Abstract—Hybrid automatic retransmission request (HARQ) is widely adopted in wireless communication systems to improve the link level and system level throughput. Adaptively selecting retransmission based on the received packet can further improve the system performance by reducing the retransmission overhead. To this end, this paper first introduces the concept of instantaneous packet information (IPI). Then, a new metric is proposed to evaluate the quality of received packets based on IPI. The most important feature of the IPI metric is its in-sensitivity to channel fading. Therefore, it can be used in practice, where channel varies significantly. Furthermore, this paper introduces an IPI-metric based HARQ (IPI-HARQ) featuring an on-demand adaptive retransmission. In IPI-HARQ, a receiver is able to evaluate the quality of a received packet and estimate the desired amount of retransmission when the packet has decoding errors. The retransmission adaption in IPI-HARQ includes two aspects: how many symbols and which part of the packet should be retransmitted. The difference between the current IPI and an target IPI is used to estimate how many bits to be retransmitted. The set of bits that can most efficiently increase IPI is selected to be retransmitted and identified by an indicator vector accompanying the NACK feedback. Clustering is used in IPI-HARQ to limit the feedback overhead. The benefit of IPI-HARQ scheme lies in the fine estimation of required retransmission amount and selection of bits with respect to the quality of current received packet. Simulation results show that the proposed IPI-HARQ scheme clearly outperforms conventional HARQ scheme in terms of throughput.

Index Terms—Instantaneous Packet Information, Instantaneous Symbol Information, Retransmission, HARQ, Chase Combining

I. INTRODUCTION

Hybrid automatic retransmission request (HARQ) is an efficient technology to improve the link and system level throughput in wireless communication systems. In HARQ, when a packet has decoding errors, a retransmission request is sent from the receiver and then the retransmitted packet is combined with the those received from previous transmissions of the same packet to improve the likelihood of successful decoding. It can be seen in [1] that a well-designed HARQ scheme will improve the throughput effectively. Two widely used HARQ schemes are based on Chase Combining (HARQ-CC) [2] and Incremental Redundancy (HARQ-IR) [3]. In HARQ-CC, the entire packet is retransmitted regardless of the quality of the received signals. In HARQ-IR, the retransmission amount is also fixed, though it is smaller compared to that in HARQ-CC and equals to a given percentage of original packet length.

It has been recognized that adaptive retransmission can further improve the throughput. There already exist some adaptive HARQ schemes in literature. In [4], adaptation has been used to adjust the modulation level with respect to the quality of retransmission link. However, the quality of the signals received in previous transmissions is not considered. In [5][6], sub-packet cyclic redundancy check (CRC) or forward error code (FEC) were utilized to figure out the corrupted sub-packet precisely and request for retransmission. Sub-packet based schemes can reduce the retransmission amount, but system overhead will be introduced when utilizing sub-packet encoding and CRC bits.

To adaptively select the retransmission based on the quality of already received packet signals, it is necessary to have an instantaneous signal quality metric for packet level evaluation before the decoder. A candidate signal quality metric is the reliability metric proposed in [7][8], which utilizes a posteriori bit probability ratio as the criteria to evaluate the quality of each received bit. Reliability-based (RB) HARQ schemes identify bits of a received packet with low reliability for retransmission. Furthermore, the reliability metric is extended to packet level measurement in [9], where RB-based HARQ is formulated as a threshold based decision making problem. Retransmission amount is estimated by comparing packet level reliability with a given threshold. However, the analysis and simulation results in [9] show that the reliability metric is sensitive to channel variation and thus has a limitation for practical application. Since the channel varies significantly, it is impossible to pre-store all the threshold for infinite channel realizations.

Therefore, a fading in-sensitive metric should be sought after for adaptive retransmission in HARQ. In [10][11][12], a fading-insensitive link quality metric has been proposed, which is based on the mutual information (MI). In [13], the MI metric shows its accuracy in mapping from MI to block error rate (BLER) and its application in retransmission amount prediction according to the MI increment. Given its statistic feature, the MI metric is capable of selecting modulation and coding schemes in fast link adaption for the first transmission.
However, the MI metric cannot follow the instantaneous channel variation and provide the signal quality evaluation for a received packet. Furthermore, the MI metric cannot identify which part of the received signal should be retransmitted.

Inspired by the MI metric, an instantaneous packet information (IPI) based signal quality metric is proposed in this paper. It will be shown in the following that, similar to the MI metric, the IPI metric is also fading-insensitive and thus can be used to evaluate BLER performance in a variety of channel conditions. In addition, this paper introduces an IPI metric based On-Demand adaptive HARQ scheme (IPI-HARQ). In IPI-HARQ, a receiver is able to evaluate the quality of a received packet and estimate the retransmission amount for success decoding when the initial packet transmission has decoding errors. Apart from estimating how many bits should be retransmitted, IPI metric is also used to choose the set of bits, that can improve the IPI most efficiently, to be retransmitted. In an ideal IPI-HARQ scheme, the positions of those bits are identified by an indication vector, which is sent back to the transmitter together in the NACK packet. To reduce the feedback overhead introduced by indication vector and enable practical applications, a bit clustering scheme is used in IPI-HARQ, which limits the feedback overhead with limited performance degradation. Simulation results show that the proposed IPI-HARQ scheme outperforms conventional HARQ schemes in terms of throughput.

The rest of the paper is organized as follows. The system model and the definition of IPI are presented in Section II. In section III, we will introduce the proposed IPI based On-Demand Adaptive HARQ scheme. A detailed comparison between IPI-HARQ and chase combining based conventional HARQ in terms of process and combination scheme is given in this section. Simulation results are given in Section IV, which reveal clear throughput gain of IPI-HARQ scheme. Conclusions are drawn in Section V.

II. INSTANTANEOUS PACKET INFORMATION

To introduce the concept of instantaneous packet information, a wireless packet transmission system is considered. Each packet consists of \( L \) symbols and is generated by encoding \( I \) information bits with a rate \( R \) FEC encoder and mapping the coded bits into \( M \)-order QAM symbols. The length of encoded bits in each packet, \( J \), and the number of symbols in a packet, \( L \), are \( J = I/R \) and \( L = J/\log_2 M \), respectively. Let \( b = [b_1, ..., b_J] \) and \( x = [x_1, ..., x_L] \) denote the bit-level and symbol-level transmission packet.

The received symbol corresponding to the \( l \)-th transmitted symbol \( x_l \) is given by

\[
y_l = h_l x_l + n_l,
\]

where \( h_l \) is channel impulse response and \( n_l \) is complex additive white Gaussian noise with zero mean and variance \( \sigma^2 \). When unit transmission power is assumed, signal to noise ratio (SNR) on the \( l \)-th symbol is expressed as

\[
\gamma_l = \frac{|h_l|^2}{\sigma^2}.
\]

The system model (1) assumes that each received symbol undergoes a flat fading channel, which is the case for broadband systems using orthogonal frequency division multiplex (OFDM) or narrowband systems.

A. Mutual Information and Instantaneous Symbol Information

In a communication system, the mutual information between a transmitted symbol \( x \in \mathcal{X} \) and received symbol \( y \in \mathcal{Y} \) with respect to a given SNR \( \gamma \) in AWGN channel is [14]

\[
I(\mathcal{X}; \mathcal{Y}|\gamma) = \mathbb{E}_Y \left[ \sum_{x} p(x|y, \gamma) \log \frac{p(x|y, \gamma)}{p(x|\gamma)} \right],
\]

where \( \mathcal{X} \) contains all possible QAM symbols, \( \mathcal{Y} \) spans all possible complex channel output, and \( p(x|y, \gamma) \) is the symbol-level likelihood function with respect to received signal \( y \) and SNR \( \gamma \). Defining the instantaneous symbol information (ISI) for a given received symbol \( y \) and SNR \( \gamma \) as

\[
ISI(y, \gamma) = \sum_{x} p(x|y, \gamma) \log \frac{p(x|y, \gamma)}{p(x|\gamma)}.
\]

Hereafter, mutual information can be rewritten as

\[
I(\mathcal{X}; \mathcal{Y}|\gamma) = \mathbb{E}_Y [ISI(y, \gamma)].
\]

From Eq. (5), the classical mutual information is a statistic averaging of ISI for a given SNR \( \gamma \) over all possible channel outputs, while the ISI is the instantiation of mutual information given a particular noise realization. ISI can be interpreted as an indicator on how much information amount is contained in a received symbol.

B. Instantaneous Bit Information and Soft Value

Each \( M \) order QAM symbol \( x \) consists of \( \log_2 M \) bits \( \{b_m|m = 1, ..., \log_2 M\} \). Similar to the definition of MI, an instantaneous bit information (IBI) metric can be defined to quantify the information amount each bit of a received symbol contains, which is given by

\[
IBI_m(y, \gamma) = p(b_m = 0|y, \gamma) \log \frac{p(b_m = 0|y, \gamma)}{p(b_m = 0|\gamma)} + p(b_m = 1|y, \gamma) \log \frac{p(b_m = 1|y, \gamma)}{p(b_m = 1|\gamma)}.
\]

The subscript \( m \) indicates the \( m \)-th bit of the received symbol \( y \). Substitute Eq.(6) into Eq.(4) and replace \( x \) by \( \{b_m\} \), we can get

\[
ISI(y, \gamma) = \sum_{m=1}^{\log_2 M} IBI_m(y, \gamma).
\]

Eq. (7) is intuitive in that the information amount of the received \( M \)-array QAM symbol \( y \) is the sum of the information amount contained in each of its \( M \) bits.

In wireless receivers, \( a \ priori \) log likelihood ratio (LLR) for each bit in a received symbol \( y \) is calculated as the per bit soft value. It is used as the input to decoder. The value and distribution of soft values within a packet have been used in RB-HARQ to evaluate the signal quality and determine retransmission adaptively [7]. However, this metric is fading...
to measure the decoding quality for an received coding block.

C. The Instantaneous Packet Information Metric

The information amount of a received coding block with \( L \) symbols and a total of \( J \) coded bits can be evaluated by the summation all IBIs obtained from each of its \( J \) bits. To normalize against the number of bits in the code block, the sum of IBIs is further averaged over the number of bit, which results in the instantaneous packet information (IPI) as

\[
IPI(y) = \frac{1}{J} \sum_{l=1}^{J} \sum_{m=1}^{L} IBI_{m}(y_{l}, \gamma_{l}),
\]

whose value ranges in \([0, 1]\). If all \( L \) symbols of the coding block use the same QAM symbol order \( M \), based on (7), IPI can also be defined as

\[
IPI = \frac{1}{(\log_{2} M) L} \sum_{l=1}^{L} ISI(y_{l}, \gamma_{l}) = \frac{1}{(\log_{2} M) L} IPI_{s}(y),
\]

where

\[
IPI_{s}(y) = \frac{1}{L} \sum_{l=1}^{L} ISI(y_{l}, \gamma_{l})
\]

is the symbol level IPI. Symbol-level IPI is equal to \( M \) times of bit-level IPI for a given coding block using uniform QAM symbol order. When \( L \) goes to \( \infty \), the following corollary exists for symbol-level IPI.

**Corollary 2.1.** When \( L \to \infty \), \( IPI_{s} \) of a given packet \( y \) is equal to the expectation of mutual information with respect to the distribution of SNR \( f(\gamma) \).

\[
\lim_{L \to \infty} IPI_{s}(y) = \int_{0}^{\infty} I(X; Y|\gamma)f(\gamma)d\gamma
\]

Due to the average operation within modulation symbol for the bit-level IPI calculation, it is independent to modulation order. IPI can be used to evaluate the received code block quality. Fig. 2 shows that the curves of packet error rate (PER) vs. IPI with respect to convolutional codes (R=1/3, 1/2, 2/3). The PER decreases with the increasing of IPI. For a given IPI, lower coding rate leads to lower PER. More importantly, the comparison between curves under AWGN channel and Rayleigh fading channel are presented. It is shown that for a given IPI, the corresponding PER under AWGN channel is similar to the PER under Rayleigh fading channel, which indicates that IPI metric is fading-insensitive. This is an important feature, which enables the IPI to be used in practical system to evaluate the received coding block quality and adaptively determine the retransmission. The assumption is that the coded block is long enough (> 1024) and a long interleaver is used. Since the same curve can be used for a specific code rate and block length combination, regardless of what is the channel condition. As the possible combinations of code rate and block length are limited in many wireless systems, only a finite number of curves need to be stored.

The IPI definition is similar to the mutual information based link quality metrics in [10][11][12]. However, those metrics
employ the statistic nature of the classical definition of mutual information, hence can only be used for link quality measure instead of the received code block quality measure. As the latter requires the consideration of instantaneous noise realization within a packet, which is exactly what IPI differs from the metrics in [10][11][12]. Given the IPI of the currently received code block and the targeted PER, how much additional IPI is needed from the HARQ retransmission can be obtained from Fig. 2. This observation is used in the next section to design an on-demand adaptive retransmission for HARQ using the IPI metric.

III. INSTANTANEOUS PACKET INFORMATION BASED ON-DEMAND ADAPTIVE RETRANSMISSION

In this section, an IPI based adaptive retransmission scheme for HARQ is introduced, in which the IPI is used to evaluate the received signal block quality and estimate the retransmission amount. For comparison purpose, the conventional HARQ-CC scheme is briefly described in the following.

A. HARQ-CC scheme

In HARQ-CC scheme, the entire packet is retransmitted if the original transmission fails the decoding process. On the receiver side, soft values from the original transmission and retransmission are combined to construct new soft values as the decoder input. Packet retransmission stops until either the packet is successfully decoded or the number of retransmission exceeds the maximum limit. A general HARQ-CC process is shown in Fig 3.

B. Adaptive IPI-HARQ Scheme

Conventional HARQ-CC and HARQ-IR schemes do not support adaptive retransmission. In HARQ-CC, the entire packet is transmitted in each retransmission. While less data is retransmitted in HARQ-IR, the retransmission lengths are predefined in the protocol. However, the actual additional information need for correct decoding varies depending on

The retransmission percentage $\rho$ will be obtained by

$$
\rho = \left\{ \begin{array}{ll}
\frac{IPI_{\text{current}} - IPI_{\text{target}}} {IPI_{\text{current}}} & \text{if } IPI_{\text{current}} < IPI_{\text{target}}, \\
\rho_0 & \text{if } IPI_{\text{current}} \geq IPI_{\text{target}},
\end{array} \right.
$$

where $\rho_0$ is a constant percentage when $IPI_{\text{current}}$ exceeds...
IPI_{target} but there is still decoding error happening. In this case, to statistically reduce PER, a given minimum percentage of packet will be retransmitted to increase IPI. The length of retransmission in terms of the number of bits, J_{re\_trans} is adaptively determined by

\[ J_{re\_trans} = \lceil \rho J \rceil, \]  

(17)

the smallest integer greater or equal to \( \rho J \).

When \( J_{re\_trans} < J \), it is also necessary to determine which part of the original packet should be retransmitted. In the following, an ideal scheme is introduced first, followed by a practical suboptimal scheme.

Defining \( p_b = [p_1, p_2, ..., p_J] \) as a bit-level indicator vector, in which

\[ p_j = \begin{cases} 1 & \text{if the } j\text{th bit will be retransmitted;} \\ 0 & \text{if the } j\text{th bit will not be retransmitted;} \end{cases} \]  

(18)

With regard to (17),

\[ \sum_{j=1}^{J} p_j = J_{re\_trans} \]  

(19)

Bit-level indicator vector is obtained according to the IBI of each bit within current received packet. Due to the saturation characteristics of large soft value amplitudes as shown in Fig. 1, to most efficiently increase IPI, bits with the lowest absolute soft values should have higher priority than others in the retransmission. For bits with high IBI, retransmission will not lead to efficient increase of the IPI. Consequently, the process of obtaining bit-level indicator vector \( p_b(z+1) \) is to select the \( J_{re\_trans} \) bits with the lowest accumulated soft value.

Let \( SV_{i,j}^{(z)} \) denote the accumulated soft value of the \( j\)-th bit of the \( i\)-th packet after the \( z\)-th transmission. It is given by

\[ SV_{i,j}^{(z)} = \begin{cases} SV_{i,j}^{(z-1)} + SV_{i,j}^{(z)}, & \text{if } p_j^{(z)} = 1, \\ SV_{i,j}^{(z-1)}, & \text{if } p_j^{(z)} = 0. \end{cases} \]  

(20)

In the above equation, \( SV_{i,j}^{(z)} \) is the soft value of the \( j\)-th bit solely obtained from the \( z\)-th transmission. Correspondingly, the IPI of the \( i\)-th packet after the \( z\)-th retransmission is given by

\[ IPI^{(z)}(i) = \frac{1}{J} \sum_{j=1}^{J} g(SV_{i,j}^{(z)}). \]  

(21)

During the NACK feedback stage, indicator vector \( p_b \) will be transmitted. Note that the size of \( p_b \) is \( J \) and the number of 1s in \( p_b \) is \( J_{re\_trans} \), where \( J \) and \( J_{re\_trans} \) is determined by the received IPI. Hence, the number of all possible \( p_b \)s satisfying Eq. (19) is \( (J_{re\_trans})^J \). Correspondingly, we can use \( \log_2 (J_{re\_trans}) \) bits to describe the the bit-level indication vector \( p_b \) [15]. Hereby, feedback overhead introduced by the transmission of bit-level indication vector is \( \log_2 (J_{re\_trans}) \). The feedback overhead grows with the increasing packet size, which diminishes or eventually cancels out the possible throughput gain achieved by adaptive retransmission. In order to reduce this system overhead introduced by the feedback of indication vector, bit clustering is applied within each packet. In the bit clustering scheme, a packet of bits is grouped into \( K \) clusters. Instead of retransmitting individual bits with lowest accumulated soft value amplitudes, a cluster of bits is the minimum unit to be retransmitted. Therefore, the retransmission amount will become \( K_{re\_trans} = \lceil \rho K \rceil \) clusters. Thus, the indication vector has a fixed length of \( K \) and the feedback overhead is controlled.

The bit clustering scheme is also naturally supported in real systems. In multiple carrier systems using OFDM, a code block will span several OFDM symbols. The symbols occupying the same subcarrier blocks are likely to experience the same channel fading condition and hence can be grouped into one cluster.

IV. SIMULATION RESULTS

In this section, we will compare IPI-HARQ scheme with HARQ-CC scheme in terms of throughput. Based on renewal-reward theorem [4], the throughput of HARQ schemes is defined as:

\[ \eta = \frac{E\{\mathcal{R}\}}{E\{\mathcal{T}\}} \]  

(22)

where \( \mathcal{R} \) is defined as the random reward, i.e., the transmitted information when a packet is successfully transmitted; \( \mathcal{T} \) denotes the cost, i.e., the total transmitted blocks for a successful packet transmission. When using OFDM system to evaluate the performance, we will use a unified definition of reward and cost, which is the number of successfully transmitted OFDM symbols and the total number of OFDM symbols employed in the whole HARQ process.

A 64-FFT based OFDM modulation system will be adopted. Meanwhile, multi-path channel with delay profile of [1 5 7] and power profile of [0 -8 -10]dB is considered. As shown in Fig.5, every OFDM symbol is divided into \( K \) clusters aligning in frequency domain. In this paper, \( K \) is set to 16 and \( IPI_{target} \) is set according to the criteria that packet error rate is equal to \( 1 \times 10^{-3} \). With regard to Fig.2, \( IPI_{target} \) is set to 0.73, 0.85 and 0.95 with respect to the coding rate of 1/3, 1/2 and 2/3. The constant retransmission percentage \( \rho_0 \) is set to 1/16.
Adaptively selecting retransmission based on the quality of current received packet can improve the system throughput in wireless communication systems. In this paper, a fading insensitive link quality metric—Instantaneous Packet Information metric is proposed. Based on IPI-metric, we can evaluate the quality of received packet and estimate the on-demand retransmission amount according to the difference between $IPI_{target}$ and $IPI_{current}$. An IPI-metric based HARQ scheme (IPI-HARQ), featuring with variable retransmission amount and an indication vector assisted NACK, is proposed. Simulation results show that the proposed HARQ scheme is able to provide a significant throughput gain over conventional HARQ scheme.

**V. Conclusion**

The average transmission times of proposed IPI-HARQ scheme is close to that of HARQ-CC scheme with regard to a large range of IPI. It means that even though fewer information is retransmitted, IPI-HARQ will guarantee transmission quality without introducing more transmission times.

**References**