shown to be tight although fairly simple. These bounds are quite general and can be easily applied to other network models. Future research will adapt these bounds for the analysis of more complicated networks and the joint analysis of additional network layers.

Appendix A
Proof of Theorem 1

To prove the theorem we distinguish between two different cases based on the power of the strongest interferer. For this purpose we define the strongest interferer power:

$$W_{\max} = \max_j W_j \tag{52}$$

and use the law of total expectation to write the ERD, (4), as

$$\begin{aligned} R(\lambda) = &\lambda \cdot \Pr(W_{\max} > \delta) E\left[\log_2\left(1+\frac{S}{\sigma^2+I}\right)\middle| W_{\max} > \delta\right] \\ &+\lambda \cdot \Pr(W_{\max} \le \delta) E\left[\log_2\left(1+\frac{S}{\sigma^2+I}\right)\middle| W_{\max} \le \delta\right]. \end{aligned} \tag{53}$$

Note that the impact of δ on the two terms in (53) is different. For large enough values of δ the first term (the case of $W_{\max} > \delta$) is very small and can be neglected. On the other hand, the second term ($W_{\max} \le \delta$) becomes small for small values of δ. Thus, it is important to choose a proper value of δ that balances the two terms. Yet, to simplify the bound, we lower bound the first term with 0 and the second term using the Jensen's inequality, resulting with:

$$\begin{aligned} R(\lambda) &\ge \lambda \cdot \Pr(W_{\max} \le \delta) E\left[\log_2\left(1+\frac{S}{\sigma^2+I}\right)\middle| W_{\max} \le \delta\right] \\ &\ge \lambda \cdot \Pr(W_{\max} \le \delta) E\left[\log_2\left(1+\frac{S}{\sigma^2+E\left[I|W_{\max} \le \delta\right]}\right)\right]. \end{aligned} \tag{54}$$

In this case, the Jensen inequality implies that given the average power of a (conditionally) Gaussian interference, the worst interference is the one with constant power. This property was already observed and used in various works (see for example [33], [34]). Note that even though the first term in (53) was lower bounded by 0, it is still important to choose δ that will make the bounds as tight as possible.

Let $\Phi_\delta(s)$ be the characteristic function that corresponds to the conditional distribution of I given $W_{\max} \le \delta$. The characteristic function of an aggregate interference, measured at the middle of a circular guard zone of radius A within a 2-dimensional PPP with density λ and a fading variable K is (using [35], [36] and the superposition property, [37], see also [38]):

$$\Phi(s) = \exp\left(-\lambda E\left[\int_A^\infty 1 - e^{-sKr^{-\alpha}}\right] 2\pi r dr\right). \tag{55}$$

where A, K can be random variables. Therefore, the characteristic function of the conditional aggregate interference, measured at a probe receiver with guard zone of $\left(\frac{\delta}{V\cdot\rho(Y)}\right)^{-\frac{1}{\alpha}}$ is:

$$\Phi_\delta(s) = \exp\left(-\lambda E\left[\int_{\left(\frac{\delta}{V\cdot\rho(Y)}\right)^{-\frac{1}{\alpha}}}^\infty \left(1-e^{-sV\cdot\rho(Y)t^{-\alpha}}\right) 2\pi t dt\right]\right). \tag{56}$$

Since the maximum interference contributed by a single interferer is bounded, the first moment of the conditional aggregate interference exists and can be calculated by:

$$\begin{aligned} E\left[I|W_{\max} \le \delta\right] &= -\frac{d}{ds}\ln\left(\Phi_\delta(s)\right)|_{s=0} \\ &= 2\pi\lambda E\left[V\cdot\rho(Y)\int_{\left(\frac{\delta}{V\cdot\rho(Y)}\right)^{-\frac{1}{\alpha}}}^\infty t^{-\alpha+1} dt\right] \\ &= 2\pi\lambda E\left[V\cdot\rho(Y)\left.\frac{t^{-\alpha+2}}{-\alpha+2}\right|_{\left(\frac{\delta}{V\cdot\rho(Y)}\right)^{-\frac{1}{\alpha}}}^\infty\right] \\ &= \left(\frac{2\pi}{\alpha-2}\right) E\left[V^{\frac{2}{\alpha}}\right] E\left[\rho^{\frac{2}{\alpha}}(Y)\right] \lambda\delta^{1-\frac{2}{\alpha}} \\ &= \left(\frac{2\pi}{\alpha-2}\right) L_\alpha \lambda\delta^{1-\frac{2}{\alpha}} \end{aligned} \tag{57}$$

where the last equality used the definition:

$$L_\alpha \triangleq E\left[V^{\frac{2}{\alpha}}\right] E\left[\rho^{\frac{2}{\alpha}}(Y)\right]. \tag{58}$$

The average number of interferers which violate the power threshold condition is:

$$\begin{aligned}\overline{N}(\lambda,\delta) &= \sum_k \Pr\left(W_k \geq \delta\right) \\ &= \sum_k \Pr\left(V_k \rho(Y_k) X_k^{-\alpha} \geq \delta\right) \\ &= \sum_k E\left[\Pr\left(X_k \leq \left(V_k \rho(Y_k)\delta^{-1}\right)^{\frac{1}{\alpha}} | V_k, \rho(Y_k)\right)\right] \\ &\overset{(a)}{=} E\left[\pi\lambda\left(V_k\rho(Y_k)\delta^{-1}\right)^{\frac{2}{\alpha}}\right] \\ &\overset{(b)}{=} \pi\lambda\delta^{-\frac{2}{\alpha}} E\left[V^{\frac{2}{\alpha}}\right] E\left[\rho^{\frac{2}{\alpha}}(Y)\right] \\ &\overset{(c)}{=} \pi\lambda\delta^{-\frac{2}{\alpha}} L_\alpha \end{aligned} \quad (59)$$

where equality (a) used the probability of a node to be inside a circle for a PPP distribution, equality (b) used the independency between r_k, V_k and $\rho(Y_k)$ and equality (c) used (58). For the PPP distribution:

$$\Pr(W_{\max} \leq \delta) = e^{-\overline{N}(\lambda,\delta)} = e^{-\pi\lambda\delta^{-\frac{2}{\alpha}} L_\alpha}. \quad (60)$$

Substituting (60) and (59) in (54) results in the bound:

$$\begin{aligned} R(\lambda) &\geq R_{\text{LB}}(\lambda,\delta) \\ &= \lambda e^{-\pi L_\alpha \lambda \delta^{-\frac{2}{\alpha}}} E\left[\log_2\left(1+\frac{\rho(Y)\cdot Y}{\sigma^2+\frac{2\pi}{\alpha-2}L_\alpha\lambda\delta^{1-\frac{2}{\alpha}}}\right)\right]. \end{aligned} \quad (61)$$

The lower bound in (61) holds for any value of δ. The best bound can be obtained by maximizing (61) with respect to δ. However, this approach turns out to be too complicated. Instead we use $\delta_\star$, given in Lemma 2 below, to produce a simpler lower bound. The value of $\delta_\star$ is motivated by the optimization of the bound with respect to both δ and λ. The resulting bound is $R_{\text{LB}}(\lambda) = R_{\text{LB}}(\lambda,\delta_\star)$ and substituting (63) into (61) completes the proof. □

Lemma 2: The values $\delta_\star$ and $\lambda_\star$ that solve the joint optimization problem:

$$\delta_\star, \lambda_\star = \arg\max_{\delta,\lambda} R_{\text{LB}}(\lambda,\delta) \quad (62)$$

satisfy the relation:

$$\delta_\star = \left(\left(\frac{\alpha}{\alpha-2}\right)\pi L_\alpha \lambda_\star\right)^{\frac{\alpha}{2}} \quad (63)$$

Proof of Lemma 2: We write the optimization problem in two stages by adding an external optimization on the expectation over the interference:

$$R_{\max} = \max_{\delta,\lambda} R_{\text{LB}}(\lambda,\delta) = \max_c \max_{\{\delta,\lambda: E[I]=c\}} R_{\text{LB}}(\lambda,\delta). \quad (64)$$

The proof of the Lemma only needs to consider the internal optimization:

$$\begin{aligned}\delta_\star, \lambda_\star = \arg\max_{\delta,\lambda}\ & \lambda e^{-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}} \\ & \cdot E\left[\log_2\left(1+\frac{\rho(Y)\cdot Y}{\sigma^2+\frac{2\pi}{\alpha-2}L_\alpha\lambda\delta^{1-\frac{2}{\alpha}}}\right)\right] \\ \text{s.t. } & \frac{2\pi}{\alpha-2}L_\alpha\lambda\delta^{1-\frac{2}{\alpha}} = c \end{aligned} \quad (65)$$

which can be simplified to:

$$\delta_\star, \lambda_\star = \arg\max \lambda e^{-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}} \text{ s.t. } \frac{2\pi}{\alpha-2}L_\alpha\lambda\delta^{1-\frac{2}{\alpha}} = c. \quad (66)$$

Define the following Lagrange function to be maximized:

$$\Lambda(\lambda,\delta) \triangleq \lambda e^{-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}} + \upsilon\lambda\delta^{1-\frac{2}{\alpha}} \quad (67)$$

where υ is a Lagrange multiplier. Any local optimum must satisfy:

$$\frac{\partial\Lambda(\lambda,\delta)}{\partial\lambda}=0 \rightarrow e^{-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}}\left(1-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}\right)+\upsilon\delta^{1-\frac{2}{\alpha}}=0 \quad (68)$$

and

$$\begin{aligned}\frac{\partial\Lambda(\lambda,\delta)}{\partial\delta}=0 \rightarrow \lambda\delta^{-1-\frac{2}{\alpha}}\Bigg(&\frac{2\pi L_\alpha\lambda}{\alpha}e^{-\pi L_\alpha\lambda\delta^{-\frac{2}{\alpha}}} \\ &+\delta\left(1-\frac{2}{\alpha}\right)\upsilon\Bigg)=0. \end{aligned} \quad (69)$$

Substituting υ from (68) into (69), leads to (63). □

Appendix B
Proof of Theorem 2

Each receiver cancels its M nearest transmitters, and basically creates a geometrical guard-zone with a random radius around itself. Denoting by r the guard zone radius, its distribution, assuming a density of λ, is given by, [39]:

$$f_r(r) = \frac{2(\pi\lambda)^M}{(M-1)!} r^{2M-1} e^{-\pi\lambda r^2}. \quad (70)$$

We use (55) to write the following characteristic function of the aggregate interference given a guard zone r:

$$\Phi(s) = \exp\left(-\lambda\int_r^\infty E\left[1-e^{-sV\cdot\rho(Y)t^{-\alpha}}\right]2\pi t dt\right) \quad (71)$$

and the expectation of the aggregate interference given r will be:

$$\begin{aligned} E[I|r] &= -\frac{d}{ds}\Phi(s)\Big|_{s=0} \\ &= 2\pi\lambda E\left[V\cdot\rho(Y)\int_r^\infty t^{-\alpha+1}dt\right] \\ &= \frac{2\pi}{\alpha-2}E[V]E[\rho(Y)]\lambda r^{2-\alpha}. \end{aligned} \quad (72)$$

Using (70) leads to:

$$\begin{aligned} E\left[r^{2-\alpha}\right] &= \frac{2(\pi\lambda)^M}{(M-1)!}\int_0^\infty r^{2M+1-\alpha}e^{-\pi\lambda r^2}dr \\ &= (\pi\lambda)^{\frac{\alpha}{2}-1}\frac{\Gamma\left(M+1-\frac{\alpha}{2}\right)}{\Gamma(M)} \end{aligned} \quad (73)$$

where the second equality used $M > \frac{\alpha-2}{2}$. Using the chain rule for expectations and substituting (73) into (72) leads to:

$$\begin{aligned} E[I] &= E_r[E[I|r]] \\ &= \frac{2\pi^{\frac{\alpha}{2}}}{\alpha-2}E[V]E[\rho(Y)]\lambda^{\frac{\alpha}{2}}\frac{\Gamma\left(M+1-\frac{\alpha}{2}\right)}{\Gamma(M)}. \end{aligned} \quad (74)$$

From (74), we see that the expectation on the aggregate interference exists and therefore we can use Jensen's inequality to bound the ERD:

$$R^M(\lambda) = \lambda E\left[\log_2\left(1+\frac{\rho(Y)\cdot Y}{\sigma^2+I}\right)\right] \geq \lambda E\left[\log_2\left(1+\frac{\rho(Y)\cdot Y}{\sigma^2+E[I]}\right)\right]. \quad (75)$$

For the last step we use Kershaw's inequality, [30], [31] for $2<\alpha<4$:

$$\frac{\Gamma(M+1-\frac{\alpha}{2})}{\Gamma(M)} \leq \left(M-\frac{\alpha}{4}\right)^{1-\frac{\alpha}{2}} \quad (76)$$

and for $4\leq\alpha<6$:

$$\frac{\Gamma(M+1-\frac{\alpha}{2})}{\Gamma(M)} \leq \frac{\left(M-\frac{1}{2}-\frac{\alpha}{4}\right)^{2-\frac{\alpha}{2}}}{M-1}. \quad (77)$$

Substituting (76) or (77) into (74) and using (75) completes the proof. □

APPENDIX C
PROOF OF COROLLARY 1

In order to prove Corollary 1 we have to show that Jensen's inequality, used in (75) and Kershaw's inequality, used in (76) and (77), are tight for $M\to\infty$.

For Jensen's inequality, we find the variance of the aggregate interference measured at a probe receiver guarded by a geometrical guard-zone of r:

$$\begin{aligned}\text{Var}(I|r) &= \frac{d^2}{ds^2}\ln(\Phi(s))\Big|_{s=0} \\ &= 2\pi\lambda E\left[V^2\cdot\rho^2(Y)\int_r^\infty t^{-2\alpha+1}dt\right] \\ &= \frac{2\pi}{2\alpha-2}E\left[V^2\right]E\left[\rho^2(Y)\right]\lambda r^{2-2\alpha}. \end{aligned} \quad (78)$$

Using (70) leads to:

$$\begin{aligned}E\left[r^{2-2\alpha}\right] &= \frac{2(\pi\lambda)^M}{(M-1)!}\int_0^\infty r^{2M+1-2\alpha}e^{-\pi\lambda r^2}dr \\ &= (\pi\lambda)^{\alpha-1}\frac{\Gamma(M+1-\alpha)}{\Gamma(M)} \end{aligned} \quad (79)$$

for $M>\alpha-1$. We will also use

$$E\left[r^{4-2\alpha}\right] = (\pi\lambda)^{\alpha-2}\frac{\Gamma(M+2-\alpha)}{\Gamma(M)} \quad (80)$$

for $M>\alpha-2$. Using the low of total variance leads to

$$\begin{aligned}\text{Var}(I) &= E_r[\text{Var}(I|r)] + \text{Var}(E[I|r]) \\ &= \frac{2(\pi\lambda)^\alpha}{2\alpha-2}E\left[V^2\right]E\left[\rho^2(Y)\right]\frac{\Gamma(M+1-\alpha)}{\Gamma(M)} \\ &\quad + \left(\frac{2}{\alpha-2}E[V]E[\rho(Y)]\right)^2(\pi\lambda)^\alpha \\ &\quad \cdot\left(\frac{\Gamma(M+2-\alpha)}{\Gamma(M)} - \left(\frac{\Gamma\left(M+1-\frac{\alpha}{2}\right)}{\Gamma(M)}\right)^2\right) \end{aligned} \quad (81)$$

where the last equality used (72), (74), (80) and the substitution of (79) into (78).

Directly from Stirling's formula, [40], we obtain the following Gamma functions asymptotic ratio:

$$\lim_{M\to\infty}\frac{\Gamma(M+a)}{\Gamma(M+b)}\cdot M^{b-a} = 1. \quad (82)$$

Thus, $E[I]$ scales as $M^{1-\frac{\alpha}{2}}$ and $\text{Var}(I)$ is scales as $M^{1-\alpha}$. Hence,

$$\lim_{M\to\infty}\text{Var}(M^{\frac{\alpha}{2}-1}\cdot I) = 0 \quad (83)$$

and we can state

$$\lim_{M\to\infty}M^{\frac{\alpha}{2}-1}\cdot I = \frac{2\pi^{\frac{\alpha}{2}}}{\alpha-2}E[V]E[\rho(Y)]\lambda^{\frac{\alpha}{2}}. \quad (84)$$

Therefore, $M^{\frac{\alpha}{2}-1}\cdot I$ converges to a deterministic constant, and Jensen's inequality is tight.

The Gamma functions asymptotic ratio, (82), also implies that the two sides of Kershaw's inequality used in (76) and (77) converges to equality for $M\to\infty$. □

REFERENCES

[1] Y. George, I. Bergel, and E. Zehavi, "Novel lower bound on the ergodic rate density of random ad-hoc networks," in *Proceedings of IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, 2013, pp. 575–578.

[2] N. Abramson, "The ALOHA system - another alternative for computer communications," in *AFIPS computer conference*, 1970, pp. 281–285.

[3] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An ALOHA protocol for multihop mobile wireless networks," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 421–436, 2006.

[4] R.T.B. Ma, V. Misra, and D. Rubenstein, "An analysis of generalized slotted-ALOHA protocols," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 936–949, 2009.

[5] S. Toumpis and A.J. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 736–748, 2003.

[6] S. Weber, J.G. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Transactions on Communications*, vol. 58, no. 12, pp. 3593–3604, 2010.

[7] S. Weber, J.G. Andrews, and N. Jindal, "The effect of fading, channel inversion, and threshold scheduling on ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4127–4149, 2007.

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

[9] A. Hunter, J. Andrews, and S. Weber, "Transmission capacity of ad hoc networks with spatial diversity," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 5058–5071, 2008.

[10] E.S. Sousa and J.A. Silvester, "Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 5, pp. 762–771, 1990.

[11] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[12] M. Haenggi, "Outage, local throughput, and capacity of random wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4350–4359, 2009.

[13] K. Stamatiou, J.G. Proakis, and J.R. Zeidler, "Channel diversity in random wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 7, pp. 2280–2289, 2010.

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

[15] D. Mandelbaum, "An adaptive-feedback coding scheme using incremental redundancy," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 388–389, 1974.

[16] R. Comroe and D. Costello Jr, "ARQ schemes for data transmission in mobile radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 472–481, 1984.

[17] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the gaussian collision channel," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1971–1988, 2001.

[18] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA," in *Vehicular Technology Conference (VTC)*. IEEE, 2001, vol. 3, pp. 1829–1833.

[19] H. Holma and Antti. Toskala, *Hsdpa/Hsupa For Umts*, Wiley Online Library, 2006.

[20] F. Wang, A. Ghosh, C. Sankaran, P. Fleming, F. Hsieh, and S. Benes, "Mobile WiMAX systems: performance and evolution," *IEEE Communications Magazine*, vol. 46, no. 10, pp. 41–49, 2008.

[21] S. Sesia, I. Toufik, and M. Baker, "LTE–the UMTS Long Term Evolution," *From Theory to Practice, published in*, vol. 66, 2009.

[22] A. Rajanna and M. Kaveh, "Analysis of cooperative HARQ with opportunistic routing," in *Proceedings of IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, 2013, pp. 76–80.

[23] R.W. Wolff, *Stochastic modeling and the theory of queues*, vol. 14, Prentice hall Englewood Cliffs, NJ, 1989.

[24] D. Stoyan, W.S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, John Wiley & Sons Inc., 1995.

[25] D. Tse and P. Viswanath, "Fundamentals of wireless communications," *lecture notes, University of California, Berkeley, Spring*, 2004.

[26] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, 1999.

[27] E. Malkamaki and H. Leib, "Coded diversity on block-fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 771–781, 1999.

[28] A. Guillen I., Fabregas and G. Caire, "Coded modulation in the block-fading channel: coding theorems and code construction," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 91–114, 2006.

[29] W. Gautschi, "Some elementary inequalities relating to the Gamma and incomplete Gamma function," *J. Math. Phys*, vol. 38, no. 77-81, pp. 127, 1959.

[30] D. Kershaw, "Some extensions of W. Gautschi's inequalities for the Gamma function," *Mathematics of Computation*, vol. 41, no. 164, pp. 607–611, 1983.

[31] F. Qi and L. Losonczi, "Bounds for the ratio of two Gamma functions," *Journal of Inequalities and Applications*, p. 204, 2010.

[32] M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Dover publications, 1965.

[33] I. Bergel and H. Messer, "Optimization of CDMA systems with respect to transmission probability, part I: mutual information rate optimization," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2075–2083, 2008.

[34] I. Bergel and H. Messer, "Optimization of CDMA systems with respect to transmission probability, part II: signal to noise plus interference ratio optimization," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2084–2093, 2008.

[35] J. Venkataraman, M. Haenggi, and O. Collins, "Shot noise models for outage and throughput analyses in wireless ad hoc networks," pp. 1–7, 2006.

[36] SB Lowen and MC Teich, "Power-law shot noise," *IEEE Transactions on Information Theory*, vol. 36, pp. 1302–1318, 1990.

[37] E.N. Gilbert and H.O. Pollak, "Amplitude distribution of shot noise," *Bell Syst. Tech. I*, vol. 39, pp. 333–350, 1960.

[38] Yaniv George and Itsik Bergel, "The spectral efficiency of slotted CSMA ad-hoc networks with directional antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3799 –3809, 2012.

[39] S. Srinivasa and M. Haenggi, "Distance distributions in finite uniformly random networks: Theory and applications," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 2, pp. 940–949, 2010.

[40] W. Feller, "A direct proof of Stirling's formula," *The American Mathematical Monthly*, vol. 74, no. 10, pp. 1223–1225, 1967.