Abstract—Incremental redundancy hybrid automatic repeat request (IR-HARQ) scheme used in recent wireless communication standards such as 3GPP LTE provides higher benefits in terms of system capacity and robustness. To map these benefits into system level evaluations is of critical importance. Traditional link abstraction techniques are usually designed for a fixed bandwidth and do not present a generic solution for the variable bandwidth assignment to the users. Therefore in this paper we propose a generic methodology for modeling the IR-HARQ link performance in LTE for the system level simulators. The proposed scheme allows for the arbitrary bandwidth assignments while at the same time reduces the storage requirement for the complex operations of IR-HARQ link abstraction. We apply the proposed methodology to a wide variety of the modulation and coding schemes (MCS) in LTE for a highly frequency selective channel and show that the proposed model provides accurate results.

I. INTRODUCTION

Modeling the performance of hybrid automatic repeat request (HARQ) for the link abstraction has gained quite an attention from the research community for its use in the system level simulators to reduce the simulation time and the computational complexity of the physical layer. However modeling the information combining from several HARQ rounds can be complex for the case of incremental redundancy (IR) HARQ where new additional redundancy bits are transmitted along with some of the previously transmitted bits in each of the HARQ round. In [1] the authors have presented method of simple HARQ (i.e. no IR-HARQ) modeling in the OFDM systems using exponential effective SINR mapping (EESM) and mutual-information based effective SINR mapping (MIESM). In [2] the authors have presented a recursive method using EESM for HARQ modeling of an OFDM-based system. An interesting work was presented in [3] where the authors showed that a reduced set of reference curves can be used for modeling the link performance of HARQ in general. But their requirement of placing the mutual information of the received bits in to the virtual circular buffer (replicating the process of rate-matching) is a computational overhead for the link abstraction. The concept of accumulated mutual information (AMI) for the HARQ link abstraction is presented in [4]. In [5] the authors have used the concept of AMI from [4] for the HARQ link abstraction in 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE). The limiting point in most of these approaches is that they do not address the case of a variable bandwidth assignment to the users and presented techniques are only implementable for some specific bandwidth assignment. Further these techniques (except [3]) require a very large number of reference curves for modeling the performance of HARQ which increases the storage requirement for link abstraction. However this paper presents a simple, robust and effective methodology for modeling the performance of IR-HARQ in the case of an arbitrary bandwidth assignment and also it uses a very reduced set of reference curves. The further discussions in this paper are in the context of LTE but the methodology can be applied to other cellular standards as well.

In the rest of the paper Section II presents the system model in the framework of LTE, Section III presents the bit wise mutual information combining for LTE IR-HARQ in frequency flat channels and explains how these can be used for the frequency selective channels. Section IV presents the proposed link abstraction methodology for IR-HARQ link abstraction in frequency selective channels and for variable bandwidth assignments. Also Section IV explains how the reference curves for the link abstraction can be reduced. Section V presents the overall summary of the proposed link abstraction in bullet points and Section VI presents the results of IR-HARQ link abstraction by using the proposed methodology. Finally Section VII concludes the paper with a short summary of the contribution of this work.

II. SYSTEM MODEL

In LTE the bandwidth is assigned to a user in terms of physical resource blocks (PRBs) for the duration of one subframe (1 millisecond). In each of the PRB and subframe there are 12 subcarriers (separated by 15 KHz) and either 14 or 12 OFDM symbols depending on the normal or extended cyclic prefix respectively. This gives rise to a two-dimensional (time, frequency) grid in which each block of the grid is referred as resource element and is capable of transmitting a distinct information symbol. In this paper we consider the baseline configuration of LTE, i.e., transmission mode 1, where a single antenna eNodeB (acronym of base station in LTE) communicates with a single antenna UE. The received signal at the UE for the duration of a subframe is given by,

$$y = Hx + z$$  \hspace{1cm} (1)$$

where $y$ is the $J \times 1$ received vector, $J$ is the number of resource elements in a subframe which can be calculated (ignoring the control and reference symbols) as $J = 12 \times NB_{RB} \times SYM$, where 12 are the frequency subcarriers in a
PRB, \( N_{RB} \) is the number of assigned PRBs to the UE and \( SYM \in \{14,12\} \) is the number of symbols in one subframe. \( \mathbf{H} \) is a \( J \times J \) diagonal matrix with zeros on the non-diagonal indices whereas the diagonal values \( [\mathbf{H}]_{jj} \in \mathbb{C}^{1 \times 1} \) symbolize the flat Rayleigh fading SISO channel of \( j \)-th resource element from the eNodeB to the UE. It can be modeled as independent identically distributed (i.i.d) zero mean circularly symmetric complex Gaussian (ZMCSCG) random variable with the variance of 0.5 per dimension. \( \mathbf{z} = J \times 1 \) ZMCSCG white noise of variance \( N_0 \) at the UE. \( \mathbf{x} \) is a \( J \times 1 \) vector of complex symbols which assumed to be independent and of variances \( \sigma^2 \) belonging to a discrete M-QAM constellations where \( M \in \{4,16,64\} \).

III. IR-HARQ IN LTE

The received signal is demodulated and decoded at the UE. For the signal to be decoded correctly in the frequency flat channel it is required that the capacity of the channel is greater than the data rate (Shannon’s channel coding theorem). However for the discrete QAM constellation and the capacity achieving turbo codes the probability of error-free decoding depends on the average bitwise mutual information (BMI) at the input of the turbo decoder [6]. The decoding shall only be successful when

\[
\text{BMI} > r_c
\]

where \( r_c \) is the channel code rate. However if (2) is not satisfied after the initial HARQ round, the UE sends a not-acknowledgment (NACK) signal to the eNodeB using physical uplink control channel (PUCCH). On receiving the NACK from the UE, eNodeB retransmits the packet using the next redundancy version\(^1\). Upon receiving the retransmission, the receiver combines the signals from both of the rounds (i.e., adds log-likelihood ratios (LLR) of the bits which are repeated in both rounds and updates the LLRs of the newly received bits) and tries to decode again. For the decoding to be successful (2) has to be satisfied and the overall BMI after two rounds can be given as [4][5],

\[
\text{BMI} = (N_Q - N_1) \cdot \text{BMI}(\gamma_1) + (N_Q - N_2) \cdot \text{BMI}(\gamma_2) + (N_1 + N_2 - N_Q) \cdot \text{BMI}(\gamma_1 + \gamma_2)
\]

Where \( N_1, N_2, N_Q \) represent the number of transmitted bits in the first round, in the second round and total number of received bits in the buffer after 2 rounds respectively. \( \gamma_1, \gamma_2 \) represents the signal to interference plus noise ratio (SINR) of the frequency flat channel experienced during round 1 and 2 respectively. The first term in (3) represents the bits which are exclusively received in the first round, the second term represents the bits which are exclusively received in the second round and the last term represents the bits which are repeated in both of the rounds. If the packet is still decoded in error after round 2, then UE sends a NACK signal to the eNodeB again which in turn retransmits the data using the next redundancy version. This process can go on until fourth HARQ round and if the packet is still decoded in error then it is finally discarded and no more retransmissions are performed. The generalized expression for (3) can be written as,

\[
\text{BMI} = \sum_{t=1}^{T} (N_Q - N_t) \cdot \text{BMI}(\gamma_t) + \left( \sum_{t=1}^{T} (N_t - N_Q) \right) \cdot \text{BMI} \left( \sum_{t=1}^{T} (\gamma_t) \right)
\]

Where \( T \in \{1, 2, 3, 4\} \) represents the HARQ rounds and \( N_Q \) represents the total number of the received bits after \( T \) HARQ rounds. It is important to note here that (4) is valid only when the channel is frequency flat during each HARQ round but can vary from one round to another. Whereas for the case when the channel is frequency selective during each HARQ round it is not possible to write an analytical expression like (4) because of the requirement to track the resource elements on which a certain bit has traveled in all of the HARQ rounds. This problem arises due to the structure of the contention free random interleaver used in LTE. However for the successful decoding, the condition presented in (2) has to be satisfied irrespective of the channel model used thus motivating our efforts towards the modeling of IR-HARQ link abstraction for the frequency selective channels.

IV. PROPOSED LINK ABSTRACTION FOR IR-HARQ

Link abstraction casts a multi-state frequency selective channel on to an equivalent single-state frequency flat channel in two steps, 1) by compressing the SINRs corresponding to the multiple channel qualities into an effective SINR value and 2) by mapping this effective SINR value on to the pre-computed performance curve to obtain the link quality in terms of block error rate (BLER) [9][10]. For the compression, non-linear functions (i.e., exponential, logarithmic, mutual information etc.) are used whereas for the mapping, additive white Gaussian noise (AWGN) performance curves for each type of modulation and coding scheme (MCS) are used. The important point in the first step is to choose the compression function and the correct criterion for combining the multiple channel qualities from multiple transmission rounds. For this purpose we shall use the mutual information based effective SINR mapping as it does not require the calibration of adjustment factors [11][12]. For the compression of multiple channel qualities from multiple HARQ rounds, we store them in each of the round and then take an overall average of accumulated qualities to calculate \( \gamma_{eff} \) such as,

\[
\gamma_{eff} = \beta_1 \left[ \frac{1}{J \times T} \sum_{j=1}^{J} \sum_{t=1}^{T} \left( \frac{\gamma_{t,j}}{\beta_2} \right) \right],
\]

where \( \beta_1 \) and \( \beta_2 \) are the calibration factors. For MIESM these are normally equal to unity thus these will be neglected in the rest of the paper. The mutual information for the received

\(^1\)LTE has four different redundancy versions (RV) which are created by reading out the bits from a different starting point in the virtual circular buffer. The details can be found in [7][8]
signal $y_j$ can be calculated as,
\[ I_M(\gamma_{t,j}) = \log_2 M - \frac{1}{M} \sum_{t, j \in x} \mathcal{E}_{x_{t,j}} \log_2 \left( \frac{\sum_{t, j \in x} \exp \left[ - |\gamma_{t,j}(x_{t,j} - x'_{t,j}) + z_{t,j}|^2 \right]}{\exp \left[ - |z_{t,j}|^2 \right]} \right) \]

Then the averaged bitwise mutual information can be calculated as,
\[ \text{BMI} = \frac{I_M(\gamma_{eff})}{\log_2 M} \]

where $\gamma_{t,j}$ is the received SINR of received signal $y_j$ on the $j$-th resource element in $t$-th HARQ round and BMI is the averaged bitwise mutual information. $\chi$ is the set of the QAM constellation with $|\chi| = M$ points and $z_{t,j} \in \mathcal{CN}(0,1)$.

After calculating the averaged BMI for all of the $J$ resource elements, the condition of (2) is applied to check if the decoding will be successful or not such as,
\[ \text{BMI} > r_Q(MCS) \]

where $r_Q(MCS)$ represents the effective code rate after $T$ number of rounds. If the condition in (9) is not met then the decoding is not successful and processing for the next HARQ round is performed. In the case of successful decoding (i.e., (9) holds true) the second step of the link abstraction can be performed to obtain the exact BLER from the AWGN curves corresponding to the used MCS and the effective code rate $r_Q$.

A. Effective Code Rate

The effective code rate after $T$ HARQ rounds must be calculated very carefully because it should be able to reflect the effect of the repeated bits and the new bits after every HARQ round. For the simplicity of the link abstraction we propose to calculate the effective code rate as if there was no bit repetition in the consecutive HARQ rounds and then correct the effective code rate for the overlapped bits. The overall code rate after $T$ rounds can be calculated as,
\[ r^\dagger_Q = \frac{\prod_{t=1}^{T} r_t}{\sum_{t=1}^{T} \prod_{s=1, s \neq t}^{T} r_s} \]

where $r_t$ represents the channel code rate used in $t$ round. For example after the third HARQ round, i.e., $t = 3$ we can obtain $r^\dagger_Q$ using equation (10) as
\[ r^\dagger_Q = \frac{r_1 \cdot r_2 \cdot r_3}{r_2 \cdot r_3 + r_1 \cdot r_3 + r_1 \cdot r_2} \]

However $r^\dagger_Q$ does not account for the bits which are repeated in the consecutive HARQ rounds and can not be used for link abstraction as it will over estimate the resulting BLER. This is due to the increased diversity of the HARQ round over a new channel in each HARQ round. Therefore we need to adjust the effective code rate considering the number of coded repeated bits in each of the HARQ round. This can be accomplished by counting the number of new and repeated bits in each of the consecutive HARQ round for each of the MCS. This information can be obtained from the rate matching algorithm of LTE and then this information is to be stored in the link abstraction module for the 29 different MCS and 4 HARQ rounds in each of the MCS. For a fixed number of resource block allocation to the users this approach does not pose a huge burden on the link abstraction module. However for the variable resource block assignments, the transport block size changes in LTE and thus this information about the repeated and new bits changes as well. Which means that adjusting the $r^\dagger_Q$ for an arbitrary resource block assignment requires a huge effort during the pre-processing of the link abstraction and also makes the link abstraction more complex in practice. To avoid this problem we propose rather a simple approach to not use the number of repeated and new bits for the correction of $r^\dagger_Q$ but the percentage of the overlapping region.

The difference in using the percentage and the actual numbers is that the proportion of the overlapped region is independent of the size of the TBS making it independent of the resource block assignment. The proportion of the overlapping region can be calculated as,
\[ \xi(MCS, t) = \frac{\sum_{t=1}^{T} (N_t) - N_Q}{\sum_{t=1}^{T} (N_t)} \]

To show that $\xi(MCS, t)$ is independent of $NB_{RB}$ and it is only a function of the MCS and $t$, we calculated $\xi(MCS, t)$ over a wide range of resource block assignment for MCS 0-27 with the help of LTE rate matching. In Figure 1 we plot the value of $\xi(MCS, t)$ for the different resource block assignments and for two different MCS. It can be seen that $\xi(MCS, t)$ does not changes for different resource block assignments and remains constant. Further in Table I we show mean and variance of $\xi(MCS, t)$ for the second, third and the fourth HARQ rounds for MCS 0 to 27 over PRB assignment from zero to twenty five with the step size of two PRBs\(^2\). Table

\(^2\)We show results till 25 PRBs as 5MHz is the maximum allowed bandwidth in our LTE link level simulator.

![Fig. 1. $\xi(MCS, t)$ in the second, third and fourth HARQ round for MCS 0 and 9 for different PRB Assignment - It is clear that for different bandwidth assignment the value of $\xi(MCS, t)$ does not changes.](image-url)
and this $r_Q$ is appropriate to be used in (9) for the link abstraction.

**B. Reference Performance Curves**

The second step in the link abstraction is to map the effective SINR value on to a pre-computed additive white Gaussian noise (AWGN) performance curves for each type of MCS. These AWGN performance curves reflect mainly the turbo codes performance which is highly dependent on the code block size (known as transport block size TBS in LTE). In LTE the TBS is a function of the bandwidth assigned to a specific user which is assigned in terms of PRBs where each PRB consists of 180 KHz of bandwidth. LTE has the operating bandwidth from 1.4 MHz to 20 MHz i.e. from 6 PRBs to 100 PRBs and there are 29 different MCS available in the LTE. A user can be assigned minimum of 1 PRB and maximum of 100 PRBs which means that for the link abstraction it is required to have 100x29=2900 different AWGN performance curves for all possible combinations of PRB assignments and MCS. For the IR-HARQ retransmissions this requirement is further increased to accomodate the possible 4 HARQ rounds, i.e., the number of required performance curves becomes 4x2900=11600. Please note, this number corresponds to the case when the same MCS is used for all of the IR-HARQ rounds. And this number will be even higher when a different MCS is allowed to be used for the retransmissions. Clearly this requirement is quite impractical and should be avoided in order to generalize the link abstraction for the variable PRB assignment.

In our observation of the performance curves we found that for the variable PRB assignment, the slope of the AWGN reference curve becomes different, but for the case of same PRB assignment the slope of the performance curve remains constant for all of the MCS. In other words, the reference curves are shifted versions of each other. We exploit this feature to reduce the number of required reference curves. In LTE three different modulations are used: QPSK (MCS 0-9), 16-QAM (MCS 10-16) and 64-QAM (17-29). We calculate the performance curve for each of the mentioned modulation type for the code rate 1/3 (which is referred as the mother code rate $r_m$). Then it is required to calculate the accurate shift for the performance curve with respect to the current MCS. We propose to calculate the proper shift $R$ in the reference curve with the help of effective code rate ($r_Q$) and the mother code rate ($r_m$). This facilitates the link abstraction process as we do not need to know anything which is not already available. The shift in the performance curves translates to the gain/loss in mutual information for the specific modulation scheme.

Since LTE uses bit interleaved coded modulation (BICM), therefore we use BICM-based BMI from (8) for calculating the shift $R$ in dBs as shown in Figure 2.

$$R(dB) = I_M^{-1}(r_m) - I_M^{-1}(r_Q).$$  \(14\)

We use the fact that for the capacity achieving channel codes with long enough code block size, the normalized mutual information and channel code rate can be used interchangeably, therefore, using $r_Q$ and $r_m$ we calculate the difference (or shift) in the performance of both code rates. Then using this difference one can be directly mapped on to another. In our approach we calculate $R$ and apply this to shift the reference curve corresponding to the mother code rate.
So in this manner we reduce the required reference curves from 29 to 3 only for a fixed bandwidth assignment. For the variable bandwidth assignment, the slope of the reference curves is different. Ideally the BLER starts coming down at the same average SNR but then its not an immediate waterfall for low bandwidth assignment. To account for this factor we propose to use an offset $\Delta$ (in dB) at the 10% of the BLER for each type of bandwidth assignment with respect to the maximum bandwidth assignment as shown in Figure 3. Unfortunately for the different possible combinations of code block size and coding rate $\Delta$ has to be calculated numerically from the link level simulator and there is no analytical expression available to calculate it. But good point is that $\Delta$ has to be calculated only once and only for the code rate 1/3. Once it is calculated then the mother code rate for specific bandwidth assignment can be adjusted with respect to the stored (maximum bandwidth assignment) reference curve, such as,

$$r_m = r_{mm} + \Delta (NB_{RB})$$  \hspace{1cm} (15)

where $r_{mm}$ represents the reference curve for the mother code rate corresponding to the maximum number of assigned resource blocks and $\Delta (NB_{RB})$ represents the offset for the $NB_{RB}$ assigned resource blocks. It is clear that by using our proposed approach the required number of AWGN performance curves is reduced from the gigantic count of 11600 (corresponding to all possible combinations of PRB assignment and MCS) to just three which is a significant improvement in state-of-the-art methodologies. This reduces the storage requirement for the link abstraction. Moreover our proposed method can also be used for the scenarios when the retransmission for HARQ rounds is done using a different coding rate.

V. SUMMARY

For the link abstraction of IR-HARQ for variable bandwidth assignment it is required to have

- 3 AWGN mother code reference curves corresponding to the maximum bandwidth assignment,
- Table I which has $\xi(MCS,t)$ values for all HARQ rounds,
- A table of $\Delta(NB_{RB})$ for different number of PRB assignment.

Then the following steps,

- Generate the frequency selective channel for assigned bandwidth (for each HARQ round)
- Determine the received SINR across each of the resource element and across each of HARQ round
- Use (5) to calculate $\gamma_{eff}$ using the received SINR from all HARQ rounds
- Use (4) to calculate the averaged BMI across all of the resource elements
- Calculate $\tau_Q$ using (13)
- Use BMI value to check the condition of (9), if decoding fails then go back to step 1 for another HARQ round
- if the $NB_{RB}$ is different than $\max(NB_{RB})$ then apply $\Delta(NB_{RB})$ to $r_m$
- Calculate $R$ using (14) and add it to $r_m$
- Determine BLER corresponding to $\gamma_{eff}$ from $r_m$ reference curve

VI. RESULTS

For the validation of our proposed scheme, we used Eurecom’s OpenAirInterface link level simulator which implements 3GPP LTE Release 8.6 physical layer [13], [14], [15] with 5 MHz bandwidth and 25 physical resource blocks (PRB). We show the results for a highly frequency selective Rayleigh channel model with 8-taps and for the bandwidth assignment of 5 MHz. The simulations are performed for a very large number of different channel realizations and for a very large number of noise realizations during each of the channel realization. For HARQ we calculate the BLER based on all of the previous transmissions and for each retransmission we generate a new channel to provide the channel diversity. The reference curves corresponding to the mother code rate are shown in Figure 4. These curves correspond to the case when all of the 25 available PRBs are assigned to a user. Then we apply the link abstraction as is summarized in Section V. Figure 5 shows the shift of mother code rate reference curve for different HARQ rounds for MCS 16 which corresponds to the 16 QAM modulation in LTE. The solid line curve with stars corresponds to the mother code rate $r_m$ and the solid lines with circles are the shifted performance curves for each round. The points around the solid lines are the abstracted experimental BLER points from the link simulator. We see from the results that the shifts for each round, i.e., $R_t$, $t = 1, \ldots , 4$ are very well calculated. Further in Figure 6 we show the mean squared error (MSE) at 10% BLER points of Rayleigh channel model with respect to the shifted AWGN curves for a wide variety of MCS. It can be seen that the MSE is generally very small but it is slightly higher for the last round in some of the cases. However if we wish to reduce

\footnote{http://www.openairinterface.org/}
the MSE further then it is only possible by performing the optimization on the calibration factors of (5).

Fig. 4. Reference AWGN curves for LTE Systems using QPSK, 16QAM and 64QAM Modulations

Fig. 5. LTE MIESM IR HARQ MCS 16 (16 QAM) For All Four Transmissions

Fig. 6. Mean Squared Error (MSE) at 10% BLER points of Rayleigh channel model with respect to the shifted AWGN curves for different MCS in LTE after 4 HARQ rounds

VII. CONCLUSION
We have presented a novel IR-HARQ link abstraction methodology for the variable bandwidth assignment in LTE which is simple, robust and requires low storage. We showed that if our approach is used then some of the important factors required for the link abstraction become independent of the bandwidth assignment. Also we showed that how the required number of reference curves for the link abstraction can be significantly reduced from 11600 to 3 with a very small extra apriori effort. In other words our proposed model is more robust and practical. Further we showed with results that the proposed method is very well designed to model the performance of IR-HARQ in LTE and provides very accurate shifts for the performance curves.

ACKNOWLEDGMENT
Eurecom’s Research is supported in part by its industrial partners: Swisscom, Thales, SFR, Orange France, ST Ericsson, SAP, BMW Group, Cisco, Monaco Telecom, and Symantec. This work is also funded in part by the French project LTENOW and by the European project Newcom#

REFERENCES