

RESEARCH

Open Access

Active interference cancellation-aided QoS-aware distributed ARQ for cognitive radios with heterogeneous traffics

Juan Liu¹, Wei Chen^{1*}, Zhigang Cao¹ and Ying Jun (Angela) Zhang²

Abstract

Relay-assisted Automatic Repeat reQuest (ARQ), which allows the relays instead of the source to carry out retransmissions of packets erroneously received at the destination, is an efficient cooperative transmission technique to improve link reliability with high spectrum efficiency in wireless networks. In cognitive radio (CR) systems, however, this relay-assisted ARQ could induce large latency, since a successful transmission of a packet may consume two idle timeslot or temporal spectrum holes. To overcome this limitation, an active interference cancellation (IC)-aided distributed relay-assisted ARQ method is proposed in this article to serve heterogeneous elastic traffics with diverse quality of service (QoS) demands in CR systems. Specifically, by applying our recently proposed distributed beamforming-based IC method in the physical layer, cognitive relays can exploit spatial spectrum holes (SSHs) to retransmit while keeping the interference to primary users at a tolerable level. Meanwhile, at the MAC layer, two scheduling policies, namely probabilistic and queue-length-based scheduling, are proposed to obtain an efficient allocation of temporal and SSHs to direct transmissions and retransmissions. The performance of the proposed schemes are analyzed and optimized by adjusting the scheduling parameters. As a result, the secondary users can obtain significant QoS gains, as validated by theoretical and simulation results.

Keywords: Cognitive radio, Distributed ARQ, Interference cancellation, Opportunistic scheduling, Markov chain

1 Introduction

As an efficient cooperative transmission technique, distributed relay-assisted Automatic Repeat reQuest (ARQ) can improve the link reliability of wireless links in wireless relay networks [1-4]. In contrast to traditional ARQ where the source is responsible of retransmitting packets erroneously received at the destination [5], distributed relay-assisted ARQ allows the relays of the source-to-destination (S-D) link to retransmit erroneous packets in a distributed way. Here, the relays are involved only when they receive Negative ACKnowledgement (NACK) packets sent by the destination, indicating it cannot successfully decode the source packet. In this way, high cooperative diversity gain can be achieved at the cost of little spectral efficiency loss.

Recently, some efforts have been devoted to applying distributed ARQ in cognitive radio (CR) systems [6-8] to improve the reliability of secondary links, and thus the utilization of scarce and valuable spectrum bands [9,10]. Meanwhile, queueing analysis methods have been applied to study the throughput and/or delay performances of buffered cognitive relay systems [9,11,12]. The previous work mainly focuses on efficient utilization of idle timeslots, or the so-called Temporal Spectrum Holes (TSHs) [13], to improve the Quality of Service (QoS) of secondary users (SUs). That is, both the SU source and relays access the channel only in idle timeslots that are not currently being occupied by PUs to avoid causing interference to them. In such a system, a successful packet delivery may require two idle timeslots, one for the S-D direct transmission and one for the retransmission carried out by the relays. This, however, would inevitably reduce the spectrum efficiency and/or induce unacceptably large latency, since TSHs are randomly spaced in time and very scarce especially when the primary system is heavily loaded.

*Correspondence: wchen@tsinghua.edu.cn

¹State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology (TNList) and Department of Electronic Engineering, Tsinghua University, Beijing, China
Full list of author information is available at the end of the article

To exploit new transmission opportunities for cooperative transmissions in CR systems, we have proposed an active interference cancellation (IC) method in [14]. Specifically, by forming a virtual antenna array, the SU relays can exploit the so-called spatial spectrum holes (SSHs) and transmit at the same time with PUs [13,15]. Our previous work has also developed two IC-based opportunistic scheduling schemes to schedule available TSHs and SSHs between the SU source and relays for efficiently serving delay-limited real-time traffics and loss-sensitive burst traffics, respectively [16,17]. Therein, we have assumed that perfect channel state information (CSI) can be acquired. We have also assumed that there is no direct S-D path, and thus the relays need to participate in every packet transmission. In contrast to the previous work, we consider in this article a more general scenario, where CSI could be imperfect and an S-D link exists.

This article proposes an IC-aided distributed ARQ method, referred to as ICaD-ARQ, to transmit heterogeneous elastic traffics with diverse QoS requirements in CR systems. Specifically, the source accesses an idle timeslot to transmit a packet to the destination. If the packet has not correctly been received by the destination node as notified via a NACK packet, the relays that correctly decode the source packet will retransmit this packet. Together with the source node, the decoding relays can form a distributed zero-forcing beamformer to transmit the erroneous packet simultaneously with PUs while completely cancelling the interference to active PUs. When perfect CSI is not available, our proposed IC method applies robust beamforming to exploit SSHs with the interference to the primary receivers maintained under a tolerable level.

To serve elastic traffics with diverse QoS demands, we then propose two practical ICaD-ARQ protocols, referred to as ICaD-ARQ-PS and ICaD-ARQ-QS, respectively, which apply probabilistic scheduling and queue-length-based Scheduling schemes to carry out link-aware retransmissions. Specifically, with ICaD-ARQ-PS, an idle timeslot is allocated for retransmission with a given probability and retransmissions are carried out according to truncated ARQ. On the other hand, the ICaD-ARQ-QS scheme schedules the SU nodes to transmit according to the length of the retransmission queue. For both schemes, we analyze the throughput and delay performance, which in turn enables us to derive the optimal scheduler configurations.

Our theoretical analysis is validated by simulation results, which show that the proposed schemes perform well in perfect and imperfect CSI scenarios. One can also see that ICaD-ARQ-PS can be applied to transmit elastic traffics with no delay constraint. While the ICaD-ARQ-QS scheme can achieve a tradeoff between the average throughput and the end-to-end delay by adjusting the

queue threshold. In practical scenarios, the parameter should be selected based on the physical layer outage probability which can be estimated using training packets.

The remainder of this article is organized as follows. Section 2 presents the system model. Section 3 introduces the ICaD-ARQ method and briefly describes the beamforming-based IC approach for exploiting SSHs. In Sections 4 and 5, we propose the ICaD-ARQ-PS and ICaD-ARQ-QS schemes and analyze their performance, respectively. Simulation results and conclusions are presented in Sections 6 and 7, respectively. Throughout this article, the following notations will be used. For a vector \mathbf{x} , \mathbf{x}^* and \mathbf{x}^T denote its conjugate transpose and transpose, respectively. $|\mathbf{x}|$ denotes the magnitude of a complex number \mathbf{x} . \mathcal{C}^l is used to denote a l -dimensional complex space. $\Pr\{e\}$ denotes the probability of the event e . For an event e , $\mathcal{I}(e)$ is an indicator function, defined as $\mathcal{I}(e) = 1$ when e is true and $\mathcal{I}(e) = 0$ otherwise.

2 System model

2.1 Network description

We consider a cognitive relay network with $K+2$ SU nodes (a source s , K relays, and a destination d), each equipped with one single antenna, as depicted in Figure 1. A set of potential relays, denoted by $\mathcal{K} = \{r_1, \dots, r_K\}$, can help the source to transmit to the destination if necessary. Assume that the S-D communication pair dynamically accesses one channel that is authorized to PUs. It is supposed that the S-D pair gains the using right of this channel through competition or negotiation with other secondary pairs. However, the concrete competition or negotiation scheme is out of the scope of this article. It is assumed that one co-channel primary link is active in the neighborhood, as shown in Figure 1. It is reasonable when the SUs are in the coverage of one primary link. In this system, the time resource is partitioned into discrete timeslots

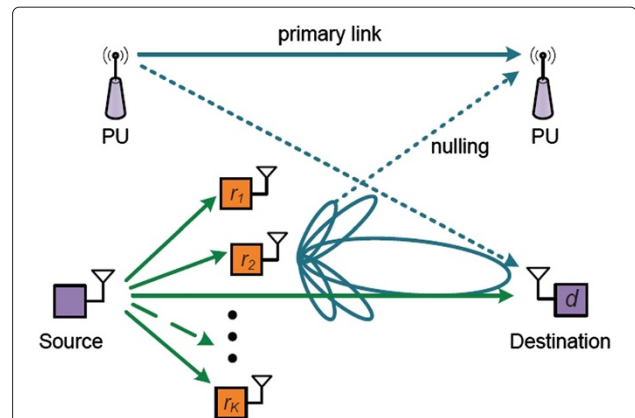


Figure 1 A cognitive relay network utilizing distributed beamforming for interference cancellation.

and all secondary transmissions are assumed to be slot synchronized. A timeslot consists of two phases: spectrum sensing and data transmission. In the sensing phase, the SUs detect the channel to see if the channel is occupied by the PUs. Then, based on the sensing results, SUs transmit their own data in the data transmission phase.

Suppose that the channel occupancy of PUs follows a discrete-time two-state Markov chain. A similar channel occupancy model of PUs was used in [18]. Let $\Theta[t]$ denote the status of channel occupancy by PUs at timeslot t , i.e., idle ($\Theta[t] = 0$) and busy ($\Theta[t] = 1$). Given that the PUs of the primary link are silent in the current timeslot, the PUs occupy the channel with the probability α and keep silent with the probability $1 - \alpha$ in the next timeslot, respectively. Similarly, the channel occupied by PUs in the current timeslot will be released with the probability β and occupied with the probability $1 - \beta$ in the next timeslot, respectively. Accordingly, the transition matrix of channel occupancy by PUs can be given by

$$\mathbf{P}_O = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}, \quad (1)$$

the (i, j) th element of which denotes the transition probability $\Pr\{\Theta[t + 1] = j - 1 | \Theta[t] = i - 1\}$ for all $i, j \in \{1, 2\}$. The activity of PUs on the channel in the long term can be characterized by the steady probability of \mathbf{P}_O . The stationary probabilities of state idle and state busy are thus given by

$$\begin{aligned} P_I &= \Pr\{\Theta[t] = 0\} = \frac{\beta}{\alpha + \beta}, \\ P_B &= \Pr\{\Theta[t] = 1\} = \frac{\alpha}{\alpha + \beta}. \end{aligned} \quad (2)$$

2.2 Channel model and outage probability

Let $h_{a,b}$ denote the channel coefficient between nodes a and b . Assuming block Rayleigh fading, $h_{a,b}$ stays constant during each timeslot while varying independently and identically in different timeslots. The channel coefficient $h_{a,b}$ can be characterized by a complex Gaussian random variable with zero-mean and variance $\sigma_{a,b}^2$. Likewise, we denote z_a to be the additive white Gaussian noise (AWGN) at the receiver of node a . Throughout this article, the AWGN at every receiver can be characterized by a complex Gaussian random variable with zero-mean and variance σ_0^2 . Assume that SUs transmit a packet in a timeslot at a data rate R bit/s/Hz.

To efficiently exploit spectrum resources, we consider to apply an incremental decode-and-forward relaying scheme based on distributed beamforming in the physical layer. The channel model and outage probability are presented in the sequel.

2.2.1 Channel model

When the source accesses a TSH to broadcast a packet, the destination and relay r_i receive

$$Y_d = h_{s,d}X_s + z_d, \quad (3)$$

$$Y_{r_i} = h_{s,r_i}X_s + z_{r_i}, \quad r_i \in \mathcal{K}, \quad (4)$$

respectively, where X_s denotes the source signal transmitted at a power P . Each of K relays receives the source packet, and will send back an ACKnowledgment (ACK) or NACK packet immediately, if it succeeds or fails to decode the packet. The ACK/NACK packets are sent by K relays over orthogonal subchannels of the feedback channel, such as different slots and spreading codes, etc. In this way, the collision of ACK/NACK messages from multiple relays can be avoided. The decoding relay set is denoted by \mathcal{D} and the number of decoding relays is denoted by k ($k \in \mathbb{N}$).

If an NACK packet from the destination is received, the source and decoding relays will jointly form a distributed beamformer to retransmit the packet in an idle or busy timeslot. Suppose that the packet is retransmitted at the same power P . Let $\mathbf{h} = [h_{s,d}, h_{r_1,d}, \dots, h_{r_k,d}]^t \in \mathcal{C}^{k+1}$ and $\mathbf{g} = [g_s, g_{r_1}, \dots, g_{r_k}]^t \in \mathcal{C}^{k+1}$ denote the channel coefficient vector and the beamforming weight vector, respectively. In a retransmission, the destination receives

$$Y_d = \mathbf{h}^\dagger \mathbf{g} X_r + X_\xi + z_d, \quad (5)$$

where X_r and X_ξ denote the transmitted signal by SU relays and the received interference signal from the primary transmitter to the SU destination, respectively. When accessing an idle timeslot to retransmit, there is no interference between SUs and PUs and hence $X_\xi = 0$. When accessing a busy timeslot to retransmit, the primary transmitter causes an interference signal X_ξ at the SU destination, degrading the quality of the received signal. Suppose that the primary link has an n_t -antenna transmitter and an n_r -antenna receiver. Let $\mathbf{h}_l = [h_{s,l}, h_{r_1,l}, \dots, h_{r_k,l}]^t \in \mathcal{C}^{k+1}$ denote the channel coefficient vector from the SU source and relay $r_i \in \mathcal{D}$ to the l th primary receive antenna ($l \in \{1, \dots, n_r\}$).

2.2.2 Outage probability

Let $f_{s,d}$ denote the outage probabilities that a packet is not successfully delivered on the S-D link. The S-D transmission is unsuccessful if the channel capacity is $C_{s,d} = \log_2 \left(1 + \frac{|h_{s,d}|^2 P}{\sigma_0^2} \right)$ falls below the data rate R . Accordingly, the outage probability is given by

$$f_{s,d} = \Pr\{C_{s,d} < R\} = 1 - \exp \left(-\frac{2^R - 1}{\text{SNR}_{s,d}} \right), \quad (6)$$

where $\text{SNR} = \frac{P}{\sigma_0^2}$ is the average transmitted SNR. Meanwhile, any relay r_i that receives the packet becomes a decoding relay with the probability $\Pr\{r_i \in \mathcal{D}\} = \exp\left(-\frac{2^R-1}{\text{SNR}\sigma_{s,r_i}^2}\right)$. Assuming that channel gains $|h_{s,r_i}|^2$ ($r_i \in \mathcal{K}$) are independently exponentially distributed, we have

$$\Pr\{\mathcal{D}\} = \prod_{r_i \in \mathcal{D}} \exp\left(-\frac{g_1}{\sigma_{s,r_i}^2}\right) \prod_{r_i \notin \mathcal{D}} \left[1 - \exp\left(-\frac{g_1}{\sigma_{s,r_i}^2}\right)\right],$$

where $g_1 = \frac{2^R-1}{\text{SNR}}$. When the source and decoding relays \mathcal{D} form a distributed beamformer to retransmit, the outage probability is given by

$$f_{bf}(\mathcal{D}) = \Pr\left\{C_{bf} = \log_2\left(1 + \frac{|\mathbf{h}^\dagger \mathbf{g}|^2}{\sigma_0^2 + \mathcal{I}(\Theta[t]=1)P_\xi}\right) < R\right\}. \quad (7)$$

where P_ξ denotes the interference power from the primary transmitter to the SU destination. Let \mathbf{g}_i^* and \mathbf{g}_b^* denote the optimal beamforming weight vector when accessing idle and busy timeslots, respectively. Accordingly, the outage probabilities of retransmitting in idle and busy timeslots are denoted by $f_{bf}^i(\mathcal{D})$ and $f_{bf}^b(\mathcal{D})$, respectively.

3 The basic ICaD-ARQ method

3.1 Scheme description

We first describe the basic idea of the ICaD-ARQ scheme. In the sensing phase of each timeslot t , the SU nodes listen to the channel to detect whether it is occupied by the PUs. If timeslot t is idle, then the source accesses it by broadcasting a new packet to the destination and K potential relays. The destination will immediately send back an ACK packet to the source if it correctly decodes the packet. Otherwise, it sends back a NACK packet to notify a transmission failure. At the same time, each of the K relay nodes tries to decode the packet. Those who correctly decode the packet become the so-called decoding relays that will participate in retransmitting the packet if the packet is not correctly received by the destination. In contrast to traditional cognitive relay systems where SU nodes can only transmit in idle timeslots when PUs are not active, in the proposed scheme, the source and decoding relays utilize busy timeslots to do the retransmission, thus leaving more idle slots for the source to transmit new packets. In particular, once a NACK packet from the destination is received, the source and decoding relays store the corresponding data packet in their buffers and wait for a busy timeslot to retransmit it. In this way, the reliability of the S-D link can greatly be improved. The flow chart of ICaD-ARQ is presented in Figure 2. It is assumed that ACK and NACK packets are delivered through a feedback channel quickly and reliably after each transmission.

To practise the ICaD-ARQ scheme, we should estimate the optimal beamforming weight vector \mathbf{g}_b^* for accessing busy timeslots. In the sequel, we first review our previously proposed IC approach, which adopts zero-forcing beamforming to completely cancel the interference to PUs when perfect CSI is available [14]. Subsequently, we review the robust beamforming scheme to handle imperfect-CSI scenarios.

3.2 Distributed beamforming-based IC

In this section, we briefly introduce how to obtain the optimal beamforming weight vector \mathbf{g}_b^* in perfect-CSI and imperfect-CSI scenarios, respectively. In this article, we assume that SUs cause no interference to PUs when perfect CSI can be obtained and they should suppress the interference to PUs below a threshold when perfect CSI is not available.

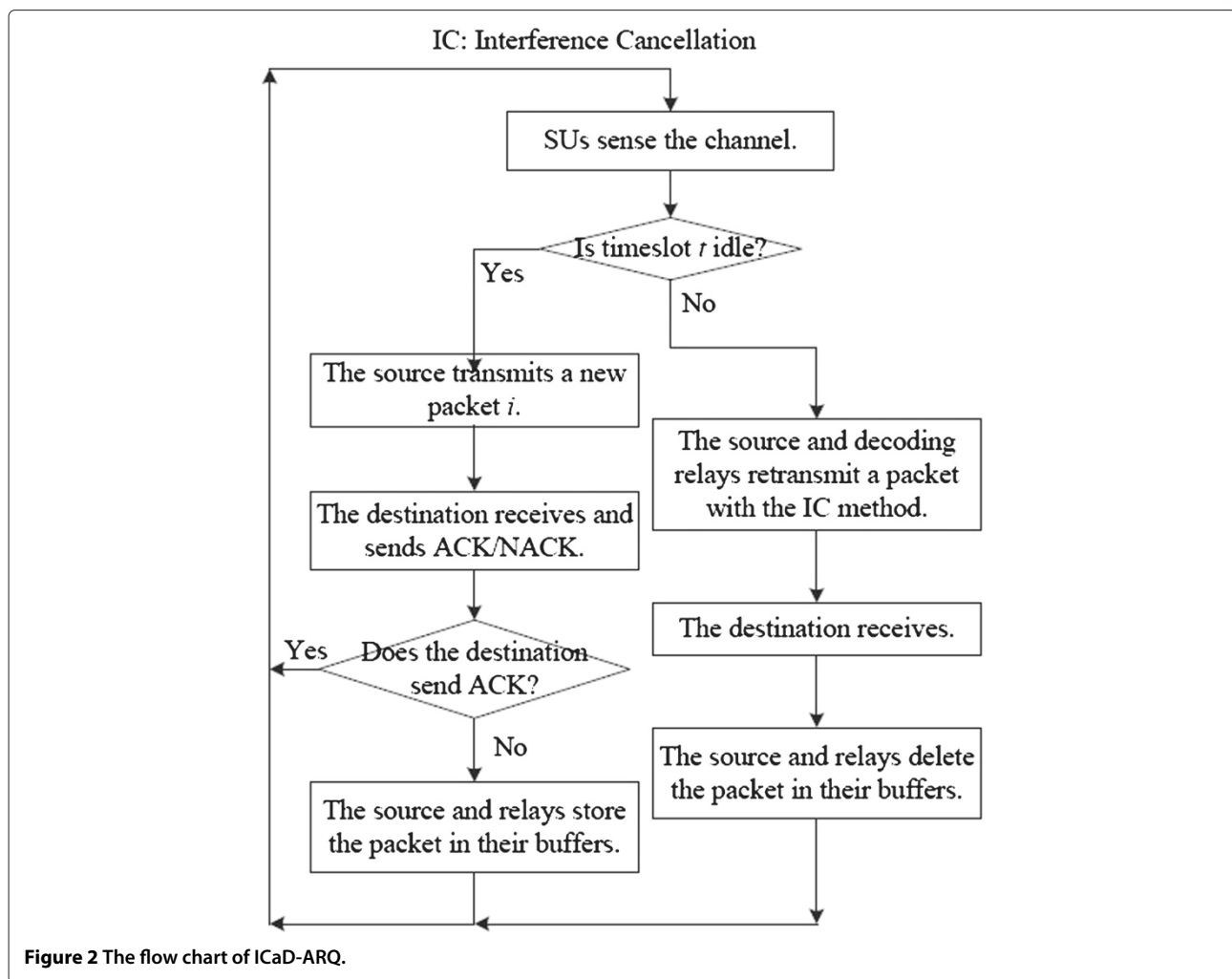
3.2.1 Perfect CSI

With perfect CSI, the zero-forcing beamformer formed by SU relays is able to null the interference from the SUs to all l primary receive antennas. From our previous work [14], we have $\mathbf{g}_b^* = \sqrt{P} \frac{\mathbf{P}_{\mathcal{V}^\perp} \mathbf{h}}{\|\mathbf{P}_{\mathcal{V}^\perp} \mathbf{h}\|}$ when accessing a busy timeslot. Here, $\mathbf{P}_{\mathcal{V}^\perp} = \mathbf{I} - \mathbf{H}(\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger$ is the orthogonal projector onto the orthogonal complementary subspace, \mathcal{V}^\perp , of the subspace $\mathcal{V} = \text{span}\{\mathbf{h}_1, \dots, \mathbf{h}_{n_r}\}$, where $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_{n_r}]$ denotes the channel coefficient matrix. Then, substituting \mathbf{g}_b^* in Equation (7) yields the outage probability $f_{bf}^b(\mathcal{D})$. From [14], zero-forcing beamforming can be applied to eliminate the interference from SU relays to PUs, when the number of decoding relays is larger than or equal to the number of receive antennas, i.e., $k \geq n_r$. Otherwise, the SUs cannot realize a zero-forcing beamformer and the retransmissions will cause interference to PUs inevitably. To avoid interference to PUs in the perfect CSI scenarios, the source and less than n_r relays should keep silent. In this case, the retransmission is always considered to be unsuccessful and accordingly the outage probability is counted as $f_{bf}^b(\mathcal{D}) = 1$ for $|\mathcal{D}| < n_r$.

3.2.2 Imperfect CSI

In general, SU relays can only obtain imperfect CSI due to estimation errors. Based on imperfect CSI, SU relays should adopt a robust beamformer to assure the interference to PUs below a threshold.

By modeling estimation errors as additive complex Gaussian noise, we have $\mathbf{h} = \hat{\mathbf{h}} + \Delta \mathbf{h}$ and $\mathbf{h}_l = \hat{\mathbf{h}}_l + \Delta \mathbf{h}_l$, where $\hat{\mathbf{h}}$ and $\hat{\mathbf{h}}_l$ are the channel estimates known at the SU relays, and $\Delta \mathbf{h}$ and $\Delta \mathbf{h}_l$ are the estimation errors. When accessing busy timeslots to retransmit, SU relays shall suppress the interference to PUs so that the received interference power at the l th primary receive antenna is below a threshold, denoted by I_l . From [19], the optimal



beamforming weight vector \mathbf{g}_b^* is the solution to the following optimization problem

$$\begin{aligned} \max_{\|\mathbf{g}\| \leq P} \quad & \Pr \left\{ |\mathbf{h}^\dagger \mathbf{g}|^2 \geq \gamma \right\} \\ \text{s.t.} \quad & \Pr \left\{ |\mathbf{h}_l^\dagger \mathbf{g}|^2 \leq I_l \right\} \geq \epsilon_l, \quad \forall l, \end{aligned} \quad (8)$$

where $\gamma = (2^R - 1)(\sigma_0^2 + P_\xi)$ and ϵ_l denote the probability threshold at which level the interference is self-controlled by SU relays. The details of computing \mathbf{g}_b^* can be referred to [19].

In the same way, by substituting \mathbf{g}_b^* in Equation (7), we can estimate the outage probability $f_{bf}^*(\mathcal{D})$. In general, it is not trivial to theoretically analyze the outage probability $f_{bf}^b(\mathcal{D})$. Fortunately, they can be estimated using training packets in practical scenarios. These physical layer parameters play an important role in selecting the optimal scheduling parameters, as shown in the following sections. By applying the distributed beamforming-based IC technique, the SUs can retransmit erroneous packets in

busy timeslots, thus sparing idle timeslots for the S–R and S–D transmissions. To further improve the whole system performance, we will further introduce two ICaD-ARQ protocols, which adopt opportunistically probabilistic and queue-length-based scheduling schemes in the link layer, respectively.

4 ICD-ARQ-PS and performance analysis

4.1 Scheme description

In the ICaD-ARQ scheme, the SUs retransmit erroneous packets in busy timeslots only. However, this cannot assure link reliability when the primary system is lightly loaded, since no sufficient SSHs can be utilized to carry out just one retransmission for each erroneous packet. To deal with this issue, we propose a more efficient ICaD-ARQ-PS scheme. In contrast to ICaD-ARQ, ICaD-ARQ-PS allows SU to exploit TSHs to retransmit if necessary and schedules retransmissions according to truncated ARQ. We denote by ω the proportion of idle timeslots allocated for retransmissions. Thus, $(1 - \omega)P_i$

and ωP_I of idle timeslots are assigned for direct transmissions and retransmissions, respectively. Let us denote by $S[t]$ whether an idle timeslot t is assigned to retransmission. $S[t] = 1$ and $S[t] = 0$ indicate that idle timeslot t is allocated to retransmitting an erroneous packet and transmitting a new packet, respectively. Thus, if timeslot t is idle and $S[t] = 0$, the source transmits a new packet to the destination. If the packet fails to be decoded, it will be stored at buffers and be retransmitted later. Otherwise, the source and decoding relays form a distributed beamformer to transmit a packet in their retransmission buffers. By applying truncated ARQ, the retransmission procedure repeats till a successful transmission is confirmed via an ACK feedback, or the number of retransmission n reaches its maximum one allowed, say N . Notice that ICaD-ARQ-PS with $\omega = 0$ and $N = 1$ is reduced to ICaD-ARQ.

Then, we will analyze the throughput performance of ICaD-ARQ-PS and discuss the optimal scheduling parameters for throughput maximization in the following two sections.

4.2 Throughput analysis

As stated above, it is sensed that PUs occupy the channel with the probability P_B and are absent with the probability $P_I = 1 - P_B$ by perfect spectrum sensing. By applying probabilistic scheduling, a proportion $1 - \omega$ of idle timeslots are allocated to transmit new packets. A packet is dropped if it cannot correctly be decoded at the destination after $N + 1$ transmissions (one direct transmission and N retransmissions). Let p_{ls} denote the average packet drop rate. The average throughput can then be given by

$$T = (1 - \omega)P_I(1 - p_{ls}). \quad (9)$$

Let $p_{rs}^{(n)}(\mathcal{D})$ denote the average outage probability that the packet is still erroneously decoded at the destination after n retransmissions given that previous n transmissions were unsuccessful, when decoding relays \mathcal{D} are involved. Since \mathcal{D} is a random set, the packet drop rate p_{ls} can be computed as

$$p_{ls} = f_{s,d} \cdot \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \prod_{n=1}^N p_{rs}^{(n)}(\mathcal{D}). \quad (10)$$

The average number of retransmissions given that the packet needs to be retransmitted is given by

$$\begin{aligned} \tilde{N}_r &= \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \left(\sum_{n=1}^N n \cdot (1 - p_{rs}^{(n)}(\mathcal{D})) \prod_{i=0}^{n-1} p_{rs}^{(i)}(\mathcal{D}) \right. \\ &\quad \left. + N \cdot \prod_{n=1}^N p_{rs}^{(n)}(\mathcal{D}) \right) = \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \sum_{n=0}^{N-1} \prod_{i=0}^n p_{rs}^{(i)}(\mathcal{D}). \end{aligned} \quad (11)$$

Here, we introduce $p_{rs}^{(0)}(\mathcal{D}) = 1$ for notational convenience. Since the amount of spectrum consumed cannot exceed the amount of spectrum allocated for retransmissions, we have

$$(1 - \omega)(1 - P_B)p_{ds}\tilde{N}_r \leq \omega(1 - P_B) + P_B. \quad (12)$$

When ARQ is applied, the success or failure of the n th transmission ($n = 0, 1, \dots, N$) depends only on the current channel state. Hence, $p_{rs}^{(n)}(\mathcal{D})$ are the same for all $n = 1, \dots, N$, i.e., $p_{rs}^{(n)}(\mathcal{D}) = p_{rs}(\mathcal{D})$. As stated above, a retransmission could be carried out in an idle timeslot with the probability $\eta \triangleq \frac{\omega P_I}{\omega P_I + P_B}$ and in a busy timeslot with the probability $1 - \eta$. Thus, the average outage probability is given by

$$p_{rs}(\mathcal{D}) = \eta f_{bf}^i(\mathcal{D}) + (1 - \eta)f_{bf}^b(\mathcal{D}), \quad (13)$$

where the outage probability $f_{bf}^i(\mathcal{D})$ can be obtained by substituting \mathbf{g}_i^* in Equation (7). Here, $\mathbf{g}_i^* = \sqrt{P} \frac{\mathbf{h}}{\|\mathbf{h}\|}$ and $\mathbf{g}_i^* = \sqrt{P} \frac{\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}$ for perfect and imperfect CSI scenarios, respectively. Hence, we have $p_{ls} = f_{s,d} \cdot \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} (p_{rs}(\mathcal{D}))^N$ and $\tilde{N}_r = \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \sum_{n=0}^{N-1} (p_{rs}(\mathcal{D}))^n$. Throughout this article, we mainly consider carrying out retransmissions according to the regular ARQ protocol. Notice that the above throughput analysis can also be extended to scenarios, where hybrid ARQ is applied to run retransmissions. In the Appendix, we derive the closed form expression of $p_{rs}^{(n)}(\mathcal{D})$ for regular and hybrid ARQ protocols in CR systems, where the source employs neighboring nodes as potential relays.

As mentioned above, the probability $p_{rs}^{(n)}(\mathcal{D})$ is a function of ω . Hence, the packet drop rate p_{ls} given by (10) and the average number of retransmissions \tilde{N}_r given by (11) are the functions of N and ω , and can be denoted by $p_{ls}(\omega, N)$ and $\tilde{N}_r(\omega, N)$, respectively. As a result, the average throughput is also a function of ω and N , and thus denoted by $T(\omega, N)$. Therefore, by selecting appropriate ω and N , the average throughput of the ICaD-ARQ-PS scheme can be maximized, as presented in the next section.

4.3 Throughput maximization

In this section, we will formulate the optimization problem and present important properties of \tilde{N}_r for the convenience of solving the problem.

The average throughput of the proposed ICaD-ARQ-PS scheme is to be maximized subject to the total number of available spectrum holes. Mathematically, the optimization problem can be formulated as

$$\max_{\omega \in [0,1], N \in \mathbb{N}^+} T = (1 - \omega)P_I(1 - p_{ls}(\omega, N)) \quad (14)$$

$$\text{s.t.} \quad \tilde{N}_r(\omega, N) \leq \frac{\omega P_I + P_B}{(1 - \omega)p_{ds}P_I}, \quad (15)$$

where the constraint stems from the inequality (12). The optimal solution to Problem (14) is denoted by (ω^*, N^*) .

For a given ω , a certain amount of spectrum resource is allocated for retransmissions. The maximum number of retransmissions N is hence limited, since each retransmission consumes one spectrum hole. Let us define $f(\omega) = \frac{\omega P_I + P_B}{(1-\omega)P_{ds}P_I}$. The average number of retransmissions $\tilde{N}_r(\omega, N)$ is hence upper bounded by $f(\omega)$. For some ω , there may exist a largest number N_l that satisfies $\tilde{N}_r(\omega, N_l) < f(\omega) < \tilde{N}_r(\omega, N_l + 1)$. In this situation, not all spectrum resources originally allocated for retransmissions are utilized actually. To make full use of the available spectrum holes, an erroneous packet is allowed to experience up to N_l retransmissions with a probability φ , and up to $N_l + 1$ retransmissions with a probability $1 - \varphi$, respectively. By doing so, the average number of retransmissions shall satisfy

$$\varphi \tilde{N}_r(\omega, N_l) + (1 - \varphi) \tilde{N}_r(\omega, N_l + 1) = f(\omega). \quad (16)$$

Thus, for each ω , we have $\varphi = \frac{\tilde{N}_r(\omega, N_l + 1) - f(\omega)}{\tilde{N}_r(\omega, N_l + 1) - \tilde{N}_r(\omega, N_l)}$. By applying such truncated ARQ, the average throughput can thus be computed as $T(\omega) = (1 - \omega)P_I(1 - \tilde{p}_{ls}(\omega))$, where the average packet drop rate becomes

$$\tilde{p}_{ls}(\omega) = \varphi \cdot p_{ls}(\omega, N_l) + (1 - \varphi) \cdot p_{ls}(\omega, N_l + 1). \quad (17)$$

We can see that Problem (14) is a mixed integer programming one, where the parameter ω is continuous and N is discrete. Meanwhile, neither the objective function nor the constraint is convex. Hence, it is not trivial to find the optimal parameters. To facilitate the discussion on the optimal solution to Problem (14), we first study the property of $\tilde{N}_r(\omega, N)$, as presented in the following lemma.

Lemma 1. (1) Given ω , $\tilde{N}_r(\omega, N)$ is an increasing function of N . (2) Given N , $\tilde{N}_r(\omega, N)$ is a decreasing function of ω .

Proof. For any $N_2 > N_1$ ($N_1, N_2 \in \mathbb{N}^+$), we have

$$\tilde{N}_r(\omega, N_2) - \tilde{N}_r(\omega, N_1) = \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \sum_{n=N_1}^{N_2-1} \prod_{i=0}^n p_{rs}^{(i)}(\mathcal{D}) \geq 0.$$

Therefore, $\tilde{N}_r(\omega, N)$ is an increasing function of N .

With the increase of ω , more idle timeslots are utilized to retransmit packets. Since a packet is more likely to be successfully retransmitted in an idle timeslot than a busy one, the probability $p_{rs}^{(n)}(\mathcal{D})$ decreases with the increase of ω . From (11), $\tilde{N}_r(\omega, N)$ is a decreasing function of ω , since each term $\prod_{i=0}^n p_{rs}^{(i)}(\mathcal{D})$ is reduced when ω is increased. \square

From this lemma, $\tilde{N}_r(\omega, N)$ increases with the increase of N till it reaches the maximum

$$\tilde{N}_{r,inf}(\omega) = \lim_{N \rightarrow \infty} \tilde{N}_r(\omega, N), \quad (18)$$

which decreases with the increase of ω . We will exploit the monotonicity of $\tilde{N}_r(\omega, N)$ to discuss the optimal solution to Problem (14) as follows.

4.4 The optimal parameters

Intuitively, when there exist sufficient SSHs, it is optimal to retransmit erroneous packets in busy timeslots only. That is, there is no need of assigning idle timeslots for retransmissions. To quantify the sufficiency of SSHs, we define a threshold

$$\tilde{P}_B = \frac{p_{ds} \cdot \tilde{N}_{r,inf}(0)}{1 + p_{ds} \cdot \tilde{N}_{r,inf}(0)} \quad (19)$$

on the probability of PUs' presence. Let ω_1 denote the solution to the equation $\tilde{N}_{r,inf}(x) - f(x) = 0$. Let ω^+ denote the optimal solution to the problem

$$\max_{\omega \in [0, \omega_1]} T(\omega) = (1 - \omega)P_I(1 - \tilde{p}_{ls}(\omega)). \quad (20)$$

Now, we discuss the optimal parameters in two cases: $P_B \geq \tilde{P}_B$ and $P_B < \tilde{P}_B$, respectively, as presented in Theorem 1. The optimal parameters are also illustrated in Figure 3 to give more insights. In this figure, both N_l and φ are determined by ω^+ , and thus can be written as functions of ω^+ , given by $N_l(\omega^+)$ and $\varphi(\omega^+)$, respectively.

Theorem 1. (1) When $P_B \geq \tilde{P}_B$, the optimal solution to (14) is given by $\omega^* = 0$, $N^* = \infty$. (2) When $P_B < \tilde{P}_B$, the optimal solution to (14) is given by $\omega^* = \omega^+$ and $\Pr\{N^* = N_l\} = \varphi$, $\Pr\{N^* = N_l + 1\} = 1 - \varphi$, where ω^+ , N_l and φ satisfy Equation (16).

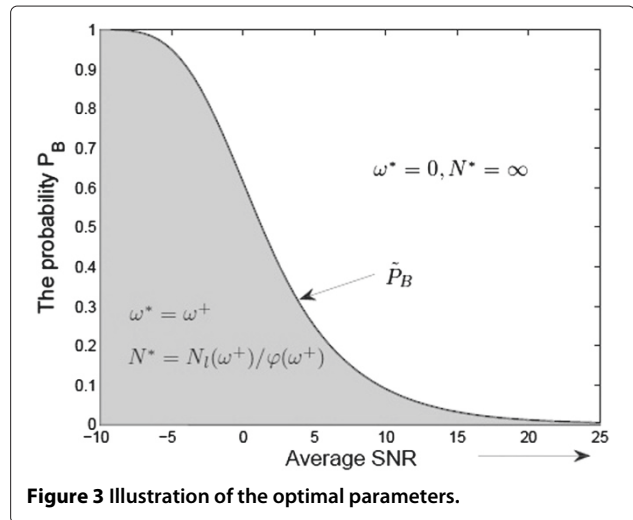


Figure 3 Illustration of the optimal parameters.

Proof. In the case when $P_B \geq \tilde{P}_B$, from Lemma 1, there exists $\tilde{N}_r(\omega, N) \leq \tilde{N}_{r,inf}(\omega) \leq \tilde{N}_{r,inf}(0) \leq f(0) \leq f(\omega)$ for all $\omega \in [0, 1]$ and $N \in \mathbb{N}^+$, since $f(\omega)$ and $\tilde{N}_{r,inf}(\omega)$ are increasing and decreasing functions of ω , respectively. That is, the maximum number of retransmission N can reach infinity, and the average packet drop rate degrades to zero, i.e.,

$$p_{ls}(\omega) = \lim_{N \rightarrow \infty} p_{ds} \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \sum_{n=0}^{N-1} \prod_{i=0}^n p_{rs}^{(i)}(\mathcal{D}) = 0.$$

Therefore, the average throughput can be given by

$$T(\omega) = \lim_{N \rightarrow \infty} T(\omega, N) = (1 - \omega)P_I,$$

which is a strictly decreasing function of ω . Thus, $\omega^* = 0$ and $N^* = \infty$ are the optimal parameters and the maximum average throughput is $T^* = P_I$.

In the case when $P_B < \tilde{P}_B$, there must exist an $\omega_1 \in (0, 1)$ that satisfies $f(0) < f(\omega_1) = \tilde{N}_{r,inf}(\omega_1) \leq \tilde{N}_{r,inf}(0)$ due to the monotonicity of $\tilde{N}_{r,inf}(\omega)$ and $f(\omega)$. Thus, we can discuss the local maximum average throughput $T(\omega)$ for $\omega \in [\omega_1, 1]$ and $\omega \in [0, \omega_1]$, respectively. When $\omega \in [\omega_1, 1]$, we have $\tilde{N}_{r,inf}(\omega) \leq \tilde{N}_{r,inf}(\omega_1) = f(\omega_1) \leq f(\omega)$. Hence, the constraint (15) always holds for all $N \in \mathbb{N}^+$. In this sense, we can apply infinite ARQ and thus have $p_{ls}(\omega) = 0$ and $T(\omega) = (1 - \omega)P_I$. Hence, the optimal solution can be searched in a reduced range $\omega \in [0, \omega_1]$. In this range, there may exist a largest integer N_l , which satisfies $\tilde{N}_r(\omega, N_l) \leq f(\omega) < \tilde{N}_{r,inf}(\omega)$. For each ω , we can find an N_l and φ which satisfy Equation (16). As a result, when $P_B < \tilde{P}_B$, we have $\omega^* = \omega^+$ and $\Pr\{N^* = N_l\} = \varphi$, $\Pr\{N^* = N_l + 1\} = 1 - \varphi$. \square

From Theorem 1, when available SSHs are sufficient, i.e., $P_B \geq \tilde{P}_B$, the optimal scheme is to transmit new packets in idle timeslots and retransmit erroneous packets in busy timeslots based on infinite ARQ. And when available TSHs are relatively sufficient, i.e., $P_B < \tilde{P}_B$, the optimal scheme is to allocate a part of idle timeslots for retransmissions that are carried out according to truncated ARQ. In this case, the optimal transmission parameters can be obtained by solving the problem (20) and Equation (16).

From the above discussion, the computational complexity for obtaining the optimal parameters can greatly be reduced compared to exhaustive search methods. First, we derive the optimal parameters in the case of $P_B \geq \tilde{P}_B$. Second, when $P_B < \tilde{P}_B$, the search range of ω is highly reduced. That is, ω^* can be found in the narrow scope $[0, \omega_1]$. Moreover, due to the unimodality of $T(\omega)$, the optimal spectrum allocation proportion ω^+ can be obtained by binary search methods [20], such as golden-section search, in $\mathcal{O}(\log(\omega_1/\delta))$ steps, where δ is the required precision. For all $\omega \in [0, \omega_1]$, $N_l(\omega)$ and $\varphi(\omega)$ can be found by solving Equation (16) via the variants of

efficient root-finding algorithms, such as bisection and Brent's methods [21]. From the proof of Theorem 1, there is $N_l(\omega_1) = \infty$. When ω is very close to ω_1 , the value of $N_l(\omega)$ could be very large. To avoid too many computations, we can set a sufficiently large integer, N_m , as an upper bound of $N_l(\omega)$ with very little performance loss, i.e., $N_l(\omega)$ can be searched from the range of $\{1, \dots, N_m\}$. The computational complexity of finding $N_l(\omega)$ is thus equal to $\mathcal{O}(\log N_m)$ by adopting binary search methods. In this way, we finally obtain the optimal parameters $\omega^* = \omega^+$, $\Pr\{N^* = N_l(\omega^+)\} = \varphi(\omega^+)$ and $\Pr\{N^* = N_l(\omega^+) + 1\} = 1 - \varphi(\omega^+)$ in $\mathcal{O}(\log(\omega_1/\delta) \cdot \log(N_m))$ steps.

5 ICaD-ARQ-QS and performance analysis

5.1 Scheme description

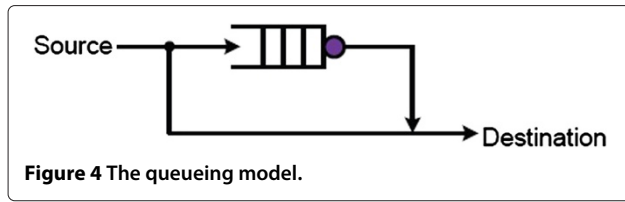
To improve the average throughput, the erroneous packets should be stored in buffers and got retransmitted by exploiting as many SSHs as possible. Conceivably, this could induce large latency to erroneous packets and thus increase the average delay. To serve elastic traffics with diverse QoS demands, we propose the ICaD-ARQ-QS scheme which schedules direct transmissions and retransmissions based on a finite buffer.

Specifically, we impose a threshold Q ($Q = 1, 2, \dots$) on the capacity of the retransmission buffer. If the number of backlogged packets in the retransmission buffer reaches the threshold Q , the source and relays will launch a retransmission in the next timeslot regardless of whether it is idle or busy. Otherwise, the source transmits new packets in idle timeslots and retransmit erroneous packets together with the relays in busy timeslots, respectively. When Q is increased from one to infinity, more and more SSHs are exploited in retransmissions, thus leading to the improved average throughput. Meanwhile, the average queueing delay that an erroneous packet may experience is increasing and thus the average delay also increases. In this sense, it is important to choose the threshold Q , when both throughput and delay are considered. Notice that when $Q = \infty$, ICaD-ARQ-QS is reduced to ICaD-ARQ.

5.2 A Markov chain model

According to the above scheme description, a part of packets are successfully delivered on the S-D link, and a part of packets that fail to be transmitted in the S-D link are successfully delivered with the help of relays. The remaining packets are dropped, since they cannot correctly be decoded by the destination on the S-D or R-D link when the channel capacity is below the data rate.

From Figure 4, the proposed incremental relaying scheme leads to a queueing process in the retransmission procedure. Let $q[t]$ denote the queue length in the end of timeslot t . The queue length is updated as $q[t] =$



$\min\{q[t-1] + a[t] - v[t], Q\}$, where $a[t]$ and $v[t]$ denote the number of packets arriving and served at the queue in the t th timeslot, respectively. The set of all possible queue sizes is denoted by $\mathcal{Q} = \{0, 1, \dots, Q\}$. The secondary transmission in this CR system is scheduled according to the status of channel occupancy by PUs and the queue length. Hence, the system state can be characterized by $s[t] = (\Theta[t], q[t])$ with $q[t] \in \mathcal{Q}$. The transition between system states can be modeled as a discrete-time two-dimensional Markov chain, as shown in Figure 5. This figure also plots the transition probabilities between the states. In particular, α indicates the probability that the PUs' state of channel occupancy transfers from idle to busy, and β means the probability that the PUs' occupancy state transfers from busy to idle. Let the symbol \bar{x} denote $\bar{x} = 1 - x$. Then, $\mu_1 = f_{s,d}$ is the probability that a packet enters the queue. It means that a packet that fails to be delivered in the S-D link will enter the retransmission queue.

The erroneous packets in the queue may be retransmitted in idle or busy timeslots by applying the distributed beamforming-based IC method, and they could be dropped due to channel fading. When decoding relays \mathcal{D} participate in the retransmission, the average failure probabilities that the retransmission is unsuccessful in an idle and busy timeslot can be computed as $f_{bf}^i = \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \cdot f_{bf}^i(\mathcal{D})$ and $f_{bf}^b = \sum_{\mathcal{D}} \Pr\{\mathcal{D}\} \cdot f_{bf}^b(\mathcal{D})$, respectively.

Specifically, we place all the states in the lexicographic order, i.e., $(0, 0), (1, 0), \dots, (0, Q), (1, Q)$. The set of states $\{(0, i), (1, i)\}$ is referred to as *level i* ($i \in \{0, 1, \dots, Q\}$).

By substituting the transition probabilities, we obtain the transition probability matrix \mathbf{P} , given by

$$\mathbf{P} = \begin{bmatrix} \mathbf{B}_0 & \mathbf{A}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_1 & \mathbf{0} \end{bmatrix}, \quad (21)$$

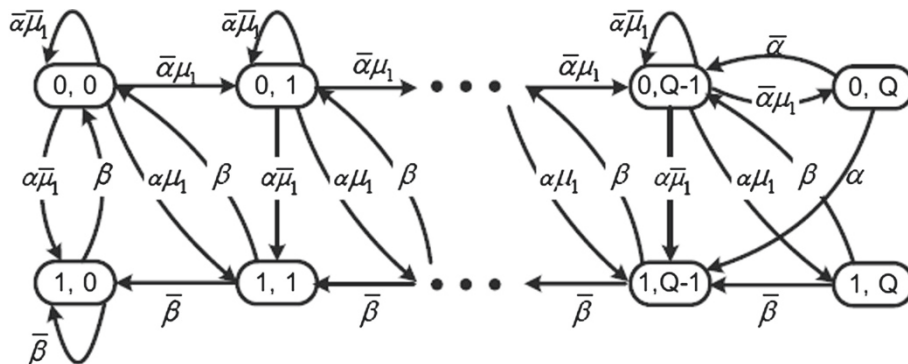
where all the blocks are 2×2 matrices. One can see that \mathbf{P} is a $2(Q+1) \times 2(Q+1)$ matrix in a block tridiagonal form. Specifically, the transition probability submatrices \mathbf{A}_2 , \mathbf{A}_1 , and \mathbf{A}_0 indicate the transition from the states of *level i* to those of *level $i-1$* , *level i* , and *level $i+1$* , respectively. \mathbf{B}_0 represents the transition between the states of *level 0*, and \mathbf{B}_1 represents the transition from the states of *level Q* to those of *level $Q-1$* . Hence, we have

$$\begin{aligned} \mathbf{A}_2 &= \begin{bmatrix} 0 & 0 \\ \beta & \bar{\beta} \end{bmatrix}, & \mathbf{A}_1 &= \begin{bmatrix} \bar{\alpha}\bar{\mu}_1 & \alpha\bar{\mu}_1 \\ 0 & 0 \end{bmatrix}, \\ \mathbf{A}_0 &= \begin{bmatrix} \bar{\alpha}\mu_1 & \alpha\mu_1 \\ 0 & 0 \end{bmatrix} \\ \mathbf{B}_0 &= \begin{bmatrix} \bar{\alpha}\bar{\mu}_1 & \alpha\bar{\mu}_1 \\ \beta & \bar{\beta} \end{bmatrix}, & \mathbf{B}_1 &= \begin{bmatrix} \bar{\alpha} & \alpha \\ \beta & \bar{\beta} \end{bmatrix}. \end{aligned}$$

Since the Markov chain characterized by \mathbf{P} is of a finite QBD type, we will analyze the system performance by matrix analytical methods [22] in the following.

5.3 Performance analysis

At first, we compute the stationary probability vector \mathbf{x} of the Markov chain \mathbf{P} . For convenience, the vector \mathbf{x} is partitioned as $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_Q]$, where \mathbf{x}_i , a 1×2 row vector, is the stationary probability vector of *level i* . Also, we denote $\mathbf{e}_i = [1, \dots, 1]^t$ to be a column vector of dimension



$2(i+1)$. When $Q = 1$, we can obtain \mathbf{x} by solving $\mathbf{x}P = \mathbf{x}$ and $\mathbf{x}\mathbf{e}_1 = 1$. By applying a matrix analytical method, we can derive the stationary distribution of the finite QBD process for $Q > 1$ [23,24].

In Figure 5, any state of the finite Markov chain can be reached from any other state. Thus, the Markov chain is ergodic. Its stationary probability vector can be obtained by solving the linear equations

$$\begin{aligned} \mathbf{x}_i &= \mathbf{v}_1 \mathbf{R}_1^i + \mathbf{v}_2 \mathbf{R}_2^{K-1-i}, \quad 0 \leq i \leq Q-1, \\ \mathbf{x}_Q &= \mathbf{x}_{Q-1} \mathbf{A}_0, \end{aligned} \quad (22)$$

where \mathbf{R}_1 is the unique minimal non-negative solution to the matrix quadratic equation $\mathbf{R} = \mathbf{A}_0 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_2$, and \mathbf{R}_2 is the unique solution to the matrix quadratic equation $\mathbf{R} = \mathbf{A}_2 + \mathbf{R}\mathbf{A}_1 + \mathbf{R}^2\mathbf{A}_0$. The coefficient vector \mathbf{v}_1 can be obtained via solving the following linear equations

$$\begin{aligned} \mathbf{v}_1 [\mathbf{R}_1^{Q-2} \mathbf{C}_0 + \mathbf{C}_1 (\mathbf{R}_2 \mathbf{A}_0 + \mathbf{A}_1 + \mathbf{A}_0 \mathbf{B}_1 - \mathbf{I})] &= \mathbf{0}, \\ \mathbf{v}_1 \left[\sum_{i=0}^{Q-1} (\mathbf{R}_1^i + \mathbf{C}_1 \mathbf{R}_2^{Q-1-i}) + (\mathbf{R}_1^{Q-1} + \mathbf{C}_1) \mathbf{A}_0 \right] \mathbf{e}_0 &= 1, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \mathbf{C}_0 &= \mathbf{A}_0 + \mathbf{R}_1 (\mathbf{A}_1 + \mathbf{A}_0 \mathbf{B}_1 - \mathbf{I}), \\ \mathbf{C}_1 &= (\mathbf{I} - \mathbf{R}_1 \mathbf{A}_2 - \mathbf{B}_0) [\mathbf{R}_2^{Q-2} (\mathbf{R}_2 \mathbf{B}_0 - \mathbf{R}_2 + \mathbf{A}_2)]^{-1}. \end{aligned} \quad (24)$$

The coefficient vector \mathbf{v}_2 is given by $\mathbf{v}_2 = \mathbf{v}_1 \mathbf{C}_1$.

When $Q \rightarrow \infty$, the vector $[\mathbf{v}_1, \mathbf{v}_2]$ becomes $[\mathbf{x}_0, \mathbf{0}]$, and the stationary probability vector of level i is reduced to

$$\mathbf{x}_i = \mathbf{x}_0 \mathbf{R}_1^i, \quad \forall i \geq 1, \quad (25)$$

where the boundary vector satisfies

$$\begin{aligned} \mathbf{x}_0 (\mathbf{B}_0 + \mathbf{R}_1 \mathbf{A}_2 - \mathbf{I}) &= \mathbf{0} \\ \mathbf{x}_0 (\mathbf{I} - \mathbf{R}_1)^{-1} \mathbf{e}_0 &= 1. \end{aligned} \quad (26)$$

In this way, we can compute the stationary distribution for the Markov chain with a finite or infinite queue capacity in an efficient and unified way.

After obtaining the stationary distribution \mathbf{x} , we then derive the average throughput and the average delay of the proposed scheme, respectively. According to the above description, the average throughput, defined as the ratio of the number of packets successful delivered and the total number of timeslots spent, can be computed as

$$T = \bar{f}_{s,d} \cdot \sum_{i=0}^{Q-1} \mathbf{x}_i(1) + \bar{f}_{bf}^i \cdot \mathbf{x}_Q(1) + \bar{f}_{bf}^b \cdot \sum_{i=1}^Q \mathbf{x}_i(2), \quad (27)$$

where $\mathbf{x}_i(j)$ denotes the stationary probability of state $(j-1, 0)$. In Equation (27), the first term expresses the average throughput on the S-D link. The second and third terms jointly give the average throughput during the retransmission procedure. In particular, the second term

means that a part of packets are successfully retransmitted in idle timeslots. The third one implies the successful retransmissions in busy timeslots.

The average delay specifies the number of timeslots that each packet experiences when it is delivered from the source to the destination. As mentioned above, a packet can successfully be transmitted from the source to the destination directly or with the help of relays. The erroneous packets may experience a queueing process. Using Little's law, the average queueing delay is equal to $D_{\text{que}} = \frac{1}{\lambda} \sum_{i=0}^Q i \cdot \mathbf{x}_i \mathbf{e}_0$, where $\lambda = f_{s,d} \bar{f}_{s,r} \cdot \sum_{i=0}^{Q-1} \mathbf{x}_i(1)$ is the arrival rate at the queue. For $Q = 1$ and $Q = \infty$, the average delay can be computed as

$$D = \frac{p_d}{p_d + p_{re}} \cdot 1 + \frac{p_{re}}{p_d + p_{re}} \cdot (1 + D_{\text{que}}), \quad (28)$$

where $p_d = \bar{f}_{s,d} \cdot \sum_{i=0}^{Q-1} \mathbf{x}_i(1)$ and $p_{re} = \bar{f}_{bf}^i \cdot \mathbf{x}_Q(1) + \bar{f}_{bf}^b \cdot \sum_{i=1}^Q \mathbf{x}_i(2)$ are proportions of successful direct transmission and retransmission, respectively. In other cases, the average delay performance will be demonstrated by simulation results.

From the above, we can see that the performances of the proposed ICaD-ARQ schemes greatly rely on the system parameters. When the ICaD-ARQ-PS scheme is applied, the average throughput is $1 - P_B$ when $P_B \geq \tilde{P}_B$, since the optimal parameters are $\omega^* = 0$ and $N^* = \infty$. Otherwise, the optimal parameters ω^* and N^* can be obtained by computations as presented in Theorem 1. Given ω , the average throughput is improved when the parameter N increases from 1 to $N_l(\omega)$. On the other hand, it grows with the increase of the average SNR till it reaches $(1 - \omega)(1 - P_B)$. When the ICaD-ARQ-QS scheme is applied, the average delay decreases with the increase of the threshold Q . While the average throughput decreases and increases with the increase of Q in the low-SNR and high-SNR regimes, respectively. When the average SNR is sufficiently high, the average throughput increases to $1 - P_B$ and the average delay is reduced to one. In this case, the proposed ICaD-ARQ-QS scheme is reduced to direct transmission, thanks to very good channel conditions.

The proposed ICaD-ARQ schemes can be extended to the case when multiple primary links exist. In this case, SUs should null the interference to any receive antenna of multiple primary links, when they utilize busy timeslots. At the same time, the interference from the transmitters of multiple primary links to the destination should be considered. Thus, the physical layer outage probability of accessing busy timeslots may be reduced. On the other hand, the MAC layer of the proposed protocols remains unchanged, since ICaD-ARQ-PS and ICaD-ARQ-QS schedule the transmission and retransmissions according to the sensing result in each timeslot

and the transmission parameters. With the known statistical channel occupancy model by PUs, we can select appropriate transmission parameters by analysis and/or simulations.

6 Simulation results

In this section, we present theoretical and simulation results to demonstrate the performance of the proposed schemes. In the following experiments, the SU source employs three cognitive nodes as relays to jointly carry out retransmissions by forming a distributed beamformer. Suppose that three SU relays lie between the source and the destination. Let $d_{a,b}$ denote the distance from node a to node b . For simplicity, we assume that the SU relays are closely located together and the maximum distance between SU relays is negligible compared to the S–R and R–D distances. Thus, the S–R and R–D distances are normalized as $d_{s,r_i} = d_{s,r} = 1, d_{r_i} = d_{r,d} = 0.8$, and the distances from SUs to the primary receiver are set to be $d_{s,p} = 2.15, d_{r,p} = 1.5d_{r,d}$. Assume that the channel coefficient follows flat Rayleigh fading, i.e., $h_{a,b} = d_{a,b}^{-2} \cdot \tilde{h}_{a,b}$ with $\tilde{h}_{a,b} \sim \mathcal{CN}(0, 1)$. We set $n_t = n_r = 1$ and assume that $\Delta\tilde{\mathbf{h}}, \Delta\tilde{\mathbf{h}}_1 \sim \mathcal{CN}(0, 0.05\mathbf{I})$. The interference threshold is set to be $I_1 = -5$ dB and the probability threshold is $\epsilon_1 = 0.95$. The average interference SNR from the primary transmitter at the SU destination is $E\{P_\xi/\sigma_0^2\} = 0$ dB. The data rate R is set to be 1 bit/s/Hz. Simulation results are obtained by running 100 simulations, each over 10^5 timeslots.

Figures 6 and 7 demonstrate the average throughput performance of the ICaD-ARQ-PS scheme in perfect and imperfect CSI scenarios, respectively. In this experiment, we set $\alpha = 0.6$ and $\beta = 0.3$. Specifically, Figure 6 plots the average throughput versus the average SNR for different

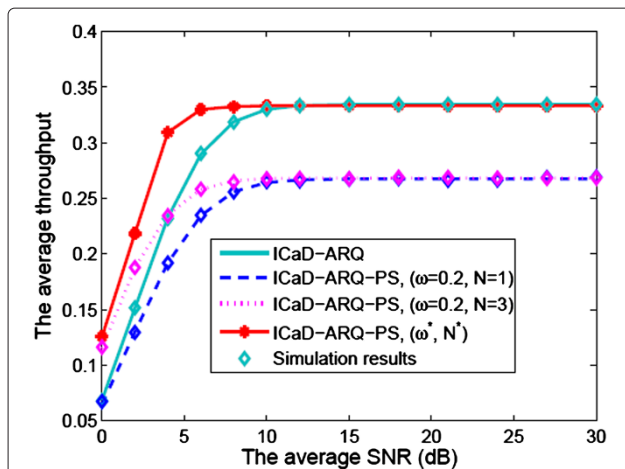


Figure 6 ICaD-ARQ-PS: the average throughput versus the average SNR with perfect CSI.

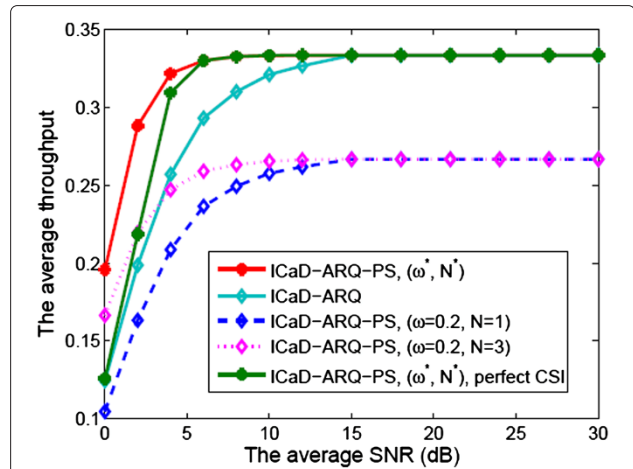
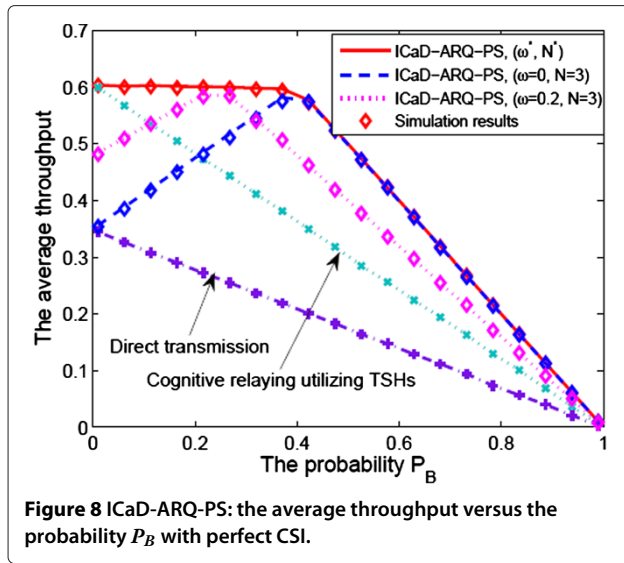


Figure 7 ICaD-ARQ-PS: the average throughput versus the average SNR with imperfect CSI.

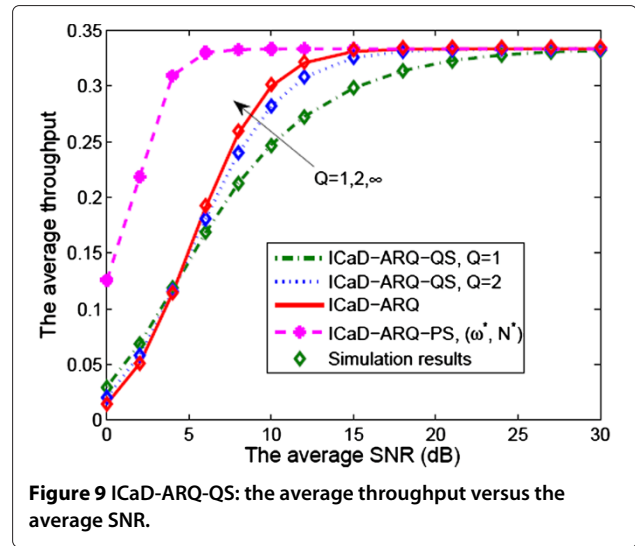
parameters ω and N with full CSI. It is observed from that the theoretical results match simulation ones very well. ICaD-ARQ-PS achieves a higher average throughput with the increase of the average SNR, since the packet drop rate is reduced. When the average SNR becomes sufficiently high, the average throughput approximately increases to $1 - P_B$. In this case, the channel condition is good enough and each packet can successfully be delivered with the help of relays by exploiting SSHs. Thus, the packet drop rate is very small and can be ignored. Given the parameter ω , the average throughput of ICaD-ARQ-PS increases with the increase of N in the low-SNR regime, since the packet drop rate is reduced by carrying out more retransmissions. In the high-SNR regime, very few packets fail to be successfully delivered in the S–D link and a smaller number of retransmissions are required. Hence, when the average SNR is larger than 12 dB, ICaD-ARQ, a special case of ICaD-ARQ-PS, is nearly the optimal scheme for throughput maximization. In this figure, one can also see that it is important to select optimal parameters ω^* and N^* to achieve the maximum throughput in different scenarios. Figure 7 presents the average throughput performance of ICaD-ARQ-PS in imperfect CSI scenarios. In this case, the ICaD-ARQ-PS scheme performs exactly as in the perfect CSI case. It can also improve the throughput by selecting optimal parameters ω^* and N^* . At the same time, it is interesting to see that in the low-SNR regime, ICaD-ARQ-PS may achieve a higher average throughput when there exist channel estimation errors compared to that in perfect CSI scenarios. This is because that in our settings, no interference should be imposed to PUs with known CSI, while PUs are supposed to endure a certain amount of interference in the imperfect CSI case. When the average SNR is higher than 10 dB, ICaD-ARQ-PS can



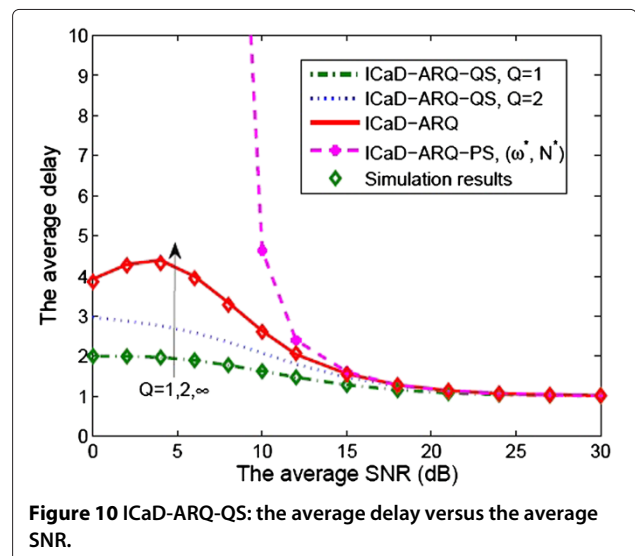
achieve the maximum average throughput by choosing the optimal parameters even when perfect CSI is not available.

Figure 8 demonstrates how the average throughput of ICaD-ARQ-PS changes with the probability P_B in the perfect CSI scenarios for the average SNR equal to 10 dB. From the previous scheme description, we can see that the throughput performance of ICaD-ARQ-PS does not rely on the realization of PUs' presence or absence on the channel but only on the probability P_B . In this experiment, we set $\alpha = P_B$ and $\beta = 1 - P_B$ without loss of generality. For comparison, we also demonstrate the throughput performance of conventional cognitive relaying and direct transmission schemes that utilize only TSHs. The average throughput of the two reference schemes linearly decrease with the increase of the probability P_B . This can easily be understood since available TSHs decreases with the increase of P_B . In contrast, given $N = 3$, the average throughput of ICaD-ARQ-PS with $\omega = 0.2$ and that with $\omega = 0$ first increase with the increase of P_B , reaches their maximum values at about $P_B = 0.3$ and $P_B = 0.5$, respectively, and then decrease to zero with the increase of P_B . The reason can be explained as follows. When P_B is relatively smaller, a larger P_B (or more accessible SSHs) leads to a significantly reduced packet drop rate. When P_B is relatively higher, the average throughput is proportional to $(1 - \omega)(1 - P_B)$, since the average packet drop rate stays constant when the allocated spectrum resource can afford three retransmissions. By selecting the optimal parameters ω^* and N^* , our proposed ICaD-ARQ-PS scheme always achieves a much higher throughput than cognitive relaying and direction transmission, since a large number of SSHs can also efficiently be utilized using the distributed beamforming-based IC method.

Figures 9 and 10 demonstrate the average throughput and delay performance of the proposed schemes,



respectively. From these two figures, one can see that the throughput and delay performance of the proposed ICaD-ARQ-QS scheme are greatly impacted by the threshold Q . In general, the average throughput increases while the average delay decreases with the increase of Q for the SNR field of interest. With a larger Q , ICaD-ARQ-QS can exploit cooperative beamforming to access more SSHs to retransmit and thus achieve a higher throughput gain. Meanwhile, each erroneous packet may experience a larger queueing delay. In this sense, it is appropriate to select a smaller Q when serving delay-sensitive traffics and a larger Q when serving delay non-sensitive traffics. Considering both throughput and latency, the proposed ICaD-ARQ-QS scheme, which can control the retransmission procedure according to the queue length, is more flexible to transmit traffics with diverse QoS requirements.



From Figures 9 and 10, ICaD-ARQ-PS achieves the maximum average throughput at the cost of the largest latency. In contrast, ICaD-ARQ-QS achieves a smaller throughput, however, can achieve different throughput-delay tradeoffs by selecting different Q . In the high-SNR regime, the channel condition is good enough and very few packets fail to be delivered on the S-D link. Hence, the average throughput and delay of the proposed ICaD-ARQ-PS and ICaD-ARQ-QS schemes are almost the same, when the average SNR is sufficiently high. We also notice that both ICaD-ARQ-PS with $\omega = 0$ and $N = 1$ and ICaD-ARQ-QS with $Q = \infty$ are reduced to the basic ICaD-ARQ scheme. Hence, by applying two different analysis methods, we have obtained the same average throughput for the ICaD-ARQ scheme, as shown in Figures 6 and 9.

7 Conclusions

In this article, we proposed an IC-aided distributed ARQ method for efficiently delivering heterogeneous elastic traffics in cognitive relay networks. By applying the ICaD-ARQ method, the SU source shall exploit TSHs to transmit to the destination. And it can form a distributed beamformer jointly with the SU relays to carry out packet retransmissions in busy timeslots without causing harmful interference to PUs. In this way, precious TSHs can efficiently be utilized for delivering more packets successfully. We also proposed two practical protocols: ICaD-ARQ-PS and ICaD-ARQ-QS, to transmit heterogeneous elastic traffics with diverse QoS demands. In particular, ICaD-ARQ-PS should be utilized to serve elastic traffics with no delay constraint. It is found that among the three schemes, ICaD-ARQ-PS can always achieve the maximum average throughput by selecting the optimal parameters ω^* and N^* in different scenarios, while ICaD-ARQ-QS can be applied to deal with elastic traffics with delay requirements. Using a matrix analytical method, we derived the average throughput and delay performance of ICaD-ARQ-QS. Simulation results confirmed our theoretical analysis and showed that ICaD-ARQ-QS can achieve different throughput-delay tradeoffs by setting different queue thresholds. Thanks to the idea of cross-layer design, our proposed schemes can always select optimal scheduling parameters and thus perform well even when perfect CSI is not available.

8 Appendix

We consider to apply the ICaD-ARQ-PS scheme in a scenario, where the source employs neighboring nodes as potential relays. In this case, we can assume that $\sigma_{s,d}^2 = \sigma_{r_i,d}^2 = 1 \ll \sigma_{s,r_i}^2 = \sigma^2$ since the source is much closer to the relays than to the destination and present a closed-form expression of the average throughput and the optimal parameters. For simplicity, we assume that

$n_t = n_r = 1$ and $P_\xi = |h_{\xi,d}^{(n)}|^2 P_{\xi_t}$, where P_{ξ_t} denotes the primary transmission power and $h_{\xi,d}^{(n)}$ denotes the interference channel coefficient that follows flat Rayleigh fading in the n th retransmission.

When the source broadcasts in an idle timeslot, the probability that a relay correctly decodes the source packet can be given by $p_1 = \Pr\{C_{s,r_i}^{(0)} \geq R\} = \exp\left(-\frac{2^R-1}{\text{SNR}\sigma^2}\right)$. The probability that k among K potential relays will participate in possible retransmissions is equal to $c_1(k) = \binom{K}{k} p_1^k (1-p_1)^{K-k}$. Then, the source and k decoding relays form a distributed beamformer to carry out at most N retransmissions.

Let us define $F(x, L) = 1 - \sum_{l=0}^L \frac{1}{l!} x^l \exp(-x)$. In the n th k -relay retransmission, the outage probability when accessing an idle timeslot can be computed as $f_{bf}^i(k) = F(x_0, k)$ with $x_0 = \frac{2^R-1}{\text{SNR}}$. When accessing a busy timeslot to retransmit, the outage probability conditioned on $|h_{\xi,d}^{(n)}|^2 = y$ is given by

$$f_{bf}^b(k, y) = 1 - \mathcal{I}(k \geq n_r) \cdot (1 - F(x(y), k - n_r)), \quad (29)$$

where $x(y) = \frac{2^R-1}{\text{SNR}} \left(1 + \frac{y P_{\xi_t}}{\sigma_0^2}\right)$ and we use $\mathcal{I}(k \geq n_r)$ to specify that a zero-forcing beamformer can be carried out only when $k \geq n_r$.

Under Rayleigh fading, the random channel power gain $y = |h_{\xi,d}^{(n)}|^2$ is exponentially distributed with the parameter λ . Let $g(y)$ denote the probability density function (pdf) of y . In the case of $k < n_r$, zero-forcing beamforming cannot be realized. Thus, we have $f_{bf}^b(k, y) = 1$ and $f_{bf}^b(k) = \int_0^\infty f_{bf}^b(k, y) g(y) dy = 1$. In the case of $k \geq n_r$, by averaging on the exponentially distributed random variable y , we have

$$\begin{aligned} f_{bf}^b(k) &= \int_0^\infty F(x(y), k - n_r) g(y) dy \\ &= 1 - \sum_{l=0}^{k-n_r} \frac{\lambda}{l!} \int_0^\infty (c_0 + c_\xi y)^l \exp[-(c_0 + c_\xi y + \lambda y)] dy \\ &= 1 - \frac{\lambda}{c_\xi + \lambda} \sum_{l=0}^{k-n_r} \frac{1}{l!} \sum_{m=0}^l \binom{l}{m} a_\xi^{l-m} b_\xi^m \sum_{r=0}^m \frac{m! c_0^{m-r} e^{-c_0}}{(m-r)!} \\ &= 1 - \frac{\lambda}{c_\xi + \lambda} \sum_{l=0}^{k-n_r} \sum_{m=0}^l \frac{a_\xi^{l-m} b_\xi^m}{(l-m)!} (1 - F(c_0, m)), \end{aligned}$$

where $c_0 = x_0$, $c_\xi = \frac{c_0 P_{\xi_t}}{\sigma_0^2}$, $a_\xi = \frac{c_0 \lambda}{c_\xi + \lambda}$, and $b_\xi = \frac{c_\xi}{c_\xi + \lambda}$. Combining these two cases, we obtain

$$\begin{aligned} f_{bf}^b(k) &= 1 - \mathcal{I}(k \geq n_r) \\ &\cdot \frac{\lambda}{c_\xi + \lambda} \sum_{l=0}^{k-n_r} \sum_{m=0}^l \frac{a_\xi^{l-m} b_\xi^m}{(l-m)!} (1 - F(c_0, m)). \end{aligned} \quad (30)$$

and thus the average outage probability

$$p_{rs}(k) = \eta \cdot f_{bf}^i(k) + (1 - \eta) f_{bf}^b(k). \quad (31)$$

When hybrid ARQ is applied, the received signals from previous transmissions are soft combined with the currently received signals using MRC rule before decision at the destination. Let $f_p(y_1, \dots, y_n)$ denote the joint pdf for independent random variables $y_i = |h_{\xi,d}^{(i)}|^2$ ($i = 1, \dots, n$). Let $\Theta_i \in \mathcal{E} = \{0, 1\}$ denote the status of PUs' presence or absence in the i th retransmission. Similarly, we use $\mathbf{E}^{(n)} = [\Theta_1, \dots, \Theta_n] \in \mathcal{E}^n$ denote a realization of the PUs' presence or absence over n retransmissions. Given $\mathbf{E}^{(n)}$ and $y_i = |h_{\xi,d}^{(i)}|^2$ for all $i = 1, \dots, n$, the instantaneous channel capacity due to MRC is

$$C_{bf}(\mathcal{D}, \mathbf{E}^{(n)}, y_1, \dots, y_n) = \log_2 \left(1 + \frac{|h_{s,d}^{(0)}|^2 P}{N_0} + \sum_{i=1}^n \frac{|h_{\xi,d}^{(i)}|^2 P}{N_0 + \mathcal{I}(\Theta_i = 1) \cdot y_i P_{\xi}} \right),$$

which depends on the channel conditions the packet has experienced in the current n th retransmission and all previous n transmissions. If $C_{bf}(\mathcal{D}, \mathbf{E}^{(n)}, y_1, \dots, y_n) < R$, the packet still cannot correctly be decoded at the destination after $n + 1$ transmissions.

By averaging over $\mathbf{E}^{(n)}$ and all y_i , the probability $p_{rs}^{(n)}(k)$ is given by

$$p_{rs}^{(n)}(k) = \sum_{\mathbf{E}^{(n)} \in \mathcal{E}^n} \Pr\{\mathbf{E}^{(n)}\} \left(\int_0^\infty \dots \int_0^\infty f_p(y_1, \dots, y_n) \times \frac{\Pr\{C_{bf}(k, \mathbf{E}^{(n)}, y_1, \dots, y_n) < R\}}{\Pr\{C_{bf}(k, \mathbf{E}^{(n)}, y_1, \dots, y_{n-1}) < R\}} dy_1 \dots dy_n \right).$$

Notice that $2|h_{s,d}^{(0)}|^2$ and $2|h_{\xi,d}^{(i)}|^2$ ($i = 1, \dots, n$) are independent chi-square random variables with

$$2v_i = \begin{cases} 2, & i = 0 \\ 2(k + 1 - \mathcal{I}(\Theta_i = 1)), & i = 1, \dots, n \end{cases} \quad (32)$$

degrees of freedom, denoted by $\chi_i^2(2v_i)$. Let us denote $\tilde{c}_0 = \frac{P}{2N_0}$ and $\tilde{c}_i = \frac{P}{2(N_0 + \mathcal{I}(\Theta_i = 1) \cdot y_i P_{\xi})}$. From [25], we can calculate the average outage probability conditioned on $\mathbf{E}^{(n)}$ and all y_i ($i = 1, \dots, n$) as

$$\begin{aligned} & \Pr\{C_{bf}(k, \mathbf{E}^{(n)}, y_1, \dots, y_n) < R\} \\ &= \Pr\left\{\sum_{i=0}^n \tilde{c}_i \chi_i^2(2v_i) < 2^R - 1\right\} \\ &= \prod_{i=0}^n (\tilde{c}_{\min}/\tilde{c}_i)^{v_i} \sum_{m=0}^\infty \delta_m \int_0^{2^R-1} g_m(\mu) d\mu, \end{aligned} \quad (33)$$

where $\tilde{c}_{\min} = \min_{i \in \{0, 1, \dots, n\}} \{\tilde{c}_i\}$, δ_m can iteratively be computed

as $\delta_m = \frac{1}{m} \sum_{l=1}^m \sum_{i=0}^n v_i (1 - \tilde{c}_{\min}/\tilde{c}_i)^l \delta_{m-l}$ for all $m = 1, 2, \dots$ with $\delta_0 = 1$, and $g_m(\mu) = \frac{\mu^{s+m-1} \exp\{-\frac{\mu}{2\tilde{c}_{\min}}\}}{(2\tilde{c}_{\min})^{s+m} \Gamma(s+m)}$ with

$s = \sum_{i=0}^n v_i$ and the Gamma function $\Gamma(s + m) = (s + m - 1)!$.

Abbreviations

ARQ: automatic repeat request; AWGN: additive white Gaussian noise; CR: cognitive radio; CSI: channel state information; IC: interference cancellation; ICaD-ARQ-PS: ICaD-ARQ with probabilistic scheduling; ICaD-ARQ-QS: ICaD-ARQ with queue-length-based scheduling; ICaD-ARQ: IC-aided distributed ARQ; NACK: negative ACKnowledgement; QoS: quality of service; SSH: spatial spectrum hole; TSH: temporal spectrum hole.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported in part by National Basic Research Program (973) of China under Grant No. 2013CB336600 and No. 2012CB316001, and NSFC Grant under Grant No. 60902001 and No. 61021001.

Author details

¹State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology (TNList) and Department of Electronic Engineering, Tsinghua University, Beijing, China. ²Department of Information Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong.

Received: 10 July 2012 Accepted: 27 November 2012

Published: 13 February 2013

References

1. JN Laneman, GW Wornell, Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inf. Theory*. **49**, 2415–2425 (2003)
2. B Zhao, MC Valenti, Practical relay networks: a generalization of hybrid ARQ. *IEEE J. Sel. Areas Commun.* **23**, 7–18 (2005)
3. I Krikidis, Distributed truncated ARQ protocol for cooperative diversity networks. *IEE Commun.* **1**, 1212–1217 (2007)
4. L Dai, KB Letaief, Throughput maximization of ad-hoc wireless networks using adaptive cooperative diversity and truncated ARQ. *IEEE Trans. Commun.* **56**, 1907–1918 (2008)
5. RA Comroe, DJ Costello, ARQ schemes for data transmission in mobile radio systems. *IEEE J. Sel. Areas Commun.* **SAC-2**, 472–481 (1984)
6. J Mitola, GQ Maquire, Cognitive radio: making software radios more personal. *IEEE Personal Commun.* **6**, 13–18 (1999)
7. FK Jondral, evolution to cognitive radio Software-defined radio-basics. *EURASIP J. Wirel. Commun. Netw.* **5**, 275–283 (2005)
8. L Lu, X Zhou, U Onunkwo, Y (G) Li, Ten years of research in spectrum sensing and sharing in cognitive radio. *EURASIP J. Wirel. Commun. Netw.* **2012**, 28 (2012)
9. O Simeone, Y Bar-Ness, U Spagnolini, Stable throughput of cognitive radios with and without relaying capability. *IEEE Trans. Commun.* **55**, 2351–2360 (2007)
10. H Boujemaa, Delay analysis of cooperative truncated HARQ with opportunistic relaying. *IEEE Trans. Veh. Technol.* **58**, 4795–4804 (2009)
11. AK Sadek, KJ Ray Liu, A Ephremides, Cognitive multiple access via cooperation: protocol design and performance analysis. *IEEE Trans. Inf. Theory*. **53**, 3677–3696 (2007)
12. I Krikidis, JN Laneman, JS Thompson, S McLaughlin, Protocol design and throughput analysis for multi-user cognitive cooperative systems. *IEEE Trans. Wirel. Commun.* **8**, 4740–4751 (2009)
13. G Zhao, J Ma, Y (G) Li, T Wu, Y Kwon, A Soong, C Yang, Spatial spectrum holes for cognitive radio with relay-assisted directional transmission. *IEEE Trans. Wirel. Commun.* **8**, 5270–5279 (2009)
14. J Liu, W Chen, Z Cao, YJ (Angela) Zhang, *A distributed beamforming approach for enhanced opportunistic spectrum access in cognitive radios*. (Hawaii, USA, 30 November–4 December 2009), pp. 1–6
15. R Zhang, Y-C Liang, Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks. *IEEE J. Sel. Topics Signal Process.* **2**, 88–102 (2008)

16. J Liu, W Chen, Z Cao, YJ (Angela) Zhang, Cooperative beamforming for cognitive radio networks: a cross-layer design. *IEEE Trans. Commun.* **60**, 1420–1431 (2012)
17. J Liu, W Chen, Z Cao, YJ (Angela) Zhang, Delay optimal scheduling for cognitive radios with cooperative beamforming: a structured matrix-geometric method. *IEEE Trans. Mob. Comput.* **11**, 1412–1423 (2012)
18. MM Rashid, J Hossain, E Hossain, VK Bhargava, Opportunistic spectrum scheduling for multiuser cognitive radio: a queueing analysis. *IEEE Trans. Wirel. Commun.* **8**, 5259–5269 (2009)
19. G Zheng, S Ma, K-K Wong, T-S Ng, Robust beamforming in cognitive radio. *IEEE Trans. Wirel. Commun.* **9**, 570–576 (2010)
20. A Antoniou, W-S Lu, *Practical Optimization: Algorithms and Engineering Applications*. (Springer, New York, 2007)
21. KE Atkinson, *An Introduction to Numerical Analysis*, 2nd edn. (Wiley, New York, 1989)
22. MF Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. (Johns Hopkins University Press, Baltimore, 1981)
23. N Akar, K Sohraby, in *Proceedings of the IEEE Infocom*, vol. 3. Finite and infinite QBD chains: a simple and unifying algorithmic approach. (Kobe, Japan, 7–12 April 1997), pp. 1105–1113
24. G Latouche, V Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. (SIAM, Philadelphia, PA, 1999)
25. PG Moschopoulos, The distribution of the sum of independent Gamma random variables. *Ann. Inst. Stat. Math. Part A*. **37**, 541–544 (1985)

doi:10.1186/1687-1499-2013-29

Cite this article as: Liu et al.: Active interference cancellation-aided QoS-aware distributed ARQ for cognitive radios with heterogeneous traffics. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:29.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com