IMPACT OF TRAINING ON THE TRANSMISSION CAPACITY OF WIRELESS AD HOC NETWORKS

Yueping Wu[*], *Raymond H. Y. Louie*[†], *Matthew R. McKay*[*]

[*]Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong

[†]Centre of Excellence in Telecommunications, School of Electrical and Information Engineering, University of Sydney, Australia

ABSTRACT

We investigate the impact of channel estimation (CE) on the performance of point-to-point transmission schemes in wireless ad hoc networks. We first derive a new closed-form expression for the outage probability when the minimum mean square error estimator is utilized for CE, and show that the training-pilot length should increase as a linear function of the network intensity to maintain a fixed CE quality. Next, we derive the optimal training-pilot length which maximizes the transmission capacity under interference-limited conditions, and show that this optimal length increases with the frame length according to a square-root law.

***Index Terms*—** Channel estimation, wireless ad hoc networks, outage probability, transmission capacity

1. INTRODUCTION

The use of channel state information (CSI) has been widely recognized as essential in achieving high performance in wireless systems, through either boosting the effective signal power, or mitigating unwanted interference. CSI is typically obtained through the use of pilot training symbols, which are known to both the transmitter and receiver. In practice, channel estimation (CE) errors occur, and the negative impact of such errors has been widely studied in simple networks, see e.g., [1, 2]. The investigation into more complex networks, such as ad hoc networks, is however quite limited. An exception is the recent work [3] which investigated the effects of CE errors on the medium access control (MAC) layer in ad hoc networks. In contrast, we will focus on studying the impact of CE errors on the physical layer.

We model the spatial distribution of the transmitting nodes as a homogeneous Poisson point process (PPP) on a 2-D plane. Besides approximating realistic network scenarios, modeling the nodes according to a PPP has the benefit of allowing network performance measures, such as the transmission capacity [4], to be obtained. The transmission capacity has been studied for a variety of transmission schemes [5] in ad hoc networks, however typically with the assumption of perfect CSI.

We consider wireless ad hoc networks comprising point-to-point transmissions, with each receiver utilizing the linear minimum mean square error (MMSE) estimator to obtain an estimate of its own communication channel. The estimated channel is then used in the standard way to aid data detection. To analyze the performance of this system, we first derive a new closed-form expression for the outage probability, and then show that the training-pilot length should scale as a linear function of the network intensity to maintain a fixed CE error variance. We then derive the optimal training-pilot length which maximizes the transmission capacity under interference-limited conditions, and this is shown to scale with the frame length according to a square-root law.

2. SYSTEM MODEL

We consider a wireless ad hoc network comprising transmitter-receiver pairs, where each transmitter communicates to its corresponding receiver in a point-to-point manner, treating all other transmissions as interference. The transmitting nodes are distributed spatially according to a homogeneous PPP of intensity $\lambda_{\rm b}$ in $\mathbb{R}^2$. Each transmitting node transmits with probability p according to a slotted ALOHA MAC protocol, and communicates with its corresponding receiving node located at a distance $r_{\rm tr}$. The final intensity of transmitting nodes is thus $\lambda = p\lambda_{\rm b}$.

Our objective is to investigate network performance measures. To obtain such measures, we invoke Slivnyak's theorem of a PPP [6] which states that it is sufficient to focus on a typical transmitter-receiver pair, denoted by index 0, with the typical receiver located at the origin. We consider a network where each node is equipped with a single antenna, and the transmitting nodes, with the exception of the typical transmitter, constitute a marked PPP [6]. This is denoted by $\Phi(\lambda) = \{(D_\ell, h_{\ell 0}), \ell \in \mathbb{N}\}$, where D_ℓ and[1] $h_{\ell 0} \overset{d}{\sim} \mathcal{CN}(0,1)$ model the location and channel respectively of the ℓth transmitting node with respect to (w.r.t.) the typical receiver. The transmitted signals are attenuated by a factor $\frac{1}{\delta + r^\alpha}$ with distance r where $\delta > 0$ and $\alpha > 2$ is the path loss exponent.

Each receiver obtains an estimate of the CSI via pilot training symbols sent from their corresponding transmitter. The transmitter-receiver channels are constant over a frame comprising L channel uses, and evolves independently from frame to frame. Each transmitter sends a frame to the corresponding receiver, which comprises training pilots of length L_T channel uses, as well as the data symbols. The training pilots are inserted at the beginning of each frame[2], and are utilized by the receiver to obtain an estimate of the channel. In the usual way, this channel estimate is then used to detect the data symbols from the corresponding transmitter, and this procedure is repeated over all subsequent frames. We now describe the CE and data detection procedure in more detail.

[1]The notation $X \overset{d}{\sim} Y$ means that X *is distributed as* Y.

[2]This is commonly used in some modern communication standards, such as the North America TDMA standard.

2.1. Channel Estimation

The $1 \times L_T$ baseband equivalent received vector $\mathbf{y}_0$ at the typical receiver, formed by concatenating the received symbols during the first L_T channel uses, is given by

$$\mathbf{y}_0 = \sqrt{\frac{L_T P}{\delta + r_{\mathrm{tr}}^{\alpha}}} h_{00}\mathbf{t}_{00} + \sum_{\ell \in \Phi(\lambda)} \sqrt{\frac{P}{\delta + |D_\ell|^{\alpha}}} h_{\ell 0}\mathbf{x}_\ell + \mathbf{n}_0 \quad (1)$$

where P is the transmission power of each symbol, $\mathbf{t}_{00}$ is a $1 \times L_T$ training symbol vector satisfying $\mathbf{t}_{00}\mathbf{t}_{00}^{\dagger} = 1$ [7], $h_{00} \overset{d}{\sim} \mathcal{CN}(0,1)$, $\mathbf{x}_\ell \overset{d}{\sim} \mathcal{CN}_{1\times L_T}(0, \mathbf{I}_{L_T})$ is a transmission vector[3] from node ℓ, and $\mathbf{n}_0 \overset{d}{\sim} \mathcal{CN}_{1\times L_T}(0, N_0\mathbf{I}_{L_T})$ is the additive white Gaussian noise (AWGN) vector.

To obtain an estimate of the channel, we utilize the low-complexity linear MMSE estimator, which is optimal among the class of linear estimators. According to the MMSE estimator [8], the best estimate of h_{00} is

$$\widehat{h}_{00} = \frac{\sqrt{\frac{L_T P}{\delta + r_{\mathrm{tr}}^{\alpha}}}\mathbf{t}_{00}\mathbf{y}_0^{\dagger}}{L_T P \left(\delta + r_{\mathrm{tr}}^{\alpha}\right)^{-1} + P\mathrm{Var}(I) + N_0} \quad (2)$$

where $I = \sum_{\ell \in \Phi(\lambda)} (\delta + |D_\ell|^{\alpha})^{-\frac{1}{2}} h_{\ell 0}\mathbf{t}_{00}\mathbf{x}_\ell^{\dagger}$ and $\mathrm{Var}(\cdot)$ denotes the variance function. The CE error can thus be expressed as $e_{00} = h_{00} - \widehat{h}_{00}$. For the MMSE estimator, it can be shown that (i) $\widehat{h}_{00}$ and e_{00} are uncorrelated, and (ii) if $e_{00} \overset{d}{\sim} \mathcal{CN}\left(0, \sigma_e^2\right)$, then $\widehat{h}_{00} \overset{d}{\sim} \mathcal{CN}\left(0, 1-\sigma_e^2\right)$.

2.2. Data Transmission

After the CE stage, the transmitter then sends its data for the rest of the frame duration. The received signal at the typical receiver during the nth channel use, for $n = L_T + 1, \ldots, L$, is given by

$$r_0[n] = \sqrt{\frac{1}{\delta + r_{\mathrm{tr}}^{\alpha}}}\widehat{h}_{00}s_0[n] \quad (3)$$
$$+ \underbrace{\sqrt{\frac{1}{\delta + r_{\mathrm{tr}}^{\alpha}}} e_{00}s_0[n] + \sum_{\ell \in \Phi(\lambda)} \sqrt{\frac{1}{\delta + |D_\ell|^{\alpha}}} h_{\ell 0}s_\ell[n] + n_0[n]}_{\text{unknown at the receiver}}$$

where $s_\ell[n]$ ($\ell \in \{0, \mathbb{N}\}$) are the independent Gaussian distributed data symbols from the ℓth transmitting node satisfying $\mathrm{E}\left[|s_\ell[n]|^2\right] = P$, and $n_0[n] \overset{d}{\sim} \mathcal{CN}(0, N_0)$ is AWGN. As the terms in (3) which are unknown at the receiver are treated as noise, an estimate of $s_0[n]$ is then formed as $\widehat{s}_0[n] = \sqrt{\delta + r_{\mathrm{tr}}^{\alpha}} \frac{\widehat{h}_{00}^*}{|\widehat{h}_{00}|^2} r_0[n]$, from which the signal-to-interference-plus-noise ratio (SINR) can be written as

$$\mathrm{SINR} = \frac{\frac{\rho}{\delta + r_{\mathrm{tr}}^{\alpha}}|\widehat{h}_{00}|^2}{\sum_{\ell \in \Phi(\lambda)} \frac{\rho}{\delta + |D_\ell|^{\alpha}}|h_{\ell 0}|^2 + \frac{\rho}{\delta + r_{\mathrm{tr}}^{\alpha}}\sigma_e^2 + 1} \quad (4)$$

where $\rho = \frac{P}{N_0}$ is the transmit signal-to-noise ratio (SNR).

[3] The Gaussian assumption for the interfering symbols is well justified. As we will show, the optimal pilot-training length is typically small compared to the frame length, thus in the optimal scenario, the majority of the frame is used for data transmission, during which the data symbols transmitted from all nodes are Gaussian distributed.

3. PERFORMANCE ANALYSIS: OUTAGE PROBABILITY AND TRANSMISSION CAPACITY

In this section, we will investigate the impact of CE on the outage probability and transmission capacity.

3.1. Outage Probability

The outage probability is defined as the probability that the mutual information falls below a data rate operating value R_{eff} data bits/node/channel use, and is given by

$$\begin{aligned} \mathrm{F}(\beta) :&= \Pr\left(\left(1 - \frac{L_T}{L}\right)\log_2(1 + \mathrm{SINR}) \le R_{\mathrm{eff}}\right) \\ &= \Pr\left(\left(1 - \frac{L_T}{L}\right)\log_2(1 + \mathrm{SINR}) \le \left(1 - \frac{L_T}{L}\right)R\right) \\ &= \Pr(\mathrm{SINR} \le \beta) \end{aligned} \quad (5)$$

where R is the code rate over the data transmission stage, $\beta = 2^R - 1$ is the SINR operating value, and the second line follows by noting that $R_{\mathrm{eff}} = \left(1 - \frac{L_T}{L}\right)R$. Based on the SINR expression in (4), we present the following lemma:

Lemma 1. *The outage probability in ad hoc networks where the MMSE estimator is utilized for CE is given by*

$$\mathrm{F}(\beta, L_T, \lambda) = 1 - \mathrm{F}_{\mathrm{suc}}^{\mathrm{SU}}(\beta, L_T, \lambda)\,\mathrm{F}_{\mathrm{suc}}^{\mathrm{I}}(\beta, L_T, \lambda) \quad (6)$$

where $\mathrm{F}_{\mathrm{suc}}^{\mathrm{SU}}(\beta, L_T, \lambda) = \exp\left(\left(1 + \frac{\delta + r_{\mathrm{tr}}^{\alpha}}{\rho}\right)\frac{\beta}{\sigma_e^2 - 1} + \beta\right)$ *and* $\mathrm{F}_{\mathrm{suc}}^{\mathrm{I}}(\beta, L_T, \lambda) = \exp\left(\frac{-\pi\lambda(\delta + r_{\mathrm{tr}}^{\alpha})\beta\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)}{(1-\sigma_e^2)^{\frac{2}{\alpha}}((1-\sigma_e^2)\delta + (\delta + r_{\mathrm{tr}}^{\alpha})\beta)^{1-\frac{2}{\alpha}}}\right)$. *In these expressions, the CE variance, σ_e^2, dependent on both λ and ρ, is given by*

$$\sigma_e^2 = \frac{1}{\frac{L_T}{(\delta + r_{\mathrm{tr}}^{\alpha})\left(\pi\lambda\delta^{\frac{2}{\alpha}-1}\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right) + \frac{1}{\rho}\right)} + 1}. \quad (7)$$

Proof. See the Appendix. □

Note that in *Lemma 1*, $\mathrm{F}_{\mathrm{suc}}^{\mathrm{SU}}(\beta, L_T, \lambda)$ can be interpreted as the outage probability of a system which, during the data transmission stage, operates in the absence of interference (i.e., single user), whilst $\mathrm{F}_{\mathrm{suc}}^{\mathrm{I}}(\beta, L_T, \lambda)$ can be interpreted as the outage probability of a system which, during the data transmission stage, operates in the absence of receiver noise (i.e., interference limited). Moreover, we note that when $\sigma_e^2 = 0$, (6) reduces to the outage probability with perfect CSI [9, Eq. (3.3)]; while when $\lambda = 0$, (6) reduces to the outage probability of a point-to-point single-user system [10, Eq. (5.54)].

In addition, we observe in *Lemma 1* that both $\mathrm{F}_{\mathrm{suc}}^{\mathrm{SU}}(\beta, L_T, \lambda)$ and $\mathrm{F}_{\mathrm{suc}}^{\mathrm{I}}(\beta, L_T, \lambda)$ decrease with σ_e^2, implying that the outage probability increases with the CE error, as expected. We can also observe that, with CE error, increasing the network intensity λ has a *dual* negative effect on the outage probability. The first effect occurs during the CE stage, where increasing λ, for a fixed training-pilot length L_T, increases the CE error. This subsequently degrades both $\mathrm{F}_{\mathrm{suc}}^{\mathrm{SU}}(\beta, L_T, \lambda)$ and $\mathrm{F}_{\mathrm{suc}}^{\mathrm{I}}(\beta, L_T, \lambda)$. Second, during the data transmission stage, increasing λ increases the multi-node interference, which subsequently degrades

$\mathrm{F}^{\mathrm{I}}_{\mathrm{suc}}(\beta, L_T, \lambda)$. Finally, we observe in (7) that the training-pilot length should increase as a linear function of the network intensity λ, to maintain a fixed error variance.

3.2. Transmission Capacity

The transmission capacity is defined as the maximum number of successfully transmitted data bits/channel use/unit area [4], and is given by

$$\mathrm{c}(\epsilon, L_T) = \left(1 - \frac{L_T}{L}\right) R\lambda(\epsilon, L_T)(1-\epsilon) \tag{8}$$

where $\lambda(\epsilon, L_T)$ is the *contention density*, defined as the inverse of $\epsilon = \mathrm{F}(\beta, L_T, \lambda)$ taken w.r.t. λ.

From (8), we can observe that increasing L_T has both a positive and a negative effect on the transmission capacity. The positive effect occurs since, as can be easily shown, the contention density $\lambda(\epsilon, L_T)$ increases with L_T. The negative effect occurs since for a fixed frame length L, the time spent for data transmission decreases with L_T. A natural question then arises as to the optimal training-pilot length L_T^* which maximizes the transmission capacity. Although obtaining an exact expression for L_T^* is difficult, progress can be made by considering asymptotic regimes. It is thus first convenient to present the following lemma for the transmission capacity in the interference-limited regime ($\rho \to \infty$) and for small outage constraints ϵ:

Lemma 2. *In the interference-limited scenario ($\rho \to \infty$) and for a sufficiently small outage constraint ϵ, the transmission capacity is given by*

$$c(\epsilon, L_T) = g(L_T)\, c_{\text{perfect}}(\epsilon) + o(\epsilon) \tag{9}$$

where $g(L_T) = \left(1 - \frac{L_T}{L}\right) \frac{L_T}{L_T + \left(1 + \left(1 + \frac{r_{\text{tr}}^\alpha}{\delta}\right)\beta\right)^{1-\frac{2}{\alpha}}}$ *is the scaling factor reflecting the effects of CE, and*

$$c_{\text{perfect}}(\epsilon) = \frac{\left(\delta + (\delta + r_{\text{tr}}^\alpha)\beta\right)^{1-\frac{2}{\alpha}} \log_2(1+\beta)\,\epsilon}{\pi\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)\left(\delta + r_{\text{tr}}^\alpha\right)\beta} \tag{10}$$

denotes the transmission capacity in an ideal scenario where no training pilot is required and the receiver has perfect CSI.

Proof. Omitted due to space limitations. □

Using *Lemma 2*, by taking derivatives, we can easily calculate the optimal training-pilot length which maximizes $c(\epsilon, L_T)$ in (9), with the result given in the following theorem:

Theorem 1. *The optimal training-pilot length which maximizes the transmission capacity in (9) is*[4]

$$L_T^* = \begin{cases} \left\lceil \widetilde{L_T^*} \right\rceil & \text{if} \quad c\left(\epsilon, \left\lceil \widetilde{L_T^*} \right\rceil\right) \le c\left(\epsilon, \left\lfloor \widetilde{L_T^*} \right\rfloor\right) \\ \left\lfloor \widetilde{L_T^*} \right\rfloor & \text{otherwise} \end{cases} \tag{11}$$

where[5]
$\widetilde{L_T^*} = \left(\sqrt{1 + \frac{\delta^{1-\frac{2}{\alpha}} L}{(\delta + (\delta + r_{\text{tr}}^\alpha)\beta)^{1-\frac{2}{\alpha}}}} - 1\right)\left(1 + \frac{(\delta + r_{\text{tr}}^\alpha)\beta}{\delta}\right)^{1-\frac{2}{\alpha}}.$

[4]Here we implicitly assume that ϵ is sufficiently small such that the $o(\epsilon)$ term in (9) can be ignored.

[5]Note that $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ are the ceiling and floor functions respectively.

From (11), we see that L_T^* increases with the frame length L, and in particular for large L, L_T^* scales as $O\left(\sqrt{L}\right)$. This implies that for large L, the fraction of the total frame length for CE, $\frac{L_T^*}{L}$, scales as $O\left(\frac{1}{\sqrt{L}}\right)$, which can be quite small. The practical interpretation is that for large L, it is preferable to dedicate a larger proportion of the frame for data transmission. A similar result was obtained in [11] for a time-division duplexing multiuser MIMO downlink scenario, where the optimal training-pilot length which maximizes the spectral efficiency was also shown to scale according to a square-root law with the frame length.

Fig. 1 plots L_T^* vs. the frame length L. We observe that the 'Analytical' curves based on (11) closely match the numerical results, and L_T^* increases sub-linearly w.r.t. L, as predicted. Moreover, we observe that it is optimal to use only a small fraction of the frame length for CE for practical networking parameters. For example, consider a typical system with a coherence bandwidth of $W_c = 500$ kHz and a coherence time of $T_c = 2.5$ ms [10]. These parameters correspond to $L = 1250$ channel uses, which from Fig. 1 corresponds to $L_T^* = 110$ when $\alpha = 3$, and thus only 8.8% of the frame is used for CE.

We will now compare the transmission capacity when using the optimal training-pilot length with an ideal (impractical) scenario where perfect CSI is obtained without the need for any training. For large L, it can be shown that

$$\begin{aligned} &\frac{c\left(\epsilon, \widetilde{L_T^*}\right)}{c_{\text{perfect}}(\epsilon)} = g\left(\widetilde{L_T^*}\right) \\ &= 1 - 2\sqrt{\left(1 + \left(1 + \frac{r_{\text{tr}}^\alpha}{\delta}\right)\beta\right)^{1-\frac{2}{\alpha}}}\sqrt{\frac{1}{L}} + o\left(\sqrt{\frac{1}{L}}\right). \end{aligned} \tag{12}$$

The key insight drawn from (12) is that for channels with a sufficiently long coherence time (i.e., the channel remains constant during the transmission of a packet of length L, whose value may be large), there is a negligible performance loss resulting from channel estimation compared to the perfect CSI scenario. This is highlighted in Fig. 2, which plots the transmission capacity ratio $\frac{c(\epsilon, L_T^*)}{c_{\text{perfect}}(\epsilon)}$ vs. frame length L for different path loss exponents α. We observe that the transmission capacity with imperfect CSI achieves a high percentage of the transmission capacity with perfect CSI for even small frame lengths.

We also observe from (12) that $\frac{c(\epsilon, L_T^*)}{c_{\text{perfect}}(\epsilon)}$ decreases with the path loss exponent when $r_{\text{tr}} \ge 1$, implying that the perfect CSI scenario is harder to approach at high path loss environments. This can also be observed in Fig. 2, where we see as predicted that the ratio between the imperfect and perfect CSI scenarios decrease with the path loss exponent.

4. APPENDIX

Applying Campbell's Theorem [6], it can be shown that $\mathrm{Var}(I) = \pi\lambda\delta^{\frac{2}{\alpha}-1}\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)$. Substituting $\mathrm{Var}(I)$ into (5), and after some algebraic manipulations, we obtain (7). We then note that $\widehat{h}_{00}$ is statistically equivalent to $\sqrt{1-\sigma_e^2}\widetilde{h}_{00}$, where $\widetilde{h}_{00} \sim \mathcal{CN}(0,1)$. By noting that $\|\widetilde{h}_{00}\|^2$ is exponentially distributed, the outage probability conditioned on the interference

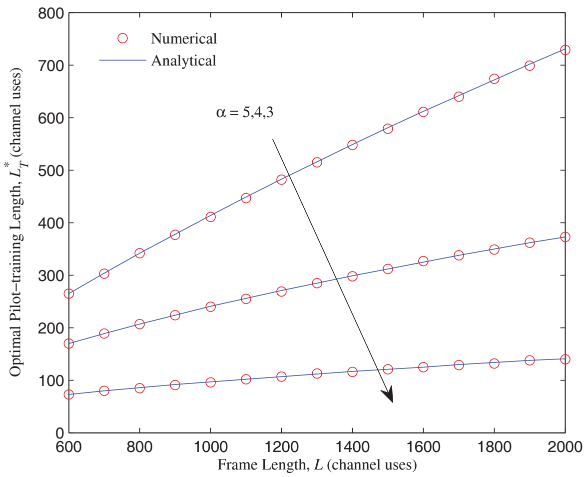

Fig. 1. Optimal training-pilot length L_T^* vs. the frame length L in the interference-limited scenario for different path loss exponent α, and with $r_{\text{tr}} = 8$ m, $\delta = 1$, $\beta = 3$, and $\epsilon = 0.01$.

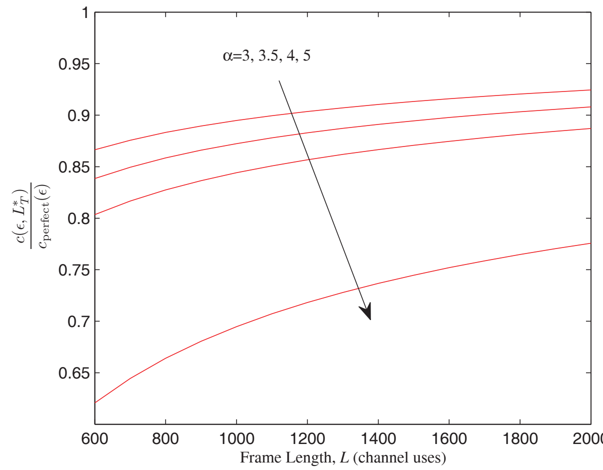

Fig. 2. Transmission capacity ratio vs. frame length L for different path loss exponents α, and with $\delta = 1$, $r_{\text{tr}} = 2.3$ m, $\epsilon = 0.2$, $\rho = 25$ dB, and $\beta = 1$.

$I_a = \sum_{\ell \in \Phi(\lambda)} (\delta + |D_\ell|^\alpha)^{-1} \|h_{\ell 0}\|^2$ is

$$\mathrm{F}(\beta|I_a) = 1 - \exp\left(-\frac{(\delta + r_{\text{tr}}^\alpha)\,\beta}{\rho\left(1-\sigma_e^2\right)}(\rho I_a + W)\right) \quad (13)$$

where $W = \frac{\rho\sigma_e^2}{\delta + r_{\text{tr}}^\alpha} + 1$. Averaging out I_a, it can be easily shown that

$$\mathrm{F}(\beta) = 1 - \exp\left(-\frac{(\delta + r_{\text{tr}}^\alpha)\,\beta}{\rho\left(1-\sigma_e^2\right)}W\right)\mathcal{L}_{I_a}(s)\Big|_{s=\frac{(\delta + r_{\text{tr}}^\alpha)\beta}{1-\sigma_e^2}} \quad (14)$$

where $\mathcal{L}_{I_a}(s)$ denotes the Laplace transform of I_a, given by [12]

$$\mathcal{L}_{I_a}(s) = \exp\left(\frac{-\pi\lambda s}{(\delta+s)^{1-\frac{2}{\alpha}}}\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)\right). \quad (15)$$

The result follows by substituting (15) into (14).

5. REFERENCES

[1] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[2] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[3] D. Chiarotto, P. Casari, and M. Zorzi, "On the impact of channel estimation errors on MAC protocols for MIMO ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3290–3300, Oct. 2010.

[4] S. P. Weber, X. Yang, J. G. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4091–4102, Dec. 2005.

[5] S. P. Weber, J. G. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3593–3604, Dec. 2010.

[6] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, John Wiley and Sons, England, 2nd edition, 1995.

[7] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.

[8] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 2nd edition, 1981.

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

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, New York, 1st edition, 2005.

[11] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[12] M. Haenggi and R. K. Ganti, "Interference in large wireless networks," *Foundations and Trends in Networking*, vol. 3, no. 2, pp. 127–248, 2009.