

Analysis of MAC-level Throughput in LTE Systems with Link Rate Adaptation and HARQ Protocols

Raffaele Bruno*, Antonino Masaracchia*,[†], Andrea Passarella* and Stefano Mangione[†]

*Institute of Informatics and Telematics (IIT)

CNR, Pisa, Italy

Email: {name.surname}@iit.cnr.it

[†]Department of Electrical Engineering

Università di Palermo, Italy,

Email: {name.surname}@unipa.it

Abstract—LTE is rapidly gaining momentum for building future 4G cellular systems, and real operational networks are under deployment worldwide. To achieve high throughput performance, in addition to an advanced physical layer design LTE exploits a combination of sophisticated mechanisms at the radio source management layer. Clearly, this makes it difficult to develop analytical tools to accurately assess and optimise the user perceived throughput under realistic channel assumptions. Thus, most existing studies focus only on link-layer throughput or consider individual mechanisms in isolation. The main contribution of this paper is a unified modelling framework of the MAC-level downlink throughput in LTE cells that caters for wideband CQI feedback schemes, AMC and HARQ protocols as defined in the LTE standard. We have validated the accuracy of the proposed model through detailed LTE simulations carried out with the ns-3 simulator extended with the LENA module for LTE.

I. INTRODUCTION

3GPP-LTE is the emerging standard for future 4G cellular systems. To achieve high throughput performance, in addition to an advanced physical layer design LTE exploits a combination of sophisticated radio resource management functionalities, such as Channel Quality Indicator (CQI) reporting, link rate adaptation through Adaptive Modulation and Coding (AMC), and Hybrid Automatic Retransmission Request (HARQ) [1]. Specifically, a base station (eNB) can simultaneously serve multiple users on orthogonal subcarriers that are grouped into frequency resource blocks (RBs). Then, each user (UE) periodically measures channel state information that is fed back to the eNB in terms of CQI reports. Typically, only aggregate CQI values are reported to reduce channel feedback information. CQI measurements are used by eNBs for scheduling and link rate adaptation on the downlink [2]. For instance, the modulation and coding scheme (MCS) is typically selected in order to maximise the data rate to the scheduled UE subject to a constraint on the error probability. Errors at the physical layer are also mitigated by exploiting automatic retransmission protocols with channel coding (HARQ). More precisely, HARQ procedures use the classical stop-and-wait algorithm, in which the eNB decides to perform a retransmission based on the exchange of ACK/NACK messages with the UE. Then, UEs try to decode the packet by combining the retransmitted copies.

Since user, cell and radio link throughputs are among the most important performance indicators that the operators adopt to assess the QoS in an LTE system [3], an extensive literature exists that investigates LTE throughput performance based on analytical models [4]–[6], simulation tools [2] or field tests [7], [8]. However, it is evident that a complex interplay exists among the various mechanisms that operate at the MAC layer to improve communication reliability and to increase data rates. This makes accurate LTE throughput analysis notably difficult. Thus, most studies limit the analysis only to the radio link throughput or consider single MAC functions in isolation [9]. Furthermore, simplified error models are typically considered that only allow deriving upper bounds for the LTE throughput [5].

In this paper we develop a unified modelling framework of the MAC-level downlink throughput that is valid for homogeneous cells [10] and Rayleigh-distributed fading. Our model simultaneously caters for CQI feedback schemes that use spectral efficiency to generate CQI, as well as AMC and HARQ protocols. Furthermore, we include in the analysis an accurate link layer abstraction model that uses the Mean Mutual Information per coded Bit (MMIB) metric to derive the physical error probability [11]. The throughput estimates of our model are accurate, as validated using the ns-3 simulator extended with the LENA module for LTE¹. Furthermore, our results confirm that the IR-HARQ mechanism is very effective in improving error correction. However, the effectiveness of the IR-HARQ scheme depends on the appropriate selection of the modulation and coding scheme of the first transmission attempt.

After a brief review of the some related work (Section II) we introduce the LTE MAC model (Section III). Then, we develop our analytical method to calculate the MAC-level downlink throughput of LTE systems in the presence of Rayleigh fading (Section IV). We validate the analysis with simulation results in two general network scenarios (Section V), and we conclude this paper with final remarks and an outlook of future work (Section VI).

¹<http://networks.cttc.es/mobile-networks/software-tools/lena/>

II. RELATED WORK

As pointed out previously, several analytical and simulation models, as well as experimental studies, have been proposed for characterising the throughput performance of LTE systems. It is out of the scope of this section to provide an extensive overview of all these studies and we only focus on reviewing analytical models that are most related to this work.

Several works are reported in the literature that focuses on analysing the bit error probability (BER) for OFDM systems under various channel configurations and in the presence of channel estimation errors. For instance, in [12] closed-form expressions for the BER performance of equalised OFDM signals in Rayleigh fading are derived for various signal constellations. The analytical results of [12] are extended in [13] to calculate the BER of an OFDM system in the presence of channel estimation errors. In [14] the BER performance of uncoded OFDM systems are analysed for Rayleigh and Rice frequency-selective fading channels in the presence of transmitter nonlinearities. Significant research efforts have been also dedicated to generalise the BER performance analysis to multiple-input multiple-output (MIMO) channels [15]–[17].

Most related to our work are the studies that focus on analysing the capacity of LTE systems with scheduling, rate adaptation and limited channel-state feedback. In [5] an upper bound is derived for the achievable throughput in LTE systems using the so-called “best-M” CQI reduction scheme and max-SNR user scheduling. Closed-form expressions for the throughput achieved in LTE systems under different schedulers (proportional fair, greedy, and round robin), multiple-antenna diversity modes and CQI feedback schemes are derived in [4], [18]. Specifically, the model in [4] applies to LTE systems that use EESM to generate CQI reports (an explanation of the EESM method is provided in Section III), while the model in [18] applies to LTE systems that generate CQI reports by simply taking an arithmetic average of the SNRs of the subcarriers. A SNR quantisation feedback scheme is also analysed in [6]. However, most of these papers assume a simplified model for the channel outage, which does not take into account HARQ procedures as specified in the LTE standard. The performance of HARQ with rate adaption for the LTE downlink is studied in [9], but only through simulations.

III. LTE MAC PROTOCOL SPECIFICATION

We now briefly describe relevant details of the LTE MAC for downlink transmissions, with special attention to frame structure, CQI feedback mechanisms and HARQ protocols. We also introduce the system model and notation, and we discuss the main assumptions that underlay our analysis.

In the downlink, LTE uses the Orthogonal Frequency Division Multiple Access (OFDMA) scheme. Specifically, in the frequency domain the system bandwidth, W , is divided into several orthogonal subcarriers. Each subcarrier has a bandwidth of 15 kHz. In the time domain, the duration of one LTE radio frame is 10 ms. One frame is divided into ten transmission time intervals (TTIs) of 1 ms each. Furthermore, each TTI consists of two 0.5 ms slots. Each slot contains either

six or seven OFDM symbols, depending on the Cyclic Prefix (CP) length. A set of twelve consecutive subcarriers over the duration of one slot is called a physical Resource Block (RB). Then, OFDMA has the ability to dynamically assign RBs to individual users. Hereafter, we denote with q the total number of RBs available over the system bandwidth.

Since the RB bandwidth is only 180 kHz, it is reasonable to assume that the channel response is frequency-flat across all the twelve subcarriers of the RB². Then, let us denote with $\gamma_{i,k}$ the SNR of the i^{th} RB of the k^{th} UE. Clearly, the statistics of the SNR depend on the channel model and the multi-antenna diversity mode of operation. As commonly adopted in other LTE models, e.g. [4], in this study we assume that the fading from the eNB to the UEs is *Rayleigh distributed*. This implies that *the SNR of each RB is an exponential random variable (RV)* [19]. Furthermore, we also assume an *homogeneous cell model* [10], i.e. the SNR is independent for different users and RBs. This also means that the SNRs of all RBs are *uncorrelated* in frequency and space, and $\gamma_{i,k}$ can be regarded as independent and identically distributed (i.i.d.) RV.

Popular methods (e.g., EESM and MIESM) that are typically used in LTE to compute CQI values rely on the concept of *effective SNR*. Basically, the UEs map the SNRs of multiple subcarriers/RBs into a single value by applying complex non-linear transformations. Then, the effective SINR is used to estimate the BLER experienced by a user and to determine the appropriate MCS, i.e. the MCS that allows the UE to decode the transport block with an error rate probability not exceeding 10%. However, the statistics of the effective SNR generated by EESM and MIESM techniques are not known in closed-form. Thus, they must be approximated or computed numerically, which makes performance analysis difficult [4], [20]. An alternative approach proposed in [21] to implement AMC capabilities is based on the *spectral efficiency*. Specifically, let us denote the with $\eta_{i,k}$ the spectral efficiency of the i^{th} RB of the k^{th} UE. Then, it holds that [22]

$$\eta_{i,k} = \log_2 \left(1 + \frac{\gamma_{i,k}}{\Gamma} \right), \quad (1)$$

where $\Gamma = -\ln(5\beta)/1.5$ and β is BLER upper bound. Now a static mapping can be determined between the spectral efficiency and the CQI index, as well as between the CQI index and the MCS value [21]. More formally, let us denote with $C_{i,k}$ the CQI index for the i^{th} RB of the k^{th} UE. Typically the value of CQI can range between 1 and L . Then, $C_{i,k} = j$ ($j = 1, \dots, L$) if $S_j \leq \eta_{i,k} \leq S_{j+1}$, with $S_0 = 0$ and $S_{L+1} = \infty$. In other words the CQI value is a quantised version of the spectral efficiency³. Closely related to the MCS selection is also the transport block (TB) size determination. More precisely, let n_k the number of RBs allocated to the

²This assumption will not hold for highly dispersive channels with long delay spread.

³Note that in the 3GPP-LTE standard, $L = 15$ and the S_j thresholds are specified in Table 7.2.3-1 of [23]. Furthermore, in the 3GPP-LTE standard the available MCS indexes are 32 but a 4-bit CQI allows selecting only 15 MCS. Thus, in practical LTE systems only a subset of available MCS is typically used.

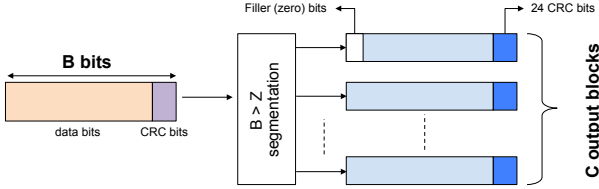


Fig. 1. Transport block segmentation.

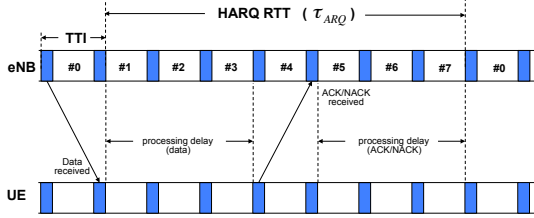


Fig. 2. HARQ processes and timing in FDD-LTE DL.

k^{th} UE during a frame. Then, the number B of bits that can be delivered in those RBs, which is called transport block, is a function of the MCS index⁴. Furthermore, if $B > Z$ (with $Z = 6144$ bits in 3GPP-LTE) the transport block is segmented into a number C of code blocks (CBs) that are independently encoded. Note that the CB size highly impacts the actual BLER performance for a given MCS [11]. Figure 1 exemplifies the transport block segmentation.

Regarding the HARQ protocol, LTE employs two types of HARQ schemes. In HARQ type-I, each encoded data frame is retransmitted until the frame passes the CRC test or the maximum number of retransmissions is reached. Erroneous frames are simply discarded. In contrast, in HARQ type-II, each transmission contains incremental redundancy (IR) about the data frame. Thus, consecutive transmissions can be combined at the receiver to improve error correction. Although our model is valid for all HARQ types, in the following we only consider HARQ type-II that is the most widely used in LTE. Note that in LTE systems retransmissions typically use the same MCS index as the initial transmission. It is also important to point out that the transmission of HARQ feedbacks (i.e. ACK/NACK messages) is not instantaneous but each received packet experiences a processing delay. According to the LTE standard, the processing delay at the receiver is about 3 ms. Thus, assuming the same delay to process data transmissions and ACK/NACK messages, the HARQ round trip time, say τ_{ARQ} , is 7 TTIs, as shown in Figure 2. For this reason, an eNB must support up to 8 parallel HARQ processes for each UE to enable *uninterrupted* communications. In this way, an eNB can continue to transmit new TBs while the UEs are decoding already received TBs.

IV. MAC-LEVEL THROUGHPUT ANALYSIS

In this section we develop the mathematical model of the MAC-level downlink throughput for a single LTE cell with n

⁴See Table 7.1.7.2.1-1 of [23] for the static mapping between TB size, MCS and number of RBs allocated to the UE.

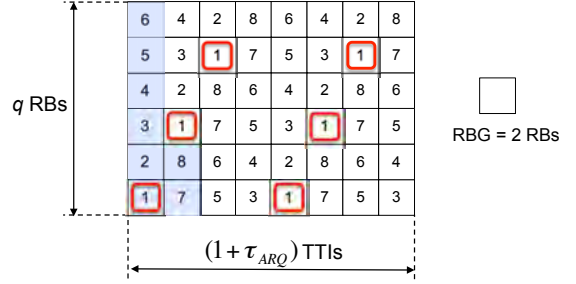


Fig. 3. RR operations with $q = 12$, $P = 2$ and $n = 8$.

randomly deployed UEs. Without loss of generality we assume asymptotic traffic conditions, i.e., infinite traffic is waiting for each user at the transmission buffer of the eNB. As discussed in Section III, the packet scheduler at the eNB is responsible for both allocating RBs to UEs every TTI, and controlling up to 8 HARQ processes per UE. Intuitively, the maximum number of HARQ processes that can be concurrently activated by the scheduler during an HARQ period is bounded by the number of times the same UE is scheduled during a τ_{arq} time interval. For the sake of simplicity, in this study we consider a Round Robin (RR) scheduler, which works by dividing the total amount of available radio resources in a fair manner among the UEs. More precisely, a RR scheduler allocates to each UE a set of consecutive resource blocks, called resource block groups (RBGs), whose size P depends on the system bandwidth [23]. Consequently the number b of RBs assigned to each UE is simply given by

$$b = \max \left\{ P, \left\lfloor \frac{q}{n} \right\rfloor \right\}. \quad (2)$$

Without loss of generality we consider a non-adaptive HARQ strategy, in which the scheduler maintains the same RBG and MCS configuration of the original TB when scheduling the retransmissions.

To illustrate the dependency between the number of times a UE is scheduled during an HARQ period, the RBG size and the total number of UEs in Figure 3 we exemplify the scheduling decisions that are cyclically performed by the RR scheduler during an HARQ period with $q = 12$, $P = 2$ and $n = 8$. As shown in the figure, each UE is scheduled six times during an HARQ period. In general, the average number of times each UE is scheduled in $(1 + \tau_{ARQ})$ TTIs is simply given by

$$n_{RR} = \frac{q(1 + \tau_{ARQ})}{n \cdot b}. \quad (3)$$

It must be noted that not all the transmission opportunities allocated by the eNB to an UE result into a successful transmission due to signal attenuation, shadowing and fading. In the following we denote with $P_e(m, k, r)$ the TB error probability at the r retransmission for the k^{th} UE when m is the adopted MCS, with $P_s(m, k, r)$ the probability that the k^{th} UE correctly decodes a TB after r retransmissions when m is the adopted MCS, and with $P_d(m, k)$ the probability

that the k^{th} UE discards a packet when m is the adopted MCS because it has reached the maximum number of failed retransmissions. In Section IV-B we provide closed-form expressions for $P_s(m, k, r)$, $P_e(m, k, r)$ and $P_d(m, k)$. Finally, to perform the throughput analysis we observe the system at the end of each successful transmission, because all the processes that define the occupancy pattern of the channel (i.e., HARQ processing delays and retransmissions) regenerate with respect to the sequence of time instants corresponding to the completion of a successful transmission. Then, it follows that the MAC-level throughput for the k^{th} UE is

$$\rho_k = n_{RR} \frac{E[TB|Succ]_k}{E[T_{ARQ}]_k}, \quad (4)$$

where $E[TB|Succ]_k$ is the average number of information bits that are delivered with a successful transmission of the k^{th} UE, and $E[T_{ARQ}]_k$ is the average time needed by an HARQ process to complete a successful transmission of the k^{th} UE. In equation (4), the multiplying factor n_{RR} is used to take into account that n_{RR} independent HARQ processes run in parallel. The following theorem provides closed-form expressions for $E[TB]_k$ and $E[T_{ARQ}]_k$.

Theorem 1: By assuming an homogenous cell with Rayleigh-distributed fading, and a RR scheduling policy

$$E[TB|Succ]_k = \sum_{j=0}^L TBS(m(j), b) [1 - P_d(m, k)] g_k[j], \quad (5a)$$

$$E[T_{ARQ}]_k = \sum_{j=0}^L \left[\sum_{r=0}^{r_{max}} (r+1)(1 + \tau_{ARQ}) P_s(m, k, r) \right] g_k[j], \quad (5b)$$

where $TBS(m(j), b)$ is a function that computes the number of data bits transmitted in b RBs using the MCS $m(j)$ ⁵, and $g_k[j]$ is the PMF of the CQI value reported by the k^{th} UE.

Proof: See Appendix. ■

In the remaining sections we develop the analysis for the unknowns in (5a) and (5b).

A. CQI feedback scheme and AMC strategy

The objective of this section is twofold. First, we develop the analytical tools to characterise the wideband CQI feedback scheme of LTE. Second, we analyse the performance of a link rate adaptation techniques based on wideband CQI reports.

Let us assume that n UEs are randomly distributed in the cell, and let d_k be the distance of the k^{th} UE from the eNB. As discussed in Section III we also assume that $\gamma_{i,k} \sim \text{Exp}(\lambda_k)$, where the rate parameter λ_k of the exponential distribution depends on the UE position. Under this assumption the statistics of the spectral efficiency for each RB can be expressed in a closed-form as given by the following Theorem.

⁵For the sake of notation brevity, we indicate with $m(j)$ the MCS that the eNB uses when the CQI value reported by a UE is equal to j .

Theorem 2: If $\gamma_{i,k} \sim \text{Exp}(\lambda_k)$ then the cumulative distribution function (CDF) of the spectral efficiency $\eta_{i,k}$ in equation (1) is computed as:

$$F_\eta(x; i, k) = \begin{cases} 1 - e^{-\lambda_k \Gamma(2^x - 1)} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \quad (6)$$

Proof: See Appendix. ■

LTE specifies different types of CQI reporting: *wideband* and *subband*. Specifically, the wideband CQI represents the SNR observed by the UE over the whole channel bandwidth, while the subband CQI represents the SNR observed by the UE over a collection of adjacent RBs. Note that a vector of CQI values should be transmitted to the eNB when using the latter feedback scheme. Thus, the subband-level feedback scheme ensures a finer reporting granularity but it also generates a higher overhead. In this study, we focus on the wideband feedback scheme and we assume that the CQI reported by the k^{th} UE, say \hat{C}_k is the *arithmetic mean* of the CQI values computed over all RBs⁶. Then, we use the spectral efficiency to generate the CQI values from the SNR measures of all RBs. The statistics of the wideband CQI are mathematically derived below.

Claim 1: The probability mass function (PMF) of the CQI value for the i^{th} RB assigned to the k^{th} UE is given by

$$g_{i,k}[j] = F_\eta(S_{j+1}; i, k) - F_\eta(S_j; i, k). \quad (7)$$

Proof: See Appendix. ■

Claim 2: In an homogenous cell the PMF of the wideband CQI value reported by the k^{th} UE is given by

$$g_k[j] = \sum_{l=qj}^{q(j+1)-1} g_k^{(q)}[l], \quad (8)$$

where $g_k^{(q)}[l]$ is the q -fold convolution of $g_{i,k}[j]$.

Proof: See Appendix. ■

As described in Section III a static mapping is typically established between the CQI value received at the eNB and the MCS for the downlink transmissions. Thus, $g_k[j]$ also characterizes the distribution of the MCS index $m(j)$ that is used by the eNB for the downlink transmission to the k^{th} UE when the reported wideband CQI is j .

B. Physical layer error model

We now conclude the analysis by introducing the physical layer error model. In this study we adopt the general approach initially proposed in [24] to accurately approximate the BLER curves of OFDMA-based wireless systems, and later specialised for the LTE case in [25]. Specifically, we assume that the *mutual information per coded bit* (MIB) of MCS m , as defined in [24], can be accurately approximated by a combination of Bessel functions of the SNR γ as follows

$$I_m(\gamma) \approx \sum_{h=1}^H \alpha_h J(\psi_h \sqrt{\gamma}), \quad (9)$$

⁶Note that an alternative solution would be to report the worst CQI value over all (or a subset of) RBs as in [5].

where H , α_h and ψ_h parameters are empirically calibrated as a function of the MCS index. Subsequently, the *mean* MIB (MMIB) value for each UE is computed by averaging the corresponding mutual information of all RBs allocated to that UE. Specifically, let $\Omega(k)$ be the set of RBs that are allocated to the k^{th} UE by the scheduler. Then, the MMIB value over the vector of SNR values for each RB assigned to the k^{th} UE when m is the adopted MCS is simply given by

$$\hat{I}_{m,k} = \frac{1}{\omega(k)} \sum_{i \in \Omega(k)} I_m(\gamma_{i,k}), \quad (10)$$

where $\omega(k)$ is the cardinality of the $\Omega(k)$ set. The non-linear nature of (9) makes an exact analysis difficult. Thus, previous studies limit the computational complexity of deriving MMIB values in multi-user scenarios by considering a quantised version of the $I_m(\gamma)$ function (9) in order to *discretise the MIB metric* [25]. More precisely, let us define a set $\mathcal{V}_m = \{\mu_m[0], \mu_m[1], \dots, \mu_m[v_m]\}$ for each MCS m such that

$$\mu_m[v] = I_m(Q_{m,v}), \quad (11)$$

where $(Q_{m,v+1} - Q_{m,v}) = \delta\gamma$ is the quantisation step size, and $Q_{m,0}$ is the minimum usable SNR for MCS m . Now, let us denote with $H_{i,m,k}$ the discrete MIB value for the i^{th} RB scheduled to the k^{th} UE when m is the adopted MCS. Similarly to the approach adopted for CQI mapping, we assume that $H_{i,m,k} = \mu_m[v]$ ($v = 0, \dots, V$) if $Q_{m,v} \leq \gamma_{i,k} \leq Q_{m,v+1}$. In other words the discrete MIB value is associated to a *range* of SNRs. It is straightforward to derive the statistics of the discretised MIB metric as follows.

Claim 3: In an homogenous cell with Rayleigh-distributed fading, the PMF of $H_{i,m,k}$ is given by

$$h_{i,m,k}[v] = F_\gamma(Q_{(v+1),m}; i, k) - F_\gamma(Q_{v,m}; i, k), \quad (12)$$

where $h_{i,m,k}[v] = Pr\{H_{i,m,k} = \mu_m[v]\}$.

Proof: See Appendix. ■

Similarly, we introduce a discrete MMIB metric, say $\hat{H}_{m,k}$, computed over the set of RBs allocated to the k^{th} UE when m is the adopted MCS. In particular, $\hat{H}_{m,k}$ can be obtained as the mean of the $H_{i,m,k}$ values over the set $\Omega(k)$. Thus, the statistics of the discretised MMIB value are derived using the same technique of Claim 2.

Claim 4: In an homogenous cell the PMF of $\hat{H}_{m,k}$ is given by.

$$h_{m,k}[v] \approx \sum_{l \in \Phi_v} h_{i,m,k}^{(\omega(k))}[l]. \quad (13)$$

where $h_{i,m,k}^{(\omega(k))}[l]$ is the $\omega(k)$ -fold convolution of $h_{i,m,k}[l]$. The definition of the Φ_v set is quite involved and is given in the proof.

Proof: See Appendix. ■

Once the MMIB value is given, a direct MMIB to BLER mapping can be used to obtain the *code block error rate*, without necessarily defining an effective SINR. Following the approach proposed in [24], the empirical BLER curve for MCS

m can be approximated with a Gaussian cumulative model as follows

$$CBLER_m(y, e) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{y - b_e}{c_e} \right) \right], \quad (14)$$

where y is the MMIB value, while b_e and c_e are parameters used to fit the Gaussian distribution to the empirical BLER curve⁷. These parameters depend on the Effective Code Rate (ECR), i.e. the ratio between the number of downlink information bits (including CRC bits) and the number of coded bits. Intuitively, the ECR value is a result of the selected TB size, MCS, and $\Omega(k)$. Then, the overall error probability for a transport block transmitted as a combination of C code blocks, each one associated with a MMIB and ECR value, can be computed as

$$TBLEER_m(y, e) = 1 - \prod_{i=1}^C (1 - CBLER_m(y_i, e_i)). \quad (15)$$

However equation (15) does not take into account the impact of an IR-HARQ mechanism that combines retransmissions to improve error correction. To generalise equation (15) for a system with incremental redundancy we adopt the same approach as in [26]. In particular, we introduce an *equivalent* MMIB metric as the average of the mutual informations per HARQ block received on the total number of retransmissions. More precisely, let us assume that the original transport block has been retransmitted r times. Then, let $(\hat{I}_{m,k}^{(0)}, \hat{I}_{m,k}^{(1)}, \dots, \hat{I}_{m,k}^{(r)})$ be the vector of MMIB values for each of these transmissions. The equivalent MMIB for the r^{th} retransmission can be computed as follows

$$\hat{I}_{m,k,r} = \frac{1}{r+1} \sum_{i=0}^r \hat{I}_{m,k}^{(i)}, \quad (16)$$

Then, the PMF of the equivalent MMIB value for the r^{th} retransmission is $h_{m,k}^{(r)}[v] = Pr\{\hat{I}_{m,k,r} = \mu_m[v]\}$. This PMF can be obtained using the same technique as in Claim 4 and it is not reported here for the sake of brevity. Similarly, we compute the effective ECR after r retransmissions, say $e^{(r)}$, by dividing the number of information bit of the original transmission with the sum of the number of coded bits of each retransmissions. Finally, by applying the law of total probability the *average* TB error probability at the r retransmission for the k^{th} UE when m is the adopted MCS can be computed as

$$P_e(m, k, r) = \sum_{v=0}^{v_m} TBLEER_m(\mu_m[v], e^{(r)}) \cdot h_{m,k}^{(r)}[v]. \quad (17)$$

Finally to evaluate $P_s(m, k, r)$ we can observe that the r^{th} retransmission of the k^{th} UE, when m is the adopted MCS, is a success only if the previous $(r-1)$ transmissions were TBs received erroneous and the r^{th} transmission is correctly

⁷Empirical BLER curves can be obtained through field measurements or detailed link-level simulations.

decoded. Hence, it immediately follows that

$$P_s(m, k, r) = \left[\prod_{i=0}^{r-1} P_e(m, k, i) \right] \times [1 - P_e(m, k, r)] . \quad (18)$$

We conclude this section by noting that the probability $P_d(m, k)$ that the k^{th} UE discards a packet when m is the adopted MCS because it has reached the maximum number of failed retransmissions is simple given by:

$$P_d(m, k) = \prod_{i=0}^{r_{max}} P_e(m, k, i) . \quad (19)$$

Finally, it is straightforward to observe that the *average* probability of discarding a packet for the k^{th} UE is computed as

$$P_d(k) = \sum_{j=0}^L P_d(m(j), k) g_k[j] . \quad (20)$$

V. PERFORMANCE EVALUATION

In this section we assess the accuracy of our throughput analysis in two different scenarios. In the first one, we assume that the mapping function that is used to convert spectral efficiency into CQI feedbacks, and then into MCS indexes is sufficiently accurate. As better explained in the following, in this condition error probabilities are typically small and retransmissions may have a negligible impact on the MAC layer throughput with respect to other protocol overheads. In the second scenario we assume that a fixed CQI is fed back to the eNB by each UE. Thus, the eNB necessarily selects a fixed MCS independently of the current channel conditions. This clearly represents a worst-case scenario, which is useful to assess the robustness of our modelling framework even when the reported CQI provides a very poor prediction of channel performance. Furthermore, it is also useful to better highlight the ability of the HARQ scheme to improve the overall system throughput without relying on per-subcarrier information.

A. Simulation setup

All the following experiments have been carried out using the ns3 packet-level simulator, which includes a detailed implementation of the LTE radio protocol stack. The main simulation parameters are summarised in Table I. Specifically, we consider an *Urban Macro* scenario, in which path loss and shadowing are modelled according to the COST231-Hata model [27], which is widely accepted in the 3GPP community. The fading is Rayleigh distributed. To limit the computation complexity of the simulator pre-calculated fading traces are included in the LTE model. Given the downlink system bandwidth (see Table I) a RBG comprises two RBs [23], i.e., $P=2$. Regarding the network topology, we considered a single cell with a varying number of static UEs, chosen in the range [1, 50]. Note that, in our settings a maximum number of 96 (i.e., $8\lfloor q/P \rfloor$) unique UEs can be scheduled within an HARQ period. Indeed, if $n > 96$ the RR period is longer than the HARQ period. All results presented in the following graphs are averaged over multiple simulation runs with different fading

traces and topology layouts. Confidence intervals are generally very tight and they are not shown in the figures if below 1%. Each simulation run lasts 300 seconds.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency (GHz)	2.14
DL bandwidth (MHz)	5
q	25
eNB TX Power (dBm)	43
CQI processing time (TTI)	2
CQI transmission delay (TTI)	4
Antenna scheme	SISO
PDCCH & PCFICH (control ch.)	3 OFDM symbols
PDSCH (data ch.)	11 OFDM symbols
n	[10, 50]
r_{max}	3

B. Results with adaptive CQI

In this section we validate the accuracy of our modelling approach by evaluating the throughput of an individual UE under varying channel conditions and congestion levels. Specifically, we assume that n UEs are randomly deployed in a cell and they are static. Then, an additional *tagged* UE is deployed at a known distance d from the center of the cell. Figure 4 shows a comparison between the model predictions and the simulation results for the MAC-level downlink throughput of the tagged UE versus its distance from the eNB and for different n values. As a first important consideration, Figure 4 proves that our analysis provides a very accurate approximation of the MAC-level throughput independently of the fading intensity. Furthermore, the results confirm that increasing the total number of UEs in the cell has the effect of reducing in a proportional manner the throughput of the tagged UE. This is due to the fact that RR is a channel-unaware scheduler that performs fair sharing of time resources among UEs. Finally, it is also important to point out that the IR-HARQ mechanism is very effective in improving error correction. As discussed in Section III the modulation and coding scheme are selected in such a way that the error probability is well below 10%. As a matter of fact, our results (not shown here) indicate that the actual error probability for the first transmission attempt is below 5% up to a distance of 1000 meters, and the probability to perform more than one retransmission is negligible.

C. Results with fixed CQI

In this second set of simulations we consider the same network scenario as in Section V-B. However, independently of the UE position the CQI feedback is assumed constant. Figure 5 shows a comparison between the model predictions and the simulation results for the MAC-level downlink throughput of the tagged UE versus its distance from the eNB when $n = 12$. Note that twelve is the maximum number of UEs such that $n_{RR} = 8$. In other words, with $n = 12$ all the UEs are scheduled during one TTI and 8 HARQ process needs to be managed simultaneously. As expected a high value for the

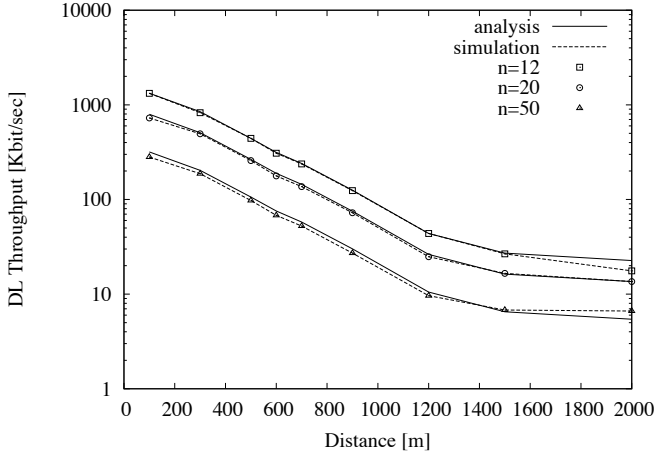


Fig. 4. *Adaptive CQI*: Comparison of analytical and simulation results for the MAC-level throughput of a tagged UE versus its distance from the eNB and the total number of UEs in the cell.

reported CQI results into the use of a very efficient MCS, which provides a high data rate at the cost of high vulnerability to channel fading. Consequently, the tagged UE obtains a high throughput when close to the eNB, but the throughput performance rapidly degrades as it gets farther from the eNB. On the contrary a low reported CQI provides a much more stable throughput performance due to the robustness of the selected MCS. However, in this case we must compromise between robustness and efficiency. To quantify this trade-off in Figure 6 we show the probability to discard a packet as given by formula (20) in the same network configurations of Figure 5. Interestingly, we can observe that there is a critical distance after which the $P_d(k)$ probability rapidly increases up to the value of one. Hence, after this critical distance even the IR-HARQ scheme becomes incapable of controlling the error probability. Furthermore, results in Figure 6(b) indicate that our model underestimates the actual $P_d(k)$ for less reliable MCSs, while results in Figure 6(a) indicate that our model overestimates the actual P_d for robust MCS.

VI. CONCLUSIONS

In this paper we have developed an analytical framework to estimate the MAC-level downlink throughput in a LTE system, which carefully models practical mechanisms of the MAC layer of the LTE technology. As a matter of fact, LTE systems achieve high communication reliability by adopting a combination of link adaptation and error correction schemes. This study is a first attempt to tackle the complexity of modelling the interplay between these mechanisms and to obtain a realistic evaluation of the throughput performance at the MAC level. Our results confirm that the IR-HARQ mechanism is very effective in improving error correction. However, the effectiveness of the IR-HARQ scheme depends on the appropriate selection of the modulation and coding scheme of the first transmission attempt.

Although there is still the need for detailed system-level simulations, we believe that the proposed analytical approach

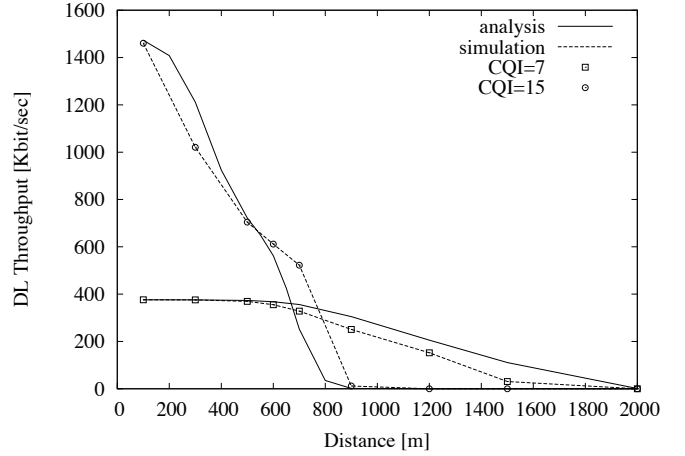


Fig. 5. *Fixed CQI*: comparison of analytical and simulation results for the MAC-level throughput of a tagged UE versus its distance from the eNB for different CQI values and $n = 12$.

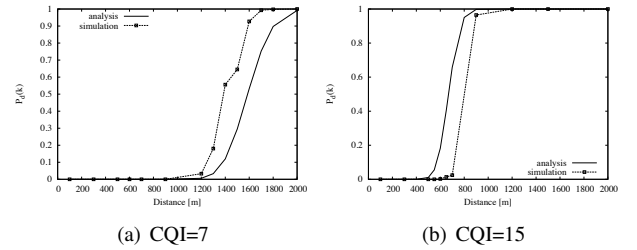


Fig. 6. *Fixed CQI*: comparison of analytical and simulation results for the probability of discarding a packet for a tagged UE versus its distance from the eNB for different CQI values and $n = 12$.

will be useful to an LTE system designer for dimensioning the LTE system and configuring the optimal set of radio MAC parameters. Furthermore, there are several possible avenues of future research. Specifically, we are working to include in the analysis channel-aware scheduling strategies (e.g., the proportional fair scheduler), and additional radio resource management aspects of LTE (e.g., the ARQ retransmission functionality of the RLC layer). Finally, we also intend to generalise our analysis to incorporate correlation in frequency and space.

APPENDIX

Proof of Theorem 2: By definition, the CDF of the spectral efficiency in equation (1) is computed as

$$\begin{aligned}
 F_{\eta}(x; i, k) &= Pr\{\eta_{i,k} \leq x\} \\
 &= Pr\left\{\log_2\left(1 + \frac{\gamma_{i,k}}{\Gamma}\right) \leq x\right\} \\
 &= Pr\left\{1 + \frac{\gamma_{i,k}}{\Gamma} \leq 2^x\right\} \\
 &= Pr\{\gamma_{i,k} \leq \Gamma(2^x - 1)\}.
 \end{aligned} \tag{21}$$

Given that $\gamma_{i,k} \sim Exp(\lambda_k)$ it follows that

$$F_{\eta}(x; i, k) = \int_0^{\Gamma(2^x - 1)} \lambda_k e^{-\lambda_k y} dy. \tag{22}$$

Solving the integral in (22) provides the formula in (6).

Proof of Claim 1: Following the procedure described in [21] to derive the CQI index from the corresponding spectral efficiency we can write that

$$\begin{aligned} g_{i,k}[j] &= Pr\{C_{i,k} = j\} \\ &= Pr\{S_j \leq \eta_{i,k} \leq S_{j+1}\} \\ &= Pr\{\eta_{i,k} \leq S_{j+1}\} - Pr\{\eta_{i,k} \leq S_j\}. \end{aligned} \quad (23)$$

By substituting (6) in (23) we obtain (7), and this concludes the proof.

Proof of Claim 2: The widband CQI for the k^{th} UE is defined as follows

$$\widehat{C}_k = \left\lfloor \frac{1}{q} \sum_{i=1}^q C_k \right\rfloor. \quad (24)$$

Let us denote with Z_k the sum of the CQI values for each RB, i.e. $Z_k = \sum_{i=1}^q C_{i,k}$. Note that Z_k is a discrete random variable and $Z_k \in [0, qL]$ since $C_{i,k} \in [0, L]$. Let $z_k[j]$ be the PMF of Z_k . It is a well-known results that the pdf of the sum of i.i.d. RVs is the convolution of the pdf of the constituent RVs. Thus, it follows that $z_k[j] = g_k^{(q)}[j]$, i.e. the q -fold convolution of $g_{i,k}[j]$. Now we can write that

$$\begin{aligned} g_k[j] &= Pr\{\widehat{C}_k = j\} \\ &= Pr\left\{\left\lfloor \frac{1}{q} Z_k \right\rfloor = j\right\} \\ &= Pr\{qj \leq Z_k < q(j+1) - 1\} \\ &= \sum_{l=qj}^{q(j+1)-1} Pr\{Z_k = l\}. \end{aligned} \quad (25)$$

By substituting $g_k^{(q)}[l]$ in (25) we obtain (8), and this concludes the proof.

Proof of Claim 3: By definition $H_{i,m,k} = \mu_m[v]$ is $Q_{m,v} \leq \gamma_{i,k} \leq Q_{m,v+1}$. The probability $h_{i,m,k}[v]$ that $H_{i,m,k}$ is equal to $\mu_m[v]$ is thus

$$\begin{aligned} h_{i,m,k}[v] &= Pr\{Q_{m,v} \leq \gamma_{i,k} \leq Q_{m,v+1}\} \\ &= \int_{Q_{v,m}}^{Q_{(v+1),m}} f_\gamma(x; i, k) dx. \end{aligned} \quad (26)$$

Solving the integral in (26) provides the formula in (12).

Proof of Claim 4: Let us denote with $Z_{m,k} = \sum_{i \in \Omega(k)} H_{i,m,k}$. Thus, $Z_{m,k}$ is a discrete random variable which takes on any value equal to the sum of all possible $\omega(k)$ -combinations of the V elements in the set \mathcal{V}_m . Then, it holds that

$$\widehat{H}_{m,k} = \frac{Z_{m,k}}{\omega(k)}. \quad (27)$$

However, the analysis is complicated by the fact that the MMIB-to-BLER mapping is defined only over the discrete set \mathcal{V}_m . Thus, approximations are needed to map $Z_{m,k}/\omega(k)$ onto one of the $\mu_m[v]$ values. First of all, we observe that the PMF of $Z_{m,k}$ is the convolution of the PMFs of the individual constituents because $H_{i,m,k}$ RVs are independent. Now, let us define with Φ_v the set of all attainable values for $Z_{m,k}$ that are

between $\omega(k)\mu_m[v]$ and $\omega(k)\mu_m[v+1]$. Then, we approximate the probability $h_{m,k}[v]$ that $H_{i,m,k}$ is equal to $\mu_m[v]$ as

$$\begin{aligned} h_{m,k}[v] &= Pr\{\widehat{H}_{m,k} = \mu_m\} \\ &\approx Pr\{\mu_m[v] \leq \frac{Z_{m,k}}{\omega(k)} \leq \mu_m[v+1]\} \\ &= \sum_{l \in \Phi_v} h_{i,m,k}^{(\omega(k))}[l]. \end{aligned} \quad (28)$$

In other words, in formula (28) the discrete MMIB value is associated to a range of values for the variable $Z_{m,k}$.

Proof of Theorem 1: By exploiting the law of total probability we can write that

$$\begin{aligned} E[TB|Succ]_k &= \sum_{j=0}^L E[TB|Succ, MCS=m(j)] \\ &\quad Pr\{Succ|MCS=m(j)\} Pr\{\widehat{C}_k = j\} \\ &= \sum_{j=0}^L TBS(m(j), b) [1 - P_d(m, k)] g_k[j], \end{aligned} \quad (29)$$

where \widehat{C}_k is the CQI value reported by the k^{th} UE. Similarly, by considering the HARQ protocol behaviours

$$\begin{aligned} E[T_{ARQ}]_k &= \\ &= \sum_{j=0}^L E[T_{ARQ}|MCS=m(j)]_k Pr\{\widehat{C}_k = j\} \\ &= \sum_{j=0}^L \left[\sum_{r=0}^{r_{max}} E[T_{ARQ}|MCS=m(j), n_{rtx} = r]_k \right. \\ &\quad \left. Pr\{n_{rtx} = r|MCS=m(j), \} \right] Pr\{\widehat{C}_k = j\}. \end{aligned} \quad (30)$$

By considering the HARQ timing as illustrated in Figure 2 it immediately follows that

$$E[T_{ARQ}|MCS=m(j), n_{rtx} = r]_k = (1+r)(1+\tau_{ARQ}). \quad (31)$$

Finally, by substituting (31) in (30) we obtain (5b).

REFERENCES

- [1] M. Rinne and O. Tirkkonen, "LTE, the radio technology path towards 4G," *Computer Communications*, vol. 33, pp. 1894–1906, 2010.
- [2] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.
- [3] "Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E UTRAN): Definitions," TS 32.450, Version 9.1.0 Release 9, June 2010.
- [4] S. Donthi and N. Mehta, "An Accurate Model for EESM and its Application to Analysis of CQI Feedback Schemes and Scheduling in LTE," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3436–3448, October 2011.
- [5] N. Varanese, J. Vicario, and U. Spagnolini, "On the Asymptotic Throughput of OFDMA Systems with Best-M CQI Feedback," *IEEE Wireless Communications Letters*, vol. 1, no. 3, pp. 145–148, June 2012.

- [6] J. Leinonen, J. Hamalainen, and M. Juntti, "Capacity Analysis of Downlink MIMO-OFDMA Resource Allocation with Limited Feedback," *IEEE Transactions on Communications*, vol. 61, no. 1, pp. 120–130, January 2013.
- [7] A. Elnashar and M. El-Saidny, "Looking at LTE in Practice: A Performance Analysis of the LTE System Based on Field Test Results," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 81–92, September 2013.
- [8] V. Buenestado, J. Ruiz-Aviles, M. Toril, S. Luna-Ramirez, and A. Mendo, "Analysis of Throughput Performance Statistics for Benchmarking LTE Networks," *IEEE Communications Letters*, vol. 18, no. 9, pp. 1607–1610, September 2014.
- [9] Z. He and F. Zhao, "Performance of HARQ With AMC Schemes in LTE Downlink," in *Proc. of IEEE CMC'10*, 2010, pp. 250–254.
- [10] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 506–522, February 2005.
- [11] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, "Link performance models for system level simulations of broadband radio access systems," in *Proc. of IEEE PIMRC'05*, 2005.
- [12] M.-X. Chang and Y.-T. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Transactions on Communications*, vol. 1, no. 4, pp. 721–732, October 2002.
- [13] P. Tan and N. Beaulieu, "Effect of channel estimation error on bit error probability in ofdm systems over rayleigh and rician fading channels," *IEEE Transactions on Communications*, vol. 56, no. 4, pp. 675–685, April 2008.
- [14] P. Banelli, "Theoretical analysis and performance of ofdm signals in nonlinear fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 2, pp. 284–293, March 2003.
- [15] H. Bolcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 225–234, February 2002.
- [16] E. Lo, P. Chan, V. Lau, R. Cheng, K. Letaief, R. Murch, and W. Mow, "Adaptive Resource Allocation and Capacity Comparison of Downlink Multiuser MIMO-MC-CDMA and MIMO-OFDMA," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 1083–1093, March 2007.
- [17] H. Li, G. Ru, Y. Kim, and L. H., "Ofdma capacity analysis in mimo channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4438–4446, September 2010.
- [18] S. Donthi and N. Mehta, "Joint Performance Analysis of Channel Quality Indicator Feedback Schemes and Frequency-Domain Scheduling for LTE," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3096–3109, September 2011.
- [19] Z. Lin, P. Xiao, and B. Vucetic, "SINR distribution for LTE downlink multiuser MIMO systems," in *Proc. of IEEE ICASSP'09*, April 2009, pp. 2833–2836.
- [20] H. Song, R. Kwan, and J. Zhang, "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.
- [21] 3GPP, "Conveying MCS and TB size via PDCCH," R1-081483, September 2010.
- [22] H. Seo and B. Lee, "Proportional-fair power allocation with CDF-based scheduling for fair and efficient multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 978–983, May 2006.
- [23] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA): Physical layer procedures (Release 9)," TS 36.213 V9.3.0, September 2010.
- [24] IEEE 802.16 Broadband Wireless Access Working Group, "Evaluation Methodology for P802.16m-Advanced Air Interface," IEEE 802.16m-08/004r2, 2008.
- [25] M. Mezzavilla, M. Miozzo, M. Rossi, N. Baldo, and M. Zorzi, "A Lightweight and Accurate Link Abstraction Model for the Simulation of LTE Networks in Ns-3," in *Proc. of ACM MSWiM '12*, 2012, pp. 55–60.
- [26] WiMAX Forum, "WiMAX System Evaluation Methodology," V2.1, July 2008.
- [27] COST Action 231, "Digital mobile radio future generation systems," Final Report - EUR 18957, 1999.