QoS-Constrained Energy Efficiency of Cooperative ARQ in Multiple DF Relay Systems

Yulin Hu, James Gross, and Anke Schmeink

Abstract

In this paper, we study the higher-layer performance as well as the "throughput, delay, energy consumption" trade-off problem for multi-relay assisted cooperative automatic repeat request (C-ARQ) protocols. We study a practical scenario where only the average channel state information is available at the source and relays. We consider four multi-relay C-ARQ protocols, and derive closed-form expressions for the transmission delay distribution, the energy consumption and the higher-layer queuing performance. Furthermore, we analyze the QoS-constrained energy efficiency performances of the protocols. Our analysis is validated by simulations. In addition, we evaluate the system performance under these C-ARQ protocols and for different topologies. We conclude several guidelines for the design of efficient C-ARQ protocols. Finally, a simple extension of the studied C-ARQ protocols is proposed, which improves the QoS-constrained energy efficiency by 4%.

Index Terms

C-ARQ, cooperative communications, DF relaying, effective capacity, energy efficiency, multiple relays.

I. Introduction

Relaying is a fundamental technique to improve the performance of a wireless link. It is based on deploying a third transceiver such that the destination likely receives a much stronger signal forwarded by the relay in comparison to the direct transmission by the source. This leads to an improvement in the coverage and the transmission reliability of the link [1]–[3]. Two major relaying principles have been established: In amplify-and-forward (AF) the analog samples of the received signal are simply amplified and passed on. No decoding of the baseband samples is performed. Contrarily, if the relay decodes the signal first and then forwards it, this is referred to as decode-and-forward (DF) relaying. Obviously, this involves processing of the received signal by the relay nodes' physical layer, which is not the case for AF relaying.

Based on these very mature relaying techniques, more recently cooperative communications has attracted a lot of research interest. In cooperative communications the fundamental assumption is that there are multiple potential relays. Hence, by spending some control overhead, the best positioned relay or a set of best positioned relays are identified to (jointly) forward the signal to the destination. This boosts the system performance as it exploits cooperative diversity [4] which is a form of spatial/multi-user diversity. Many different techniques have been proposed how to specifically exploit this cooperative diversity. If for example the source has instantaneous channel state information (CSI), [5] proposes to apply distributed space-time block coding schemes at the source and relays. The paper shows that this leads to a higher spectral efficiency in comparison to pure relaying. Under the same assumption of CSI at the source, [6], [7] let each relay transmit separately for a certain time fraction. By optimizing the transmission duration of each relay, the outage probability [6] and capacity [7] can then be improved in comparison to pure relaying without cooperation. In general, cooperative communications is well known to effectively increase the throughput and reliability in case that multiple relays are present [8], [9]. However, this fundamentally comes at the price of an increased complexity, energy consumption and hardware cost.

On top of the above mentioned cooperative communications, there is the possibility to further improve the reliability by combining the cooperation principle with automatic repeat request (ARQ) protocol mechanisms. These combined schemes are referred to as cooperative-ARQ (C-ARQ) protocols. A C-ARQ protocol lets relay(s) retransmit the data packet to the destination when the initial transmission fails instead of having the source retransmit the packet. Hence, the reliability of the system is improved due to cooperative diversity. Furthermore, if the destination can combine the received signals of the initial transmission and retransmissions, the reliability is additionally boosted by the combined channel qualities. It has been shown over the last few years [10]–[15] that - from a pure physical-layer perspective - C-ARQ protocols have an outstanding throughput and reliability performance in comparison to pure relaying and cooperative communications.

Nevertheless, as with cooperative communications in general, C-ARQ protocols lead to a higher system complexity and energy requirement. Few works attempt to address this trade-off. For instance, from a pure physical-layer perspective [16] shows that C-ARQ has higher energy efficiency than the non-cooperative ARQ schemes under certain channel conditions. In addition, [17], [18] show that the energy efficiency of C-ARQ has a strong dependence on the number of retransmissions. Apart from the additional energy consumption, C-ARQ protocols also lead to a more complex framing-structure which likely introduces additional stochastic delays to the information flow from the source to the destination. While studies exist which have been investigating the delay of C-ARQ protocols [14], [19], [20], they are typically limited to the physical layer and do not take higher-layer queuing effects into account. However, a joint delay analysis is particularly relevant for delay-sensitive higher-layer applications. It typically comes with a limited arrival rate (instead of the often studied full-buffer assumption) but also with constraints regarding the reception delay and the violation probability of this target delay (for instance, for voice flows a typical scenario is to keep the delay below 200 ms while the delay violation probability is around 5% depending on the used codec). Summarizing, C-ARQ protocols actually introduce a tripartite trade-off among throughput, delay and energy consumption. However, to date this trade-off is only understood from a physical layer perspective and a more complete analysis of this trade-off with respect to higher-layer performance is missing.

In the following, we present such an analysis of this trade-off. Our analysis is based on the *effective capacity*, which is a well-known performance metric that accounts for transmission and queuing effects in (wireless) networks [21]. The effective capacity characterizes the (maximum) arrival rate of a flow to a queuing system and relates the stochastic characterization of the service of the queuing system to the queue-length or delay constraints of the flow. It has been widely applied to the analysis of wireless systems, however, no effective capacity analysis of C-ARQ protocols has been published so far to the best of our knowledge. In addition to the effective capacity analysis of C-ARQ protocols, we further evaluate the associated energy efficiency, by considering the ratio between the effective capacity and the energy consumption of the considered C-ARQ protocols. In the following we refer to this ratio as the effective energy efficiency (EEE).

Based on these performance metrics, we study four C-ARQ protocols representing different choices with respect to the integration of the direct link into the protocol as well as the dynamic choice of relays for forwarding. These different protocols are considered under quite practical system assumptions where the channel knowledge is not perfect and ARQ feedback overhead is considered. Based on this set-up we provide the following contributions: (1) A closed-form expression of the outage probability of multi-DF-relay two-hop transmission is derived. (2) For the above four C-ARQ protocols, closed-form expressions of the transmission delay distribution, the energy consumption and the approximation of effective capacity are derived and later on validated by simulation. In addition, we derive the outage probabilities of the C-ARQ protocols with truncations. (3) By numerical analysis we conclude a set of guidelines for the design of efficient C-ARQ protocols as follows: i. Direct link transmission can usually be ignored by the C-ARQ protocols unless a specific topology of the nodes is given. ii. There exists an optimal number of relays deployed in the system when maximizing the EEE. iii. Dynamic relay selection should be enabled as it leads to a better efficiency. iv. Overhead should be spent to acquire the exact number of relays present in the system.

The rest of the paper is organized as follows. Section II describes the system model as well as four investigated C-ARQ protocols. We introduce the EEE measure in Section III and investigate the EEE performance for the C-ARQ protocols in Section IV. Section V validates our theoretical analysis and evaluates the C-ARQ performance. Moreover, in Section V we propose a minor C-ARQ protocol variant which achieves a higher EEE. We conclude the paper in Section VI.

II. System Model and studied C-ARQ Protocols

In this section, we first describe the system, and then present the models for channel behavior and energy consumption. At last, the investigated C-ARQ protocols are introduced.

A. General System Description

We consider a DF relaying scenario which contains a source S, a destination D and J DF relays $\mathrm{R}_1, \mathrm{R}_2, ..., \mathrm{R}_J$ as schematically shown in Fig. 1. The distances between relays are assumed to be much shorter than the distances either from the source to relays or from the relays to the destination.

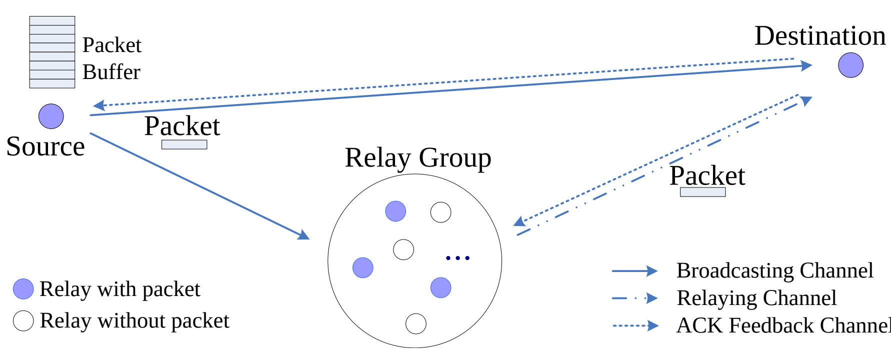

Fig. 1. Example of the considered multiple relay system scenario.

The entire system operates in a slotted fashion where time is divided into frames of length T_f. Depending on the exact operational rules of the protocol, these frames are either broadcasting frames, forwarding/relaying frames or feedback frames with acknowledgment (ACK) or negative-ACK (NACK). In a broadcasting frame, the source takes a fixed amount of data ρ out of its buffer and then broadcasts this packet to the destination and relays. Due to the random channel behavior, a random amount of relays is able to decode the broadcasted packet successfully and will forward the packet to the destination in the subsequent forwarding frame. Once the destination received the packet (either successfully or not) it gives feedback through an ACK/NACK frame. We assume this feedback to be error-free. Due to C-ARQ protocols, a packet is retransmitted by relays if the initial transmission fails. The destination is assumed to apply maximal ratio combining (MRC) during the retransmissions[1] (the combination of MRC and ARQ is also known as Hybrid-ARQ II [22]). This results in a significantly stronger signal at the destination.

At the source, a constant data flow originates with arrival rate r bits per frame. The transmission of this data is subject to QoS requirements $\{d, \mathrm{Pr_d}\}$ (e.g., for audio or video applications) where d stands for a maximum tolerable delay and $\mathrm{Pr_d}$ denotes the delay violation probability. Data that cannot be transmitted immediately at the source is put into a first-in-first-out buffer. Note finally that relays do not queue data, they only store the currently transmitted data packet.

B. Channel, SNR and Energy Consumption Models

Channels are considered to experience Rayleigh blockfading where they are invariant during one frame but vary independently from one frame to the next. We denote the channel gains from the source to the destination, from the source to relay j and from relay j to the destination in frame i by $h^2_{\mathrm{S,D},i}$, $h^2_{\mathrm{S},j,i}$ and $h^2_{j,\mathrm{D},i}$ $(j=1,2,...,J)$. The corresponding average channel gains are $\bar{h}^2_{\mathrm{S,D}}$, $\bar{h}^2_{\mathrm{S},j}$ and $\bar{h}^2_{j,\mathrm{D}}$. In addition, we denote P_{tx} as the transmit power at the source and each relay, and denote P_{fb} as the transmit power at the destination for feedback. The noise power is denoted by σ^2. Also, we assume no interference to be present. Hence, the instantaneous signal-to-noise ratio (SNR) from the source to the destination in frame i is $\gamma_{\mathrm{S,D},i}{=}P_{\mathrm{tx}}h^2_{\mathrm{S,D},i}/\sigma^2$. Similarly, the instantaneous SNRs in frame i from the source to relay j and from relay j to the destination are $\gamma_{\mathrm{S},j,i}=P_{\mathrm{tx}}h^2_{\mathrm{S},j,i}/\sigma^2$ and $\gamma_{j,\mathrm{D},i}{=}P_{\mathrm{tx}}h^2_{j,\mathrm{D},i}/\sigma^2$. As multiple relays forward the same packet in a forwarding frame, the destination obtains a joint instantaneous SNR as the sum of the instantaneous SNRs of links from these relays as $\gamma_{\xi_i,\mathrm{D},i}{=}\sum_{R_j\in\xi_i}\gamma_{j,\mathrm{D},i}$, where ξ_i is the set of relays which forwards/reforwards the packet during frame i.

Moreover, by applying MRC over the initially transmitted and retransmitted signals a cumulative joint SNR is obtained at the destination, which refers to the sum of the joint instantaneous SNRs in the initial

[1]Compared with the SNR of a relaying frame, the SNR of a direct link is significantly lower and can be treated as negligible. To simplify the derivation, we only consider combining the signals transmitted/retransmitted by relays.

transmission and retransmissions. This cumulative joint SNR is given by $\gamma_{\text{C-ARQ}} = \sum_{i\in\pi} \gamma_{\xi_i,\mathrm{D},i}$, where π is a set of indices of frames during which the packet is forwarded and re-forwarded by relays.

Recall that we assume only the average CSI to be available at the source and each relay. To account for this, we consider an error model for the transmitted data at relays and the destination. Given an SNR γ, at most $T_{\mathrm{f}} B \log_2 (1+\gamma)$ bits can be conveyed correctly per frame, where B is the bandwidth. Hence, a currently transmitted/retransmitted packet of size ρ is successfully received if the instantaneous/cumulative SNR is above the threshold $\gamma^* = 2^{\rho/BT_{\mathrm{f}}} - 1$. On the other hand, if the instantaneous/cumulative SNR is lower than the threshold, an outage occurs which leads to a retransmission by the corresponding C-ARQ protocol.

Regarding the energy consumption, we adopt the energy consumption model introduced in [23], [24]. In general, we consider two effects that contribute differently to the energy consumption. The first one is the energy consumption due to transmission. The other one is the basic energy consumption which is spent for signal processing, battery backup, cite cooling and so on. We assume that the basic energy consumption at the source, each relay and the destination are the same while the corresponding power is denoted by P_{c}.

C. Multi-relay C-ARQ Protocols

C-ARQ is a link-level ARQ protocol that exploits relay(s) for retransmission. In particular, a multi-relay C-ARQ protocol requires multiple relays to retransmit a packet if the initial transmission or the previous retransmission fails. In this work, we study four multi-relay C-ARQ protocols which have different behaviors regarding the trade-off between throughput, delay and energy consumption. The protocols differ in the way they deal with the initial transmission, and in the way they compose the group of retransmission relays.

1) Proactive/Reactive Behavior: A proactive protocol activates relays to proactively forward the packet in the initial transmission as well as retransmissions. Hence, the initial transmission of a proactive protocol is a relay assisted transmission consisting of a broadcasting frame and a relaying frame. Contrarily, a reactive protocol uses a direct link transmission as the initial transmission and requires relays to forward the packet only if the direct link transmission fails. In other words, reactive protocols let the destination try to decode the packet by a one-frame direct link transmission. If the direct link transmission succeeds, as shown at the top right of Fig. 2 the transmission delay of this packet is only two frames while saving transmit power at the relays. However, if the direct link transmission fails, the reactive protocol introduces

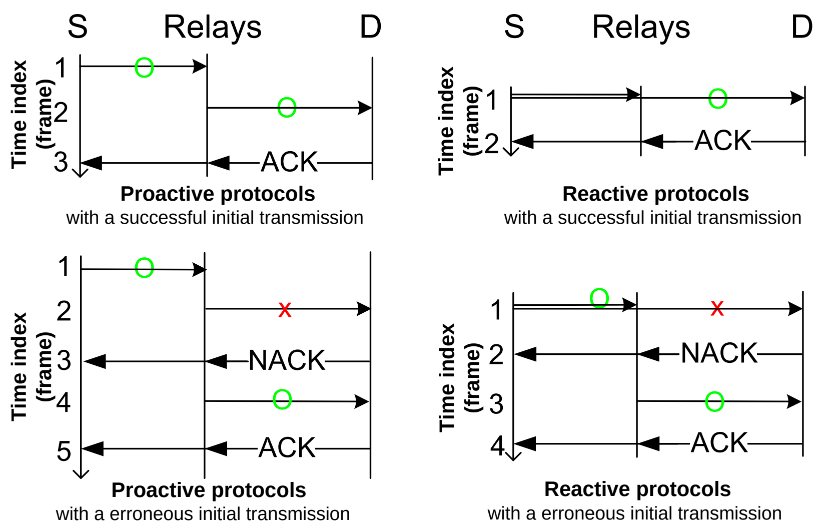

Fig. 2. The difference between proactive protocols and reactive protocols.

an additional delay and consumes additional energy.

2) Static/Dynamic Relaying Behavior: The second category of C-ARQ protocols concerns the number of relays participating during the retransmission frames. After the source has broadcasted a packet, not necessarily all the relays decode the packet successfully due to channel fading. Let us call these relays which decode the packet from the source correctly as active relay set and call the rest as passive relay set. Obviously, only active relays are able to participate in the first retransmission. As shown in Fig. 3

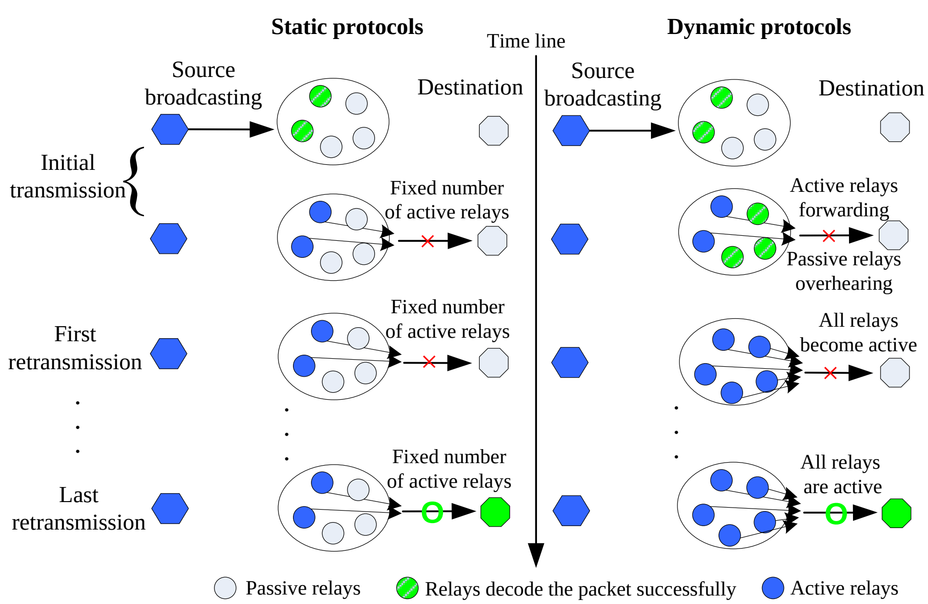

Fig. 3. The difference between static protocols and dynamic protocols.

(left), static protocols only exploit these active relays for retransmission and let the passive relays idle. Hence, the number of active relays under a static protocol is fixed during the retransmission process of a packet. On the other hand, dynamic protocols require all the passive relays to overhear the packet when the active relays forward the packet to the destination. Therefore, more relays might join the set of the active relays later on in the retransmission rounds, as shown in Fig. 3 (right). Regarding the trade-off between throughput, delay and energy consumption, a dynamic protocol likely has a higher throughput and shorter latency while consuming more energy due to a higher number of (initially passive) relays participating in the retransmission process after a while.

In the rest of the paper, we will investigate the performance of the four resulting combinations of the different categories: dynamic proactive (DP) protocol, dynamic reactive (DR) protocol, static proactive (SP) protocol and static reactive (SR) protocol, which are actually the combinations of the above two presented protocol categories.

III. Effective Energy Efficiency

In this section, we first briefly review the effective capacity framework, then we introduce a higher-layer energy efficiency metric called Effective Energy Efficiency (EEE) which captures the trade-off among throughput, delay and energy consumption. Finally, we analyze mathematically the EEE performances of the C-ARQ protocols.

A. *Effective Service Capacity*

The effective service capacity is an approximation for the queue length distribution of a queuing system with random service process. Denote the service process, i.e., the amount of bits that effectively leave the queue at frame i, as s_i, the cumulative service process is $S_i = \sum_{n=0}^{i} s_n$. Assume that the queue is stable as the average service rate is larger than the average arrival rate. Hence, the random queue length Q_i at frame i converges to the steady-state random queue length Q. To characterize the long-term statistics $\Pr\left\{Q\right\}$ of the queue length, the framework of effective service capacity gives us an upper bound $\Pr\left\{Q > x\right\} \leq K \cdot e^{-\theta^* \cdot x}$, where K is the probability that the queue is non-empty and θ^* is the so called QoS exponent. Based on [25], for a constant rate source with r bits per frame, the exponent θ^* has to fulfill the constraint $r < \Lambda\left(-\theta^*\right)/\theta^*$, where $\Lambda\left(\theta\right)$ is the log-moment generating function of the cumulative service process S_i defined as:

$$\Lambda\left(\theta\right) = \lim_{i \to \infty} \frac{1}{i} \log \mathrm{E}\left[e^{\theta \cdot \left(S_i - S_0\right)}\right]. \tag{1}$$

The ratio $\Lambda(-\theta)/\theta$ is called the effective service capacity. Denote by D_i the random queuing delay of the head-of-line bit during frame i. If the constant arrival rate at the source is r, with a queue length of $Q = q$, a current delay of the head-of-line bit is given by $D = q/r$. This yields the following approximation for the steady-state delay distribution:

$$\Pr\{D > d\} \leq K \cdot e^{-\theta^* \cdot r \cdot d}. \tag{2}$$

If s_i is independent and identically distributed (i.i.d.), a convenient simplification is to obtain the log-moment generating function via the law of large numbers. Hence, the effective service capacity is given by [26]:

$$\frac{\Lambda(-\theta)}{\theta} = \lim_{i\to\infty} \frac{1}{i \cdot \theta} \log \mathrm{E}\left[e^{-\theta \cdot s_i}\right] = \mathrm{E}[s_i] - \frac{\theta}{2}\mathrm{Var}[s_i]. \tag{3}$$

Therefore, the queuing performance of the system is determined by the mean and the variance of the random increment of the service process given by s_i. The maximum arrival rate at the source r^* that can be supported by the random service process for given QoS requirements $\{d, \mathrm{Pr_d}\}$ is obtained by combining the constraint $r < \Lambda(-\theta^*)/\theta^*$ and (2) (upper bounding K by 1):

$$r^*|_{\{d,\mathrm{Pr_d}\}} \approx \frac{\mathrm{E}[s_i]}{2} + \frac{1}{2}\sqrt{(\mathrm{E}[s_i])^2 + \frac{2\ln(\mathrm{Pr_d})}{d} \cdot \mathrm{Var}[s_i]}, \tag{4}$$

In the following, we call the maximum source rate as Maximum Sustainable Data Rate (MSDR) and refer to MSDR as a metric for the queuing performance. Based on (4), the major challenge for determining the MSDR is to obtain the mean and variance of the increment of the service process for the different C-ARQ protocols.

B. Energy Consumption of C-ARQ

Based on the energy model in Section II-B, the energy consumption contains two parts: the basic energy consumption and the energy consumption due to transmission. The average total consumed energy per frame is then obtained by (recall that K is the probability that the queue is non-empty):

$$\bar{\Psi} = K \cdot \bar{\Psi}_\mathrm{T} + (J+2) \cdot P_\mathrm{c} T_\mathrm{f}, \tag{5}$$

where $(J+2) \cdot P_\mathrm{c} T_\mathrm{f}$ is the basic energy consumption over the duration of one frame of all nodes including the source, the destination and J relays. Obviously, all the studied C-ARQ protocols have the same basic energy consumption. Furthermore, $\bar{\Psi}_{\mathrm{T}}$ is the average energy consumption over one frame due to data transmission (when the queue is non-empty). $\bar{\Psi}_{\mathrm{T}}$ depends on the C-ARQ protocol, as static and dynamic protocols have different numbers of relays participating in the retransmission process.

Furthermore, in general C-ARQ protocols require different numbers of broadcasting frames, relaying frames and ACK/NACK frames for transmitting/retransmitting a packet. From an asymptotic perspective, these different types of frames appear alternately with different probabilities. In other words, these frame types occupy different proportions over a longer time span of operation. For static relaying protocols, we denote by p_{S}, p_{R} and p_{D} the percentages of broadcasting frame, relaying frame (including initial transmission and retransmissions) and ACK/NACK frame. Therefore, the average transmit energy consumption per frame is obtained as $(p_{\mathrm{S}}P_{\mathrm{tx}} + \mathrm{E}[\Omega]p_{\mathrm{R}}P_{\mathrm{tx}} + p_{\mathrm{D}}P_{\mathrm{fb}})T_{\mathrm{f}}$, where Ω denotes the random number of active relays with expectation $\mathrm{E}[\Omega]$.

Under a dynamic protocol, the number of active relays varies in the retransmission process. Recall that under a dynamic protocol, initially some relays become active by decoding the broadcasted (from the source) packet successfully while the other relays (passive relays) become active later on by overhearing the forwarded packet from the initial active relays. As the distances between relays are assumed to be much shorter than the distances from the source to relays, all the links between relays are significantly more reliable than the source-relay links. Hence, all the passive relays are very likely to decode the packet successfully by overhearing one retransmission. Then, all the relays become active for retransmission. Therefore, the percentage of relaying frames of a dynamic protocol can be modeled by $p_{\text{i-R}}$ and $p_{\text{a-R}}$ which are the percentages of initial active relays retransmitting frames and all relays retransmitting frames. Hence, the average transmit energy consumption[2] (per frame) of dynamic protocols can be modeled as $(p_{\mathrm{S}}P_{\mathrm{tx}} + \mathrm{E}[\Omega]p_{\text{i-R}}P_{\mathrm{tx}} + Jp_{\text{a-R}}P_{\mathrm{tx}} + p_{\mathrm{D}}P_{\mathrm{fb}})T_{\mathrm{f}}$, where Ω is the number of the initial active relays.

Summarizing, the average energy consumption due to transmission of the C-ARQ protocols is given by:

$$\bar{\Psi}_{\mathrm{T}}=\begin{cases} \left(p_{\mathrm{S}}+\mathrm{E}[\Omega]p_{\mathrm{R}}\right)P_{\mathrm{tx}}T_{\mathrm{f}}+p_{\mathrm{D}}P_{\mathrm{fb}}T_{\mathrm{f}}, & \text{static;} \\ \left(p_{\mathrm{S}}+\mathrm{E}[\Omega]p_{\text{i-R}}+Jp_{\text{a-R}}\right)P_{\mathrm{tx}}T_{\mathrm{f}}+p_{\mathrm{D}}P_{\mathrm{fb}}T_{\mathrm{f}}, & \text{dynamic.} \end{cases} \tag{6}$$

As each node is assumed to operate with the same transmit power, $\bar{\Psi}$ is fully determined by the variables p_{S}, p_{R} (or $p_{\text{i-R}}$ and $p_{\text{a-R}}$), p_{D} and $\mathrm{E}[\Omega]$. We will derive these variables in Section IV-C for each protocol.

[2]This expression is rather realistic under the assumed scenario where the relays are very close to each other. It will be an upper bound for the energy consumption if the relays are more spread out.

C. Effective Energy Efficiency (EEE)

In this work, the (higher-layer) QoS-constrained energy efficiency of the C-ARQ protocols are investigated by considering the ratio between the MSDR (under QoS constraints $\{d, \mathrm{Pr_d}\}$) and the related average energy consumption. We refer to this ratio as EEE. Recall that the MSDR is the maximum arrival rate at the source that can be supported by the system under certain QoS constraints and is derived by upper bounding K by 1. Hence, the related energy consumption of the system when the arrival rate at the source equals the MSDR can be obtained based on (5) by considering $K = 1$. Therefore, the EEE is given based on (4) and (5) by:

$$\Phi = \frac{0.5\,\mathrm{E}\left[s_i\right] + 0.5\sqrt{\left(\mathrm{E}\left[s_i\right]\right)^2 + \frac{2\ln(\mathrm{Pr_d})}{d}\cdot \mathrm{Var}\left[s_i\right]}}{\bar{\Psi}_{\mathrm{T}}(p_{\mathrm{S}}, p_{\mathrm{R}}, p_{\mathrm{D}}, \mathrm{E}[\Omega]) + (J+2)\cdot P_{\mathrm{c}} T_{\mathrm{f}}}. \tag{7}$$

The EEE differs from physical-layer energy efficiency metrics in that it directly demonstrates how effective the consumed energy is on providing the delay-constrained service (where we also take the queuing performance of the system into account). For example, let us consider a system which consumes low energy but supports a high data rate for a service with certain delay constraint. However, the delays of most of the transmitted packets violate the constraint of the service. Therefore, the EEE of the system is rather low while the physical-layer energy efficiency has a high value. In addition, if system A has a higher EEE than system B, this means that system A is able to use less energy to support the same service (with the same data arrival rate and the same QoS requirements) compared to system B or that while consuming the same energy system A is able to transmit more bits under the QoS requirements.

As can be concluded from (7), the major challenge to obtain the EEE becomes the derivation of the mean and variance of the service process increments s_i together with the energy consumption variables of $\bar{\Psi}_{\mathrm{T}}$. We deal with the derivation of these variables in the next section.

IV. EEE PERFORMANCE OF C-ARQ PROTOCOLS

Under C-ARQ protocols, the transmission (including initial transmission and retransmissions) of a packet with size ρ takes a random amount of frames. Denote this random number by τ. Considering the service process increment s_i at the source and assuming it to start at frame i, it takes the form $s_i = 0, s_{i+1} = 0 \ldots, s_{i+\tau} = \rho$ over the time span of the τ frames until successful transmission (as the packet is taken out of the source queue only if it was successfully transmitted). However, in order to simplify the analysis, we adopt a service model [25] where the incremental service process equals instead

$s_i = \rho/\tau, s_{i+1} = \rho/\tau, \ldots, s_{i+\tau} = \rho/\tau$. The advantage of this model is that we only need to determine the statistics of τ in order to obtain the mean and variance of the service process increments. This comes at the price of being an approximation, however, we validate in Section V that this assumption is not impacting the system performance significantly.

In addition to the mean and variance of the service process increments, based on the distribution of τ, we can also obtain the percentages of the frame types for each C-ARQ protocol. Therefore, the fundamental challenge of investigating the EEE performances of C-ARQ protocols is to derive the distribution of τ. In the following, we first analyze the outage probability of the initial transmission, and based on this derive the distribution of τ for each protocol. Then, we derive the mean and variance of the service process increment and the energy consumption variables based on the distribution of τ. As a consequence, the EEE performance can be finally determined. The methodology of the EEE analysis of each C-ARQ protocol

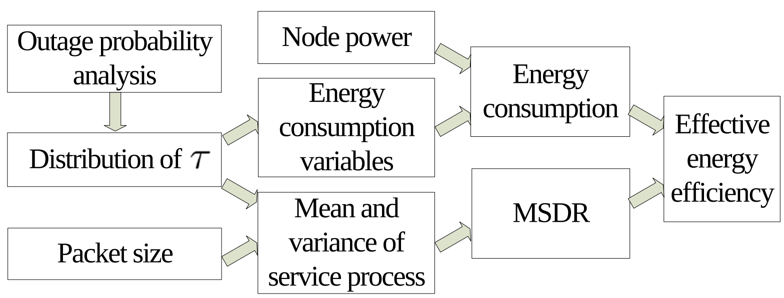

Fig. 4. EEE analysis methodology of the considered C-ARQ protocols.

is shown in Fig. 4.

A. Outage Probabilities of the Initial Transmission

The initial transmission of a reactive protocol is the direct link transmission. With Rayleigh block-fading channels, the outage probability of the direct link is given by:

$$\mathrm{Pr}_0 = 1 - \exp\left(-\gamma^* \sigma^2 / 2\bar{h}^2_{\mathrm{S,D}} P_{\mathrm{tx}}\right). \tag{8}$$

Recall that γ^* is the SNR threshold subject to the packet size ρ.

For a proactive protocol, the initial transmission is a multi-relay assisted two-frame transmission. Our previous work [27] has shown that the outage probability of the single two-frame transmission is given by:

$$\mathrm{Pr}^{\mathrm{out}}_{\mathrm{Initial}} = \sum\nolimits_{n=0}^{J} \mathrm{Pr}_2(n) \cdot \mathrm{Pr}_{\mathrm{B}}\left(n; J, \mathrm{Pr}_1\right), \tag{9}$$

where the number of active relays n is a binomially distributed random variable. Recall that J is the total number of relays deployed in the system. Denote by $\mathrm{Pr_B}\left(n; J, \mathrm{Pr}_1\right)$ the probability density function of n, hence we have:

$$\mathrm{Pr_B}\left(n; J, \mathrm{Pr}_1\right) = \begin{pmatrix} J \\ n \end{pmatrix} \cdot \left(1 - \mathrm{Pr}_1\right)^n \cdot \left(\mathrm{Pr}_1\right)^{J-n}, \tag{10}$$

where Pr_1 is the outage probability of the link form the source to relay j and is given by:

$$\mathrm{Pr}_1 = 1 - \exp\left(-\gamma^* \sigma^2 / 2\bar{h}^2_{\mathrm{S},j} P_{\mathrm{tx}}\right). \tag{11}$$

In (9), $\mathrm{Pr}_2(n)$ is the outage probability of the relaying frame with n active relays. Due to the MRC, the combined SNR at the destination results from the superposition of several fading signals, which leads to a Gamma-distributed random variable for the joint SNR [27]. Hence, $\mathrm{Pr}_2(n)$ is equivalent to the cumulative distribution function $\mathrm{F}\left(\gamma^*; n, \beta\right)$ of the Gamma-distributed random variable:

$$\begin{aligned} \mathrm{Pr}_2(n) &= \mathrm{F}\left(\gamma^*; n, \beta\right) = \Pr(\gamma \leq \gamma^*; n, \beta) \\ &= \begin{cases} 1 - \sum_{j=0}^{n-1} \dfrac{1}{j!} \left(\dfrac{\gamma^*}{\beta}\right)^j e^{-\frac{\gamma^*}{\beta}}, & n > 0; \\ 1, & n = 0, \end{cases} \end{aligned} \tag{12}$$

where β is the scaling parameter of the gamma distribution and is given by $\beta = 2\sum_{j=1}^{J} P_{\mathrm{tx}} \bar{h}^2_{j,\mathrm{D}} / J\sigma^2$. Both the gamma distribution and the binomial distribution are approximations based on the topology simplification (recall that the distances between relays are assumed to be significantly shorter than the distances either from the source to relays or from relays to the destination). However, we have shown in [27] that both these distributions are indeed appropriate approximations even if the relays are more separated[3].

B. Distribution of τ and Outage Probability of C-ARQ

The number of retransmission frames consists of re-broadcasting frames and re-forwarding frames. The source re-broadcasts the packet if no relay decodes the packet correctly while a re-forwarding frame occurs if the previous forwarding/re-forwarding failed. Hence, the random variable τ (the number of transmission and retransmission frames) depends on two random variables k_1 and k_2 which are the number of broadcasting frames (including initial broadcasting and re-broadcastings) and the number of forwarding

[3]The errors of the approximations are no longer negligible if the maximal distance between relays is larger than half of the source-relays distance [27].

frames (including initial forwarding and re-forwardings). Taking into account the feedback frames, the expressions of τ of the proactive and reactive protocols are:

$$\tau_{k_1,k_2} = \begin{cases} 2k_1+2k_2-1, & \text{proactive protocols;} \\ 2k_1+2k_2, & \text{reactive protocols.} \end{cases} \tag{13}$$

Assuming the packet to be successfully decoded at some relay after k_1 broadcasting attempts, the probability of this equals $\mathrm{Pr}_1^{J(k_1-1)} \cdot \mathrm{Pr}_\mathrm{B}\left(n; J, \mathrm{Pr}_1\right)$. Similarly, if the destination finally decodes the packet correctly based on combining the signals transmitted/retransmitted k_2 times from n active relays, the probability of this under a static proactive protocol equals $\mathrm{F}\left(\gamma^*; n(k_2-1), \beta\right) - \mathrm{F}\left(\gamma^*; nk_2, \beta\right)$.

For a dynamic protocol, the computation of the successful transmission during the relaying phase is actually more complex to derive, as one needs to consider the changing number of relays during the retransmission attempts. If the initial transmission of a dynamic proactive protocol succeeds (which means $k_2 = 1$) and the number of active relays is n, then the probability that this happens equals $\mathrm{F}\left(\gamma^*; 0, \beta\right) - \mathrm{F}\left(\gamma^*; n, \beta\right)$. Otherwise, the destination finally decodes the packet correctly based on combining the signals transmitted once by n active relays and retransmitted $k_2 - 1$ ($k_2 > 1$) times from all the J relays. This probability is equal to $\mathrm{F}\left(\gamma^*; J(k_2-2)+n, \beta\right) - \mathrm{F}\left(\gamma^*; J\left(k_2-1\right)+n, \beta\right)$. Therefore, the probability mass function (PMF) of τ under the static proactive protocol and the dynamic proactive protocol are given by (14) and (15).

$$\Pr\{\tau_{\mathrm{SP}} = 2k_1+2k_2-1\} = \sum\nolimits_{n=1}^{J} \left\{ \mathrm{Pr}_1^{J(k_1-1)} \cdot \mathrm{Pr}_\mathrm{B}\left(n; J, \mathrm{Pr}_1\right) \cdot \left[\mathrm{F}\left(\gamma^*; n(k_2-1), \beta\right) - \mathrm{F}\left(\gamma^*; nk_2, \beta\right)\right] \right\}. \tag{14}$$

$$\begin{aligned} &\Pr\{\tau_{\mathrm{DP}} = 2k_1+2k_2-1\} \\ &= \sum\nolimits_{n=1}^{J} \left\{ \mathrm{Pr}_1^{J(k_1-1)} \cdot \mathrm{Pr}_\mathrm{B}\left(n; J, \mathrm{Pr}_1\right) \cdot \left[\mathrm{F}\left(\gamma^*; \max\left\{J(k_2-2)+n, 0\right\}, \beta\right) - \mathrm{F}\left(\gamma^*; J\left(k_2-1\right)+n, \beta\right)\right] \right\}. \end{aligned} \tag{15}$$

Similarly, the PMF of τ under the static reactive protocol and the dynamic reactive protocol are given by Equations (16) and (17).

$$\Pr\{\tau_{\mathrm{SR}} = 2(k_1+k_2)\} = \begin{cases} 1 - \mathrm{Pr}_0, & k_1 = 1; \\ \mathrm{Pr}_0 \sum\nolimits_{n=1}^{J} \left\{ \mathrm{Pr}_1^{J(k_1-1)} \cdot \mathrm{Pr}_\mathrm{B}\left(n; J, \mathrm{Pr}_1\right) \cdot \left[\mathrm{F}\left(\gamma^*; n\left(k_2-1\right), \beta\right) - \mathrm{F}\left(\gamma^*; nk_2, \beta\right)\right] \right\}, & k_1 > 1. \end{cases} \tag{16}$$

$$\Pr\{\tau_{\text{DR}}{=}2(k_1{+}k_2)\}$$
$$=\begin{cases} 1-\Pr_0, & k_1=1; \\ \Pr_0\sum_{n=1}^{J}\left\{\Pr_1^{J(k_1-1)}\cdot \Pr_{\text{B}}\left(n;J,\Pr_1\right)\cdot\left[\text{F}\left(\gamma^*;J(k_2-2)+n,\beta\right)-\text{F}\left(\gamma^*;J(k_2-1)+n,\beta\right)\right]\right\}, & k_1>1. \end{cases} \tag{17}$$

Based on the distribution of τ, for all the C-ARQ protocols we can immediately obtain the residual outage probability of a packet after τ transmission and retransmissions frames:

$$\Pr_{\text{C-ARQ}}^{\text{out}}(\tau)=1-\sum_{k_1=1}^{\lceil\tau/2\rceil}\sum_{k_2=\lceil\tau/2\rceil-k_1}^{\lceil\tau/2\rceil}\Pr\left[\tau_{k_1,k_2}\right], \tag{18}$$

where $\lceil\cdot\rceil$ is the rounding function (rounding to the next highest integer). In fact, (18) shows the outage probability of C-ARQ protocols with limited number of retransmissions (also known as the truncated retransmission). We do not focus on the truncation of C-ARQ protocols in this work. However, note that based on (18) one can consider the maximization of the MSDR or the EEE of a truncated C-ARQ protocol.

C. Energy Consumption

Recall that variables $\text{E}[\Omega]$, p_{S}, p_{D}, and p_{R} of static protocols as well as $p_{\text{i-R}}$ and $p_{\text{a-R}}$ of dynamic protocols determine the overall energy consumption of the system (see Section III-B). The expected size of the initial active relay set $\text{E}[\Omega]$ is equivalent to the expected number of relays which decode the packet correctly after a single broadcasting frame. In addition, Ω is a binomially distributed random variable [27]. Hence, $\text{E}[\Omega]$ is obtained by:

$$\text{E}[\Omega]=J\cdot(1-\Pr_1). \tag{19}$$

Second, for all the protocols the percentage of the broadcasting frames is given by:

$$p_{\text{S}}=\sum_{k_1,k_2\in\mathbb{N}}\frac{k_1}{\tau_{k_1,k_2}}\cdot\Pr\left[\tau_{k_1,k_2}\right]. \tag{20}$$

Third, the percentages of relaying frames and ACK/NACK frames under the proactive protocols and the reactive protocols need to be analyzed separately. From Fig. 2 we observe that there is a one-to-one correspondence between the relaying frames and ACK/NACK frames in each proactive protocol. Hence, the percentages of relaying frames and ACK/NACK frames are the same, and are given by:

$$p_{\mathrm{R}} = p_{\mathrm{D}} = (1 - p_{\mathrm{S}})/2. \tag{21}$$

For the reactive C-ARQ protocols, the corresponding percentages of the relaying frames for the different relay sets are:

$$p_{\text{i-R}} = \sum\nolimits_{k_1,k_2 \in \mathbb{N}} \frac{1}{\tau_{k_1,k_2}} \cdot \Pr\left[\tau_{k_1,k_2}\right], \tag{22}$$

$$p_{\text{a-R}} = \sum\nolimits_{k_1,k_2 \in \mathbb{N}} \frac{k_2 - 1}{\tau_{k_1,k_2}} \cdot \Pr\left[\tau_{k_1,k_2}\right]. \tag{23}$$

Finally, the percentage of the ACK/NACK frames of reactive protocols p_{D} is obtained by:

$$p_{\mathrm{D}} = 1 - p_{\mathrm{S}} - p_{\text{i-R}} - p_{\text{a-R}}. \tag{24}$$

D. Mean and Variance of the Service Process Increments

To derive the mean and variance of the service process increments, we first clarify the difference and the relationship between the probability mass function (PMF) of τ and the PMF of the service process increment s_i. Recall that if we take randomly one packet from all the packets that have been transmitted the probability that this packet has a transmission delay τ is $\Pr\left[\tau_{k_1,k_2}\right]$. This is the PMF of τ. Based on our service process model, the corresponding service process increment of this packet has value ρ/τ and lasts τ frames. In other words, when the service process increments show up with a certain value ρ/τ, they show up in a group with size τ. This group behavior needs to be accounted for. Therefore, the PMF of s_i can be obtained based on the PMF of τ by scaling the probability for each possible value of τ. Hence, the PMF of s_i is given by $\Pr(s_i = \frac{\rho}{\tau_{k_1,k_2}}) = \frac{\Pr\left[\tau_{k_1,k_2}\right] \cdot \tau_{k_1,k_2}}{\sum_{k_1,k_2 \in \mathbb{N}} \tau_{k_1,k_2} \Pr\left[\tau_{k_1,k_2}\right]}$, where the numerator is the expected value of τ_{k_1,k_2} given by $\mathrm{E}\left[\tau_{k_1,k_2}\right]$. Then, the mean and variance of the service process increments result as:

$$\begin{aligned} \mathrm{E}\left[s_i\right] &= \frac{1}{T_{\mathrm{f}}} \cdot \sum_{k_1,k_2 \in \mathbb{N}} \frac{\rho}{\tau_{k_1,k_2}} \cdot \Pr(s_i = \frac{\rho}{\tau_{k_1,k_2}}) \\ &= \frac{1}{T_{\mathrm{f}}} \cdot \sum_{k_1,k_2 \in \mathbb{N}} \frac{\rho}{\tau_{k_1,k_2}} \cdot \frac{\Pr\left[\tau_{k_1,k_2}\right] \cdot \tau_{k_1,k_2}}{\mathrm{E}\left[\tau_{k_1,k_2}\right]} = \frac{1}{T_{\mathrm{f}}} \cdot \frac{\rho}{\mathrm{E}\left[\tau_{k_1,k_2}\right]}, \end{aligned} \tag{25}$$

$$\mathrm{Var}\left[s_i\right] = \frac{\rho^2}{T_{\mathrm{f}} \cdot \mathrm{E}\left[\tau_{k_1,k_2}\right]} \sum_{k_1,k_2 \in \mathbb{N}} \frac{\Pr\left[\tau_{k_1,k_2}\right]}{\tau_{k_1,k_2}} - \left(\mathrm{E}\left[s_i\right]\right)^2. \tag{26}$$

By now, we have obtained all the components of the EEEs for the four C-ARQ protocols. Therefore, the MSDR and EEE can finally be obtained from (4) and (7).

V. Numerical Results and Discussion

In this section, by simulation we first validate our analytical expressions and then evaluate the C-ARQ performance. For all validations and evaluations, we consider the following parameterization of the system model: We randomly deploy a certain number (five) of relays in a circle with radius $R = 20$ m, while the distances of broadcasting and relaying links are both set to 200 m (we vary the system topology in Section V-C). We assume that the center frequency is 2 GHz and the frame length is 20 ms. In addition, we set $P_{\text{tx}} = 20$ dBm, $P_{\text{fb}} = 15$ dBm, $P_{\text{c}} = 17$ dBm and noise power $\sigma^2 = -95$ dBm, respectively. We utilize the well-known COST231 model for calculating the path-loss and utilize the Rayleigh distribution for obtaining the channel fading for both theoretical and simulation values. In the evaluation, we mainly vary the following parameters: packet size, the number of relays deployed in the system and the location of the destination. Finally, the MSDR and the EEE are the performance metrics we mainly consider.

A. Validation of the Analysis

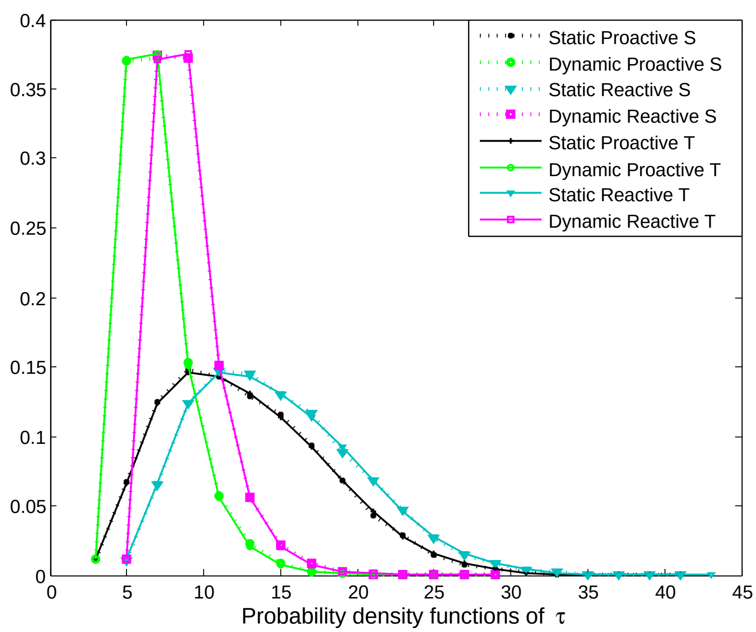

Fig. 5. Validation of the distributions of τ of the considered C-ARQ protocols (packet size 85 bits).

First, we observe a very good match of the simulation results in comparison to the theoretical model in Fig. 5 where the results are obtained from the model parameterization introduced at the beginning of this section while the packet size being fixed to 85 bits. This validates the distribution of τ as derived in Equations (14) - (17).

Next, we validate the analytical expressions for the mean and variance of the service process increment as well as the resulting MSDR and EEE, which are shown in Fig. 6-a, 6-b, 6-c and 6-d. The results are obtained by using the same model parameterization as Fig. 5 while varying the packet size. Again we observe a very good match between the simulation results and the theoretical analysis as derived in (20) - (26).

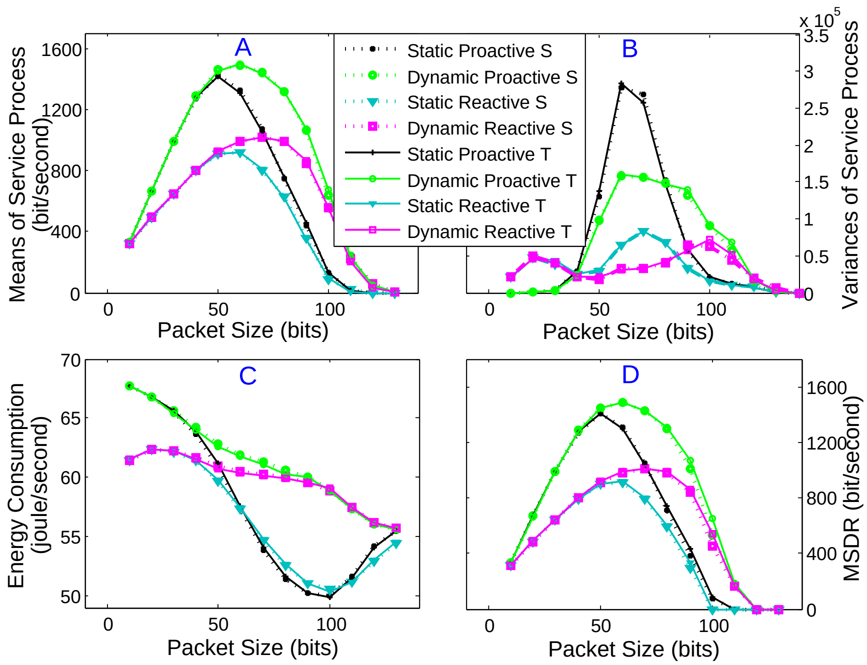

Fig. 6. The mean and variance of service process increments together with the energy consumption and MSDR of the C-ARQ protocols.

Concluding, Fig. 5 and Fig. 6 are strong indications that the derived analytical models fit the real system behavior well, despite several approximation assumptions that were introduced for the derivations. In the following we continue to provide the simulation results as well, while the focus of the next subsections is to derive guidelines for the design of efficient C-ARQ protocols.

B. Higher-Layer Performance of the C-ARQ Protocols

In this subsection we present results regarding the MSDR and EEE of the considered C-ARQ protocols. First, it is already shown in Fig. 5 that the probability density function of random variable τ is more concentrated for dynamic protocols than for static protocols. This indicates that dynamic protocols are more time-efficient than static protocols.

Second, we observe in Fig. 6-B that the dynamic protocols have a lower variance of the service process increment in comparison to the static protocols. Also, we observe in Fig. 6-B that dynamic protocols have relatively lower variance values. In addition, we find that all the variance curves have global maxima for distinct packet sizes ($\rho > 50$ bits). The reason is as follows. As the variance of ρ/τ is increasing in ρ^2, the variance decreases for smaller packet sizes. However, small packets also lead to a higher transmission success probability which results in a low expected value of τ. As the packet size increases, the (expected) value of τ increases since the destination needs more retransmissions to finally decode the packet successfully. This finally leads to a low value of the variance of ρ/τ. Summarizing, the variance of ρ/τ actually demonstrates a trade-off between ρ and τ. This is the reason why the variance curves first increase and then decrease as the packet size increases. Note that the variance of each reactive protocol has also a local maximum for small packet sizes (of about 20 bits). These local maxima are caused by the direct link transmission. The initial transmission of a reactive protocol is a direct link transmission which has a much lower SNR in comparison to relay-assisted transmissions. Small packet sizes let the packet be decoded successfully at the destination relying on just a direct link transmission. Hence, the probability increases strongly that $\tau = 2$. As ρ slightly increases, the direct link transmission still very likely leads to a successfully transmission, which means that τ does not increase. Therefore, the variance of ρ/τ slightly increases. However, as ρ continues to increase, the probability of a successful direct link transmission significantly decreases. Hence, more and more packets need to be retransmitted by the relays, which leads to a much larger τ. As a result, the variance of ρ/τ then decreases.

Third, we observe from Fig. 6-C that static protocols and dynamic protocols consume almost the same energy for either big or small packet sizes. As we know, the static protocols and the dynamic protocols differ in the number of retransmission relays. In other words, they actually have the same procedure of initial transmission. At the same time, outages and retransmissions seldomly occur for small packet sizes. Hence, they have similar energy consumption for small packet sizes. On the other hand, for rather large packet sizes, no relay decodes it correctly. Therefore, there is no relay forwarding. Instead, the source is constantly retransmitting the packet. That is the reason why all the energy consumption curves finally converge as the packet size increases.

Fourth, compared to reactive protocols, proactive protocols consume more energy but have higher MSDRs (shown in Fig. 6-D). This raises the question which of these protocols are more efficient, which we address in Fig. 7 with respect to a scenario where the distances from the source to relays and from

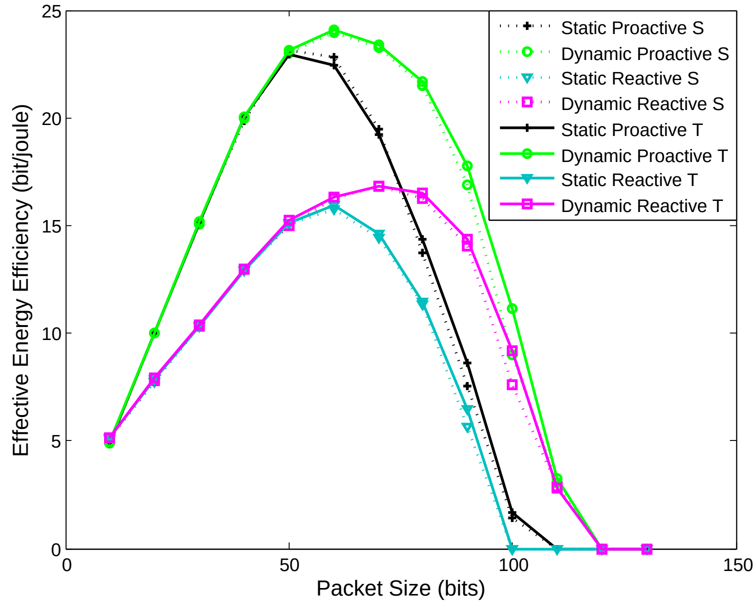

Fig. 7. The EEE performance comparison among the protocols for QoS parameters $\{d, \mathrm{Pr_d}\} = \{20T_f, 10^{-1}\}$.

relays to the destination are the same. We find that the dynamic proactive protocol is the most energy-efficient one. In addition, a dynamic protocol always significantly outperforms a static protocol no matter whether it is combined with a reactive protocol or a proactive protocol.

C. Evaluation of the Impact of the Relay Number on the System Performance

As the multi-relay C-ARQ protocols exploit the diversity of multiple relays, the performance of the protocols are subject to the number of relays present in the system. In addition, the performance of reactive protocols are influenced by the channel quality of the direct link which is further subject to the distance between the source and the destination. Hence, the considered protocols are expected to show different performances while varying the system topology. To evaluate these performances is important for the design of effective C-ARQ systems.

1) Evaluation of the System Performance on Relay Numbers: We show in Fig. 8 and 9 the MSDR and the EEE for a varying number of relays present in the system. We plot the corresponding Shannon-capacity-based data rates (SDR) and Shannon-capacity-based energy efficiencies (SEE) as comparison cases, i.e. the physical layer throughput and energy efficiency of the system.

We learn that deploying more relays in the system monotonically increases the MSDR. However, regarding EEE, deploying more relays is not always beneficial. Instead, as shown in Fig. 9 the EEE curves appear to have a convex behavior in the number of the relays. Moreover, we observe significant gaps between MSDRs as well as between EEEs and SEEs. All the EEE curves (dotted lines) are zero

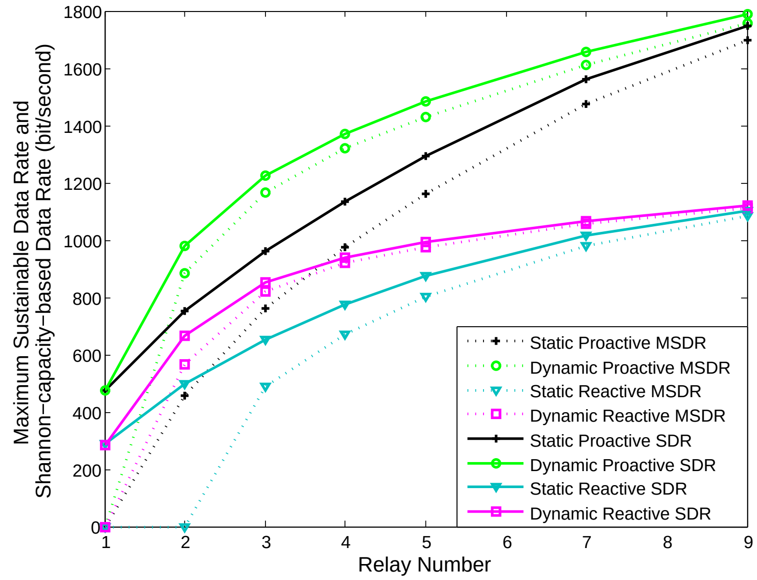

Fig. 8. The comparison between the MSDR and Shannon-capacity-based data rate (SDR) of the protocols while varying the number of relays deployed in the system. The QoS parameters are set to $\{d,\mathrm{Pr_d}\}$ =$\{20T_f$, $10^{-3}\}$.

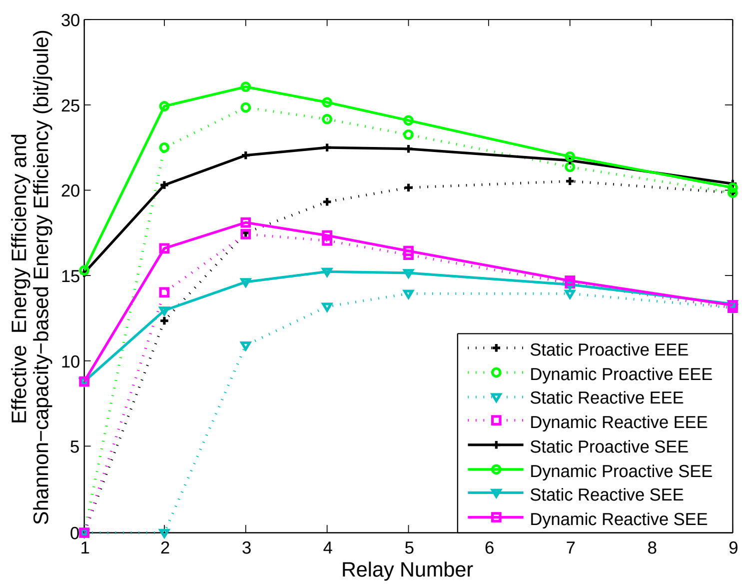

Fig. 9. The comparison between the EEE and Shannon-capacity-based energy efficiency (SEE) of the protocols while varying the number of relays deployed in the system. The QoS parameters are set to $\{d,\mathrm{Pr_d}\}$ =$\{20T_f$, $10^{-3}\}$.

when the system only has one relay as shown in Fig. 9. This means that the system cannot support the corresponding QoS requirements and hence the MSDR becomes 0. In other words, with only one relay the system consumes energy but cannot transmit any packets under the QoS requirements. At the same time, none of the SEE curves (solid lines) has value zero when the relay number is one. Hence, if QoS has to be provided by the system, the SEE is not a good metric to evaluate the efficiency of the system. Instead, the EEE has to be considered, as EEE directly demonstrates how effective the system spends energy for the QoS requirements of interest.

2) Performance Loss due to Relay Number Uncertainty: Previous evaluations (e.g., in Fig. 7) show that the packet size is one of the major concerns for maximizing the EEE. Actually, when the number of relays in the system varies, we find that the optimal packet size changes accordingly. However, if the source only knows that there is a relay group (at least one relay) in the system but does not know the exact number of the relays, the source could decide the packet size according to the pessimistic assumption that there is only one relay in the system in order to make the transmission reliable. This leads to a performance loss of the system, if there are in fact more relays. We are interested in this performance loss and illustrate this in Fig. 10. The curves with solid lines represent the situations that the source decides the packet size according to the pessimistic assumption while the other curves with dotted lines represent the situations that the source knows the exact relay number. The figure shows that the performance loss is about 20% for a system with two relays and is about 40% for a system which has nine relays. Therefore, we conclude that a C-ARQ protocol should spend the required overhead to determine the exact number of relays in the system first.

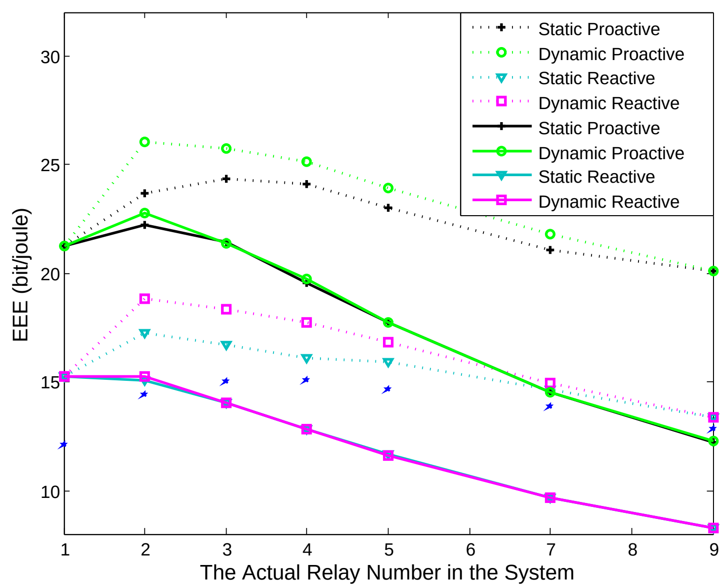

Fig. 10. EEE Performance Loss due to relay number uncertainty.

3) Optimality of Dynamic Protocols for Different Topologies: Based on the above numerical results, the dynamic protocols always show better EEE performance than the static protocols. Hence, we further compare the two dynamic protocols in the following. Since a reactive protocol uses the direct link transmission as the initial transmission, the performance of it is subject to the system topology, especially the distance of the direct link $d_{\text{S-D}}$. To compare the dynamic reactive protocol with the dynamic proactive

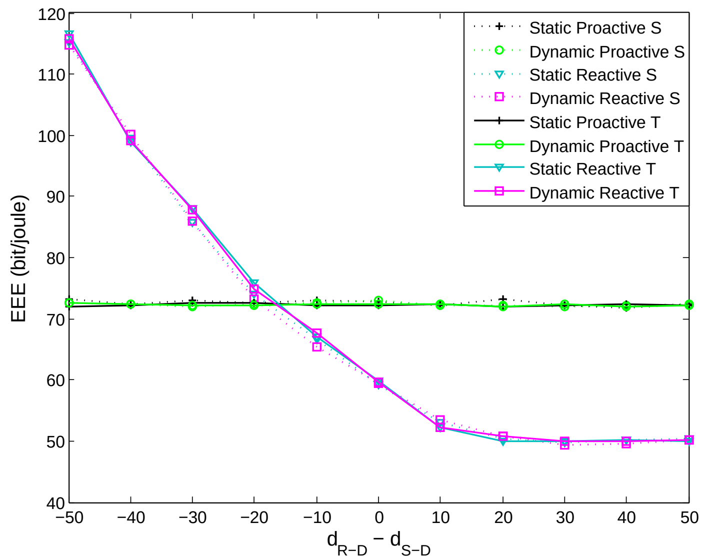

Fig. 11. The EEE performance comparison between the dynamic reactive protocol and the dynamic proactive protocol while varying the location of the destination. $d_{\mathrm{S-R}}$ and $d_{\mathrm{R-D}}$ are set to 50m and 100m. At the same time, $d_{\mathrm{S-D}}$ varies. The x-axis is the distance gap between $d_{\mathrm{R-D}}$ and $d_{\mathrm{S-D}}$, which indicates whether the destination is more close to the relays or to the source.

protocol, we vary the system topology in the following way: While fixing the locations of the source and relays and fixing the distance $d_{\text{R-D}}$ from the center of relay group to the destination, we vary the location of the destination by circularly moving it around the relay group. We obtain the comparison between the dynamic proactive protocol and the dynamic reactive protocol which is shown in Fig. 11. We observe that even at the zero point of the X-axis ($d_{\text{R-D}} - d_{\text{S-D}} = 0$), where the source-destination distance is the same as the relays-destination distance, using direct link as the initial transmission attempt (reactive protocols) has still a lower performance than using relaying (proactive protocols).

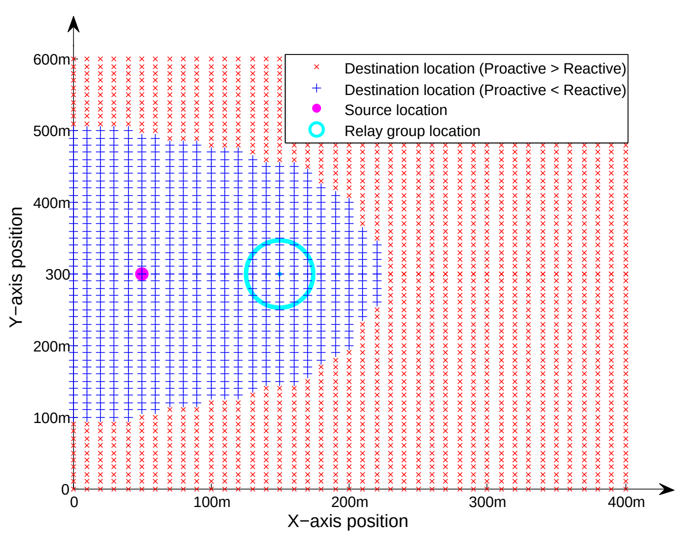

Fig. 12. The EEE performance comparison between proactive protocols and reactive protocols, where only the locations of the source and relays are fixed.

Spatially, we show the comparison between proactive protocols and reactive protocols again in Fig. 12.

The reactive protocols are superior over the proactive protocols only if the destination is located in a rather small range around the source.

D. An Improved C-ARQ Protocol Design

Based on the evaluation of the above four protocols, we find that static protocols consume less energy while dynamic protocols allow for higher MSDR. This indicates that the EEE is likely to be improved by trading-off the characteristics of the two schemes. Therefore, we propose a simple extension of the

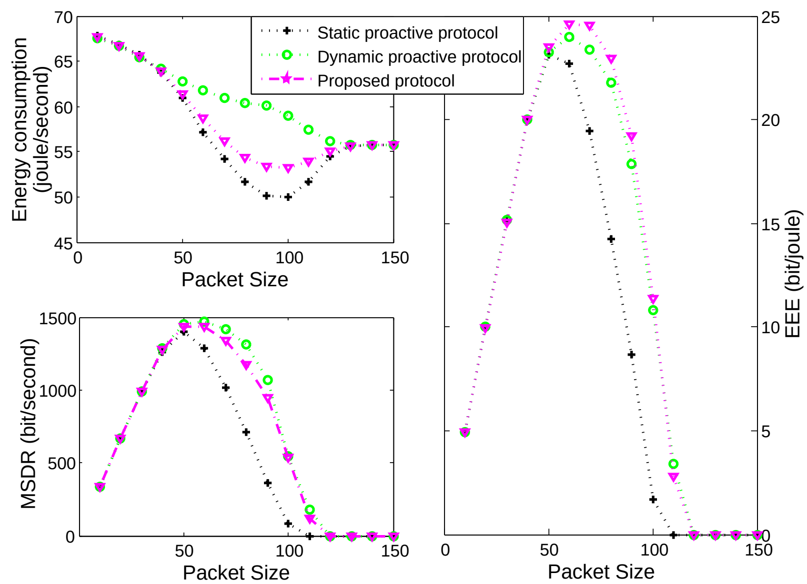

Fig. 13. The energy consumption, MSDR and EEE performance of proposed protocol with $\mathrm{Pr}_* = 0.8$.

discussed C-ARQ protocols. Different from the dynamic protocols which require all the passive relays to overhear the forwarding frame and then participate in the retransmission, we introduce a probability Pr_*, ($\mathrm{Pr}_* \in [0, 1]$). In the proposed protocol, if a passive relay decodes the packet successfully by overhearing, it participates in the retransmission with probability Pr_*. Hence, the performance of the proposed protocol is similar to a static protocol when the value of Pr_* is close to 0, and similar to a dynamic protocol if Pr_* is close to 1. In numerical experiments (which is not shown here), we observe that while varying the Pr_* from 0 to 1 the EEE performance of the trade-off protocol initially outperforms the static protocols and later on even exceeds the dynamic protocol performance, before the EEE decreases again for large values of Pr_*. In Fig. 13 where Pr_* is set to 0.8, we show that the proposed protocol has slightly lower MSDR but much lower energy consumption than the proactive protocols. As a result, the proposed protocol improves the EEE performance by 4% (when the packet size is around 70 bits).

VI. Conclusion

In this paper, we investigated the higher-layer performances as well as the "throughput, delay, energy consumption" trade-off of four C-ARQ protocols in a multi-relay system. Following this approach, a closed-form expression of the outage probability of the multi-DF-relay two-hop transmission was derived for analyzing the initial transmission of the C-ARQ protocols. Moreover, the distribution of the transmission delay and the energy consumption were also derived for each protocol. Furthermore, we derived a tight approximation of the effective capacity by determining the mean and variance of the corresponding service process increments based on considering the distribution of τ (the number of frames for successful packet reception). Finally, we analyzed the higher-layer energy efficiency of the C-ARQ protocols by considering the EEE which is the ratio between the effective capacity and the energy consumption of the considered C-ARQ protocols.

Through numerical results, we validated our analytical model. In addition, we concluded several guidelines for the design of efficient C-ARQ protocols from the numerical analysis of our analytical framework: (1) Direct link transmission can usually be ignored by the C-ARQ protocols unless a specific topology of the nodes is given. (2) There exists an optimal number of relays deployed in the system when maximizing the EEE. (3) Dynamic relay selection should be enabled as it leads to a better efficiency. (4) Overhead should be spent to acquire the exact number of relays present in the system. At last, we proposed a simple multi-relay C-ARQ protocol which is a trade-off between dynamic protocols and static protocols. We showed that the proposed protocol improves the EEE performance by some additional 4%.

References

[1] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, Sep. 1979, pp. 572–584.

[2] G. Neonakis Aggelou and R. Tafazolli, "On the relaying capability of next-generation GSM cellular networks," *IEEE Pers. Commun.*, vol. 8, no. 1, Feb. 2001, pp. 40–47.

[3] R. Pabst, B. H. Walke, D. Schultz, et al., "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, Sep. 2004, pp. 80–89.

[4] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, Oct. 2004, pp. 74–80.

[5] J. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, Oct. 2003, pp. 2415–2425.

[6] C. Sun, J. Cheng, and M. Taromaru, "Achieving the lowest outage probability for multi-phase cooperative wireless networks," in *Proc. IEEE Information Theory Workshop on Information Theory for Wireless Networks*, Bergen, Norway, Jul. 2007, pp. 1–5.

[7] P. Elia, K. Vinodh, M. Anand, and P. Vijay Kumar, "D-MG tradeoff and optimal codes for a class of AF and DF cooperative communication protocols," in *Proc. IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007, pp. 681–685.

[8] J. Laneman, D. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, Dec. 2004, pp. 3062–3080.

[9] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, Sep. 2005, pp. 3037–3063.

[10] M. Dianati, X. Ling, K. Naik, and X. Shen, “A node-cooperative ARQ scheme for wireless ad hoc networks,” *IEEE Trans. Veh. Technol*, vol. 55, no. 3, May 2006, pp. 1032–1044.

[11] 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, Nov. 2008, pp. 1907–1918.

[12] I. Cerutti, A. Fumagalli, and P. Gupta, “Delay models of single-source single-relay cooperative ARQ protocols in slotted radio networks with poisson frame arrivals,” *IEEE/ACM Trans. Netw.*, vol. 16, no. 2, Apr. 2008, pp. 371–382.

[13] E. Zimmermann, P. Herhold and G. Fettweis, “On the performance of cooperative relaying protocols in wireless networks,” *Eur. Trans. Telecomm.*, vol. 16, no. 2, Jan. 2005, pp. 5–16.

[14] S. Lee, W. Su, S. Batalama, and J. Matyjas, “Cooperative decode-and-forward ARQ relaying: Performance analysis and power optimization,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, Aug. 2010, pp. 2632–2642.

[15] J. Ramis and G. Femenias, “Cross-layer QoS-constrained optimization of adaptive multi-rate wireless systems using infrastructure-based cooperative ARQ,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, May 2013, pp. 2424–2435.

[16] J. Alonso-Zarate, E. Stavrou, A. Stamou, P. Angelidis, L. Alonso, and C. Verikoukis, “Energy-efficiency evaluation of a medium access control protocol for cooperative ARQ,” in *Proc. IEEE International Conference on Communications*, Kyoto, Japan, June 2011.

[17] I. Stanojev, O. Simeone, Y. Bar-Ness, and D. H. Kim, “Energy efficiency of non-collaborative and collaborative hybrid-ARQ protocols,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, Jan. 2009, pp. 326–335.

[18] M. Maaz, P. Mary, and M. Helard, “Energy efficiency analysis in relay assisted hybrid-ARQ communications,” in *Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, Sydney, Australia, Sep. 2012, pp. 2263–2268.

[19] H. Boujemaa, “Delay analysis of cooperative truncated HARQ with opportunistic relaying,” *IEEE Trans. Veh. Technol*, vol. 58, no. 9, Nov. 2009, pp. 4795–4804.

[20] M. Maaz, P. Mary, and M. Helard, “Delay outage probability in block fading channel and relay-assisted hybrid-ARQ network,” *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, Apr. 2014, pp. 129–132.

[21] D. Wu and R. Negi, “Effective capacity: a wireless link model for support of quality of service,” *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, July 2003, pp. 630–643.

[22] Y.-M. Wang and S. Lin, “A modified selective-repeat type-II hybrid ARQ system and its performance analysis,” *IEEE Trans. Commun.*, vol. 31, no. 5, May 1983, pp. 593–608.

[23] F. Richter, A. Fehske, and G. Fettweis, “Energy efficiency aspects of base station deployment strategies for cellular networks,” in *Proc. IEEE Vehicular Technology Conference*, Anchorage, USA, Sep. 2009.

[24] H. Kim, C.-B. Chae, G. de Veciana, and R. Heath, “A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, Aug. 2009, pp. 4264 –4275.

[25] A. Kumar, D. Manjunath, and J. Kuri, *Communication Networking: An Analytical Approach*, Elsevier, 2004.

[26] B. Soret, M. Aguayo-Torres, and J. Entrambasaguas, “Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, June 2010, pp. 1901 –1911.

[27] Y. Hu and J. Gross, “On the outage probability and effective capacity of multiple decode-and-forward relay system,” in *Proc. IFIP Wireless Days*, Dublin, Ireland, Nov. 2012.