

# A REALISTIC PERFORMANCE ANALYSIS FOR PRACTICAL CHANNEL-AWARE SCHEDULING

*Pengcheng Zhan, Ramesh Annavajjala, A. Lee Swindlehurst, Todd Chauvin*

## ABSTRACT

It is well-known that opportunistic transmission schemes are sum-capacity optimal, in the Shannon sense, for symmetric cellular networks with single-antenna transceivers. However, for a practical system operating at non-negligible error probability, the correctly delivered throughput (or spectral efficiency) is a more useful figure-of-merit. In this paper, we present a mathematical model for the average throughput of a practical wireless system with channel-aware user scheduling in a symmetric Rayleigh fading cellular system. By employing an accurate model for the block error probability with turbo coding, and accounting for channel estimation errors, channel feedback quantization, and feedback delay, we present closed-form expressions for the average spectral efficiency of the downlink scheduler.

**Index Terms**— Channel-aware scheduler, cross-layer design, multi-user diversity, imperfect channel estimation.

## 1. INTRODUCTION

Optimal utilization of the wireless spectrum requires efficient and intelligent radio resource allocation. Recently, the idea of opportunistically allocating transmission resources has gained wide-spread popularity in the context of a multi-user cellular system. The results of [1, 2] show that scheduling the best user, in terms of instantaneous received signal-to-noise ratio (SNR), achieves the Shannon sum-capacity for both the uplink and downlink of a symmetric (equal average received SNR for all users) cellular network with single-antenna transceivers. In addition to employing channel codes with long block lengths, in order to achieve capacity, the channel state information (CSI) has to be perfect at both the user terminals (UTs) and the base station (BS).

In this paper, we consider the downlink performance of a practical wireless system comprised of a single BS and multiple UTs that schedules transmission resources over a given interval to the one UT that has the best channel conditions [3]. Our system model is different from the existing approaches [1, 2, 3] in the following aspects:

- The system employs a finite number of modulation and coding schemes (MCS).
- Short block length channel codewords are considered with non-negligible block error rate (BLER).
- A limited number of pilots are assumed for channel estimation.

---

Pengcheng Zhan is with the Dept. of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602. e-mail: ezhan@byu.edu

Ramesh Annavajjala and Todd Chauvin are with ArrayComm LLC, San Jose, CA. e-mail: ramesh.annavajjala@gmail.com, chauvin@arraycomm.com

A. Lee Swindlehurst is with the Dept. of Electrical Engineering and Computer Science, University of California, Irvine, Irvine, CA 92697. e-mail: swindle@uci.edu

- The effect of UT mobility is explicitly modeled.
- A limited number of feedback bits are assumed for reporting channel quality.

While Shannon capacity is impractical to attain, by adjusting the number of bits per symbol and the code rate for various channel conditions, Shannon capacity can be closely approached in the vicinity of operating SNRs. To close the gap between Shannon capacity and the achievable system throughput, a number of different MCS schemes are usually implemented at the Forward Error Correction (FEC) block in a typical wireless communication system. For example, in the 802.16e standard [4], a number of MCS, ranging from QPSK rate 1/2 to 64QAM rate 5/6, are supported in the channel coding module. Depending on the data payload and channel allocation, the block length of the codewords can be relatively short, e.g. it can be as few as 96 QAM symbols (2 slots, 48 bytes) for 64QAM rate  $\frac{2}{3}$  in 802.16e [4]. Limited block lengths further hinder the achievement of Shannon capacity and its effect should be addressed when the system throughput is quantified. It is also reasonable to assume that practical communication systems operate at a relatively high BLER, e.g. 1% or 10% in a WiMAX system. The aforementioned aspects suggest that a better figure of merit when evaluating system performance would be the correctly delivered throughput, or spectral efficiency. For a number of purposes including message decoding, user selection, etc., CSI has to be obtained with a certain degree of accuracy at both the BS and UTs. The most common way to obtain channel estimates is through embedded pilot symbols, but their use also creates overhead in the system. When a dedicated control channel is available to feedback the UT's channel or SNR estimate to the BS, the finite precision of the quantizer and scarce feedback channel bandwidth limit the number of bits that are available for channel quality reporting. Even without inaccuracies due to estimation error and feedback quantization, when the network consists of high mobility users, the channel can be largely de-correlated by the time the BS receives the CSI report.

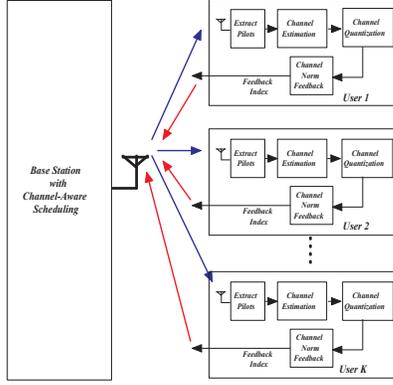
This paper is concerned with the analysis of the spectral efficiency of a realistic communication system. The practical issues discussed in the above paragraph including channel estimation, user mobility, finite feedback precision, etc., will be discussed. The remainder of the paper is organized as follows. Section 2 introduces the system model, and presents a piecewise log-linear model for the block error rate as a function of the instantaneous received SNR. In Section 3.1, we analyze the throughput performance of a symmetric network with perfect CSI at the scheduler side. Section 3.2 addresses different causes of imperfect CSI and their impacts on the system throughput. Results and discussions are given in Section 4 and conclusions in Section 5.

## 2. SYSTEM MODEL AND ASSUMPTIONS

The opportunistic scheduler under investigation is depicted in Fig. 1. We focus on the downlink scheduling problem, where based on CSI

No. Slots	QPSK $\frac{1}{2}$	QPSK $\frac{3}{4}$	16QAM $\frac{1}{2}$	16QAM $\frac{3}{4}$	64QAM $\frac{1}{2}$	64QAM $\frac{2}{3}$	64QAM $\frac{3}{4}$
1	(1.12, 4.02)	(2.12, 1.77)	(3.99, 1.53)	(9.49, 0.59)	(10.64, 0.42)	(22.49, 0.21)	(14.23, 0.06)
2	(1.12, 5.71)	(2.38, 2.76)	(3.99, 1.96)	(7.54, 0.46)	(11.94, 0.68)	(25.24, 0.35)	(31.86, 0.28)

**Table 1.** Parameters for the log-linear BLER model.



**Fig. 1.** Block diagram of a practical downlink scheduler.

for all the users, the Base Station (BS) selects the optimal user and allocates the RF resource to it for a number of time slots. A single antenna is assumed for both the BS and UTs. As illustrated in Fig. 1, at the user side, channel estimation is performed using the embedded pilot symbols. A quantized instantaneous SNR value per user is then fed back to the BS for scheduling purposes. As discussed in Section 1, in this system setup, the finite pilot power, limited number of pilot symbols, finite feedback word length, Doppler spread, and feedback delay will all have the effect of limiting the accuracy of the channel SNR information available to the scheduler. Thus, the scheduling decision based on this information will likely cause a certain loss in system spectral efficiency. Hereafter, we use  $K$  to denote the number of users in the downlink, the random variable  $\Gamma_j$  to represent the  $j$ -th user's instantaneous SNR. An iid Rayleigh fading scenario is assumed throughout the paper. In other words, we assume a symmetric network, where each user has the same average SNR.

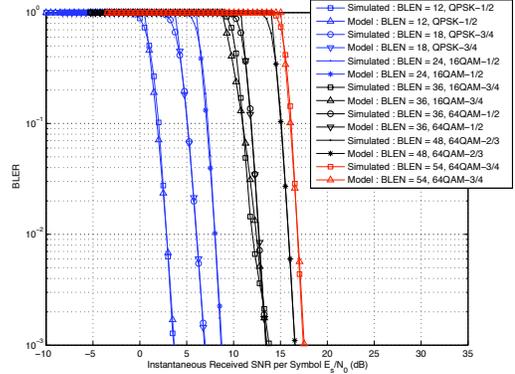
The applications of FEC encoders and Automatic Repeat Request (ARQ) techniques in communication systems make analytical BLER evaluation difficult. Extensive simulations performed at ArrayComm for each MCS class in an AWGN channel reveal the feasibility of modeling the system's BLER behavior with a log-linear model. In the linear domain, for the  $j$ -th MCS, the BLER can be written as:

$$\text{PER}_j = \begin{cases} 1 & \gamma < \gamma_{T,j} \\ e^{-b_j \cdot (\gamma - \gamma_{T,j})} & \gamma \geq \gamma_{T,j} \end{cases} \quad (1)$$

for some choice of the model parameters  $\gamma_{T,j}$  and  $b_j$  for each MCS. These parameters can be determined by solving the following optimization problem for each MCS scheme:

$$\{\gamma_{T,j}, b_j\} = \arg \max_{\gamma_{T,j}, b_j} \sum_{k=1}^M \left\{ (\gamma_{k,j} - \gamma_{T,j})^2 \cdot \mathbf{1}(\gamma_{k,j} \geq \gamma_{T,j}) + (1 - \ln P_{k,j})^2 \cdot \mathbf{1}(\gamma_{k,j} < \gamma_{T,j}) \right\}, \quad (2)$$

where  $(\gamma_{k,j}, P_{k,j})$  is the  $k$ -th pair of SNR and BLER measurements for the  $j$ -th MCS collected by simulating the AWGN channel (coded with or without ARQ depending on the specific system setup). The obtained parameters  $\gamma_{T,j}$  and  $b_j$  for the Convolutional Turbo Code



**Fig. 2.** Comparison of the piecewise log-linear BLER model against simulations.

(CTC) encoder in a WiMAX system are listed in Table 1. Curve fit result given in Fig. 2 shows that the model in (1) very accurately characterizes the system BLER behavior. The term BLEN in the legend of Fig. 2 denotes the codeword length of the FEC block in bytes.

As discussed previously, adaptively switching between various MCS schemes based on instantaneous channel SNR will improve the system spectral efficiency. The thresholds that control the switching can be chosen to guarantee that the system's BLER operating point is strictly lower than a pre-determined value  $P_T$ . Using the model in (1), the threshold for the  $j$ -th modulation scheme can be chosen as:  $\gamma_j = -\frac{\ln(P_T)}{b_j} + \gamma_{T,j}$ . When the instantaneous maximum SNR of the system is determined and found to be within the range of  $[\gamma_j, \gamma_{j+1}]$ , the corresponding user will be assigned the entire channel bandwidth and the  $j$ -th MCS scheme. For simplicity of notation, use  $\gamma_0 \equiv 0$  and  $\gamma_{N+1} \equiv \infty$ .

In general, the spectral efficiency of interest can be written as:

$$\bar{T} = \mathcal{E} \left( C_j (1 - \text{PER}(j)) \cdot \mathbf{1}(\gamma_j \leq \gamma_{\max} < \gamma_{j+1}) \right), \quad (3)$$

where  $C_j = \log_2 M_j$ ,  $M_j$  is the number of constellation points for the  $j$ -th MCS scheme,  $\text{PER}(j)$  is the Packet Error Rate model (used interchangeably with BLER) for the  $j$ -th MCS class, and  $\mathbf{1}(\bullet)$  is an indicator function, which equals 1 when its parameter is true, or 0 otherwise. As can be observed from (3),  $\bar{T}$  depends on the MCS switching thresholds (corresponding to BLER operating point  $P_T$ ), BLER model parameters ( $b_j$  and  $\gamma_{T,j}$ ), the accuracy of the CSI (how  $\gamma_{\max}$  is determined), and so on.

### 3. SPECTRAL EFFICIENCY ANALYSIS

#### 3.1. Perfect Channel State Information

When perfect CSI can be assumed, the BS makes scheduling decisions based on a collection of all the users' instantaneous SNRs. In this case, the system spectral efficiency is completely determined by the distribution of the maximum SNR among the users, i.e.  $\gamma_{\max} =$

$$\begin{aligned}
F(a, b, c) &\equiv \int_0^a e^{-x} Q_1(\sqrt{bx}, c) dx \ \& \ G(a, b, c) \equiv \int_a^\infty e^{-x} \cdot Q_1(\sqrt{bx}, c) \cdot dx \\
P(j) &\equiv \int_{\gamma_j}^{\gamma_{j+1}} f_{\hat{\gamma}_{\max}}(\gamma_h) \cdot \int_{\gamma_{T,j}}^\infty e^{-b_j \cdot (\gamma_g - \gamma_{T,j})} \cdot f_{\Gamma_g|h}(\gamma_g) \cdot d\gamma_g \cdot d\gamma_h = \sum_{l=0}^K (-1)^{l+1} \binom{K}{l} \cdot \\
&\left\{ \left[ F\left(\frac{l \cdot \gamma_{j+1}}{\bar{\gamma}_h}, \frac{p_{1,j}}{l}, q_{1,j}\right) - F\left(\frac{l \cdot \gamma_j}{\bar{\gamma}_h}, \frac{p_{1,j}}{l}, q_{1,j}\right) \right] - t_j \left[ F\left(\frac{l \cdot \gamma_{j+1}}{\bar{\gamma}_h}, \frac{p_{2,j}}{l}, q_{2,j}\right) - F\left(\frac{l \cdot \gamma_j}{\bar{\gamma}_h}, \frac{p_{2,j}}{l}, q_{2,j}\right) \right] \right\}, \ 1 \leq j < N \\
P(N) &= \sum_{l=0}^K (-1)^{l+1} \binom{K}{l} \cdot \left[ G\left(\frac{l\gamma_N}{\bar{\gamma}_h}, \frac{p_{1,N}}{l}, q_{1,N}\right) - t_N \cdot G\left(\frac{l\gamma_N}{\bar{\gamma}_h}, \frac{p_{2,N}}{l}, q_{2,N}\right) \right] \tag{4} \\
p_{1,j} &= \frac{2|\rho_{eq}|^2}{(1-|\rho_{eq}|^2)}, \ q_{1,j} = \sqrt{\frac{2\gamma_{T,j}}{(1-|\rho_{eq}|^2)\bar{\gamma}_g}}, \ t_j = e^{b_j\gamma_{T,j}} \cdot \frac{e^{-\frac{|\rho_{eq}|^2 \cdot |h|^2}{(1-|\rho_{eq}|^2)\Omega_h} \cdot \frac{b_j(1-|\rho_{eq}|^2)\bar{\gamma}_g}{1+b_j(1-|\rho_{eq}|^2)\bar{\gamma}_g}}}{1+b_j(1-|\rho_{eq}|^2)\bar{\gamma}_g} \\
p_{2,j} &= \frac{2|\rho_{eq}|^2|h|^2}{(1-|\rho_{eq}|^2)\Omega_h(1+b_j(1-|\rho_{eq}|^2)\bar{\gamma}_g)}, \ q_{2,j} = \sqrt{\frac{2\gamma_{T,j}(1+b_j(1-|\rho_{eq}|^2)\bar{\gamma}_g)}{(1-|\rho_{eq}|^2)\bar{\gamma}_g}}, \ \rho_{eq} = \rho_d \cdot \rho_e \tag{5}
\end{aligned}$$

$\max(\Gamma_1, \Gamma_2, \dots, \Gamma_K)$ . As is well known, the Cumulative Distribution Function (CDF) of  $\gamma_{\max}$  is given by [5]:

$$F_{\gamma_{\max}}(\mu) = (1 - e^{-\frac{\mu}{\bar{\gamma}}})^K,$$

where  $\bar{\gamma}$  is the average SNR for each user, i.e., the system operating SNR point. By taking the derivative and carrying out the expansion, the Probability Density Function (PDF) of  $\gamma_{\max}$  is given as:

$$f_{\gamma_{\max}}(\mu) = \sum_{l=0}^K (-1)^{l+1} \cdot \binom{K}{l} \cdot \frac{l}{\bar{\gamma}} \cdot e^{-\frac{l\mu}{\bar{\gamma}}}. \tag{6}$$

After some complex mathematical manipulations, the supported spectral efficiency can be expressed as [6]:

$$\begin{aligned}
\bar{T} &= \sum_{j=0}^{N-1} C_j (F_{\gamma_{\max}}(U_j) - F_{\gamma_{\max}}(L_j)) + C_N \cdot (1 - F_{\gamma_{\max}}(L_N)) \\
&- \sum_{j=0}^{N-1} C_j \sum_{l=0}^K J_l \cdot e^{b_j \gamma_{T,j}} \cdot W_{j,l} \left( e^{-\frac{U_j}{W_{j,l}}} - e^{-\frac{L_j}{W_{j,l}}} \right) \\
&+ C_N \sum_{l=0}^K J_l \cdot e^{b_j \gamma_{T,j}} \cdot W_{N,l} e^{-\frac{L_N}{W_{N,l}}}
\end{aligned}$$

where  $J_l \equiv (-1)^{l+1} \frac{l}{\bar{\gamma}} \binom{K}{l}$  and  $W_{j,l} \equiv \frac{1}{b_j + \frac{l}{\bar{\gamma}}}$ .

## 3.2. Imperfect Channel State Information

### 3.2.1. Pilot-aided Channel Estimation and Feedback Delay

In this scenario, we assume the CSI is estimated using the MMSE method and the pilots available at the UT side, and then fed back to the BS to make the scheduling decision. The channel de-correlation due to Doppler spread and feedback delay is also addressed in this section.

The relation between the true channel  $g(t)$  and the estimated channel  $h(t)$  is found to be [7]:

$$g(t - \tau) = \rho_d(\tau) \rho_e \sqrt{\frac{\Omega_g}{\Omega_h}} h(t) + \varsigma(t), \tag{7}$$

where  $\varsigma(t) \sim \mathcal{CN}(0, (1 - |\rho_d(\tau)|^2 |\rho_e|^2) \Omega_g)$  and is independent of  $h(t)$ .  $\rho_e = \sqrt{\frac{\bar{\gamma}_p}{1 + \bar{\gamma}_p}}$  denotes the normalized correlation between the true and estimated channel gains with  $\bar{\gamma}_p$  representing the post-processing pilot SNR.  $\rho_d(\tau)$  characterizes the channel de-correlation

due to the Doppler spread, and when isotropic scattering is assumed, it can be written as  $\rho_d(\tau) = \mathcal{J}_0(2\pi f_d \tau)$ , where  $\mathcal{J}_0(\cdot)$  is the zero-th order Bessel function of the first kind and  $f_d$  is the Doppler spread of the user due to mobility.  $\Omega_g$  and  $\Omega_h$  are the variances of the true and estimated channel respectively. The scheduling decision is made based on the estimated maximum SNR  $\hat{\gamma}_{h,\max}$ . Besides the estimation inaccuracy due to finite power and finite duration pilots, another effect of channel estimation is the spectral efficiency loss due to pilot insertion. Taking all these aspects into account, the spectral efficiency can be expressed as [6]:

$$\bar{T} = \frac{N_d}{N_p + N_d} \mathcal{E} (C_j (1 - \text{PER}(j)) \cdot \mathbf{1}(\gamma_j \leq \hat{\gamma}_{h,\max} < \gamma_{j+1})),$$

where  $N_d$  and  $N_p$  are respectively the number of data and pilot symbols in one frame, over which the channel is assumed to be constant. Using the functions defined in (4), where  $Q_1(\cdot, \cdot)$  is the Marcum Q function, we can write  $\bar{T}$  in closed-form [6]:

$$\bar{T} = \frac{N_d}{N_d + N_p} \cdot \left( \sum_{j=0}^{N-1} C_j \cdot P(j) + C_N \cdot P(N) \right), \tag{8}$$

where the related coefficients are defined in (5).

### 3.2.2. Finite Level of Quantization Feedback Effect

In addition to the channel estimation error and feedback delay, the limited feedback channel bandwidth and finite precision of the quantizer further reduce the accuracy of the CSI at the BS. If we use the notation  $\hat{\gamma}_h$  to denote the received maximum quantized SNR,  $\gamma_h$  to represent the maximum estimated user SNR without quantization, and  $\gamma_g$  to denote the true maximum user SNR, the system throughput will be a joint function of  $\gamma_g$ ,  $\gamma_h$  and  $\hat{\gamma}_h$ , and the spectral efficiency evaluation needs to be averaged over them all and can be written as [6]:

$$\begin{aligned}
\bar{T} &= \sum_{j=1}^N C_j \cdot \frac{N_d}{N_p + N_d} \cdot \\
&\mathcal{E} \{ (1 - \text{PER}(j, \gamma_g, \gamma_h, \hat{\gamma}_h)) \cdot \mathbf{1}(\gamma_j \leq \hat{\gamma}_h < \gamma_{j+1}) \}.
\end{aligned} \tag{9}$$

After some mathematical derivations, the spectral efficiency in this case is found to be almost identical to (8), except that  $\gamma_j$  and  $\gamma_{j+1}$  in (4) need to be replaced with  $\gamma'_j$  and  $\gamma'_{j+1}$ , which are defined as follows. Let  $\hat{\gamma}_j$  denote the smallest quantized value that is greater

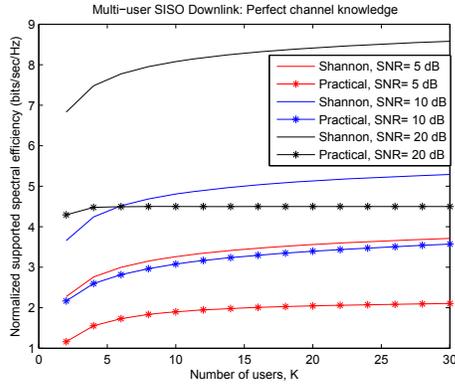


Fig. 3. Average spectral efficiency with perfect CSI.

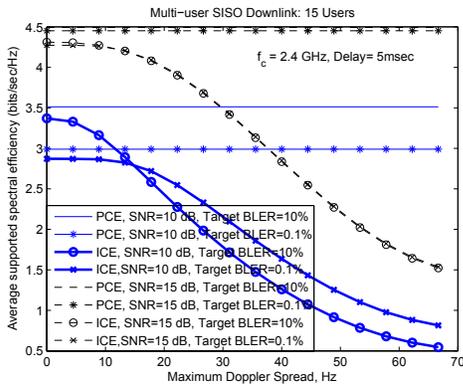


Fig. 4. Combined effects of imperfect CSI and Doppler on spectral efficiency.

than or equal to  $\gamma_j$ , and  $\hat{\gamma}_{j+1}$  denote the largest quantized value that is smaller than or equal to  $\gamma_{j+1}$ . Then  $\gamma'_j$  is defined as the lower bound of the quantization range for  $\hat{\gamma}_j$ , and  $\gamma'_{j+1}$  is defined as the upper bound of the quantization range for  $\hat{\gamma}_{j+1}$ .

#### 4. RESULTS AND DISCUSSIONS

Some results based on the above derived expressions are given in this section. The parameters used to generate these plots are given in the figures. Fig. 3 shows the performance gap between the ideal Shannon capacity and the supported spectral efficiency, under the assumption that the CSI at the BS is perfect. It can be clearly observed that, as the number of the users in the network increases, the Shannon capacity increases, but the practical supported spectral efficiency saturates at high SNR. This saturation is mainly due to the finite constellation size and can only be improved by packing more bits into each symbol. Fig. 4 considers the imperfect CSI induced by the feedback delay and Doppler spread as well as the finite pilot power. For a fixed feedback delay of 5ms, the larger the Doppler spread, the bigger the loss in spectral efficiency. The role of pilot power in the spectral efficiency is clear from the figure as well. Fig. 5 compares the spectral efficiency of a scheduler with perfect CSI versus one without. It can be seen from the figure that, within a wide range of average SNRs, the multiuser diversity gain is observed

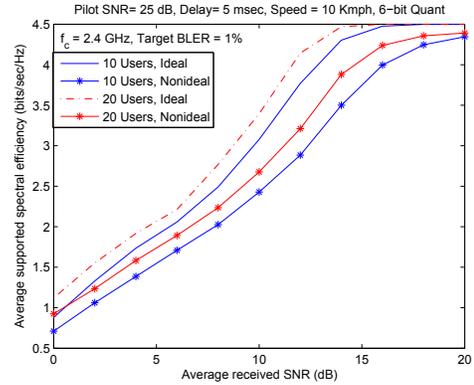


Fig. 5. Role of imperfect CSI, Doppler and limited feedback on spectral efficiency.

as the gap between the set of curves denoting different numbers of users. But as the average SNR increases, both curves saturate, and the multiuser diversity gain is no longer obvious.

#### 5. CONCLUSION

In this paper, we presented an analytical framework for the realistic performance of a practical channel-aware scheduler. We have proposed a simple log-linear model for the BLER of CTC operating at non-negligible BLER with a finite number of MCS classes. By taking into account the impact of channel estimation errors, Doppler spread due to user mobility, and the finite number of feedback bits available for channel quality reporting, we have derived closed-form expressions for the average spectral efficiency of a multi-user downlink scheduler on symmetric Rayleigh fading channels. Our results showed that coarse granularity of the number of MCS levels limits the achievable spectral efficiency, whereas imperfect channel estimation, coupled with Doppler and limited feedback, might lead to scheduling a user who does not have the best instantaneous channel quality, thereby causing an additional loss in spectral efficiency.

#### 6. REFERENCES

- [1] R. Knopp and P.A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *IEEE International Conference on Communications*, June 1995, vol. 53, pp. 331–335.
- [2] D.N. Tse, "Optimal power allocation over parallel gaussian broadcast channels," in *Proceedings of IEEE International Symposium on Information Theory*, 1997, p. 27.
- [3] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," in *IEEE Transactions on Information Theory*, June 2002, vol. 48, pp. 1277–1294.
- [4] IEEE., "Standard 802.16e-2005. part16: Air interface for fixed and mobile broadband wireless access systems amendment for physical and medium access control layers for combined fixed and mobile operation in licensed band," Dec 2005.
- [5] H. A. David, *Order Statistics*, Wiley, 1970.
- [6] P. Zhan, R. Annavajalla, A. Lee Swindlehurst, and T. Chauvin, "Realistic performance analysis for downlink multi-user diversity," in Preparation.
- [7] M. Schwartz, W. Bennett, and S. Stein, *Communications Systems and Techniques*, Wiley-IEEE Press, 1995.