Abstract—This paper presents a link error prediction model capable of accurately predicting the block error ratio in OFDM systems with hybrid ARQ. The joint modeling of the channel coding and the retransmissions solves many of the issues with present models. Our model is based on accumulated conditional mutual information, mutual information effective SNR averaging, and simulated AWGN performance curves. As a consequence, our model is fading-insensitive and takes into account the non-ideality of rate matching. The model can be applied to any OFDM system that utilizes rate matching as a means to achieve adaptive modulation and coding and HARQ. We demonstrate the application of our model in exemplarily LTE and validate its accuracy by simulation. The whole simulation environment together with the results of this paper are made available for download from our homepage.

I. INTRODUCTION

Modern communication systems use Hybrid Automatic Repeat Request (HARQ) on top of the Physical (PHY) layer to compensate for incorrectly decoded packets. In HARQ systems, an incorrectly received packet is retransmitted and all transmissions of this packet are jointly decoded. Depending on how the retransmitted packets are generated, we distinguish between: (i) Chase Combining (CC), in which each retransmission is a repetition of the original transmission, and (ii) Incremental Redundancy (IR), in which each retransmission adds new bits [2, 3].

In order to simulate complete networks with many links between base stations and mobile terminals, Link Error Prediction (LEP) models that accurately abstract the PHY procedures at low complexity are required [4]. Such models predict the link performance for each individual link and enable (i) the assessment of performance at system level at reduced computational complexity and (ii) at the mobile terminal, the design of better link adaption algorithms.

While LEP models that incorporate HARQ do already exist, the model proposed here addresses issues that are currently not considered by both analytical models (i)-(v) and models based on fitting of simulation results (vi):

(i) Bit repetition due to low coding rates being obtained from a higher rate mother code is currently not considered in analytical models [5, 6]. Since modern systems employ Effective Code Rates (ECRs) much lower than that of the mother code (4.3 times lower than the mother code rate in the case of LTE), the predicted performance significantly deviates from the real performance.

(ii) It has been mentioned in [7, 8] that a packet could be separated in CC and IR parts. However, since the transmitted bits are interleaved and IR combining is done at bit level, there is no simple way of differentiating the two parts, which is why this is currently not done in models. By utilizing repetition ratios in our model, the need of identifying overlapping bits in the signal and calculating their individual SNRs is eliminated.

(iii) In analytical models, OFDM is either typically not accurately modeled or just flat fading is considered [5, 9, 10]. Since OFDM systems exploit frequency diversity by means of channel coding, a model taking the OFDM signal structure into account is required for actual systems.

(iv) Results from models for modulation orders higher than 4-QAM and retransmission numbers higher than one are not provided [5–10]. Since modern systems, such as LTE, support modulations up to 64-QAM and up to three retransmissions, a good LEP model has to be validated also for these cases.

(v) Exponential Effective Signal to Interference and Noise Ratio Mapping (EESM)-based methods, such as [6], require extensive calibration. In our model we instead apply Mutual Information Effective SNR Mapping (MIESM) as SINR averaging method. MIESM has the advantage of not requiring calibration at all and additionally being fading-insensitive [11–13].

(vi) In contrast to our model, in simulation-based models, such as [9], the complexity of tuning the model parameters increases exponentially with the number of retransmissions.

Addressing the above issues, this paper presents a unified LEP model for HARQ that is based on Accumulated Conditional Mutual Information (ACMI) and MIESM. Although the numerical results are presented for LTE, our model can accurately predict the Block Error Ratio (BLER) of any OFDM system that applies rate-matching to adjust the channel coding rate and to implement HARQ.

The remainder of this paper is organized as follows. In Section II we present the information-theoretic view of the different types of HARQ and how these types are applied in a system that uses a one-step rate matching process to both adapt the ECR and generate data packets with different redundancy from the same data. Section III describes how the BLER is
modeled. In Section IV we apply our model to LTE, for which we show performance results in Section V. We conclude the paper in Section VI.

II. HARQ

In this section, the concept of ACMI is presented for the different HARQ schemes. We find that we can model the rate-matching process as a concatenation of an inner code and an outer repetition code. In this section we assume that all symbols transmitted in one frame experience the same SNR. This assumption will not be required in the next section.

Depending on the type of HARQ being employed, the ACMI $I_*$ of the combination of several HARQ blocks is calculated differently [14]. In the case of CC, in which the same bits are retransmitted $M$ times, the effective receive SNR is increased with every retransmission. Thus, the ACMI $I_*$ can be expressed as

$$I_*^{CC} (\gamma) = I \left( \sum_{m=0}^{M} \gamma_m \right), \quad (1)$$

where $\gamma_m$ denotes the SNR of the $m$-th retransmission ($m = \{0, 1, \ldots, M\}$ with $m = 0$ corresponding to the initial transmission) and $I$ is the BICM capacity [15], calculated as explained in (6) below.

For IR, if a parity-priority (IR-PP) scheme is used, only new parity bits are sent in subsequent retransmissions, thereby directly increasing the ACMI $I_*$

$$I_*^{IR-PP} (\gamma) = \sum_{m=0}^{M} I (\gamma_m). \quad (2)$$

If systematic priority (IR-SP) is applied, each retransmission contains the systematic bits of the original message as well as new parity bits. In this case $I_*$ results in a combination of (1) and (2)

$$I_*^{IR-SP} (\gamma) = \frac{D}{C} I \left( \sum_{m=1}^{M} \gamma_m \right) + \left(1 - \frac{D}{C}\right) \sum_{m=1}^{M} I (\gamma_m), \quad (3)$$

where $D$ is the number of data (systematic) bits and $C$ the total number of bits sent in each transmission.

A. HARQ by means of rate-matching

In a one-stage rate-matching process (such as the one applied in LTE) the generation of the different data block versions for the IR HARQ is completely integrated. For a specific target ECR $r_{\text{eff}}$, different values of $m$ result in different bit selections, that is, different HARQ retransmissions (Figure 1).

Because of the finite length of the mother code (rate $r_c$), it is necessary that for $C > D/r_c$, bits are repeated, where $D$ is the number of data bits and $C$ the resulting number of coded bits (Figures 1 and 2). In order to apply our ACMI-based model, we represent the combined received block as a concatenation of IR (represented by a channel code with variable rate $r_m$) and CC (represented by a repetition code with rate $1/N_{\text{rep}}^m$), as shown in Figure 2.

Thus, after the $m$-th retransmission, the original $D$ bits are coded into $C (m + 1)$ bits, which comprise $C_{\text{IR}}^m$ unique bits from the mother code and $C_{\text{RR}}^m$ repeated bits:

$$C (m + 1) = C_{\text{IR}}^m + C_{\text{RR}}^m; \quad C_{\text{IR}}^m = D/r_m, C_{\text{RR}}^m \geq 0. \quad (4)$$

The numeric values of $C_{\text{IR}}^m$ and $C_{\text{RR}}^m$ depend on the specific rate matching process. An example for LTE is given in Section IV.

III. OUTAGE PROBABILITY AND PERFORMANCE METRIC

In this section, the concepts presented in Section II are applied to derive the outage probabilities for AWGN and frequency selective channels.

A. AWGN model

If a capacity-approaching channel code with suitably long blocklength is used, it is well known that the BLER can be approximated by the Mutual Information (MI) outage probability [16–18]. In the case of a system with HARQ, equivalent expressions can be derived by using ACMI. Under this assumption, the outage probability $\varepsilon$ is the probability $\mathbb{P}$ that $I_* < D$. Thus, in the AWGN case we obtain:

$$\varepsilon (\gamma, m, D, C, n) = \mathbb{P} \left[ \frac{C_{\text{IR}}^m}{n} \cdot I_n \left( N_{\text{rep}}^m \cdot \gamma \right) < D \right], \quad (5)$$

where $n$ is the number of bits per symbol of the applied Modulation and Coding Scheme (MCS), and $I_n$ is the BICM capacity for that modulation, expressed as [15]:

$$I_n (\gamma) = n - E_{Y \Gamma}$$

$$\Gamma = 2 \frac{1}{2} \sum_{i=1}^{n} \sum_{b=0}^{1} \sum_{z \in \mathcal{A}_b^i} \sum_{\hat{\alpha} \in \mathcal{A}_b^i} \log_2 \left( \exp \left( -\left| Y - \sqrt{\gamma} (\hat{\alpha} - z) \right|^2 \right) \right)$$

$$A \ldots \text{the set of } 2^n \text{symbols (the complete symbol alphabet)}$$

$$A_b^i \ldots \text{the set of symbols for which bit } i \text{ equals } b$$

$$Y \ldots \text{CN}(0, 1) \ , \quad (6)$$

where $\hat{\alpha}$ cycles through the whole symbol alphabet and $\hat{\alpha}$ just through the set of symbols for which bit $i$ equals $b.$
B. Frequency selective model

By applying the MIESM method it is possible to compress a vector $\gamma$ of SNR values into an AWGN-equivalent effective SNR. This is accomplished in a more robust manner than with other methods, such as EESM, which require a much more precise calibration [11, 19].

The AWGN-equivalent SNR is obtained by stacking the subcarrier SNR row vectors of each transmission in a vector $\gamma$ (7), which is then non-linearly averaged using MIESM (8):

$$\gamma = \left[ \gamma_0, \gamma_1, \cdots, \gamma_M \right]$$  \hspace{1cm} (7)

$$\gamma_{\text{eff}} (\gamma) = I_n^{-1} \left\{ \frac{1}{N_{\text{SCs}}} \sum_{\gamma_i \in \gamma} I_n (\gamma_i) \right\}$$  \hspace{1cm} (8)

Here, $N_{\text{SCs}}$ is the total number of subcarriers and $\gamma_{\text{eff}}$ is the resulting AWGN-equivalent SNR. As in (5), the outage probability $\varepsilon$ can be calculated as:

$$\varepsilon (\gamma, m, D, C, n) = \mathbb{P} \left[ \frac{\gamma}{n} \cdot I_n (N_{\text{rep}} \cdot \gamma_{\text{eff}}) < D \right]$$  \hspace{1cm} (9)

C. Non-ideality of rate matching

In order to consider the non-ideal behaviour of the channel coding and the loss in performance due to the rate matching process, we perform BLER simulations in Additive White Gaussian Noise (AWGN) channels. The thereby obtained AWGN BLER curves are used as an approximation for the outage probability $\varepsilon$:

$$\varepsilon (\gamma_{\text{AWGN}}, m, D, C, n) \approx \text{BLER}_{\text{AWGN}} (r_m, n, \gamma_{\text{AWGN}})$$  \hspace{1cm} (10)

Here, $\gamma_{\text{AWGN}}$ are appropriate SNR values to obtain a BLER curve in the range of approximately $10^{-3}$ to 1. The AWGN SNR $\gamma_{\text{AWGN}} = N_{\text{rep}} \gamma_{\text{eff}}$ is given by the SNR gain $N_{\text{rep}}$ due to the repetition coding and the MIESM averaging of the frequency selective SNR distribution $\gamma$ in (8) to obtain $\gamma_{\text{eff}}$. Note that for applying the approximation (10) we have to precalculate one BLER curve for each combination of modulation order $n$ and outer channel coding rate $r_m$. The actual number of necessary curves depends on the number of MCSs and the rate matching algorithm of the specific system.

IV. APPLICATION TO LTE

In this section, we show how to apply our LEP model in LTE. Equivalently, the same method can be used to apply our model to other OFDM systems. The 3GPP LTE standard [20] defines which MCSs, named Channel Quality Indicators (CQIs), can be used by the LTE system (Table I).

In LTE, the values for $r_m$ cannot be directly calculated from the definition of the rate-matcher [21]. However, by using the implementation of the rate matching process in our LTE link level simulator [22], equivalent puncturing matrices applied to the mother code of rate $r_c = 1/3$ can be obtained for each of the HARQ retransmissions. The outer turbo coding rate $r_m$ and the inner repetition coding rate $1/N_{\text{rep}}$ are then easily calculated from the puncturing matrices. For LTE, $r_m$ and $N_{\text{rep}}$ are shown in Figures 3 and 4.

In Figure 3, we see that the outer turbo coding rate drops as more retransmissions are available at the receiver ($m$ increases). Furthermore, we see that for CQIs smaller than

<table>
<thead>
<tr>
<th>CQI</th>
<th>Modulation</th>
<th>D/ECR</th>
<th>bits/symb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4-QAM</td>
<td>13.13</td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>4-QAM</td>
<td>8.53</td>
<td>0.23</td>
</tr>
<tr>
<td>3</td>
<td>4-QAM</td>
<td>5.31</td>
<td>0.38</td>
</tr>
<tr>
<td>4</td>
<td>4-QAM</td>
<td>3.32</td>
<td>0.60</td>
</tr>
<tr>
<td>5</td>
<td>4-QAM</td>
<td>2.28</td>
<td>0.88</td>
</tr>
<tr>
<td>6</td>
<td>4-QAM</td>
<td>1.70</td>
<td>1.18</td>
</tr>
<tr>
<td>7</td>
<td>16-QAM</td>
<td>2.71</td>
<td>1.48</td>
</tr>
<tr>
<td>8</td>
<td>16-QAM</td>
<td>2.09</td>
<td>1.91</td>
</tr>
<tr>
<td>9</td>
<td>16-QAM</td>
<td>1.66</td>
<td>2.41</td>
</tr>
<tr>
<td>10</td>
<td>64-QAM</td>
<td>2.29</td>
<td>2.75</td>
</tr>
<tr>
<td>11</td>
<td>64-QAM</td>
<td>1.81</td>
<td>3.32</td>
</tr>
<tr>
<td>12</td>
<td>64-QAM</td>
<td>1.54</td>
<td>3.90</td>
</tr>
<tr>
<td>13</td>
<td>64-QAM</td>
<td>1.33</td>
<td>4.52</td>
</tr>
<tr>
<td>14</td>
<td>64-QAM</td>
<td>1.17</td>
<td>5.12</td>
</tr>
<tr>
<td>15</td>
<td>64-QAM</td>
<td>1.08</td>
<td>5.55</td>
</tr>
</tbody>
</table>
TABLE II

<table>
<thead>
<tr>
<th>m</th>
<th>AWGN</th>
<th>PdB</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-QAM</td>
<td>0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>16-QAM</td>
<td>0.04</td>
<td>0.40</td>
</tr>
<tr>
<td>64-QAM</td>
<td>0.07</td>
<td>0.30</td>
</tr>
</tbody>
</table>

four and for higher retransmission numbers \( m \), the outer turbo coding rate saturates at the mother code rate \( r_c = 1/3 \). An ECR lower than 1/3 is achieved in our model by decreasing the inner repetition code rate, thereby increasing the SNR gain \( N_{\text{rep}} \) of the repetition code (see Figure 4).

V. PERFORMANCE

We evaluated the accuracy of our model in extensive single-user link level simulations utilizing our open source Vienna LTE simulator [22]. AWGN and time-correlated ITU Pedestrian-B [23–25] channel models were employed and the simulated BLER was compared to the one predicted by our model.

For each of the 15 LTE MCSs, the BLER curves from the simulation and from our model are compared at 10% BLER, which is the target for the link adaptation. This target is known to lead to near-optimal performance [17]. Figure 5 shows a comparison for CQI 6, with the 10% BLER points marked by a dot. The SNR values of the 10% BLER points for different retransmission numbers are shown for AWGN and ITU Pedestrian-B channels in Figures 6 and 7.

In the case of 4-QAM and 16-QAM transmissions, our modeled performance is very close to the simulated performance. Only in the case of 64-QAM transmission with more than one retransmission, a deviation is observed. Table II shows the average deviation between model and simulation. In order to quantify the impact of the lower accuracy in the case of 64-QAM and more than one retransmission, we have to quantify the probability of these cases occurring in a real system. For doing so, we generated an SINR distribution of a cell (shown in Figure 8) by considering the macroscopic pathloss and antenna gain parameters listed in Table III.

In such a cell, our simulations show that 64-QAM modulation is only used in approximately 11% of all transmissions. In 0.04% of all transmissions, 64-QAM modulation is employed and one retransmission is required for correct decoding. In none of the total of 8.6·10^6 simulated transmissions, 64-QAM-modulated packets required more than one retransmission to be decoded correctly. We thus conclude that the deviation of our model will have a negligible impact on the link error prediction.

Note that our LEP model does not consider the following two specific aspects of LTE:

(i) At very small codeblock sizes the performance of the LTE channel code degrades. Thus the BLER performance becomes dependent on the codeblock size. In our simulations we always used the longest possible codeblock length that fits in a maximum bandwidth of 5 MHz, as the well-established ITU power delay profiles show a periodicity in their frequency correlation properties for larger bandwidths [27].

(ii) In LTE, the maximum codeblock length is 6 144 bits. Larger codeblocks are segmented prior to turbo encoding. This effect could be easily modeled by calculating an effective BLER_{\text{eff}} = (\text{BLER})^s, where \( s \) is the number of codeblock segments. Since we did not consider codeblock segmentation
in our model, we chose for our simulations the largest possible bandwidth $B \leq 5$ MHz that results in a codeblock size $\leq 6144$.

VI. CONCLUSIONS

We introduced a model that is capable of predicting the BLER of OFDM transmissions with HARQ. Our model is based on accumulated conditional mutual information, mutual information effective SNR mapping, and link level simulations to adjust for the non-ideality of the rate-matching process. The model does not require calibration and is fading-insensitive. Complexity-wise, it requires the precalculation of AWGN BLER curves for each accumulated effective code rate and BICM curves for each modulation order. Both can be achieved offline and do not influence runtime complexity. In LTE, this corresponds to the calculation of 26 BLER curves and three BICM curves for the 15 modulation and coding schemes and allows BLER prediction for up to three retransmissions.

The accuracy of our proposed link error prediction model is validated in AWGN/ITU Pedestrian-B channels and an SINR distribution matching that of a cell. All data, tools and scripts are available online in order to allow other researchers to reproduce our results [1].

ACKNOWLEDGMENTS

The authors would like to thank the LTE research group for continuous support and lively discussions. This work has been funded by the Christian Doppler Laboratory for Wireless Technologies for Sustainable Mobility, KATHREIN-Werke KG, and A1 Telekom Austria AG. The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

REFERENCES