Qi, Y., Hoshyar, R., Imran, M. A., and Tafazolli, R. (2011) H2-ARQ-relaying: spectrum and energy efficiency perspectives. IEEE Journal on Selected Areas in Communications, 29(8), pp. 1547-1558.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/136441/

Deposited on: 8 February 2017

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

H^2-ARQ-Relaying: Spectrum and Energy Efficiency Perspectives

Yinan Qi, *Member, IEEE*, Reza Hoshyar, *Member, IEEE,* Muhammad Ali Imran, *Member, IEEE* and Rahim Tafazolli, *Member, IEEE*

***Abstract*—In this paper, we propose novel Hybrid Automatic Repeat re-Quest (HARQ) strategies used in conjunction with hybrid relaying schemes, named as H^2-ARQ-Relaying. The strategies allow the relay to dynamically switch between amplify-and-forward/compress-and-forward and decode-and-forward schemes according to its decoding status. The performance analysis is conducted from both the spectrum and energy efficiency perspectives. The spectrum efficiency of the proposed strategies, in terms of the maximum throughput, is significantly improved compared with their non-hybrid counterparts under the same constraints. The consumed energy per bit is optimized by manipulating the node activation time, the transmission energy and the power allocation between the source and the relay. The circuitry energy consumption of all involved nodes is taken into consideration. Numerical results shed light on how and when the energy efficiency can be improved in cooperative HARQ. For instance, cooperative HARQ is shown to be energy efficient in long distance transmission only. Furthermore, we consider the fact that the compress-and-forward scheme requires instantaneous signal to noise ratios of all three constituent links. However, this requirement can be impractical in some cases. In this regard, we introduce an improved strategy where only partial and affordable channel state information feedback is needed.**

***Index Terms*—Cooperative systems, channel state information, throughput, energy efficiency.**

I. Introduction

PACKET oriented data transmission has long been used in wireless communication systems, where a major concern is how to control transmission errors caused by the time-varying channel and noise. One approach to this problem is the application of Automatic Repeat re-Quest (ARQ), or its advanced version Hybrid-ARQ (HARQ). Moreover, various distributed network architectures based on relaying techniques are emerging to provide uniform high rate coverage for future wireless systems. It is natural to combine HARQ and the relaying schemes in a complementary way to further exploit the degrees of freedom of the wireless media and enjoy improved performance.

In addition to amplify-and-forward (AF) and decode-and-forward (DF) [1]-[3], compress-and-forward (CF) has recently drawn considerable attention [4]-[11]. Unlike DF, which benefits from transmit diversity, in CF the relay compresses and forwards its observation to the destination and thus provides receive diversity [12]. Some advanced relaying schemes are proposed in [6] and [13]-[15], where the conventional mechanisms are combined in a hybrid fashion and some improvements are identified.

However an effective relaying scheme by itself is not sufficient to prevent packet loss and therefore, retransmission techniques based on ARQ [16] and HARQ can be used to circumvent this problem. In [17] and [18], the ARQ techniques are combined with the relaying techniques, where the diversity-multiplexing-delay tradeoff and the combination of adaptive cooperation and ARQ, respectively, are analyzed. Their work has been extended to HARQ in [19]-[20]. In [21], the energy consumption of a DF based HARQ system is analyzed. However, in these previous works, HARQ is deployed with a non-hybrid forwarding scheme, usually DF, and hence referred as H-ARQ-Relaying. From [6], we know that the outage behavior can be greatly improved if the relay adaptively switches between different forwarding schemes. Further improvement can be expected by combining a hybrid relay system with HARQ. This combined strategy is called H^2-ARQ-Relaying since both the retransmission protocols and the relaying schemes are hybrid. Some preliminary work on H^2-ARQ-Relaying is done in [22], where the relay system can switch between AF and DF and exhibits significant frame error rate improvement.

The existing analysis on the combination of the HARQ mechanisms with the relay systems suffers from several limitations and drawbacks: 1) the performance of the AF/DF scheme is fundamentally limited by the source-relay link and it also suffers from the bandwidth loss due to the use of repetition codes in AF. 2) No throughput analysis for a hybrid relay system has been provided in an analytical framework. 3) Although the impact on the total energy, including both transmission and circuitry energy, has been investigated in a DF based system, it has not been addressed for hybrid systems. In addition, the work in [21] has several constraints, such as the same distance between three nodes, uniform power allocation for the source and the relay. These constraints are removed in this work to exploit the optimal performance with more flexibility. In this paper, we focus on the hybrid CF/DF based H^2-ARQ-Relaying strategy, although the AF/DF one is also addressed briefly. Only incremental redundancy (INR) based HARQ is considered as it offers higher throughput [23]. The proposed strategies and their spectrum and energy efficiency analysis are the main contributions of this work.

We notice that the CF technique requires global instantaneous channel state information (CSI) at the relay in order to perform ideal Wyner-Ziv coding [24]-[25]. We propose a

Manuscript received October 10, 2010; revised February 13, 2011.

Yinan Qi, Muhammad Ali Imran and Rahim Tafazolli are with the Centre for Communication Systems Research, University of Surrey, United Kingdom, GU2 7XH, e-mail: (Yinan.Qi@surrey.ac.uk).

Reza Hoshyar is with National Semiconductor research lab, California, USA.

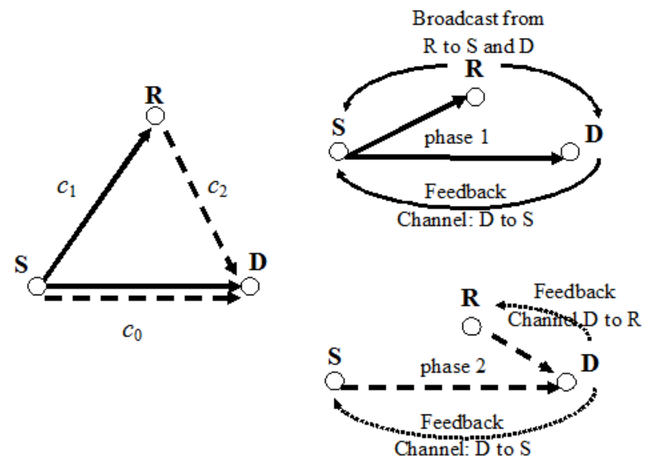

Fig. 1. Half-duplex relay system model with feedback

practical strategy requiring a very small number of extra feedback bits which carry only partial CSI. Although the amount of feedback information is greatly reduced, this approach performs close to the original H^2-ARQ-Relaying strategy with full CSI. The rest of the paper is organized as follows. The system model and assumptions are introduced in the next section. The H^2-ARQ-Relaying strategies are proposed and their performance in terms of spectrum and energy efficiency is studied in section III and IV, respectively. In section V, the practical strategy is analyzed. Simulation results are discussed in Section VI and conclusions are drawn in the final section.

We employ the following notations and definitions: vectors and matrices are represented by bold letters, e.g. $\mathbf{X}$; a random m-vector is a column vector with m independent and identical distributed (i.i.d.) random elements; sequences of scalar x and vector $\mathbf{X}$ are designated as $\underline{x}_n = (x_1, ..., x_n)$ and $\underline{\mathbf{X}}_n = (\mathbf{X}_1, ..., \mathbf{X}_n)$, respectively, and the truncated sequence is defined as $\underline{x}_{l,m} = (x_l, ..., x_m)$.

II. The hybrid relay system model

We introduce the system model in Fig.1, where S, R and D denote the source, relay and destination, respectively, and c_0, c_1 and c_2 are the S-D, S-R and R-D channel gains, respectively, modeled by circularly symmetric complex Gaussian (CSCG) distribution with unit variance and zero mean. Some important assumptions and definitions are clarified as follows:

1) A complete transmission lasts for T seconds, where $T \leq T_{max}$ and is optimized as a design parameter. A half-duplex model is assumed, where a complete frame consists of a relay-receive phase with duration αT and a relay-transmit phase with duration $(1 - \alpha T)$
2) Block fading is assumed. We also assume that the global CSI is known at the relay but not the source. Under such CSI availability assumptions, a theoretical bound of the CF/DF based strategy is derived as a benchmark. We will propose effective solutions to relax the CSI requirements at the relay. However, this CSI assumption could still be valid as the CF technique is a viable solution when the R-D link is strong. It allows for the S-D link CSI feedback from the destination. The knowledge of the R-D link can be obtained either through channel reciprocity or feedback. Acquiring the S-R CSI at the relay is trivial. Following the same line of argument, provision of the global CSI for the source is possible only when the source has a strong link to the destination, which is of no interest as there would not be any need for the relay to cooperate. There could be other possibilities for CSI assumptions. However, those are out of the scope of this work.
3) The global CSI is also available at the destination, which can be achieved through channel estimation and a negligible cost of rate [6]. Perfect frequency and time synchronizations between the source, destination and relay are assumed. However, phase synchronization is not assumed due to implementation difficulty [12].
4) If needed, the subscripts and superscripts denote nodes and time phases, respectively.

The source transmits n_w bits to the destination by sending a message $w \in \{1, ..., 2^{n_w}\}$. We define the following codebooks: $\mathbf{X}_s^1(w)$ and $\mathbf{X}_s^2(w)$ are random αn_r- and $(1-\alpha)n_r$-vectors, respectively, and $\mathbf{X}_{r,DF}(w)$ and $\mathbf{X}_{r,CF}(\nu)$ are random $(1-\alpha)n_r$-vectors, where n_r is the frame length and ν is the compression index. We assume independent CSCG distribution with zero mean for all the elements of the defined codebooks.

In phase 1, the source transmits $\mathbf{X}_s^1(w)$ and the received signals at the relay and the destination are

$$\mathbf{Y}_r = \frac{c_1}{\sqrt{K_t d_1^{\zeta/2}}}\mathbf{X}_s^1(w) + \mathbf{Z}_r, \mathbf{Y}_d^1 = \frac{c_0}{\sqrt{K_t d_0^{\zeta/2}}}\mathbf{X}_s^1(w) + \mathbf{Z}_d,$$

respectively, where d is the distance, ζ is the path loss exponent, K_t is a constant indicating the physical characteristics of the link and the power amplifier as in [21] and [28], and $\mathbf{Z}_r$ and $\mathbf{Z}_r$ are the additive noise vectors whose elements are modeled by i.i.d. CSCG distribution with zero mean. During phase 2, depending on the decoding status, the relay takes different actions:

Relay Decoding Success (DF Mode): the relay and the source transmit $\mathbf{X}_{r,DF}(w)$ and $\mathbf{X}_s^2(w)$, respectively and the destination receives

$$\mathbf{Y}_d^2 = \frac{c_0}{\sqrt{K_t d_0^{\zeta/2}}}\mathbf{X}_s^2(w) + \frac{c_2}{\sqrt{K_t d_2^{\zeta/2}}}\mathbf{X}_{r,DF}(w) + \mathbf{Z}_d.$$

The destination decodes w jointly from $\mathbf{Y}_d^1$ and $\mathbf{Y}_d^2$.

Relay Decoding Failure (CF or AF Mode): In CF mode, the relay quantizes $\mathbf{Y}_r$ into an intermediate bin index u, which is then compressed to index ν [7], [23]. In contrast to DF mode, the relay now forwards $\mathbf{X}_{r,CF}(\nu)$. At the destination, $\mathbf{X}_{r,CF}(\nu)$ is firstly decoded by treating $\mathbf{X}_s^2(w)$ as noise. Once ν is obtained, with the help of the side information, the estimation of the relay's observation, denoted as $\hat{\mathbf{Y}}_r$, is reconstructed. The destination then subtracts the known part from $\mathbf{Y}_d^2$ and uses $\bar{\mathbf{Y}}_d^2 = \mathbf{Y}_d^2 - \frac{c_2}{\sqrt{K_t d_2^{\zeta/2}}}\mathbf{X}_{r,CF}(\nu)$, $\mathbf{Y}_d^1$ and $\hat{\mathbf{Y}}_r$ to jointly decode w.

AF mode is nothing different but replacing CF with AF and its signal model is described in Appendix A. In the following sections, we only focus on the CF/DF based strategy.

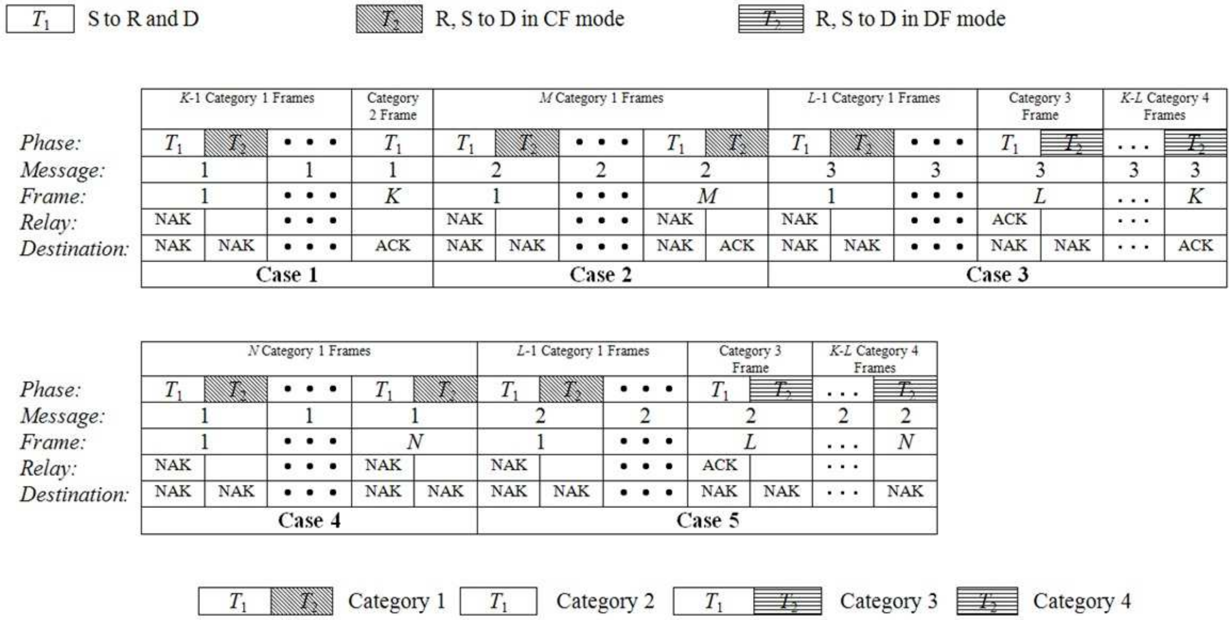

Fig. 2. H^2-ARQ-Relaying Strategy (4 frame categories defined)

III. A CF/DF BASED H^2-ARQ-RELAYING STRATEGY

When retransmission protocols are considered, we can expand the defined codebooks to $\mathbf{X}^1_{s,n}(w)$, $\mathbf{X}^2_{s,n}(w)$, $\mathbf{X}_{r,DF,n}(w)$ and $\mathbf{X}_{r,CF,n}(\nu_n)$, for $1 \leq n \leq N$, where N is the maximum retransmission limit. As depicted in Fig.1, the feedback channels are used to convey the decoding status at the relay and the destination. During phase 1 of the n-th frame, the source broadcasts $\mathbf{X}^1_{s,n}(w)$. If the message is successfully decoded by the destination, where the decoding is based on all buffered signals, an acknowledgement (ACK) message is sent back to the source and the relay and the transmission of w is finished. If the destination detects errors, it broadcasts a not acknowledgement (NAK) message. Once the relay receives the NAK message, it tries to decode through a joint processing of $\underline{\mathbf{Y}_{r}}_{n}$. Different actions are envisaged based on its decoding status:

Relay Decoding Success: An ACK is broadcast by the relay to indicate that it is working in DF mode. The relay and the source transmit $\mathbf{X}_{r,DF,n}(w)$ and $\mathbf{X}^2_{s,n}(w)$, respectively. Afterwards, the source keeps silent during phase 1 of the following frames to save the energy. It only transmits with the relay in a more efficient cooperative manner during phase 2 until the source receives ACK or the retransmission limit N is reached.

Relay Decoding Failure: The relay switches to CF mode and broadcasts a NAK message. At the end of phase 2, the destination attempts to decode by joint processing $\underline{\mathbf{Y}^1_{d}}_{n}$, $\underline{\hat{\mathbf{Y}}_{r}}_{n}$ and $\underline{\bar{\mathbf{Y}}^2_{d}}_{n}$. Upon successful decoding, an ACK is sent back; otherwise, the destination issues a NAK message to start a new frame if $n \leq N$. If $n = N$, HARQ failure is announced.

Fig.2 sketches all the possible scenarios. In the first one, w is successfully decoded during phase 1 of the K-th frame after $(K-1)$ CF operations, in case 2, successful decoding occurs after exact M CF operations, and in case 3, DF mode is activated in the L-th frame but the destination does not correctly decode w until the K-th frame. Fig.2 also depicts two scenarios where HARQ failure is announced when the relay works in CF or DF mode, respectively. Accordingly, 4 frame categories are defined in Fig.2. Table I gives the definition of frame categories. Here we assume that frames are spaced enough along the time domain or the frequency domain to ensure independency of channel gains.

TABLE I
FRAME CATEGORIES

Category	Name	Definition
1	CoF(C)	Complete Frames consisting of two phases, operating in CF mode in phase 2
2	BrF	Broadcasting Frames consisting of phase 1 only
3	CoF(D)	Complete Frames consisting of two phases, operating in DF mode in phase 2
4	RtF(D)	Relay-transmit Frames consisting of phase 2 only, operating in DF mode
5	CoF(A)	Complete Frames consisting of two phases, operating in AF mode in phase 2
6	CoF(Di)	Complete Frame consisting of two phases with direct transmission mode in both phases

A state diagram is presented in Fig.3. B_n stands for the state where the source is ready to broadcast w in the n-th frame. The state D_n indicates that DF mode is activated at the end of phase 1 of the n-th frame. $DF_{m,n}$ is the state where the system has entered DF mode in the n-th frame and cooperative transmission in phase 2 has been repeated m times but the destination still detects errors. C_n defines the state in which the relay is ready to conduct CF during phase 2 of the n-th frame. S and F_a are the successful decoding state and HARQ failure state, respectively.

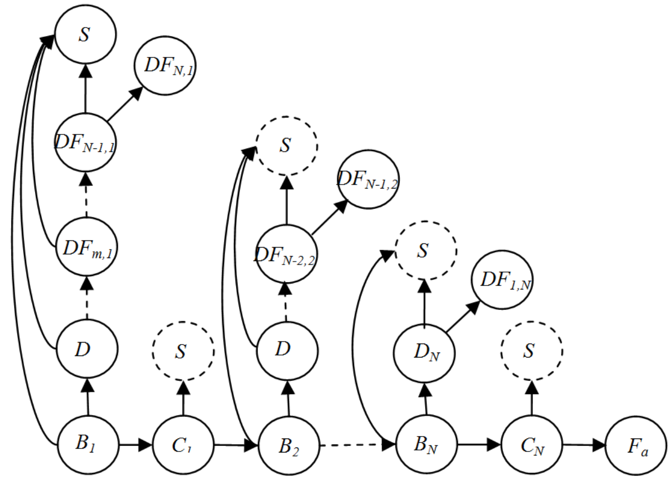

Fig. 3. State Diagram of the CF/DF based H^2-ARQ-Relaying Strategy

IV. H^2-ARQ-Relaying Performance Analysis

In order to investigate the spectrum and energy efficiency of the H^2-ARQ-Relaying strategies, the outage probability needs to be derived. Now we assume delay- and error-free feedback links for the feedback information.

A. Outage Analysis

For the CF/DF based strategy, only the first 4 frame categories in Table I are valid. We define the valid frame category set as $\mathcal{J} = \{1, 2, 3, 4\}$. The achievable rates (ARs) in different frame categories are different and derived in Appendix B. Let $R^{(t)}$ denote the AR at the destination in a category t frame, where $t \in \mathcal{J}$, after n frames the AR can be expressed as

$$R_d(n) = \sum_{i=1}^{n} R^{(t_i)}, \tag{1}$$

where t_i is the category attribute of the i-th frame. Define $\underline{t}_n = (t_1, ..., t_n)$ as the frame category indication sequence (fcis). We also define $\upsilon_t(\underline{t}_{m,n}) \overset{def}{=} \sum_{i=m}^{n} 1_{\{t_i=t\}}$ to indicate the total number of category t frames from frame m to n, where $1_{\{t_i=t\}}$ is an indication function whose value is 1 when $t_i = t$ and 0 otherwise. The AR at the relay after k transmissions is given as

$$R_r(k) = \sum_{i=1}^{k} R_{r,i} = \sum_{i=1}^{k} \alpha B \log_2 \left(1 + \frac{\gamma_{1,i} E_{b,s}^1 n_w}{\alpha T N_0 B K_t d_1^{\zeta}} \right), \tag{2}$$

where $E_{b,s}^1$ is the energy per bit at the source during phase 1, $\gamma_{j,i} = |c_{j,i}|^2$ for $0 \le j \le 2$, B is the bandwidth, and N_0 is noise power.

If n_w bits are transmitted in one frame, we can define the event $\mathcal{A}(\boldsymbol{\upsilon}, k) \overset{def}{=} \{R_d(n) \le n_w/T, R_r(k) \le n_w/T\}$, and the outage probability $q(\boldsymbol{\upsilon}, k) = \Pr(\mathcal{A}(\boldsymbol{\upsilon}, k))$, where $\boldsymbol{\upsilon}(\underline{t}_n) = (\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4)$ with $\upsilon_t = \upsilon_t(\underline{t}_n)$ and $n = \sum_{t=1}^{4} \upsilon_t$. It is straightforward to show that under the block fading assumption, for any fcis pair $\underline{t}_n$ and $\underline{t}'_n$, the outage probabilities are the same as long as $\boldsymbol{\upsilon}(\underline{t}_n) = \boldsymbol{\upsilon}(\underline{t}'_n)$. The outage probability $q(\boldsymbol{\upsilon}, k)$ can be calculated by resorting to the 2-dimensional characteristic function of $R_d(n)$ and $R_r(k)$,

$$\begin{aligned}\Psi(\mathbf{s}, n, k) &= \mathbb{E}\left\{ e^{-s_1 R_d(n) - s_2 R_r(k)} \right\} \\ &= \mathbb{E}\left\{ \prod_{i=1}^{k} e^{-s_1 R^{(t_i)} - s_2 R_{r,i}} \prod_{i=k+1}^{n} e^{-s_1 R^{(t_i)}} \right\} \\ &= \prod_{t \in \mathcal{J}} \left\{ \phi_t(\mathbf{s}) \right\}^{\upsilon_t(\underline{t}_k)} \prod_{t \in \mathcal{J}} \left\{ \bar{\phi}_t(s_1) \right\}^{\upsilon_t(\underline{t}_{k+1,n})},\end{aligned} \tag{3}$$

where $\mathbf{s}$ is $(s_1 \ s_2)$, and

$$\phi_t(\mathbf{s}) = \mathbb{E}\left\{ e^{-s_1 R^{(t)} - s_2 R_r} \right\}, \bar{\phi}_t(s_1) = \mathbb{E}\left\{ e^{-s_1 R^{(t)}} \right\}.$$

The outage probabilities can be expressed by the Laplace inversion formula of $\Psi(\mathbf{s}, n, k)$ and approximated in [26],

$$\begin{aligned}q(\boldsymbol{\upsilon}, k) &= \frac{1}{(2\pi j)^2} \int_{d_2 - j\infty}^{d_2 + j\infty} \int_{d_1 - j\infty}^{d_1 + j\infty} \frac{\Psi(\mathbf{s}, n, k) e^{\frac{s_1 n_w}{T} + \frac{s_2 n_w}{T}}}{s_1 s_2} \mathrm{d}s_1 \mathrm{d}s_2 \\ &\approx \sum_{i=1}^{M} \sum_{j=1}^{M} K_i K_j \frac{\Psi(\mathbf{z}, n, k)}{z_i z_j},\end{aligned} \tag{4}$$

where d_1 and d_2 are proper constants, z_i are the poles of the Padé rational function, $\mathbf{z}$ is $(z_i T/n_w \ z_j T/n_w)$, K_i are the corresponding residues and M is an arbitrary integer that determines the approximation accuracy and is chosen as 10 in this paper. The closed-form expressions of $\phi_t(\mathbf{z})$ and $\bar{\phi}_t(z_i T/n_w)$ are difficult to obtain but we can resort to three-dimensional numerical integration. For instance

$$\bar{\phi}_t(z_i T/n_w) = \iiint_{\mathbb{R}^+} e^{\left(-z_i R^{(t)} T/n_w\right)} p(\boldsymbol{\gamma}) \mathrm{d}\gamma_0 \mathrm{d}\gamma_1 \mathrm{d}\gamma_2, \tag{5}$$

where $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2)$, and $p(\boldsymbol{\gamma})$ is the joint distribution of γ_0 to γ_2.

AF/DF based H^2-ARQ-Relaying: The above approximation can be easily extended to the AF/DF based strategy. The corresponding valid frame category set is $\mathcal{J} = \{2, 3, 4, 5\}$, where in frame category 5, CF is replaced by AF. The AR at the destination in AF mode is derived in Appendix A.

B. State and Transition Probabilities

For two event sets e_1 and e_2, if $(e_1 \subseteq e_2)$, $\Pr(e_1, e_2)$ is equal to $\Pr(e_1)$. In the state diagram, the transition probability to state j from any of its adjacent incoming state i is

$$\begin{aligned}\Pr\left(state\ i \rightarrow state\ j\right) &= \Pr\left(state\ j | state\ i\right) \\ &= \frac{\Pr\left(state\ j\right)}{\Pr\left(state\ i\right)}.\end{aligned} \tag{6}$$

In the state diagram shown in Fig.3, all the states have this property except the state *S*. Therefore we can firstly calculate the state probabilities and then obtain the transition probabilities based on (6).

We start from B_n and initialize $\Pr(B_1) = 1$. $\Pr(B_n)$ can be easily obtained as $q\left((n-1, 0, 0, 0), n-1\right)$. It follows that

$\Pr(F_a)$ is $q\left((N,0,0,0),N\right)$. When the relay detects errors at the end of phase 1 of the n-th frame, the state probability is $\Pr(C_n) = q\left((n-1,1,0,0),n\right)$; otherwise, the system moves to state D_n and the corresponding probability is given by

$$\begin{aligned}\Pr(D_n)=&\Pr\left(\mathcal{A}\left(\left(n-1,1,0,0\right),n-1\right)\right)\\&-\Pr\left(\mathcal{A}\left(\left(n-1,1,0,0\right),n\right)\right)\\&=q\left((n-1,1,0,0),n-1\right)-q\left((n-1,1,0,0),n\right).\end{aligned}$$

When the system is in the state $DF_{m,n}$, it implies that $(n+m-1)$ frames consisting of $(n-1)$ CoF(C), 1 CoF(D) and $(m-1)$ RtF(D) have been transmitted. Therefore, the probability is

$$\begin{aligned}\Pr(DF_{m,n}) = &q\left((n-1,0,1,m-1),n-1\right)\\&-q\left((n-1,0,1,m-1),n\right).\end{aligned}$$

There are three routes stemming from B_n. As the probabilities of all the state transitions emanating from a single state must add up to 1, we have $\Pr(B_n \to S) = 1 - \Pr(B_n \to D_n) - \Pr(B_n \to C_n)$. $\Pr(C_n \to S)$ and $\Pr(DF_{m,n} \to S)$ can be obtained in the same way.

C. Spectrum Efficiency Analysis

We analyze the spectrum efficiency in terms of the throughput. Using the state and transition probabilities, we can apply the renewal-reward theorem [23] to evaluate the throughput by investigating the random reward Φ and average airtime $\mathcal{T}$. We first investigate the successful decoding events depicted in Fig.2. T_1 stands for the average time associated with case 1 in Fig.2 and is given by

$$\begin{aligned}&\mathrm{T}_1 = T\sum_{K=1}^{N} P_{BS}(K)\left((K-1)+\alpha\right),\\&P_{BS}(K) = \Pr(B_K)\Pr(B_K \to S).\end{aligned}$$

For case 2, where the system is able to decode after M CF operations, the average time associated is

$$\mathrm{T}_2 = T\sum_{M=1}^{N} P_{CS}(M)M, P_{CS}(M) = \Pr(C_M)\Pr(C_M \to S).$$

The associated average time of case 3 is

$$\begin{aligned}&\mathrm{T}_3 = T\sum_{L=1}^{N}\sum_{K=L}^{N} P_{DS}(L,K)\left(L+(K-L)(1-\alpha)\right),\\&P_{DS}(L,K) = p^{DS}(L,K),\end{aligned}$$

where $p^{DS}\left(L,K\right) = \Pr(D_L)\Pr(D_L \to S)$ when $K = L$ and is equal to $\Pr(DF_{(K-L),L})\Pr(DF_{(K-L),L} \to S)$ for $K > L$.

In the above cases, the decoding is successful and the average reward Φ is the number of successfully transmitted bits n_w. When outage events happen, Φ is equal to zero. The outage events may happen in CF mode in case 4 and the average time is $\mathrm{T}_4 = T\Pr(F_a)N$. In contrast, if the outage events occur in DF mode, the average time is given by

$$\begin{aligned}&\mathrm{T}_5 = T\sum_{L=1}^{N} P_{DF}^{out}(L)(L+(N-L)(1-\alpha)),\\&P_{DF}^{out}(L) = \Pr(DF_{N-L+1,L}).\end{aligned}$$

Since the outage events in DF and CF modes are mutually exclusive, the outage probability is $p^{out} = \Pr(F_a) + \sum_{L=1}^{N} P_{DF}^{out}(L)$. The average random reward and the average airtime take the forms of $\mathbb{E}\left\{\Phi\right\} = n_w\left(1-p^{out}\right)$ and $\mathbb{E}\left\{\mathcal{T}\right\} = \sum_{i=1}^{5}\mathrm{T}_i$. The average throughput is

$$\mu(\alpha,T) = \frac{\mathbb{E}\{\Phi\}}{\mathbb{E}\{\mathcal{T}\}} = \frac{n_w(1-p^{out})}{\sum_{i=1}^{5}\mathrm{T}_i}. \tag{7}$$

The resulted throughput is a function of the duplexing ratio α and the activation time T. Therefore it should be maximized with respect to α and T as

$$\mu_{max} = \max_{T,\alpha}\mu(T,\alpha). \tag{8}$$

D. Energy Efficiency Analysis

Denote the transmission energy per bit at the source during phase 1, 2, and at the relay as $E_{b,s}^1$, $E_{b,s}^2$ and $E_{b,r}$, respectively. In phase 1, one node transmits and two nodes receive, the consumed energy is

$$E^1 = n_w E_{b,s}^1 + \left(P_{ct} + 2P_{cr}\right)\alpha T,$$

where P_{ct} and P_{cr} are transmission and reception circuitry power, respectively. During phase 2, if the relay decodes the message, the relay intends to transmit n_w bits and the consumed energy is

$$E_{DF}^2 = n_w E_{b,s}^2 + n_w E_{b,r} + \left(2P_{ct} + P_{cr}\right)(1-\alpha)T.$$

If not, the consumed energy is

$$E_{CF}^2 = n_w E_{b,s}^2 + sE_{b,r} + \left(2P_{ct} + P_{cr}\right)(1-\alpha)T,$$

where s is the number of bits required to transmit the compressed signal. Since s is a variable depending on the CSI of three links, if $E_{b,r}$ is fixed, the relay transmitting power $P_r = \frac{sE_{b,r}}{(1-\alpha)T}$ varies from frame to frame. In this work we assume that P_r is fixed and thus $E_{b,r}$ is adjusted by the relay in each frame. It is more suitable to evaluate P_r instead of $E_{b,r}$ in CF mode. For simplicity, we let the transmitting power of the source be the same during two phases, i.e. $\frac{n_w E_{b,s}^1}{\alpha T} = \frac{n_w E_{b,s}^2}{(1-\alpha)T}$, the overall energy consumptions per bit are

$$\begin{aligned}E_{b,DF} = &\frac{E_{b,s}^1}{\alpha} + E_{b,r} + \frac{(P_{ct}+2P_{cr})\alpha T}{n_w}\\&+\frac{(2P_{ct}+P_{cr})(1-\alpha)T}{n_w},\\E_{b,CF} = &\frac{E_{b,s}^1}{\alpha} + \frac{(P_{ct}+2P_{cr}+P_r)\alpha T}{n_w}\\&+\frac{(2P_{ct}+P_{cr})(1-\alpha)T}{n_w},\end{aligned}$$

respectively. We can take similar steps in the analysis of average airtime $\mathcal{T}$ and calculate the average consumed energy per bit in five different cases in Appendix C. The average energy per bit is $\varepsilon(T, E_{b,r}, E_{b,s}^1) = \sum_{i=1}^{5} E_{b,i}$. The optimization

TABLE II
INDICATION BITS $\mathbf{b}_n$

$\mathbf{b}_n$	$\boldsymbol{\gamma}$ regions
00	$\gamma_{0,n} \in (0, a_2), \gamma_{2,n} \in [a_1, +\infty)$
01	$\gamma_{0,n} \in (0, a_3), \gamma_{2,n} \in (0, a_1)$
10	$\gamma_{0,n} \in [a_2, a_3], \gamma_{2,n} \in [a_1, +\infty)$
11	$\gamma_{0,n} \in [a_3, +\infty)$

of $\varepsilon(T, E_{b,r}, E^1_{b,s})$ must be subject to some constraints. Here we use the outage probability as the constraint and this optimization problem can be formulated as

$$\begin{gathered}\min \varepsilon(T, E_{b,r}, E^1_{b,s}) \\ \text{subject to } p^{out} \leq p^t, \text{and } T \leq T_{max},\end{gathered} \tag{9}$$

where p^t is the target outage probability. The dependence between p^{out} and T, $E_{b,r}$ and $E^1_{b,s}$ has been shown in Appendix B in the AR expressions. This optimization problem can be solved numerically.

V. A MODIFIED STRATEGY WITH LIMITED CSI FEEDBACK

In this section, we focus on the practical issues and propose a modified H^2-ARQ-Relaying strategy with partial CSI feedback. The basic idea of this strategy is to separate the two-dimensional space (γ_0, γ_2) into different operational regions, where different relaying schemes are carried out.

A. Strategy Description

Inspired by the idea in [6], we use two indication bits $\mathbf{b}_n = (b_{1,n}\; b_{2,n})$ to designate the operational regions as shown in Table II. These two indication bits should be sent back from the destination to the relay and the source when CF mode is activated and different operations are performed accordingly during phase 2. If the S-D link is above some threshold, i.e.,$a_3 \leq \gamma_{0,n}(\mathbf{b}_n = 11)$, we carefully choose a_3 such that

$$\begin{aligned}\frac{n_w}{T} = {} & \alpha B \log_2\left(1 + a_3 \frac{n_w E^1_{b,s}}{\alpha T N_0 B K_t d_0^{\zeta}}\right) \\ & + (1-\alpha) B \log_2\left(1 + a_3 \frac{n_w E^2_{b,s}}{(1-\alpha) T N_0 B K_t d_0^{\zeta}}\right),\end{aligned}$$

and successful decoding is guaranteed at the destination even without resorting to the buffered signals. In such a case, there is no need for the relay to cooperate and it is silent during phase 2. If $\gamma_{0,n} \in [a_2, a_3]$ and $\gamma_{2,n} \in [a_1, +\infty)(\mathbf{b}_n = 10)$, the relay performs CF. As a direct consequence of Corollary 1 of [6], the AR at the destination in this non-ideal scenario is derived in Appendix B. When $\gamma_{0,n} \in (0, a_2)$ and $\gamma_{2,n} \in [a_1, +\infty)(\mathbf{b}_n = 00)$, the 'amount' of side information is insufficient for efficient CF [6]. If $\gamma_{0,n} \in (0, a_3)$ and $\gamma_{2,n} \in (0, a_1)(\mathbf{b}_n = 01)$, the quality of the R-D link is too low to support efficient transmission of the compressed signal. In either case, the relay cannot help and is kept silent. Since the direct link is of low quality now, to save the energy, the source is switched off in phase 2 of the current frame. In this frame the source only broadcasts in phase 1, thus we call it broadcasting mode.

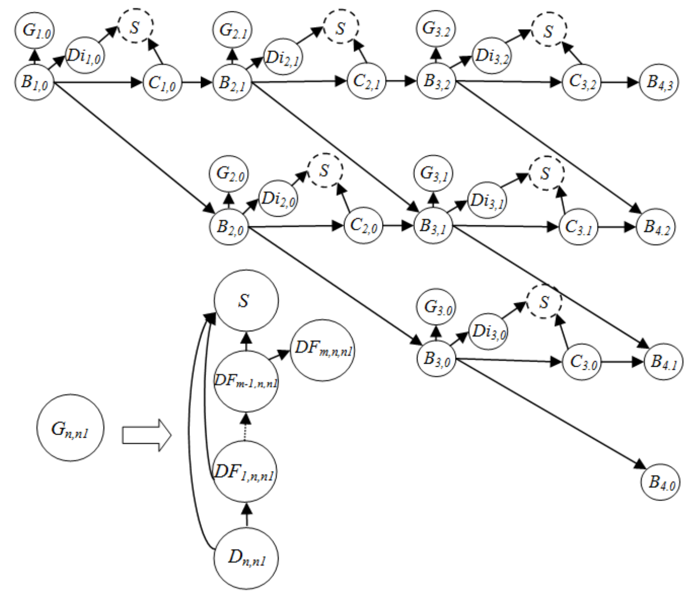

Fig. 4. State transition diagram of the modified H^2-ARQ-Relaying strategy (N=3, G_{n,n_1}: the combination state; Di_{n,n_1}: the direct transmission state)

The state diagram of the modified strategy is shown in Fig.4. B_{n,n_1} is defined as the state where the destination has not been able to decode after $(n-1)$ frames. It also implies that in n_1 frames, CF is conducted and in the rest of the frames, the system works in broadcasting mode. When the system is in the state B_{n,n_1}, upon successful decoding at the relay, the system is ready to work in DF mode and transits to D_{n,n_1}. Afterwards, when category 4 frames are repeated for m times and the destination still detects errors, the system reaches DF_{m,n,n_1}. If the relay detects errors, with $\mathbf{b}_n = 10$, the system transits to C_{n,n_1} and performs CF; with $\mathbf{b}_n = 11$, the relay knows that the direct link is strong, therefore it keeps silent during phase 2 and the system enters a new state Di_{n,n_1}; for the rest of the cases, the system transits to B_{n+1,n_1} directly. In order to keep Fig.4 concise, we use a combination state G_{n,n_1} to indicate the DF operation. It should be noted that the direct transmission is guaranteed to be successful by carefully selecting a_3.

B. Outage Analysis

The fcis of the modified strategy has two properties:

1) Each $\underline{t}_n$ can be segmented into two sub-sequences $\underline{t}_n = (\underline{t}_\kappa, \underline{t}_{\kappa+1,n})$. From frame 1 to κ, the system works in CF or broadcasting mode and $\underline{t}_\kappa$ is a sequence composed of 1 and 2; from the $(\kappa+1)$-th frame, the system works in DF or direct transmission mode, thus $\underline{t}_{\kappa+1,n}$ is formed with 3, 4 and 6. We have $\kappa = \upsilon_1 + \upsilon_2$, $n - \kappa = \upsilon_3 + \upsilon_4 + \upsilon_6$ and obviously $\upsilon_1(\underline{t}_\kappa) = \upsilon_1(\underline{t}_n)$.
2) For each $\upsilon_1(\underline{t}_n)$, there are $\mathcal{C}^{\kappa}_{\upsilon_1} = \binom{\kappa}{\upsilon_1}$ possible sub-sequences $\underline{t}_\kappa$ forming a set $\mathcal{F}_\upsilon$. However, once $\upsilon_3(\underline{t}_{\kappa+1,n})$,$\upsilon_4(\underline{t}_{\kappa+1,n})$ and $\upsilon_6(\underline{t}_{\kappa+1,n})$ are known, the corresponding sub-sequence $\underline{t}_{\kappa+1,n}$ is unique because according to the strategy description, it can only take the forms of sequence (3), (6) or $(3, 4, 4, ..., 4)$.

Now we rewrite the outage probability as an explicit function of $\underline{t}_n$: $q(\underline{t}_n, k) = \Pr(R_d(n) < n_w/T, R_r(k) < n_w/T, \underline{t}_n)$ and define $\boldsymbol{\upsilon} = (\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4, \upsilon_6)$. Under the block fading assumption, it is straight forward to come to the conclusion that if $\boldsymbol{\upsilon}(\underline{t}_n) = \boldsymbol{\upsilon}(\underline{t}'_n)$, we still have $q(\underline{t}_n, k) = q(\underline{t}'_n, k)$. It means that for any particular $\boldsymbol{\upsilon}$, each fcis in $\mathcal{F}_\upsilon$ has the same outage probability. Thus without loss of generality, we assume that for any $\underline{t}_n$, the first sub-sequence after segmentation takes the form as $\underline{t}^P_\kappa(\upsilon_1) = (1_1, 1_1, ..., 1_{\upsilon_1}, 2_{\upsilon_1+1}, 2_{\upsilon_1+2}, ..., 2_\kappa)$ for simplicity. Then the outage probability is

$$\begin{aligned} &q(\underline{t}_n, k) = q\left((\underline{t}^P_\kappa(\upsilon_1), \underline{t}_{\kappa+1,n}), k\right) \\ &= \Pr\left(R_d(n) < \frac{n_w}{T}, R_r(k) < \frac{n_w}{T}, \left(\underline{t}^P_\kappa(\upsilon_1), \underline{t}_{\kappa+1,n}\right)\right). \end{aligned}$$

The overall outage probability associated with set $\mathcal{F}_\upsilon$ can be expressed as

$$\Pr(\mathcal{F}_\upsilon) = \mathcal{C}^\kappa_{\upsilon_1} q\left(\left(\underline{t}^P_\kappa(\upsilon_1), \underline{t}_{\kappa+1,n}\right), k\right).$$

Unfortunately, this outage probability cannot be approximated in the same way described in the previous section due to the explicit argument $\underline{t}^P_\kappa(\upsilon_1)$, which introduces κ additional conditions and this fact violates the applicable condition of the 2-dimensional characteristic function.

C. State and Transition Probabilities

The state probabilities of the modified strategy can be obtained in a similar way. $\Pr(B_{1,0})$ is initialized to 1 and the state probability of B_{n,n_1} can be expressed as

$$\Pr(B_{n,n_1}) = \mathcal{C}^{n-1}_{n_1} q(\underline{t}^P_{n-1}(n_1), n-1).$$

In the state B_{n,n_1}, the source broadcasts. If only the relay successfully decodes the message, the system moves to D_{n,n_1} and its state probability is

$$\Pr(D_{n,n_1}) = \mathcal{C}^{n-1}_{n_1}\bigg(q(\underline{t}^P_n(n_1), n-1) - q(\underline{t}^P_n(n_1), n)\bigg).$$

Similarly, the state probability of C_{n,n_1} is

$$\Pr(C_{n,n_1}) = \mathcal{C}^{n-1}_{n_1} q(\underline{t}^P_{n-1}(n_1), n, \underline{\mathbf{b}}_n),$$

where we make $\underline{\mathbf{b}}_n$ explicit to indicate the operational regions. The rest of the state probabilities can be given as

$$\begin{aligned} \Pr(DF_{m,n,n_1}) = \mathcal{C}^{n-1}_{n_1}\bigg(& q\left(\left(\underline{t}^P_{n-1}(n_1), 3, 4_1, ..., 4_{m-1}\right), n-1\right) \\ & - q\left(\left(\underline{t}^P_{n-1}(n_1), 3, 4_1, ..., 4_{m-1}\right), n\right)\bigg), \end{aligned}$$

and $\Pr(Di_{n,n_1}) = q\left((\underline{t}^P_{n-1}(n_1), 6), n-1\right)$.

D. Throughput and Energy Analysis

To evaluate the average airtime, we investigate the time consumed when the system reaches a certain state. Define $e(l,m) \overset{\text{def}}{=} m + (l-m)\alpha$. $\mathrm{T}_1(K, n_1)$ is the average time associated with case 1 of the original strategy, where the destination decodes in phase 1 of the K-th frame. However, it differs in that it also depends on n_1 and is given as:

$$G_1(K, n_1) = TP(B_{K,n_1})P(B_{K,n_1} \to S)\left(e(K-1, n_1) + \alpha\right).$$

When all possible K and n_1 are taken into consideration, the overall average time is

$$\mathrm{T}_1 = \sum_{K=1}^{N}\sum_{n_1=0}^{K-1} G_1(K, n_1).$$

The consumed energy is then given as

$$\begin{aligned} E_{b,1} = \sum_{K=1}^{N}\sum_{n_1=0}^{K-1} P(B_{K,n_1})P(B_{K,n_1} \to S)\bigg(& n_1 E_{b,CF} \\ & + \frac{(K-n_1)E^1}{n_w}\bigg). \end{aligned}$$

By averaging with respect to n_1 as well, the average airtime and energy in the rest of the cases can be obtained in a similar way except $\mathrm{T}_4 = T\sum_{n_1=0}^{N} P(B_{N+1,n_1})e(N, n_1)$, which is the average airtime when HARQ failure is announced in CF or broadcasting mode. In addition, we also need to consider another case where the system ends up in direct transmission mode during the M-th frame. We have

$$\begin{aligned} &\mathrm{T}_6 = \sum_{M=1}^{N}\sum_{n_1=0}^{M-1} G_6(M, n_1) \\ &G_6(M, n_1) = T\Pr(Di_{M,n_1})\Pr(Di_{M,n_1} \to S)e(M, n_1). \end{aligned}$$

The consumed energy is

$$\begin{aligned} &E_{b,6} = \sum_{M=1}^{N}\sum_{n_1=0}^{M-1} \Pr(Di_{M,n_1})\Pr(Di_{M,n_1} \to S) \\ &\cdot\left(n_1 E_{b,CF} + \frac{(M-n_1-1)E^1}{n_w} + E_{b,CF} - \frac{P_{ct}(1-\alpha)T}{n_w}\right). \end{aligned}$$

where the last two terms represent the energy per bit of the direct transmission. The average energy per bit is $\varepsilon\left(T, E_{b,r}, E^1_{b,s}\right) = \sum_{i=1}^{6} E_{b,i}$, which is a similar optimization problem as (9). The resulted average throughput is a function of the duplexing ratio α, the activation time T and the parameters a_1, a_2 and a_3. Therefore the average throughput can be maximized as

$$\mu_{max} = \max_{T,\alpha,a_1,a_2,a_3} \frac{\mathbb{E}\{\Phi\}}{\mathbb{E}\{\mathcal{T}\}} = \frac{n_w(1-p^{out})}{\sum_{i=1}^{6}\mathrm{T}_i}. \qquad (10)$$

where

$$p^{out} = \sum_{n_1=0}^{N}\Pr(B_{N+1,n_1}) + \sum_{L=1}^{N}\sum_{n_1=0}^{L-1}\Pr(DF_{N-L+1,L,n_1}).$$

VI. Numerical results, comparisons and discussions

In this section, we provide some results for the proposed strategies and compare their performance with some benchmark ones:

1) H-ARQ with direct transmission (*DT*): The relay keeps silent throughout the transmission.

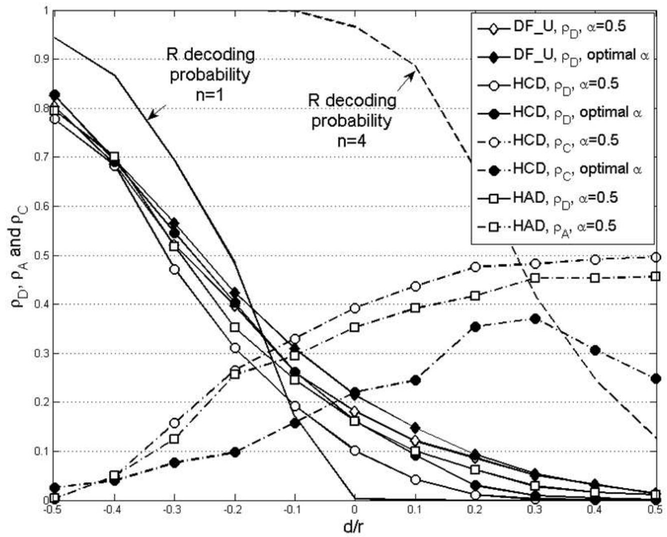

Fig. 5. ρ_D, ρ_A and ρ_C for optimal and fixed α. (e=1.2, $b/r = e/6$, $N = 4$)

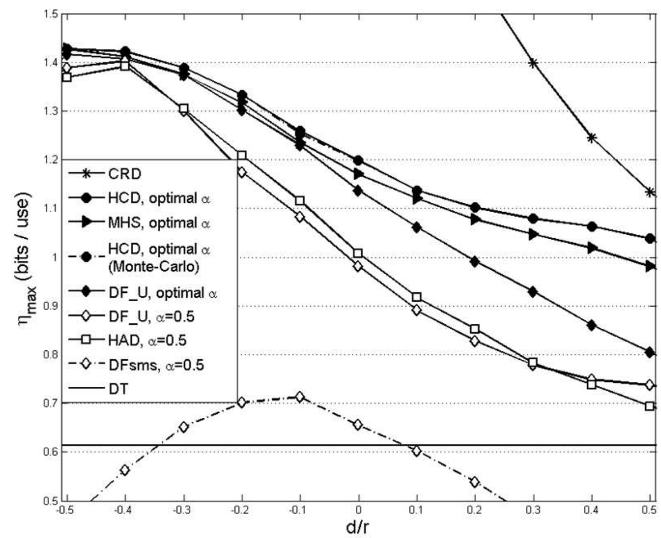

Fig. 6. Throughput μ_{max} with the ellipse model ($e = 1.2$, $b/r = e/6$, Monte-Carlo means that the curve is obtained through Monte-Carlo simulation))

2) H-ARQ-Relaying with the co-located relay and destination (*CRD*): We assume that the relay and the destination are connected by a wire such that full receive diversity is achieved.
3) H-ARQ-Relaying with conventional DF (*DF_U*): When the relay detects errors, it keeps silent during phase 2; otherwise, the relay and the source transmit simultaneously.
4) H-ARQ-Relaying with simple multi-hop strategy (*DF-sms*): There is no direct link between the source and the destination. During phase 1, the source transmits to the relay and only the relay receives and decodes. During phase 2, the source keeps silent. If the relay detects errors, it is silent; otherwise, the relay transmits to the destination. Here α is set at 0.5.
5) H^2-ARQ-Relaying with hybrid AF/DF (*HAD*): The relay performs AF when detecting errors and DF when successfully decoding. Note that in this case, α is also fixed at 0.5 and repetition coding is used when AF is performed.

The maximum retransmission limit N is set to 4. The CF/DF based strategy and its modified version are denoted as *HCD* and *MHS*, respectively. The analytical derivations of the previous sections are verified through appropriate Monte Carlo simulations and close match was observed for all the possible range of SNRs as shown in Fig.6. To avoid crowded figures, we only present the Monte Carlo results for *HCD* to show the match (here the two curves are actually overlapping). For the rest of the strategies, only analytical results are illustrated. We use the following settings: $n_w = 10$, $P_{ct} = 98mW$, $P_{cr} = 112.4mW$, $K_t = 6.05 \times 10^9$, $\zeta = 3$, $N_0 = -171dBm/Hz$, $B = 1$ and $T_{max} = 1$. The distance between the source and the destination is r. The source and the destination are assumed to be placed in the foci of the ellipse $(d/r)^2/(e/2)^2 + (y/r)^2/(b/r)^2 = 1$, for $1 < e < +\infty$, and the relay is moving along the upper side of the ellips.

A. Adaptation Ability

Here we assume that $r = 100\text{m}$, the transmit power of the source and the relay is the same and normalized so that the average SNR of the S-D link is 0dB (change of this value might change the absolute values of our results but the trends will not be affected). In order to investigate the dynamics of the hybrid systems, we define ρ_D and ρ_C as the percentages of DF and CF operation time in the overall airtime, respectively, and their expressions are described in Appendix C. These two parameters depend on two factors: the average number of DF or CF operations and the duration of phase 2. We show ρ_D and ρ_C of different strategies in Fig.5. In addition, we plot the decoding probability of the relay node when $n = 1$ and 4. It is shown that when the relay is moving towards the source (d decreases), the relay's decoding probability increases. Hence, the system is more likely to operate in DF mode and the duration of phase 2 becomes longer, leading to increased ρ_D. While d increases, the relay's decoding probability decreases and the average number of CF operations increases, which tends to increase ρ_C. However, the duration of the second phase tends to be shortened, which decreases ρ_C. Therefore ρ_C is not monotonically changing with d. Only when α is fixed, ρ_C is a non-decreasing function of d. For the AF/DF H^2-ARQ-Relaying strategy, ρ_A can be obtained in a similar manner and the same claims stand.

B. Comparisons of Spectrum Efficiency

Fig.6 depicts the throughputs of strategies with respect to d/r. For the modified strategy we resorted to the Monte-Carlo simulations. Basically, the throughputs are upper bounded by *CRD* because it is able to enjoy full receive diversity. The CF/DF based strategy is the closest one to *CRD*. It can be interpreted as follows: when the relay is moving towards the destination, the S-R link is getting less reliable and the successful decoding probability at the relay decreases. In such

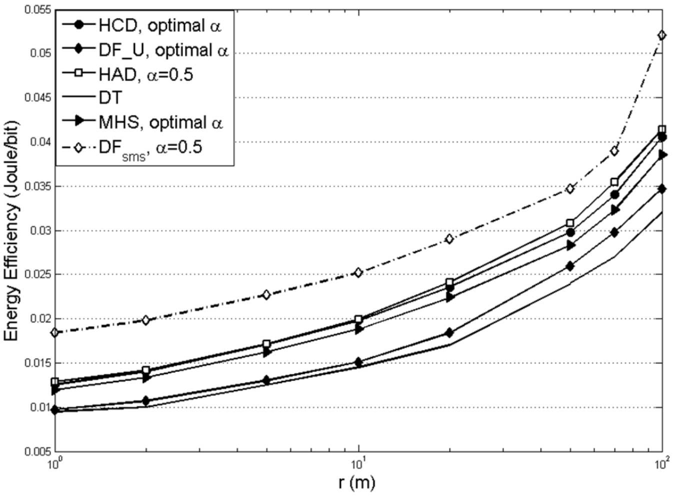

Fig. 7. Energy Efficiency in the short range (Joule per bit)

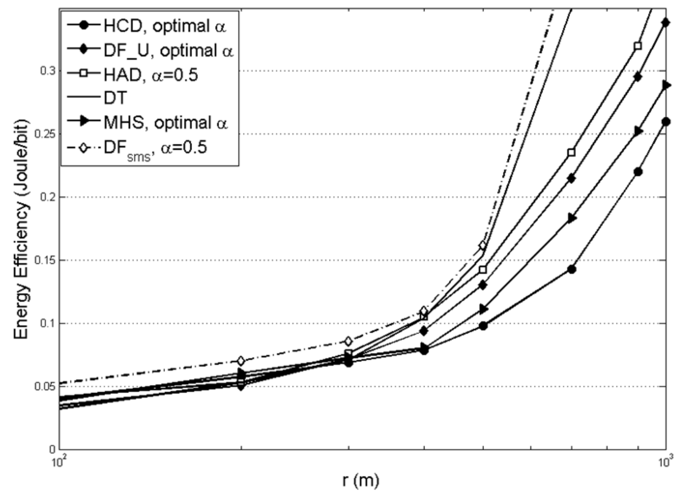

Fig. 8. Energy Efficiency in the long range (Joule per bit)

a scenario, if the system uses DF, it is more likely that the relay's decoding fails and the system operates in direct transmission mode during phase 2, thus no diversity can be achieved. In contrast, when the hybrid CF/DF scheme is used, the system is able to achieve receive diversity through CF. As far as *HAD* is concerned, it enjoys a certain level of flexibility and its throughput is improved in comparison with *DF_U* when α is fixed at 0.5. However, when the S-R link is of low quality, the relay forwards nothing but mostly its own noise. In addition, *HAD* suffers from the bandwidth loss due to the use of repetition coding, and its throughput is therefore lower than *DF_U* with optimal α. Among all the relaying strategies *DFsms* has the worst performance because without the direct S-D link, when the relay is either too close to the source or the destination, there must be a very weak link between the S-R and the R-D links and this weak link dominates and degrades the performance. Another common trend (except *DFsms*) is that when the relay moves away from the source, their throughputs decrease. For *DF_U*, this can be explained by the smaller decoding probability of the relay, which causes fewer opportunities for the relay to cooperate. For *HCD*, although CF can be used, the achievable rate depends on the quality of the S-D and S-R links as shown in (14). The average SNR of the S-R link is decreasing when the relay moves towards the destination. Hence, the throughput reduces. The AF/DF based strategies can be explained following the same line of argument.

C. Comparisons of Energy Efficiency

We optimize the energy per bit for both the source and the relay and compare the energy efficiency performance of various strategies when the relay is close to the destination (the S-D link and R-D link are vertical to each other). Fig.7 shows the situation where the S-D distance is short ($r < 100\mathrm{m}$). The direct transmission is more efficient than any of the cooperative strategies in terms of the energy per bit in short range. This is because now the circuitry energy consumption is the major part of the overall energy consumption. If the relay is activated, although the transmission energy per bit can be reduced, the overall circuitry consumption is almost doubled. Thus, to optimize the energy consumption, the relay should be switched off to save energy. The DF based strategy consumes less energy than others because in this strategy, the relay is silent when detecting errors and the circuitry energy consumption is saved, while in the hybrid strategies, the relay still transmits, leading to extra circuitry energy consumption. When the S-D distance is increased ($r > 100\mathrm{m}$) in Fig.8, circuitry energy consumption becomes minor and the relay is able to help to reduce the overall energy per bit. The CF/DF based strategy has the best energy efficiency performance. Compared with the DF based one, the energy per bit is reduced by around 20% when $r = 1000\mathrm{m}$ for a target outage probability 10^{-2}. However, this savings is under the constraint that full CSI feedback to the relay requires negligible amount of energy. When partial CSI case is considered, in small transmission range scenario the energy per bit of the *MHS* strategy is slightly better than that of *HCD*. The reason is that based on the partial CSI feedback, the *MHS* strategy is able to shut down both the source and the relay during phase 2 when the channel quality is very low to save energy, especially the circuitry energy. In the long range case, *MHS* strategy performs close to the perfect CSI case. Another interesting observation is that, without the direct S-D link, the energy consumption of *DFsms* is even higher than direct transmission. This can be easily understood: now the relay is close to the destination and the S-R link is weak. Thus *DFsms* suffers from significant loss in the AR. To achieve the target outage probability, the power of the source and relay should be increased and the increased energy consumption outweighs the savings of using the relay. It should be noted that the results in this work have been provided for a stringent target outage probability of 10^{-2}, if the systems is allowed to operate at a lower outage level around 10^{-1}, more energy savings can be expected.

VII. Conclusions and Future Works

In this paper, we developed novel H^2-ARQ-Relaying strategies, where flexible approaches are taken to combine different forwarding strategies, enabling the system to efficiently and dynamically adapt itself to the variations of the channel. The CF/DF based strategy shows significant improvement in terms of throughput and energy consumption. However, it should be emphasized that compared with direct transmission, the cooperative strategy only shows improved energy efficiency when the distance between the destination and the source is large. The main advantage of the proposed hybrid CF/DF or AF/DF based strategy is that it can achieve either transmit or receive diversity. Moreover, since the global CSI is required at the relay for CF, an effective and practical method of quantized CSI feedback is proposed. It imposes very small feedback overhead. The modified strategy is able to approach the performance of the original one while keeping the CSI feedback overhead below practically viable limits. Another issue we need to emphasize is the power model employed. For a meaningful analysis we need to strike the right balance between realistic modeling and mathematical tractability and we believe that with current model the more important aspects of realistic energy consumption are well captured. We will try to use more complicated and realistic models in our future work.

Generally speaking, in addition to the channel circumstances described in this paper, there are many other scenarios that are worth investigating. Firstly, some practical limitations do apply. For instance, in case of DF based relaying, the relay needs to be aware of all necessary parameters that an intended receiver should know in order to decode the message before forwarding it. Furthermore, the relay may need more information regarding the second hop channel, in order to devise an energy and spectrum efficient forwarding strategy. However, in most of the practical system scenarios these requirements can be met at the expense of increased signaling overhead and the relaying schemes can still be used to improve the overall system performance. Secondly, the feedback might not be error-free and this more realistic assumption will affect the performance of the schemes. Finally, the relaying scheme can be more flexible by switching between more operational modes such as silent mode, AF, DF, CF and even quantize-and-forward [6] to achieve an optimal trade-off between the throughput, complexity, energy consumption and signaling overhead. It would be of great interest to examine these issues in the future.

Appendix A

For the AF scheme, when the source is allowed to transmit during phase 2, the received signal at the destination takes the form as $\mathbf{Y}_d = \mathbf{C}X_s^1 + \mathbf{B}\mathbf{Z}$,where

$$\mathbf{Y}_d = \left(Y_d^1 \ Y_d^2\right)^T, \mathbf{Z} = \left(Z_r^1 \ Z_d \ Z_r^2\right)^T$$

$$\mathbf{C} = \left(\frac{c_0}{\sqrt{K_t} d_0^{\zeta/2}} \quad \frac{\beta_r c_1 c_2}{K_t (d_1 d_2)^{\zeta/2}} + \frac{c_0}{\sqrt{K_t} d_0^{\zeta/2}} \right)^T$$

$$\mathbf{B} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\beta_r c_2}{\sqrt{K_t}(d_2)^{\zeta/2}} & 0 & 1 \end{pmatrix}$$

and β_r is the amplification function, subject to the power constraint $\beta_r \leq \sqrt{P_r / \left(\frac{\gamma_1 n_w E_{b,s}^1}{\alpha T K_t d_1^\zeta} + N_0 B\right)}$.Using vector results in [27], the AR to the destination is

$$I(X_s^1; \mathbf{Y}_d) = B \log_2 \det \left(\mathbf{I}_2 + \left(\mathbf{C}\mathbf{C}^H\right) \left(\mathbf{B} E\left(\mathbf{Z}\mathbf{Z}^H\right) \mathbf{B}^H\right)^{-1} \right),$$

where $\mathbf{I}_2$ is a 2-by-2 identity matrix and the covariance matrix $\mathbb{E}\left\{\mathbf{Z}\mathbf{Z}^H\right\} = \mathbf{I}_3 N_0 B$. Hence, we have (11), where θ_i is the phase of the channel gain c_i, which follows uniform distribution in the region $[-\pi, \pi)$. Let $\theta = \theta_0 - \theta_1 - \theta_2$, the AR should be averaged with respect to θ as (12). According to [6], the trivial upper bound is shwon in (13). $R^{(5)}$ can be maximized by maximizing β_r when $\frac{\gamma_1}{d_1^\zeta} \leq \frac{\gamma_0}{d_0^\zeta}$, and 0 otherwise. Note that we need 1 bit feedback from the destination to notify the relay this information to optimize β_r.

Appendix B

The AR to the destination for different frame categories except category 5 can be obtained based on the results in [12]. In category 1, CF is conducted and the AR is [12],

$$\begin{aligned} R^{(1)} = & \\ & \alpha B \log_2 \left(1 + \frac{\gamma_0 n_w E_{b,s}^1}{\alpha T N_0 B K_t d_0^\zeta} + \frac{\gamma_1 n_w E_{b,s}^1}{\alpha T (N_0 B + \sigma_W^2) K_t d_1^\zeta}\right) \\ & + \bar{\alpha} B \log_2 \left(1 + \frac{\gamma_0 n_w E_{b,s}^2}{\bar{\alpha} T N_0 B K_t d_1^\zeta}\right), \end{aligned} \tag{14}$$

where σ_W^2 is the compression noise and $\bar{\alpha} = (1 - \alpha)$. The minimum number of transmitted bits, such that the compressed signal can be recovered with the smallest distortion, is given as [6], [12],

$$\begin{aligned} s &= \bar{\alpha} B T \left(H(W|Y_d^1) - H(W|Y_r)\right) \\ &= \bar{\alpha} B T \Bigg(\log_2 \left(\frac{\frac{\gamma_1 n_w E_{b,s}^1 N_0 B}{\alpha T K_t d_1^\zeta}}{\frac{\gamma_0 n_w E_{b,s}^1}{\alpha T K_t d_0^\zeta} + N_0 B} + N_0 B + \sigma_W^2 \right) \\ &\quad - \log_2 \left(\sigma_W^2\right) \Bigg), \end{aligned} \tag{15}$$

where W is an auxiliary random variable. The achievable rate in the R-D link is

$$C_0 = Blog_2 \left(1 + \gamma_2 \frac{\frac{P_r}{K_t d_2^\zeta}}{N_0 B + \frac{\gamma_0 n_w E_{b,s}^2}{\bar{\alpha} T K_t d_0^\zeta}} \right).$$

We have $s \leq \bar{\alpha} T C_0$. Thus σ_W^2 can be minimized as

$$\sigma_W^2 = \frac{\frac{\gamma_1 n_w E_{b,s}^1 N_0 B}{\alpha T K_t d_1^\zeta} + N_0 B \left(\frac{\gamma_0 n_w E_{b,s}^1}{\alpha T K_t d_0^\zeta} + N_0 B\right)}{\left(2^{\frac{\bar{\alpha} C_0}{\alpha B}} - 1\right) \left(\frac{\gamma_0 n_w E_{b,s}^1}{\alpha T K_t d_0^\zeta} + N_0 B\right)} \tag{16}$$

when equality holds. For category 2, 3 and 4, the ARs are

$$R^{(2)} = \alpha B \log_2 \left(1 + \frac{\gamma_0 n_w E_{b,s}^1}{\alpha T N_0 B K_t d_0^\zeta}\right), R^{(3)} = R^{(2)} + R^{(4)}$$

$$R^{(4)} = \bar{\alpha} B \log_2 \left(1 + \frac{\gamma_0 n_w E_{b,s}^2}{\bar{\alpha} T N_0 B K_t d_0^\zeta} + \frac{\gamma_2 n_w E_{b,r}}{\bar{\alpha} T N_0 B K_t d_2^\zeta}\right),$$

$$I(X_s^1;\mathbf{Y}_d)=B\log_2\left(1+\frac{\gamma_0 P_s}{N_0BK_td_0^\zeta}+\frac{P_s|\frac{\beta_rc_1c_2e^{-j(\theta_0-\theta_1-\theta_2)}}{K_t(d_1d_2)^{\zeta/2}}+\frac{c_0}{\sqrt{K_t}d_0^{\zeta/2}}|^2}{N_0B\left(1+\frac{\beta_r^2\gamma_2}{K_td_2^\zeta}\right)}\right), \tag{11}$$

$$R_{AF}=\frac{1}{2\pi}\int_{-\pi}^{+\pi}I(X_s^1;\mathbf{Y}_d)\mathrm{d}\theta=R^{(5)}+B\log_2\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{\frac{4\beta_r^2\gamma_0\gamma_1\gamma_2}{K_t^3(d_1d_2d_3)^\zeta}}{\left(\left(1+\frac{\gamma_0n_wE_{b,s}^1}{\alpha TN_0NK_td_0^\zeta}\right)\left(1+\frac{\beta_r^2\gamma_2}{K_td_2^\zeta}+\frac{\beta_r^2\gamma_1\gamma_2}{K_t^2(d_1d_2)^\zeta}\right)\right)^2}}\right), \tag{12}$$

$$R^{(5)}=B\log_2\left(1+\frac{\gamma_0n_wE_{b,s}^1}{\alpha TN_0BK_td_0^\zeta}+\frac{n_wE_{b,s}^1\left(\frac{\beta_r^2\gamma_1\gamma_2}{K_t^2(d_1d_2)^\zeta}+\frac{\gamma_0}{K_td_0^\zeta}\right)}{\alpha TN_0B\left(1+\frac{\beta_r^2\gamma_2}{K_td_2^\zeta}\right)}\right). \tag{13}$$

respectively.

When the modified scheme is applied, the AR has the same expression as (14) by replacing σ_w^2 with $\bar{\sigma}_W^2$. If γ_0 and γ_2 are in the region $[a_2,a_3]$ and $[a_1,+\infty)$ respectively, we can carefully choose the transmission rate of the relay as

$$\frac{s}{\bar{\alpha}T}\le\frac{\bar{s}}{\bar{\alpha}T}\le\bar{C}_0$$
$$\bar{C}_0=B\log_2\left(1+a_1\frac{\frac{P_r}{K_td_2^\zeta}}{N_0B+\frac{a_3n_wE_{b,s}^2}{\bar{\alpha}TK_td_0^\zeta}}\right)\le C_0,$$

where $\bar{s}$ can be calculated as (15) by replacing γ_0 with a_2. To calculate $\bar{\sigma}_W^2$, we can replace γ_0 and C_0 with a_3 and $\bar{C}_0$ in (16).

Appendix C

The average consumed energy per bit in five different cases for the hybrid schemes are

$$E_{b,1}=\sum_{M=1}^{N}P_{CS}(M)ME_{b,CF},$$
$$E_{b,2}=\sum_{K=1}^{N}P_{BS}(K)\left((K-1)E_{b,CF}+\frac{E^1}{n_w}\right),$$
$$E_{b,3}=\sum_{L=1}^{N}\sum_{K=L}^{N}P_{DS}(L,K)$$
$$\cdot\left((L-1)E_{b,CF}+E_{b,DF}+(K-L)\frac{E_{DF}^2}{n_w}\right),$$
$$E_{b,4}=\Pr(F_a)NE_{b,CF},$$
$$E_{b,5}=$$
$$\sum_{L=1}^{N}P_{DF}^{out}(L)\left((L-1)E_{b,CF}+E_{b,DF}+(N-L)\frac{E_{DF}^2}{n_w}\right),$$

Definition of ρ_D and ρ_C are

$$\rho_D=\frac{\bar{\alpha}}{\mathbb{E}\{\mathcal{T}\}}\Bigg(\sum_{K=1}^{N}\big((P_{BS}(K)+\Pr(DF_{N-K+1,K}))(K-1)$$
$$+P_{CS}(K)K\big)+\sum_{L=1}^{N}\sum_{K=L}^{N}P_{DS}(L,K)(L-1)+\Pr(F_a)N\Bigg)$$
$$\rho_C=\frac{\bar{\alpha}}{\mathbb{E}\{\mathcal{T}\}}\Big(\sum_{K=1}^{N}\Pr(DF_{N-K+1,K})(N-K+1)$$
$$+\sum_{L=1}^{N}\sum_{K=L}^{N}P_{DS}(L,K)(K-L+1)\Big)$$

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 247733 - project EARTH.

References

[1] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans. Signal Process.*,vol. 52, no. 2, pp. 362-371, Feb. 2004.

[2] T. E. Hunter and A. Nosratinia, "Diversity through coded cooperation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 283-289, Feb, 2006.

[3] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[4] Z. Liu, V. Stankovic and Z. Xiong, "Wyner-Ziv coding for the half-duplex relay channel," in *Proc. IEEE Int. Conf. on ASSP*, Philadelphia, PA, USA, Mar. 2005, pp. 1113-1116.

[5] R. Hu and J. Li, "Practical compress-forward in user cooperation: Wyner-Ziv cooperation," in *Proc. ISIT*, Seattle, USA, Jul. 2006, pp. 489-493.

[6] M. Katz and S. Shamai, "Relay protocols for two co-located users," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2329-2344, Jun. 2006.

[7] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong, "Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: high-rate performance analysis and code design," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4358-4379, Oct. 2006.

[8] M. Yuksel, and E. Erkip, "Multi-antenna cooperative wireless systems: A diversity-multiplexing tradeoff perspective," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3371-3393, Oct. 2007.

[9] C. T. K. Ng and A. Goldsmith, "The Impact of CSI and power allocation on relay channel capacity and cooperation strategies," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5380-5389, Dec. 2008.

[10] C. T. K. Ng, N. Jindal, A. Goldsmith, and U. Mitra, "Capacity gain from two-transmitter and two-receiver cooperation," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3822-3827, Oct. 2007.
[11] P. Rost and G. Fettweis, "Analysis of a mixed strategy for multiple relay networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 1, pp. 174-189, Jan. 2009.
[12] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2020 - 2040, Jun 2005.
[13] Y. Li, B. Vucetic, Z. Chen, and J. Yuan, "An improved relay selection scheme with hybrid relaying protocols," in *Proc. IEEE GLOBECOM*, New Orleans, USA, Nov. 2007, pp. 3704-3708.
[14] Y. Li, B. Vucetic, and J. Yuan, "Distributed turbo coding with hybrid relaying protocols," in *Proc. IEEE 19th Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC'08)*, Cannes, France, Sept. 2008, pp. 1-6.
[15] S. Serbetli and A. Yener, "Power allocation and hybrid relaying strategies for F/TDMA Ad Hoc networks," in *Proc. IEEE Int. Conf. on Communications (ICC'06)*, Istanbul, Turkey, Jun. 2006, pp. 1562-1567.
[16] L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, Third Edition, London, Academic Press, 2003.
[17] T. Tabet, S. Dusad, and R. Knopp, "Diversity-multiplexing-delay tradeoff in half-duplex ARQ relay channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3797 - 3805, Oct. 2007.
[18] L. Dai and K. Letaief, "Throughput maximization of ad-hoc wireless networks using adaptive cooperative diversity and truncated ARQ," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1907-1918, Nov. 2008.
[19] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE J. Select. Areas Commun. (JSAC)*, vol. 23, no. 1, pp. 7-18, Jan. 2005.
[20] R. Hoshyar, and R. Tafazolli, "Performance evaluation of HARQ schemes for cooperative regenerative relaying," in emphProc. IEEE Int. Conf. on Communications (ICC'09), Dresden, Germany, Jun. 2009, pp. 1-6.
[21] I. Stanojev, O. Simeone, Y. Bar-Ness, and D. Kim, "Energy efficiency of non-collaborative and collaborative hybrid-ARQ protocols," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 326-335, Jan. 2009.
[22] K. Pang, Y. Li and B. Vucetic, "An improved hybrid ARQ scheme in cooperative wireless networks," in emphProc. 68th IEEE semiannual Vehicular Technology Conf. (VTC'08-Spring), Singapore, Singapore, Sept. 2008, pp. 1-5.
[23] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1971-1988, Jul. 2001.
[24] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1-10, Jan. 1976.
[25] A. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inform. Contr.*, vol. 38, pp. 60-80, Jul. 1978.
[26] K. Singhal, J. Vlach, and M. Vlach, "Numerical inversion of multidimensional Laplace transform," in *Proc. of the IEEE*, vol. 63, no. 11, pp. 1627 - 1628, Nov. 1975.
[27] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europe Trans. Telecommun.*, vol. 10, pp. 585-596, Nov.-Dec. 1999.
[28] S. Cui, A. Goldsmith and A. Bahai, "Energy-Constrained Modulation Optimization", *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349-2360, Sep. 2005.

Yinan Qi received the B.Sc degree in Electronics and Information Theory and the M.Sc degree in mobile communications from Peking University, Beijing, China, in 2000 and 2003, respectively, and the Ph.D degree in Electronic Engineering from University of Surrey, Surrey, United Kingdom. From 2003 to 2004, he has been a research scholar in National University of Singapore. He is currently a research fellow at the Centre for Communication Systems Research (CCSR) at the University of Surrey. His main research interests include cooperative communications, coding, analysis and systematic modelling of future cellular systems, and green communications.

Reza Hoshyar received his B.S. degree in communications engineering and his M.S. and Ph.D. degrees, both in mobile communications, from Tehran University, Tehran, Iran, in 1991, 1996, and 2001, respectively. He received the top and second rank awards for his B.S. and M.S. degrees from Tehran University. From 2002 to September 2010, he has been a Research Fellow and then Senior Research Fellow in the Mobile Communications Research Group, Centre for Communication Systems Research (CCSR), University of Surrey, Surrey, United Kingdom During this period, he was been actively involved in several European projects focusing on advanced digital signal processing, coding and modulation, and scheduling techniques for relaying, and cooperative communications in OFDM/OFDMA systems. Currently he is a system design engineer at National Semiconductor research lab, California, USA. His current focus is on power line communications and high-speed 60 GHz short range wireless communication.

Muhammad Ali Imran (M'03) received his B.Sc. degree in Electrical Engineering from University of Engineering and Technology Lahore, Pakistan , in 1999, and the M.Sc. and Ph.D. degrees from Imperial College London, UK, in 2002 and 2007, respectively. He secured first rank in his BSc and a distinction in his MSc degree along with an award of excellence in recognition of his academic achievements confered by the President of Pakistan. He is currently a lecturer in the Centre for Communication Systems Research (CCSR) at the University of Surrey, UK. He has been actively involved in European Commission funded research projects ROCKET and EARTH, Mobile VCE funded project on fundamental capacity limits and EPSRC funded project India UK ATC. He is the principal investigator of EPSRC funded REDUCE project. His main research interests include the analysis and modelling of the physical layer, optimisation for the energy efficient wireless communication networks and the evaluation of the fundamental capacity limits of wireless networks.

Rahim Tafazolli (SM'09) is a professor and the Director of Centre for Communication Systems Research (CCSR), University of Surrey, United Kingdom. He has been active in research for more than 20 years and has authored and co-authored more than 360 papers in refereed international journals and conferences. Professor Tafazolli is a consultant to many mobile companies, has lectured, chaired, and been invited as keynote speaker to a number of IEE and IEEE workshops and conferences. He has been Technical Advisor to many mobile companies, all in the field of mobile communications. He is the Founder and past Chairman of IEE International Conference on 3ed Generation Mobile Communications. He is Chairman of EU Expert Group on Mobile Platform (e-mobility SRA) and Chairman of Post-IP working group in e-mobility, past Chairman of WG3 of WWRF. He is nationally and internationally known in the field of mobile communications.