Rate Selection Analysis under Semi-Persistent Scheduling in LTE Networks

Donald Parruca, James Gross
UMIC Research Center
RWTH Aachen University
Aachen, Germany
Email: {parruca, gross}@umic.rwth-aachen.de

***Abstract*—Upcoming LTE networks have basically two different modes for scheduling data in the down-link by the base station. *Dynamic scheduling* brings the advantage of exploiting instantaneous channel state information while it puts on the other hand a significant burden on the system in terms of overhead and computation requirements. Especially for small packets that show up periodically, the overhead is typically too high. Therefore, the base station can serve such packet flows by the semi-persistent scheduling mode. In this mode, a certain resource allocation is fixed to a periodic schedule. While this does not allow any longer to exploit instantaneous channel states, it requires much less overhead. In this paper, we address the problem of selecting a modulation and coding scheme for such semi-persistent scheduling grants. The problem lies here in the stochastic characterization of the resource blocks over the next few seconds while on the other hand estimating based on such a characterization the blok error rate (and hence the average goodput). We provide a novel scheme, which outperforms all previously presented schemes significantly. The underlying model that we provide can also be used for any other long-term decision in an LTE system with semi-persistent scheduling such as interference coordination, handover decision etc.**

I. Introduction

Upcoming fourth generation cellular Long Term Evolution (LTE) systems are ultimately becoming deployed nowadays with several countries already offering LTE data services to customers. Such LTE systems promise much higher data rates together with quite low latencies achieves through a complete redesign especially of the air interface. Key features of LTE systems are an increased bandwidth, orthogonal frequency division multiplexing (OFDM) transmission scheme, multiple input multiple output (MIMO) transmission capabilities, a frequency reuse of one as well as advanced coding schemes (among many other features). All these features give the system a significant flexibility to serve various different traffic types under diverse conditions. As LTE is a centralized communication system, where the base station controls and schedules all data transmissions of all terminals in its association set, the base station is also responsible for optimally configuring the links and exploiting the provided flexibility. While this allows to achieve a high spectral efficiency, the constant reconfiguration costs of the system can become a burden.

A good example of the above circumstance is the operation of the base station scheduler in the down-link. In principle, LTE allows the scheduler to serve terminals via two different modes: a dynamic scheduling mode as well as a semi-persistent mode. In the dynamic mode, terminals periodically feed back detailed information regarding the instantaneous channel states of the transmission resources. This information is used at the base station to dynamically assign transmission resources to terminals depending on the feed back. While this allows to achieve a high spectral efficiency, it also puts a significant burden on the system due to channel state acquisition, computational load for fast assignments and signalling of the assignments to the terminals. Especially for small packets this overhead is not justified [3]. Therefore, the base station can also serve terminals based on a semi-persistent mode where a fixed resource grant allows a periodic transmission of data on the same set of resource blocks with the same modulation and coding type at certain repetition intervals. This saves overhead but is less efficient in terms of throughput as this resource grant can not take the instantaneous channel states into account.

In this paper we address the problem how to dimension such semi-persistent scheduling grants. In particular we focus on the question how to select the modulation and coding scheme for such a semi-persistent scheduling grant. This is a difficult question due to two issues. Firstly, the decision has to take the stochastic variations of the resource blocks of the grant into account. As LTE systems are known to be interference-limited due to the frequency reuse of one, the stochastic variations are not only due to the fading of the signal of interest but also due to the fading of interfering signals. Furthermore, if multiple resource blocks belong to the scheduling grant then all resource blocks will be utilized by the same modulation and coding type. Hence, determining already the instantaneous block error rate is involved as different resource blocks typically have different channel qualities. While for the instantaneous channel states good (but complex) models exist to determine the block error rate, it is open how to determine the long-term expected block error rate if only the channel statistics are known. However, it is essential for selecting the modulation and coding scheme accurately.

In this paper we provide an approximate model that allows primarily to predict the average goodput of a semi-persistent scheduling grant given the selection of a particular modulation and coding scheme. This model takes the statistics of the interference-limited channel behaviour into account as well

978-1-4673-5288-8/13/$31.00

as the fact that all resource blocks of the scheduling grant are employing the same modulation and coding type. The model provides quite good results for small numbers of resource blocks in the scheduling grant but looses accuracy for bigger grants (which are nevertheless unlikely to occur as for larger data chunks the dynamic scheduling mode pays off). Based on this model, we then present a straight-forward scheme for rate selection which allows to maximize the achievable goodput under semi-persistent scheduling. Compared to various other rate selection schemes from related work, our scheme provides a significantly better performance.

The paper is structured in the following way. We initially give a very brief overview of LTE systems, the addressed problem as well as related work in Section II. Then, we develop our analytical model in the Sections III and IV. Afterwards, we present numerical evaluations of the model in Section V before we conclude the paper in Section VI.

II. Preliminaries

In this section, we first give a brief introduction to LTE networks. We then discuss the addressed problem before we summarize related work with respect to it.

A. System Description

The 3GPP Long Term Evolution (LTE) system [2] is the upcoming fourth generation of cellular networks. Its architecture consists of the (wireless) access system and the (optical) core network. In the following we only focus on the access system. Wireless access is handled by base stations (also referred to as eNB in LTE lingo) which control all data transmissions within a dedicated area. The standard defines both frequency division and time division duplexing (FDD and TDD), from which FDD is the more likely deployment scheme. Furthermore, according to the standard time is split up into so called Time to Transmission Intervals (TTIs). These are the basic 'heart beats' of the system which enable the transmission of data according to more complex scheduling decision of the base station as discussed below. The duration of one TTI is in the range of milliseconds, we set in the following $T_{\mathrm{TTI}} = 1$ [ms]. Terminals are associated to dedicated base stations and exchange data frames with the base stations in up- and down-link direction. For this a certain bandwidth is provided and several different options are specified. A quite important feature of LTE networks is that neighbouring base stations of the same operator are going to fully reuse the up-link and down-link frequency bands. Hence, LTE networks are predominantly interference-limited systems in down-link and up-link. In the following, we will focus mainly on the down-link transmission direction.

1) Physical Layer: The LTE physical layer features Orthogonal Frequency Division Multiplexing (OFDM) as basic transmission system. Thus, for down-link data transmission each base station can utilize a system bandwidth of B [Hz] which is split into N disjoint sets of frequency bands of bandwidth B/N [Hz], also referred to as Resource Blocks (RBs). They are the minimal transmission resource that can be assigned to a terminal within a TTI. A RB consists of a bundle of N_{C} subcarriers and per TTI each subcarrier features N_{S} symbols for data transmission. Hence, $N_{\mathrm{C}} \cdot N_{\mathrm{S}}$ data symbols belong to one RB per TTI. All these symbols must be modulated with the identical constellation type while also all these symbols can only convey data for one distinct terminal associated to the base station. Per TTI a RB can not be shared by multiple terminals. Furthermore, a quite important feature of LTE is *link adaptation*. This refers to the possibility for the base station to utilize different modulation and coding (MCS) types depending on the quality of the resource blocks. LTE features three different modulation types: QPSK, 16-QAM and 64-QAM as well as several different coding rates. However, in LTE all resource blocks assigned to one distinct terminal must utilize the same MCS type. This choice makes the modelling of the system performance quite involved, as we will discuss later in the paper.

The quality of a resource block depends on the signal-to-interference-and-noise ratio (SINR). Taking the interference-limitation into account, any terminal in the system experiences on each resource block a randomly varying SINR which is a result of path-loss, shadowing and fading of the signal of interest as well as the interfering signal. Denoting by $h^2_{k,j,n}(t)$ the instantaneous channel gain of resource block n from base station k to terminal j at time t, the SINR $\gamma_{j,n}(t)$ on this resource block n at terminal j and time t is defined as:

$$\gamma_{j,n}(t) = \frac{P_{s,n} \cdot h^2_{s,j,n}(t)}{P_{i,n} \cdot h^2_{i,j,n}(t) + \sigma^2}\ , \tag{1}$$

where σ^2 denotes the noise power, s the base station where the signal of interest is coming from and i the interfering base station. Furthermore, $P_{k,n}$ denotes the transmit power of base station k on resource block n. For simplicity we assume in the following that all resource blocks are utilized by all base stations with the same transmit power.

The channel gains $h^2_{k,j,n}(t)$ vary randomly in time, frequency and space (terminal locations) due to fading that is caused by the mobility of terminals and the multi-path propagation environment. A common model for the fading is to assume the Rayleigh model, i.e. the channel gains $h^2_{k,j,n}(t)$ are exponentially distributed over time with mean $\bar{h}^2_{k,j,n}(t)$. This model applies to the signal of interest as well as to the interference signals. Over shorter time spans (i.e. over time spans up to a second) the means $\bar{h}^2_{k,j,n}(t)$ can be assumed to stay constant. However, they change over longer time spans (i.e. minutes and longer).

LTE features different methods for indicating the channel states from the terminals to the base stations. For instantaneous feedback there is a dedicated control channel. Furthermore, terminals can figure out the average channel gain of the signal of interest as well as of the interfering signals from the terminal measurements of the Reference Signal Received Power (RSRP) which serves primarily as indicator when to initiate a hand-over. By analysing the RSRP in combination with the transmission power of the downlink Reference Signals

(RSs) [4] the average channel gain can be determined and fed back to the base station.

2) Medium Access Control Layer: LTE systems feature two basic modes of operation at the base station regarding the MAC. On the one hand there is *dynamic scheduling*. Here the base station assigns resource blocks from TTI to TTI to different terminals depending on their channel qualities (which are fed back to the base station). This mode promises high spectral efficiency but requires fast computations of the scheduling assignments as well as some overhead to be taken into account due to channel state feedback and signalling the assignments to the terminals. Especially for small packets, this overhead can quickly outgrow the benefit from dynamic scheduling. Therefore, a second operation mode allows the base station to fix resource block assignments in a periodic pattern. This mode is referred to as *semi-persistent scheduling*. Semi-Persistent scheduling grants some terminal j a dedicated set of resource blocks every T_P TTIs where the resource blocks are employing MCS type m. Such a periodic scheduling grant does not require instantaneous channel state information. There is also no signalling overhead to be spent apart from the initial one. However, the dynamic channel states can not be exploited which reduces spectral efficiency. Furthermore, due to the stochastic variation of the channel states, packets transmitted over persistently assigned resource blocks are more likely to be erroneous as the instantaneous channel states can not be taken into account. Hence, these packets need to be retransmitted which is usually done via the dynamic scheduling mode, i.e. every base station typically reserve some time slots for dynamic scheduling of the retransmitted packets.

B. Problem Statement and Related Work

In this paper we focus exclusively on the way to determine semi-persistent scheduling grants. The difficulty here stems from the fact that the scheduling grant needs to be chosen for the next few seconds and is hence a decision that has to be taken under uncertainty. The question that we will address is how many resource blocks are to be assigned to a terminal and which MCS scheme to employ on these resource blocks. The basic trade off that we face is that a higher MCS scheme has a higher spectral efficiency but is also more vulnerable to bad channel states stemming from the random variations. In fact, two issues make the choice of the scheduling grant a tough one. On the one hand, there is the random channel variation of the resource blocks in time and frequency due to fading. Note that the fading is actually occurring for the signal of interest but also for the interference signal. Hence, we require a stochastic model of the SINR given the average channel gains for the signal of interest and the interference signal as the semi-persistent scheduling decision relates to a longer time span over which the SINRs will vary significantly. On the other hand, these variations translate now into a certain packet error rate which depends on the choice of the modulation and coding scheme for the semi-persistent grant. Ultimately, we require a good analytical model for the average goodput that we can expect for a certain terminal if assigning it some resource blocks with MCS type m. Especially if multiple resource blocks are assigned, the fact that all resource blocks are employing the same MCS type but have differently fading SINRs makes the theoretic analysis quite difficult. In this paper, we address both issues and provide on top of our developed model for the average goodput a MCS selection scheme with good performance.

These issues have not been addressed by related work so far. Despite the importance of the joint stochastic knowledge of the assigned subcarriers, to the best of our knowledge, no work so far considers it for the MCS and RB assignment. In [3] the principle and performance of semi-persistent scheduling is discussed. To fit the block size of different packets different combinations of RB and MCS allocation combinations are used. In [5] a semi-persistent packet scheduling scheme limiting the control channel resources is presented. It uses a fixed MCS QPSK 2/3 with 2 RBs allocation for VoIP packet transmission. The selected MCS in both the previous works is selected blindly of the channel gain conditions, making these approaches subject to channel over- or under-exploitation.

In [6], the authors describe and evaluate by simulation different semi-persistent radio resource assignment algorithms aiming VoIP capacity improvement. For the downlink estimation, worst case SINR is assumed, collected from wideband measurement reports. Similar to [6] a more conservative MCS selection is proposed in [7]. In case that the reported channel quality index (CQI) (available only during the dynamic period, i.e. every 20 ms) falls the bins 7 - 9 or 10 15, the lowest CQI of the bin will be used. Such a scheme showed a 15% capacity improvement compared with linear ones as in [6]. Such feedbacks are based on the temporary channel realizations and do not give any information on the long term channel characteristics. Therefore, the MCS selection will be outdated after the duration of the coherence time, typically a few TTIs.

The MCS selection algorithms proposed in [8] assume a Nakagami-m SINR distribution. The MCS allocation is based on the minimization of the number of HARQ retransmissions, neglecting the statistics of additionally RBs assigned in the same TTI. In [9] the authors extend the work of [8] for time correlated channel conditions. In contrast to these works, we concentrate on the performance improvement since the first transmission and consider the joint statistics of the assigned RBs.

III. Basic Model Development

We start our model development by rehearsing briefly the state-of-the-art for instantaneous performance modelling of LTE links. Assuming a certain set of resource blocks is assigned to a terminal and the SINR per resource block varies (and is denoted in the following by γ_n), models exist that accurately predict the block error probability of the data transmission for a choice of MCS m. Such a model is generally referred to as link error prediction model (LEP). All known LEPs depend on a conversion of the current state of the link (i.e. the set of the γ_n) into a single figure of merit and deducing from this the error probability by a table look-up or an approximation formula. However, notice that these models do not predict the long term packet error rate as the fading varies randomly. Instead, they predict the error rate for a given set of SINR values of the resouce blocks.

There are two well-established and accurate LEP models developed for this purpose: the exponential-effective SINR Mapping (EESM) model [10] and the Mutual Information Effective SINR Mapping (MIESM) model [11]. Both models have been validated and are accepted by the community. In this paper we will work with the EESM model. In the EESM model, the resource block SINRs are mapped into a so called *effective SINR* γ_{eff} of the link which is then used to determine the block error rate $p_{\text{e}}(\gamma_{\text{eff}}, m)$ depending on the chosen MCS.

The EESM compression function is:

$$\gamma_{\text{eff}} = \beta_m \log \frac{1}{N_{\text{RB}}} \sum_{n=1}^{N_{\text{RB}}} \exp\left(-\frac{\gamma_n}{\beta_m}\right). \qquad (2)$$

Given the effective SINR and the used MCS m the block error rate is the determined based on a lookup table generated from link-level simulations. In the above equation, β_m is a constant obtained from link level simulations. It is chosen such that the BLER of the effective SINR is as close as possible to the BLER of the compressed SINR realizations.

As discussed above, the EESM model gives the corresponding BLER during the upcoming TTI that the transmission will experience. In contrast, our interest in this paper is on determining the average link goodput $R_j(m)$ for a significantly longer time instant than one TTI (in the order of seconds). In order to obtain this average link capacity, we first have to obtain a distribution of the effective SINR and derive from this the expectation of the goodput as given by the throughput multiplied by the average block error rate over all channel realizations. Formally, we are interested in the following expression:

$$\mathrm{R}_j(m) = N_{\text{C}} N_{\text{S}} N_{\text{RB}} \frac{C(m)}{T_P \cdot T_{\text{TTI}}} \int_0^\infty f_{\text{eff}}(x)(1 - p_{\text{e}}(x, m))dx \qquad (3)$$

where $f_{\text{eff}}(x)$ is the probability density function of the effective SINR γ_{eff}, $C(m)$ is the symbol weight (in bits) of MCS m [12] and N_{RB} is the number of simultaneously assigned RBs.

The difficulty of expression 3 relies on the calculation of $f_{\text{eff}}(x)$ for interference-limited systems and arbitrary numbers of resource blocks mapped into a scheduling grant. We briefly summarize up-to-date efforts from related work. In [13] the moment generating function (MGF) of the exponential effective SNR for two independently fading resource blocks was derived. Later on in [14] the MGF calculation was extended to N independent and identically distributed Rayleigh fading channels. However, due to the difficulty in obtaining a closed form solution of the effective SNR PDF an approximate function using the method of Pearson system was introduced in [15]. Although valuable these contributions take only noise-limited systems into account. As previously stated, LTE systems are mainly interference-limited and the interference statistics needs to be taken into account as well [16]. In the following we will perform such an analysis using the idea to approximate the effective SINR PDF and CDF with the help of minimum order statistics while using exact expressions for the basic SINR PDFs.

IV. Long-Term Performance Model for Persistent Scheduling

A. Basic Analysis

We start with the derivation of the PDF and the CDF of the SINR on a given resource block assuming only Ralyeigh fading to influence the signal-of-interest and the interference signal. Assuming such a fading model for both signals, the PDF of the SINR can be obtained, as for example discussed in [16]: Let X denote the random SINR variable with PDF $f_X(x)$ and CDF $F_X(x)$. Furthermore, denote the average received power of the signal of interest at the terminal by $P_{\text{S}} = P_{s,n} \cdot \bar{h}^2_{s,j,n}$ and the average received interference power by $P_{\text{I}} = P_{i,n} \cdot \bar{h}^2_{i,j,n}$ we have for the PDF:

$$f_X(x) = \left[\frac{\sigma^2}{P_{\text{I}} \cdot x + P_{\text{S}}} + \frac{P_{\text{S}} \cdot P_{\text{I}}}{(P_{\text{I}} \cdot x + P_{\text{S}})^2}\right] e^{-\frac{\sigma^2}{P_{\text{S}}} \cdot x}, \qquad (4)$$

and for the CDF:

$$F_X(x) = 1 - \frac{P_{\text{S}}}{P_{\text{I}} \cdot x + P_{\text{S}}} e^{-\frac{\sigma^2}{P_{\text{S}}} \cdot x}. \qquad (5)$$

Note that the LTE standard defines the subcarriers to have a bandwidth of 15 kHz each. As a resource block consists of 12 such subcarriers, the fading within the resource block can be considered to be flat. Hence, we model the SINR PDF of all subcarrier within a resource block by the same random variable and only consider in the following the N_{RB} resource blocks of a scheduling grant. Let $x_{n,j}$ be the random variable representing the SINR for all subcarriers of RB n at the terminal j.

In [14] it was proven that for a set of N_{RB} subcarrier SNR random variables $X_{1,j}, X_{2,j}, \cdots, X_{N_{\text{RB}},j}$ with MCS m used for transmission, the effective SNR is bounded by $Y_{1,j} \leq \gamma_{\text{eff}} \leq Y_{1,j} + \beta_m \log(N_{\text{RB}})$, where $Y_{1,j}$ is the smallest SNR within the set of the N_{RB} resource blocks. This result also holds for interference-limited channels. Hence, the effective SINR γ_{eff} is also bounded by the above interval taking the SINRs of the subcariers into account.

Through order statistics [17] we can obtain the PDF of the smallest random variable among a set of random variables. In the following we consider a generalized derivation which also applies if the transmit powers per resource block are different. Let $X_1, X_2, \ldots, X_n$ be n independent and non identically distributed random variables having CDF $F_i(x)$ and PDF $f_i(x)$. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the ordered sample of the random variables. Therefore, we have $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$. Denote by $f_{1:n}(x)$ the PDF of the minimum of the order set of the n random variables. Then the PDF of this minimum order statistic is:

$$f_{1:n}(x) = \frac{1}{(n-1)!} \operatorname{Per} \mathbf{A}, \quad x \in \mathbb{R}, \qquad (6)$$

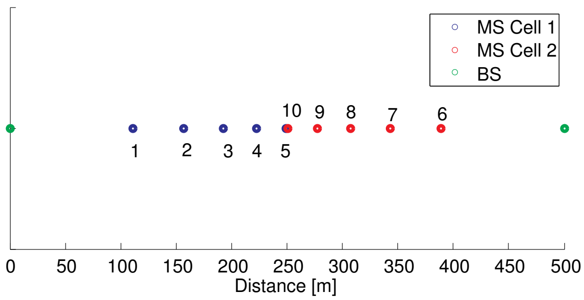

Fig. 1: MS Distribution

with

$$\mathbf{A} = \begin{array}{c} \begin{pmatrix} f_1(x) & f_2(x) & \dots & f_n(x) \\ \bar{F}_1(x) & \bar{F}_2(x) & \dots & \bar{F}_n(x) \end{pmatrix} \end{array} \begin{array}{ll} 1 & \text{rows} \\ n-r & \text{rows} \end{array}.$$

Per is the permanent operator on the matrix $\mathbf{A}$ and $\bar{F}_i(x) = 1 - F_i(x)$.

Hence, $f_{1:n}(x)$ gives us the probability that from a subset of n r.v. the minimum of the realizations will be within an interval around x. However, the real effective SINR lies within the range $[X_{1:n}, X_{1:n} + \beta \log(n)]$. As it is difficult to obtain a more precise expression for the effective SINR, we will match in the following the best value of the range that fits simulated data. In other words, we use in the following as empirical model for the effective SINR the model $X_{1:n} + \epsilon_m \beta_m \log(n)$ and determine ϵ_m empirically. This results in the following simplification of Equation 3:

$$\begin{aligned} \mathrm{R}_j(m) \quad &\approx \quad N_{\mathrm{C}} N_{\mathrm{S}} N_{\mathrm{RB}} \frac{C(m)}{T_P \cdot T_{\mathrm{TTI}}} \\ &\int_0^\infty f_{1:n}(x) \left[1 - p_e\left(x + \epsilon_m \beta_m \log(N_{\mathrm{RB}})\right)\right] dx \end{aligned} \tag{7}$$

B. Parameter Estimation of ϵ_m

In order to match the ϵ parameter as good as possible, we conducted simulations of the average goodput and obtained afterwards the ϵ parameter which matched the simulated values as good as possible (minimizing the error between analytically predicted and simulated results). For the simulation we considered a system set up with mobile stations dropped along a straight line between the serving and the interfering base station, as shown in Figure 1. Terminals were assumed to be stationary. The transmission power per BS was 20 W and a uniform power mask was used for transmission meaning that for a bandwidth of 5 MHz with 25 RBs a total of 0.8 W per RB was assumed as transmit power. The terminals were persistently assigned per TTI n different RBs with MCS m during a simulation run of 5 seconds simulated time. The total amount of data that could be sent per TTI equalled $N_{\mathrm{C}} N_{\mathrm{S}} N_{\mathrm{RB}} C(m)$. After each packet was received at the terminal, based on the EESM model the corresponding BLER was calculated leading to a decision of correct packet reception or not. Overall, this yielded an average goodput at each considered station. This value was used afterwards to match the parameter ϵ To achieve statistical confidence, we repeated these simulation runs 30 times and obtained the overall average goodput over all simulation runs (taking care of the seeds of the simulation runs, obviously). Denote this overall average goodput per terminal j as $\bar{\delta}_{m,n,j}$ where n stands for the number of resource blocks assumed in the persistent scheduling grant and m stands for the MCS chosen. We then compared these empirical goodputs with the estimations $\delta_{m,n,j}$ from Equation 8 and selected the ϵ parameter which reduced the error $E(m,n)$ considering all terminals in the system according to the following formula:

$$E(m,n) = \sum_{j=1}^{J} \sum_{n=1}^{N} |\bar{\delta}_{m,n,j} - \delta_{m,n,j}|. \tag{8}$$

The corresponding results are shown in Figure 2. In general we observe that a bigger ϵ yields a lower error between measured and predicted average goodput. Also, the bigger the number of combined resource blocks is the bigger is the error. This is important, as it limits the validity of our model, as also discussed below. For all following results, we simply set the ϵ parameter to 1.

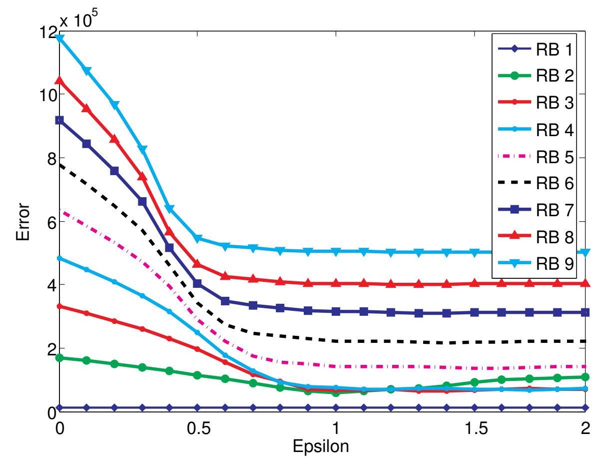

Fig. 2: Error between the predicted goodput and the obtainable goodput for different ϵ values

V. Goodput Prediction and Rate Selection - Numerical Investigation

After setting the parameter ϵ we are now in the position to evaluate the goodput prediction quality of our model. We do so by considering the same simulation setting as above (for the parameter matching) and comparing the results for the average goodput for the different terminals from simulations with our analytical model. We refer in the figures to the simulated goodputs as 'obtainable' results while the results from our model are referred to as 'estimated'. Finally, we also compare the results with an approach we only take the statistics from a single resource block into account, i.e. using the formula:

$$\begin{aligned} \mathrm{R}_j(m) \quad &= \quad N_{\mathrm{C}} N_{\mathrm{S}} N_{\mathrm{RB}} \frac{C(m)}{T_{\mathrm{TTI}}} \\ &\int_0^\infty f_X(x)(1 - p_e(x,m)) dx \end{aligned} \tag{9}$$

where $f_X(x)$ is the SINR PDF of only a single RB (and not the minimum SINR PDF among a set of N resource blocks) assigned to the MS as defined in 4. In the result plots the outcome of this model is labeled as "Single" and is considered as a rate prediction comparison scheme.

We compare the goodput results of all three variants below in Figures 3, 4 and 5 for a persistent scheduling grant of 1, 4 and 6 resource blocks for all 15 MCS types. The reassignment period of the grants is chosen to be 20 ms. We furthermore consider different terminal positions representing different combinations of average received power from the signal of interest and from the interfering signal. As it can be noticed, the proposed empirical prediction scheme has a good match with the actual rates obtained from simulations. The "Single" model and the prediction one, are exactly the same when only one RB is assigned. That is why their curves are overlapping in Figure 3. For high MCS schemes and cases when more than one RBs are simultaneously assigned, the "Single" comparison scheme overestimates the obtainable rate significantly. This happens due to two factors: firstly, for higher MCS the BLER is higher. Secondly, the stochastic properties of the effective SINR are not taken into account accurately for more than one assigned resource block. Nevertheless, the more resource blocks are assigned in a grant, the bigger is the difference between the proposed model and the observed average goodput from the simulations. The deviation becomes rather big for scheduling grants with more than 6 resource blocks (not shown here) such that we propose to only use our model up to 6 resource blocks.

Next, we discuss the proposed rate selection scheme. For each MS according to its position there is an optimal MCS to be used which is typically the same from the simulations as well as from our model. Hence, we propose to use our model to iteratively precompute the expected goodput and use the MCS scheme that maximizes it, formally defined as:

$$\mathrm{m}_j^* = \arg\max_m \{\mathrm{R}_j(m)\}. \tag{10}$$

We evaluate in the following the suitability of this rate selection scheme and compare it to simulations. In addition, we also consider two comparison schemes from related work. The first one is referred to as "CQI Feedback" scheme in the graphs. In this scheme the base stations observes the channel state feedback reports from the stations and adapts the MCS type for the persistent scheduling grant after some time span to the worst report received from the terminal [6], [7]. The second comparison scheme simply selects always the MCS scheme QPSK with coding rate $3/5$ whatever terminal position or resource block size is set for the grant. A very similar scheme has been proposed by [5].

The corresponding results for all MCS selection schemes are shown in Figures 6, 7 and 8. For all considered scenarios our proposed scheme is quite close to the obtainable goodput as observed from exhaustive simulations. For terminals farer away than the cell center it should be noted that the fixed setting of MCS type QPSK with a coding rate of $3/5$ provides a remarkably good performance while the other proposed rate

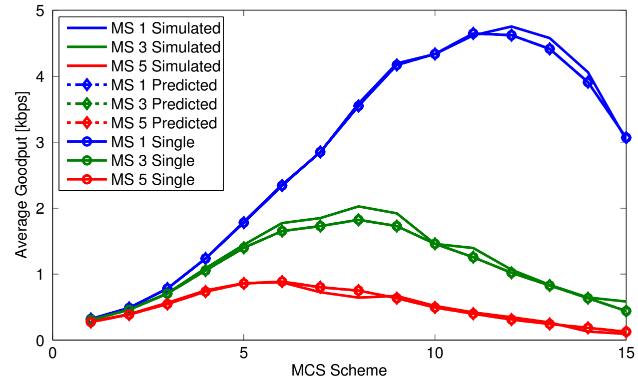

Fig. 3: Goodputs for 1 resource block only.

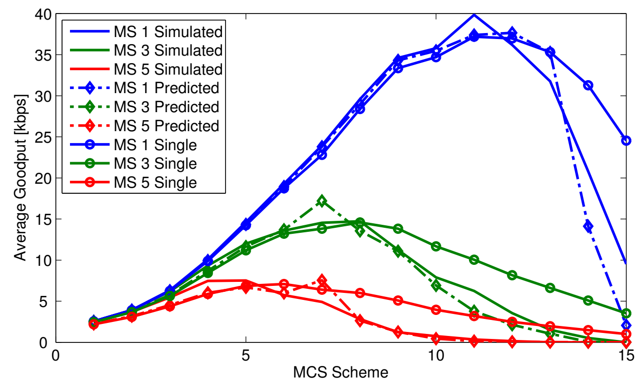

Fig. 4: Goodputs for 4 resource blocks.

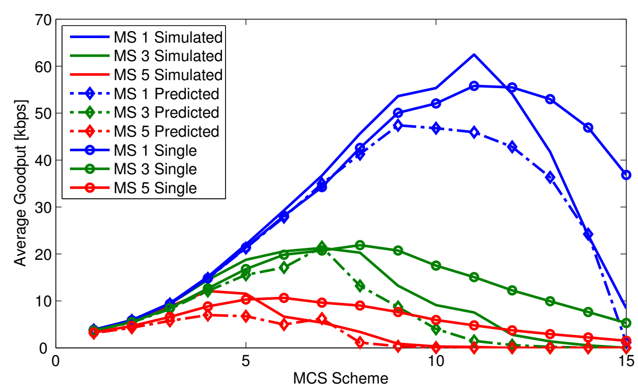

Fig. 5: Goodputs for 6 resource blocks.

selection scheme essentially is too conservative and causes too many erroneous transmissions.

VI. Conclusions and Future Work

In conclusion, we observed that for an optimal link adaptation the appropriate analytical model is needed. It needs to consider the long-term stochastic properties of all the contemporary assigned RBs. Additionally, the average goodput estimation model introduced in this paper predicts relatively well the rates obtained during simulation. Thanks to this property, the proposed MCS scheme is quite close to the obtainable goodput observed from exhaustive simulations.

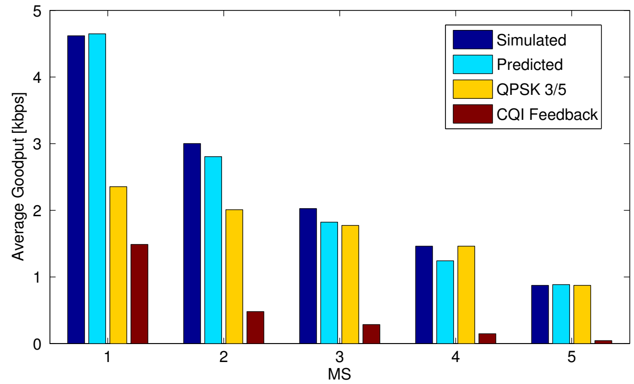

Fig. 6: Goodputs of various MCS selection approaches: 1 RB.

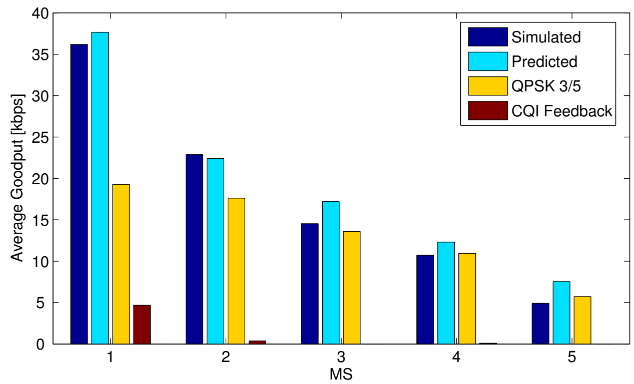

Fig. 7: Goodputs of various MCS selection approaches: 4 RBs.

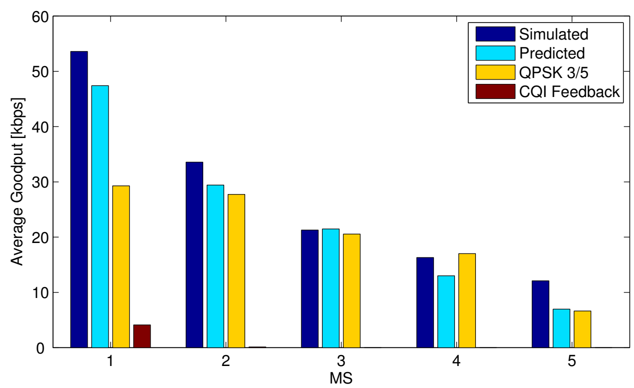

Fig. 8: Goodputs of various MCS selection approaches: 6 RBs.

Consequently, making it also perform better than existing link adaption algorithms for semi-persistent scheduling.

In the future, we would like to consider the mobility of the wireless terminals as well. In such a scenario, the persistent scheduling grant duration plays an important role on the system performance. For an optimal allocation of resources, it needs to be updated according to the mobility scenario of the terminals in the system. The identification of the right update period would be of special interest.

VII. Acknowledgements

This work was supported by the DFG Cluster of Excellence on Ultra High-Speed Mobile Information and Communication (UMIC), German Research Foundation grant DFG EXC 89.

References

[1] D. Jiang, H. Wang, E. Malkamaki, and E. Tuomaala, "Principle and performance of semi-persistent scheduling for voip in lte system," in *Proc. Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on*, Sept. 2007.

[2] 3GPP, "Overview of 3gpp release 11 v0.1.1 (2012-06)," Tech. Rep. V0.1.1, Jun. 2012.

[3] D. Jiang, H. Wang, E. Malkamaki, and E. Tuomaala, "Principle and performance of semi-persistent scheduling for voip in lte system," in *Proc. International Conference on Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007.*, Sept. 2007.

[4] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley Publishing, 2009.

[5] Y. Fan, P. Lunden, M. Kuusela, and M. Valkama, "Efficient semi-persistent scheduling for voip on eutra downlink," in *Proc. IEEE 68th Vehicular Technology Conference, 2008. VTC 2008-Fall.*, Sept. 2008.

[6] M. Muehleisen and B. Walke, "Evaluation and improvement of voip capacity for lte," in *Proc. European Wireless 2012*. Poznan, Poland: Polish Society of Theoretical and applied Electrical Engineering, Apr. 2012.

[7] L. Ding, F. Tong, Z. Chen, and Z. Liu, "A novel mcs selection criterion for VOIP in LTE," in *Proc. 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), 2011*, Sept. 2011. 1–4.

[8] C. Cho, H. Jin, N.-O. Song, and D. K. Sung, "MCS selection algorithms for a persistent allocation scheme to accommodate voip services in ieee 802.16e ofdma system," in *Proc. IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, 2009*, Sept. 2009.

[9] H. Jin, C. Cho, N.-O. Song, and D. K. Sung, "Optimal rate selection for persistent scheduling with harq in time-correlated nakagami-m fading channels," *Proc. IEEE Transactions on Wireless Communications*, vol. 10, Feb. 2011.

[10] E. Tuomaala and H. Wang, "Effective sinr approach of link to system mapping in ofdm/multi-carrier mobile network," in *Proc. 2nd International Conference on Mobile Technology, Applications and Systems, 2005*, Nov. 2005.

[11] L. Wan, S. Tsai, and M. Almgren, "A fading-insensitive performance metric for a unified link quality model," in *Proc. IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, 2006.

[12] G. T. 36.213 V8.5.0, *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures,(Release 8)*, Dec. 2008.

[13] H. Song, R. Kwan, and J. Zhang, "On statistical characterization of eesm effective snr over frequency selective channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 3955–3960, Aug. 2009.

[14] ——, "General results on snr statistics involving eesm-based frequency selective feedbacks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1790–1798, May 2010.

[15] ——, "Analytical approximations of eesm effective snr distribution using pearson system," in *Proc. Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011.

[16] F. Naghibi and J. Gross, "How bad is interference in ieee 802.16e systems?" in *Proc. IEEE European Wireless Conference (EW), 2010*, Apr. 2010.

[17] N. Balakrishnan, "Permanents, order statistics, outliers, and robustness," *Revista matemática complutense*, vol. 20, no. 1, pp. 7–107, 2007.