Adaptive Channel Quality Feedback for LTE

The Nam Pham, Xiang Xu and Rudolf Mathar
Institute for Theoretical Information Technology, RWTH Aachen University
Aachen, Germany 52072
Email: {pham, xu, mathar}@ti.rwth-aachen.de

***Abstract*—In Long Term Evolution (LTE) networks, resource allocation and link adaptation rely on the channel quality indicator (CQI), which is a quantized representation of the signal interference plus noise ratio (SINR). Since CQI is measured by user equipment (UE) and sent to eNodeB via a feedback channel, in time varying channel, feedback delay should be compensated with prediction based on previous knowledge. In this work, several known prediction schemes are studied. Moreover, an adaptive CQI feedback scheme based on the sample autocovariance function is proposed. Simulation results show the proposed prediction scheme outperforms the other known prediction schemes in terms of overall prediction accuracy and system throughput. In addition, the performance is evaluated for systems both with and without hybrid automatic repeat request (HARQ), to investigate the influence of HARQ.**

I. Introduction

To improve the throughput of LTE systems, link adaptation based on adaptive modulation and coding scheme (AMC), is utilized to adjust the coding rate to fit the current channel state information (CSI) [1]. In addition, resource allocation also plays a crucial role in LTE networks, in which the transmit (Tx) power distribution and sub-channels assignment to UEs are optimized in order to maximize downlink capacity. Both link adaptation and resource allocation rely on the 4-bit quantized CQI reported by each UE to eNodeB in the uplink. CQI indicates the downlink signal to interference plus noise ratio (SINR) at the UE side. A CQI value appropriately derived from SINR ranges from 0 to 15. It suggests to the eNodeB the highest achievable efficiency of the downlink transmission under a specific channel condition, while assuring a block error rate (BLER) lower than 10% [2].

The importance of CQI in LTE network can be seen from its influence on system throughput. Moving UEs experience time varying mobile channels. And there is an unavoidable delay between a CQI report being generated at the UE and being used at the eNodeB. The feedback delay consists of measuring delay, propagation delay and processing delay. Therefore, there is always a mismatch between reported CQIs and realtime CQIs. An overestimated CQI can lead to transmission with high data rate in a poor channel and eventually an undecodable frame, whereas an underestimated CQI can diminish the downlink transmission throughput. Furthermore, in terms of resource allocation for multi-users, a sub-channel might not be assigned to an UE with a good channel quality but to an UE with a very optimistically reported CQI. Hence, the both CQI underestimation and overestimation lead to the decline in system throughput.

This work is partially supported by UMIC, a research project in the framework of the German excellence initiative.

In case of an undecodable frame, another frame will be transmitted using HARQ to provide extra information for decoding. Without HARQ, the throughput performance is consistent with prediction accuracy. However, with HARQ, it is not necessarily the case. The reason is, every retransmission of HARQ can be considered as an improvement in signal strength. And the bandwidth efficiency of LTE does not grow linearly to the SINR. Therefore, a retransmitted larger packet can carry more information than a few smaller packets without retransmission.

In previous works, several prediction schemes at the both eNodeB and UE side are investigated for only single-user systems. On one hand, the prediction can be performed at the eNodeB when reported CQIs with delay are used to predict current CQIs. On the other hand, the prediction can be performed at the UE, where the SINRs currently available in the observed time window are used to predict the SINR after the feedback delay. Subsequently, the predicted SINRs are mapped into CQIs, which will be used at the eNodeB directly for link adaptation and resource allocation. A simple extrapolation scheme, which requires no prior knowledge of the statistics of the channel, is proved to be very accurate for low speed users. However, it has numerical problem when users move faster [3]. A short-term average studied in [4] can provide good throughput for high speed users however performs poorly for low speed users. Wiener filtering can also be used for CQI prediction [5]. However, without a good model for the autocorrelation function, only sample autocorrelation can be used. Therefore, Wiener filtering has relatively high complexity.

In this work, existing prediction schemes are firstly addressed and compared. Beyond that, a technique to adapt prediction schemes to UE speed is proposed. Using the proposed adaptation technique, the optimum prediction scheme can be chosen and thus the system performance can be improved. The impact of HARQ on channel quality indicator feedback is also investigated based on throughput and latency, which distinguish this work from most of the existing works, which mostly only focus on throughput. The performance of the proposed schemes is tested in a multi-cell multi-user environment. Simulations confirm that the adaptive prediction scheme does not only provide superb prediction accuracy and throughput performance, but also reduces transmission latency.

This paper is organized as follows: The system model is introduced in Section II. Prediction algorithms followed by adaptation method are presented in Section III. Numerical results and analysis are shown in Section IV. And finally conclusions are given in Section V.

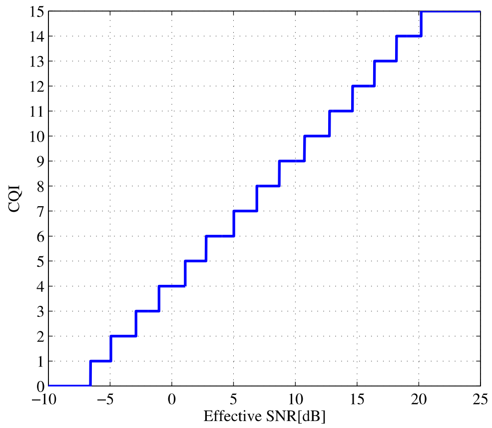

Fig. 1. SNR-CQI mapping function

II. System Model

In LTE system, the transmission resource is divided into physical resource blocks (PRB). A PRB has one slot of 0.5 ms in time domain and 180 kHz (12 sub-carriers) in frequency domain. The transmission time interval (TTI) is defined with a duration of 1 ms and it is also the time needed to transmit one subframe. A subframe consists of 2 PRB in time domain. In multi-user cellular networks, the sub-channel state presented by SINR can be measured for every subframe at the UE side, and the instantaneous received SINR on the kth sub-channel between UE m and its serving cell n at time t is given as:

$$\gamma_m(t,k) = \frac{P_n|H_{n,m}(t,k)|^2}{\sum_{j\neq n} P_j|H_{j,m}(t,k)|^2 + N_0} \quad (1)$$

where j denotes the index of neighboring cells of cell n. P_n is the Tx power of cell n, assuming the Tx power is uniformly distributed among sub-carriers and constant over time. $|H_{n,m}(t,k)|^2$ is the channel power response, and N_0 is the power of the additive white Gaussian noise (AWGN).

To reduce signaling overhead, the SINRs on a few consecutive sub-carriers, called sub-band, are compressed into an equivalent effective signal to noise ratio (SNR) in AWGN channel before feedback. The CQI messages are accordingly generated on sub-band level. The equivalence between SINR and effective SNR in AWGN channel is justified by reaching similar BLER with similar SNR. The SNR-BLER relationship in AWGN channel can be obtained by simulation using the modulation and coding schemes (MCS) specified in the CQI table Table I. The commonly used compression technique is exponential effective SINR mapping (EESM), more details about EESM can be found in [6].

Once the effective SNR is obtained, CQI can be derived using a simple stair function as showed in Fig. 1. The stair function is generated by using the effective SNR values from simulations with the constraint of BLER ≤ 0.1.

HARQ is an effective technique to combat burst errors in wireless channels. In addition, it also plays an important role in compensating channel prediction errors. In case of a transmission failure, the UE detects the error and sends back a non-acknowledgement signal to trigger the HARQ

TABLE I. The 4-bit CQI table in LTE [2]

CQI index	Modulation	Code rate × 1024	Efficiency [bit/s/Hz]
0	out of range		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16QAM	378	1.4766
8	16QAM	490	1.9141
9	16QAM	616	2.4063
10	64QAM	466	2.7305
11	64QAM	567	3.3223
12	64QAM	666	3.9023
13	64QAM	772	4.5234
14	64QAM	873	5.1152
15	64QAM	948	5.5547

mechanism. By using a hybrid of forward error correction (FEC) and automatic repeat request (ARQ), different copies of the encoded data can be sent several times and combined to provide better decoding performance. In other words, each retransmission in HARQ is equivalent to an enhancement of effective SNR.

In this work, a low-complexity HARQ simulation model proposed in [7] is adopted. The SNR gain $\Delta\gamma$ is modeled as a linear function of the number of retransmissions i, modulation order and effective code rate (ECR) ξ, which can be found in Table I. The dependency of SNR gain on i, the modulation order and ECR is given as:

$$\Delta\gamma_i = \mu_i \cdot \xi + \epsilon_i \quad (2)$$

where μ_i and ϵ_i are modeling parameters, as specified in Table II.

TABLE II. HARQ model parameters for each retransmission

Modulation	i	$\mu \cdot 10^{-2}$	ϵ
4-QAM	1	0.0804	2.89
	2	0.1628	4.57
	3	0.2006	5.62
16-QAM	1	0.0420	1.17
	2	0.8435	0.74
	3	0.9464	1.15
64-QAM	1	0.8996	-1.23
	2	1.2288	-1.71
	3	1.2728	0.15

III. Prediction Algorithms

The noise of the feedback CQI comes from different sources: The compression of SINR information from different subcarrier into a 4-bit CQI value, the inaccurate SINR estimation, as well as the feedback delay. Due to the scope of this work, SINR before compression is assumed to be noiseless, in order to isolate the effects of the feedback procedure. The prediction algorithms are dedicated to compensate the noise caused by feedback delay.

The prediction of SINR is based on a collection of past observations, which is defined as the prediction window. A general expression of prediction algorithms with a finite prediction window can be written in an autoregressive (AR) form

$$\widehat{\gamma}_m(t+\tau) = \sum_{l=0}^{p-1} w(l)\gamma_m(t-l), \quad (3)$$

where τ is the feedback delay, p is the prediction window size, w is weighting factor. For different user speeds, different weighting factors can be used to optimize the prediction.

1) Wiener filtering: The sample temporal autocorrelation of a finite observation period $[t - t_0, t]$ is defined as

$$r_{\gamma\gamma}(t, \Delta t) = \mathrm{E}\{\gamma_m(t')\gamma_m(t' + \Delta t)\}, \quad (4)$$

where $t' \in [t - t_0, t - \Delta t]$, Δt is time difference between two samples and $t_0 > \Delta t$ is essential. Sub-band index b is dropped, since the temporal correlation and frequency correlation are independent [8].

Due to the complexity of the multi-cell structure, analytical expression of the autocorrelation function (ACF) of SINR is difficult to obtain. Unlike channel impulse response, the temporal statistics of SINR in a multi-cell scenario is rarely studied in literature. In [9], the second order statistics of SINR in a multi-cell scenario are derived analytically. However, the analytical expression is very complicated and the accuracy of the model degrades as the speed of UE increases. In this work, assuming γ is a stationary process, the sample autocorrelation with larger t_0 is closer to theoretical value. Using the sample correlation, an order p Wiener filter can be obtained with the general form, using Δt equal to the feedback delay τ. The mean square error (MSE) of the predicted SINR can be written as

$$\varepsilon(t) = \mathrm{E}\left\{(\gamma_m(t + \tau) - \widehat{\gamma}_m(t + \tau))^2\right\}. \quad (5)$$

The minimum MSE (MMSE) can be found by letting the derivative be equal to zero

$$\frac{\partial \varepsilon(t)}{\partial w} = 0. \quad (6)$$

And the following equation can be obtained [5]:

$$\mathbf{R}\mathbf{w} = \mathbf{r} \quad (7)$$

where

$$\mathbf{w} = [w(0), w(1), \cdots, w(p-1)]^{\mathrm{T}}, \quad (8)$$

$$\mathbf{R} = \begin{pmatrix} r_{\gamma\gamma}(t,0) & r_{\gamma\gamma}(t,1) & \cdots & r_{\gamma\gamma}(t,p-1) \\ r_{\gamma\gamma}(t,1) & r_{\gamma\gamma}(t,0) & \cdots & r_{\gamma\gamma}(t,p-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{\gamma\gamma}(t,p-1) & r_{\gamma\gamma}(t,p-2) & \cdots & r_{\gamma\gamma}(t,0) \end{pmatrix}, \quad (9)$$

and

$$\mathbf{r} = [r_{\gamma\gamma}(t,\tau), r_{\gamma\gamma}(t,\tau+1) \cdots, r_{\gamma\gamma}(t,\tau+p-1)]^{\mathrm{T}}. \quad (10)$$

The MMSE prediction filter can be obtained by taking the inverse of $\mathbf{R}$

$$\mathbf{w}_{\text{Wiener}} = \mathbf{R}^{-1}\mathbf{r}. \quad (11)$$

Wiener filtering is more effective when the correlation is stronger. Therefore, for UEs with higher speed, the performance of Wiener filtering drops.

2) Short-term average: For UEs moving with very high speed, the autocorrelation becomes too weak and thus not helpful for prediction. In this case, an uniform weighting factor

$$w_{\text{Average}} = \frac{1}{p} \quad (12)$$

is supposed to offer a good prediction with very low computational complexity [4].

Since the purpose of short-term average is to average out fast fading but preserve slow fading, the size of the prediction window is quite important for short-term average. Window size too small is not enough to compensate for fast fading, whereas window size too large ignores slow fading.

3) Extrapolation: It is suggested in [3] that a simple extrapolation can provide a good prediction for UEs moving slowly. For very slow UEs, extrapolation can provide almost perfect prediction. However, for the UEs with higher speed, it can have severe numerical problem. In addition, extrapolation does not rely on correlation functions. Hence, the prediction window of extrapolation can be much shorter than of Wiener filtering.

4) Adaptive feedback scheme: Since each of the three aforementioned prediction schemes have their best operating range, a feedback scheme adapted to UE speed can combine their advantages. To enable the adaptive feedback scheme, the autocovariance of the SINR is studied. The autocovariance is given by

$$C_{\gamma\gamma}(t, \Delta t) = r_{\gamma\gamma}(t, \Delta t) - \mu_\gamma(t)\mu_\gamma(t + \Delta t) \quad (13)$$

where $\mu_\gamma(t)$ and $\mu_\gamma(t+\Delta t)$ are average SINR at t and $t+\Delta t$, respectively. Assuming Δt is small enough, we have

$$\mu_\gamma = \mu_\gamma(t) = \mu_\gamma(t + \Delta t) \quad (14)$$

and μ_γ can be approximated by the sample mean measured in the observation period $[t - t_0, t]$.

There are two factors influencing the variation of SINR, namely, distance related path-loss and fading process. In a very short period, the location of the UE can be treated as static. In this case, the normalized autocovariance of SINR can be very well approximated by a squared zero order Bessel function of the first kind

$$\widetilde{C}_{\gamma\gamma}(\Delta t) = J_0^2(2\pi f_{\mathrm{D}} \Delta t), \quad (15)$$

where f_{D} is the maximum Doppler frequency [9].

In practice, the autocovariance of SINR follows a similar pattern. But as the speed goes up, the static location approximation becomes less valid. Thus, the measured sample autocovariance is biased from (15), especially for larger Δt, as shown in Fig. 2. Nevertheless, the speed of UE can be roughly estimated through the sample autocovariance of SINR, and thus the best prediction scheme can be chosen accordingly. Furthermore, although the complexity is very low, the prediction accuracy of short-term average is not higher than Wiener filtering, even for extremely high speed ($v > 400$ km/h) [10]. Thus, only Wiener filtering and spline function extrapolation are considered in the adaptive feedback scheme.

Since for very small Δt, the autocovariance is a monotonic function of the moving speed, a fixed threshold can be set to

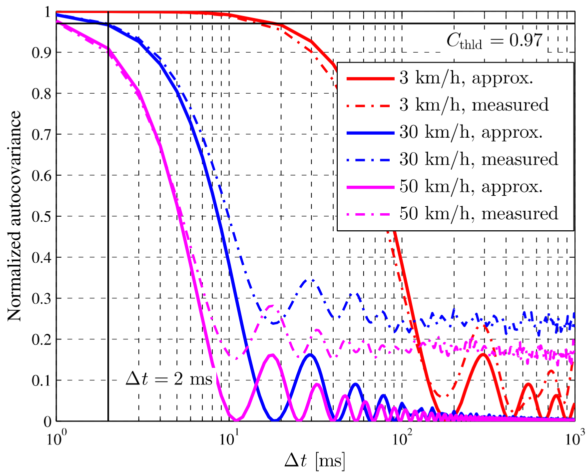

Fig. 2. Normalized sample autocovariance of SINR and its approximation for different speeds and threshold

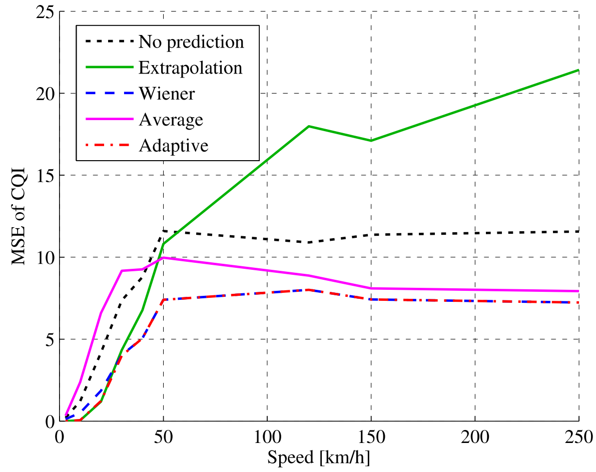

Fig. 3. MSE performance of prediction schemes for different speeds

determine the most suitable feedback scheme. In this work, the threshold is set to $C_{\text{thld}}(\Delta t) = 0.97$ at fixed $\Delta t = 2$ ms. These values are empirically calibrated through experiments. The sample covariance $C_{\gamma\gamma}(\Delta t)$ is measured for the same Δt and compared with $C_{\text{thld}}(\Delta t)$. For $C_{\gamma\gamma}(\Delta t) \geq C_{\text{thld}}(\Delta t)$, low speed is assumed and extrapolation is used, otherwise Wiener filtering is used. Therefore, the CQI is well predicted for both low speed and high speed UEs.

IV. Numerical Evaluation

Unlike the channel impulse response, the variation of SINR is associated with more facts, including the location, moving direction of the users, interference from other cells, etc. Therefore, it is difficult to analyze the performance of the prediction algorithms theoretically. In this work, the evaluation is purely simulation based.

The feedback schemes are tested in simulations with multi-cell environment. Standard simulation setup for a LTE system operating at 800 MHz is considered [11]. In total 19 eNodeBs, each equipped with 3-sector antennas, make up a network with 57 cells in hexagonal layout. 27 randomly moving UEs are dropped in a central cell for evaluation, where all the other 56 cells provide interferences. The mobility model restraint these users from leaving the serving cell. The UEs are divided into 9 different groups, according to their speeds. These speeds are given in Table III. SISO and full-buffer are assumed for the sake of simplicity. The UEs measure the SINR for every sub-frame and perform prediction. The feedback delay is set to 8 ms. Fixed feedback schemes are compared with the proposed adaptive feedback scheme. In addition, the performance of feedback without prediction is also provided as reference. Important parameters of the simulations are summarized in Table III.

TABLE III. Simulation Parameters

Channel model	Rayleigh fading channel
Total bandwidth	4.32 MHz
Sub-carrier bandwidth	15 kHz
Number of sub-carriers pro PRB	12
Number PRBs pro sub-band	4
Number of sub-bands	6
Total eNodeB transmit power	46 dBm
Inter-eNodeB distance	500 m
Observed window size	10 ms
HARQ	maximum 3 retransmissions
HARQ retransmission delay	8 ms
UE noise figure	9 dB
UE mobility model	random motion
UE's speed level	3,10,20,30,40,50,120,150,250 km/h
Path-loss model	$L = 15.3 + 18.8 \log_{10} d$
Power delay profile	Exponential
Shadow fading	Not considered
Channel knowledge	perfect
Simulation length	10,000 sub-frames

Due to the extent of this paper, we do not focus on optimal power and rate allocation algorithms, but apply constraints from LTE standards. Tx power on sub-bands is equally distributed during a TTI, and the resource allocation is simplified to sub-band assignment based on UEs signal quality, i.e. CQI. To achieve high system throughput, the best CQI scheduling scheme is adopted, where in each TTI the resource blocks are assigned successively to the user with the best CQI [12].

In Fig. 3, MSE of predicted CQI and actual CQI is compared for different speeds, where the adaptive prediction scheme is overlapped with extrapolation scheme in the low speed range (0 to around 30 km/h), and with Wiener filtering in the high speed range (above 30 km/h). And the overlapped area is the optimal operating range for those two prediction schemes. Therefore, the proposed adaptive prediction scheme can give the best overall accuracy for prediction. On the other hand, the accuracy of extrapolation decreases dramatically when the speed is higher than 50 km/h. And the short-term average is worse than no prediction for low speed, and less accurate than Wiener filtering for high speed.

The throughput performance is compared in Fig. 4, where the throughput without considering HARQ is almost consistent with the MSE result. Though extrapolation gives the poorest overall MSE, it has almost perfect prediction for low speed UEs. Therefore, its average throughput is better than using no prediction. The short-term average is only good for high speed UEs, but the advantage is insignificant. Regarding

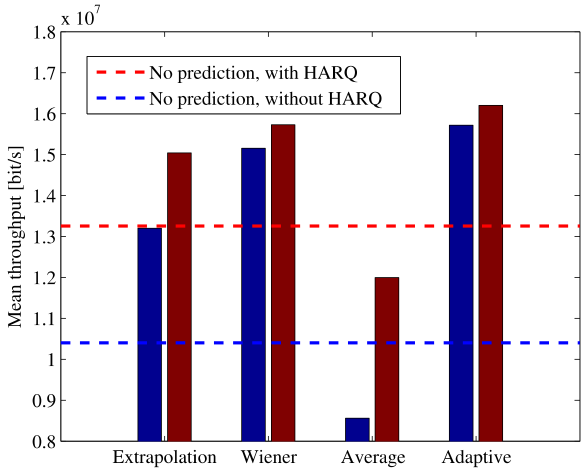

Fig. 4. Average throughput for different prediction schemes

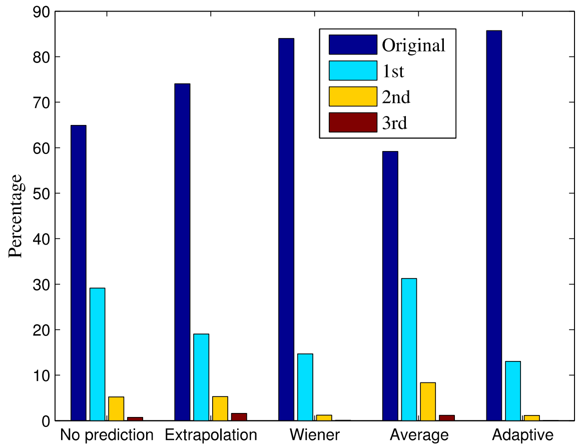

Fig. 5. Proportion of sub-frames for original transmissions and the 1st, 2nd and 3rd retransmissions

throughput, the overall performance of short-term average is even worse than no prediction. Wiener filtering has good prediction accuracy, and accordingly, the average throughput is also good. The proposed adaptive prediction scheme is the best prediction scheme in sense of average throughput, because the optimal prediction scheme is always chosen for a given speed. Since HARQ is equivalent to a SINR gain, the throughput is improved if HARQ is applied. The advantage of using HARQ is also clearly shown in Fig. 4. For all the prediction schemes, a performance improvement can be observed.

Moreover, although HARQ can improve throughput, retransmissions cause latency. More retransmissions typically mean larger latency. The proportion of sub-frames used for the original packets and the three retransmissions are listed in Fig. 5. It can be seen that for Wiener filtering and the adaptive prediction scheme, the proportions of the original transmissions and the 1st retransmissions are significantly larger than the other prediction schemes. As the number of retransmissions go up, the average latency for a successful transmission becomes larger. Therefore, another conclusion can be made here is that the adaptive prediction scheme also has the smallest latency.

V. Conclusion

In this work, a simple technique for adapting CQI feedback scheme to the moving speed of the UE is proposed. The proposed adaptation technique utilizes the property of the sample autocovariance of the SINR to estimate the moving speed of the UE, and switch between Wiener filtering and extrapolation accordingly. Since extrapolation gives very accurate estimates for low speed users whereas Wiener filtering is most effective for users with higher speed, the adaptive CQI feedback scheme offers overall the best performance in sense of both prediction accuracy of CQI and system throughput.

In addition, the role of HARQ in CQI prediction and feedback is discussed. By applying a simplified HARQ model, latency in systems with HARQ is evaluated. Simulation results confirm that the proposed adaptive feedback scheme gives smaller latency comparing to non-adaptive schemes, in the presence of HARQ.

VI. Acknowledgement

The authors thank the anonymous reviewers for their thoughtful comments.

References

[1] 3GPP TS 36.212 V8.8.0, "Evolved universal terrestrial radio access (E-UTRA): Multiplexing and channel coding," Dec. 2009.

[2] 3GPP TS 36.213 V8.8.0, "Evolved universal terrestrial radio access (E-UTRA): Physical layer procedures," Oct. 2009.

[3] M. Ni, X. Xu, and R. Mathar, "A channel feedback model with robust SINR prediction for LTE systems," in *7th European Conference on Antennas and Propagation (EuCAP)*, Gothenburg, Sweden, Apr. 2013.

[4] H. Dai, Y. Wang, C. Shi, and W. Zhang, "The evaluation of CQI delay compensation schemes based on Jakes' model and ITU scenarios," in *Vehicular Technology Conference. (VTC Fall). IEEE 76th*, Quebec, Canada, Sep. 2012.

[5] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley and Sons, 1996.

[6] S. N. Donthi and N. B. Mehta, "An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 10, pp. 3436–3448, Oct. 2011.

[7] J. C. Ikuno, M. Wrulich, and M. Rupp, "Performance and modeling of LTE h-ARQ," in *International ITG Workshop on Smart Antennas (WSA 2009)*, Berlin, Germany, Feb 2009.

[8] X. Xu and R. Mathar, "Low complexity joint channel estimation and decoding for LDPC coded MIMO-OFDM systems," in *Vehicular Technology Conference (VTC Spring), IEEE 73rd*, Budapest, Hungary, May 2011.

[9] X. Xu, T. N. Pham, and R. Mathar, "Second-order statistics of channel state information in a multi-cell mobile network with fast fading," in *IEEE 25th International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Washington DC, Sep 2014.

[10] N. Pham, "Predictive channel quality feedback for multi-user cellular networks," Bachelor thesis, Dec 2013, RWTH-Aachen University.

[11] 3GPP TS 36.814 V9.0.0, "Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects," Mar. 2010.

[12] S. Schwarz, C. Mehlfuhrer, and M. Rupp, "Low complexity approximate maximum throughput scheduling for LTE," in *Signals, Systems and Computers (ASILOMAR), the Forty Fourth Asilomar Conference on*, Pacific Grove, CA, USA, Nov 2010.