Cognitive Access Policies under a Primary ARQ process via Forward-Backward Interference Cancellation

Nicolò Michelusi, Petar Popovski, Osvaldo Simeone, Marco Levorato, Michele Zorzi

***Abstract*—This paper introduces a novel technique for access by a cognitive Secondary User (SU) using best-effort transmission to a spectrum with an incumbent Primary User (PU), which uses Type-I Hybrid ARQ. The technique leverages the primary ARQ protocol to perform Interference Cancellation (IC) at the SU receiver (SUrx). Two IC mechanisms that work in concert are introduced: *Forward IC*, where SUrx, after decoding the PU message, cancels its interference in the (possible) following PU retransmissions of the same message, to improve the SU throughput; *Backward IC*, where SUrx performs IC on previous SU transmissions, whose decoding failed due to severe PU interference. Secondary access policies are designed that determine the secondary access probability in each state of the network so as to maximize the average long-term SU throughput by opportunistically leveraging IC, while causing bounded average long-term PU throughput degradation and SU power expenditure. It is proved that the optimal policy prescribes that the SU prioritizes its access in the states where SUrx knows the PU message, thus enabling IC. An algorithm is provided to optimally allocate additional secondary access opportunities in the states where the PU message is unknown. Numerical results are shown to assess the throughput gain provided by the proposed techniques.**

***Index Terms*—Cognitive radios, resource allocation, Markov decision processes, ARQ, interference cancellation**

I. Introduction

Cognitive Radios (CRs) [3] offer a novel paradigm for improving the efficiency of spectrum usage in wireless networks. Smart users, referred to as *Secondary Users* (SUs), adapt their operation in order to opportunistically leverage the channel resource while generating bounded interference to the Primary Users (PUs) [4]–[6]. For a survey on cognitive radio, dynamic spectrum access and the related research challenges, we refer the interested reader to [6]–[9].

In a standard model for cognitive radio, the PU is a legacy system oblivious to the presence of the SU, which needs to satisfy given constraints on the performance loss caused to the PU (underlay cognitive radio paradigm [8]). Within this framework, we propose to exploit the intrinsic redundancy, in the form of copies of PU packets, introduced by the Type-I Hybrid Automatic Retransmission reQuest (Type-I HARQ [10]) protocol implemented by the PU by enabling Interference Cancellation (IC) at the SU receiver (SUrx). We introduce two IC schemes that work in concert, both enabled by the underlying retransmission process of the PU. With *Forward IC* (FIC), SUrx, after decoding the PU message, performs IC in the next PU retransmission attempts, if these occur. While FIC provides IC on SU transmissions performed in future time-slots, *Backward IC* (BIC) provides IC on SU transmissions performed in previous time-slots within the same primary ARQ retransmission window, whose decoding failed due to severe interference from the PU. BIC relies on buffering of the received signals. Based on these IC schemes, we model the state evolution of the PU-SU network as a Markov Decision Process [11], [12], induced by the specific access policy used by the SU, which determines its access probability in each state of the network. Following the approach put forth by [13], we study the problem of designing optimal secondary access policies that maximize the average long-term SU throughput by opportunistically leveraging FIC and BIC, while causing a bounded average long-term throughput loss to the PU and a bounded average long-term SU power expenditure. We show that the optimal strategy dictates that the SU prioritizes its channel access in the states where SUrx knows the PU message, thus enabling IC; moreover, we provide an algorithm to optimally allocate additional secondary access opportunities in the states where the PU message is unknown.

The idea of exploiting PU retransmissions to perform IC on future packets (similar to our FIC mechanism) was put forth by [14], which devises several cognitive radio protocols exploiting the hybrid ARQ retransmissions of the PU. Therein, the PU employs hybrid ARQ with incremental redundancy and the ARQ mechanism is limited to at most one retransmission. The SU receiver attempts to decode the PU message in the first time-slot. If successful, the SU transmitter sends its packet and the SU receiver decodes it by using IC on the received signal. In contrast, in this work, we address the more general case of an arbitrary number of primary ARQ retransmissions, and we allow a more general access pattern for the SU pair over the entire primary ARQ window. We also model the interplay between the primary ARQ protocol and the activity of the SU, by allowing for BIC. It should be noted that IC-related schemes are also used in other context, *e.g.*, decoding for graphical codes [15] and multiple access protocols [16].

Other related works include [17], which devises an oppor-

N. Michelusi and M. Zorzi are with the Department of Information Engineering, University of Padova, Italy ({michelusi,zorzi}@dei.unipd.it); P. Popovski is with the Department of Electronic Systems, Aalborg University, Denmark (petarp@es.aau.dk); O. Simeone is with the Center for Wireless Communication and Signal Processing Research (CWCSPR), New Jersey Institute of Technology, New Jersey, USA (osvaldo.simeone@njit.edu); M. Levorato is with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, USA, and with Stanford University, USA (levorato@stanford.edu).

This work is a generalization of [1], presented at the *Information Theory and Applications Workshop* in February 2011, and of [2], presented at the *49th Annual Allerton Conference on Communication, Control, and Computing* in September 2011.

Manuscript received date: Apr. 15, 2012.
Manuscript revised date: Sep. 1, 2012.

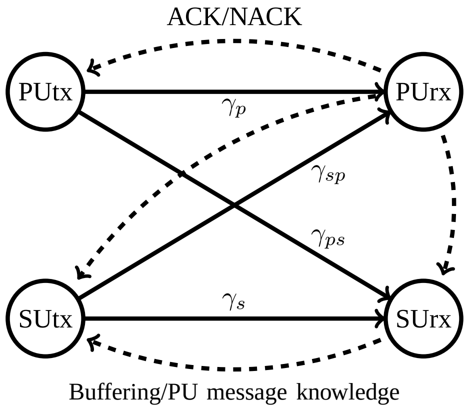


Fig. 1. System model

tunistic sharing scheme with channel probing based on the ARQ feedback from the PU receiver. An information theoretic framework for cognitive radio is investigated in [18], where the SU transmitter has non-casual knowledge of the PU's codeword. In [19], the data transmitted by the PU is obtained causally at the SU receiver. However, this model requires a joint design of the PU and SU signaling and channel state information at the transmitters. In contrast, in our work we explicitly model the dynamic acquisition of the PU message at the SU receiver, which enables IC. Moreover, the PU is oblivious to the presence of the SU.

The paper is organized as follows. Sec. II presents the system model. Sec. III introduces the secondary access policy, the performance metrics and the optimization problem, which is addressed in Sec. IV. Sec. VI presents and discusses the numerical results. Finally, Sec. VII concludes the paper. The proofs of the lemmas and theorems are provided in the appendix.

II. System Model

We consider a two-user interference network, as depicted in Fig. 1, where a primary transmitter and a secondary transmitter, denoted by PUtx and SUtx, respectively, transmit to their respective receivers, PUrx and SUrx, over the direct links PUtx→PUrx and SUtx→SUrx. Their transmissions generate mutual interference over the links PUtx→SUrx and SUtx→PUrx.

Time is divided into time-slots of fixed duration. Each time-slot matches the length of the PU and SU packets, and the transmissions of the PU and SU are assumed to be perfectly synchronized. We adopt the block-fading channel model, *i.e.*, the channel gains are constant within the time-slot duration, and change from time-slot to time-slot. Assuming that the SU and the PU transmit with constant power P_s and P_p, respectively, and that noise at the receivers is zero mean Gaussian with variance σ_w^2, we define the instantaneous Signal to Noise Ratios (SNR) of the links SUtx→SUrx, PUtx→PUrx, SUtx→PUrx and PUtx→SUrx, during the nth time-slot, as $\gamma_s(n)$, $\gamma_p(n)$, $\gamma_{sp}(n)$ and $\gamma_{ps}(n)$, respectively. We model the SNR process $\{\gamma_x(n), n = 0, 1, \dots\}$, where $x \in \{s, p, sp, ps\}$, as i.i.d. over time-slots and independent over the different links, and we denote the average SNR as $\bar{\gamma}_x = \mathbb{E}[\gamma_x]$.

We assume that no Channel State Information (CSI) is available at the transmitters, so that the latter cannot allocate their rate based on the instantaneous link quality, to ensure correct delivery of the packets to their respective receivers. Transmissions may thus undergo outage, when the selected rate is not supported by the current channel quality.

In order to improve reliability, the PU employs Type-I HARQ [10] with deadline $D \geq 1$, *i.e.*, at most D transmissions of the same PU message can be performed, after which the packet is discarded and a new transmission is performed (the PU is assumed to be backlogged). We define the *primary ARQ state* $t \in \mathbb{N}(1, D)$[1] as the number of ARQ transmission attempts already performed on the current PU message, plus the current one. Namely, $t = 1$ indicates a new PU transmission, and the counter t is increased at each ARQ retransmission, until the deadline D is reached. We assume that the ARQ feedback is received at the PU transmitter by the end of the time-slot, so that, if requested, a retransmission can be performed in the next time-slot.

On the other hand, the SU, in each time-slot, either accesses the channel by transmitting its own message, or stays idle. This decision is based on the access policy μ, defined in Sec. III. The activity of the SU, which is governed by μ, affects the outage performance of the PU, by creating interference to the PU over the link SUtx→PUrx. We denote the primary outage probability when the SU is idle and accesses the channel, respectively, as[2]

$$q_{pp}^{(\mathrm{I})}(R_p) \triangleq \Pr\left(R_p > C\left(\gamma_p\right)\right),$$
$$q_{pp}^{(\mathrm{A})}(R_p) \triangleq \Pr\left(R_p > C\left(\frac{\gamma_p}{1+\gamma_{sp}}\right)\right), \tag{1}$$

where R_p denotes the PU transmission rate, measured in bits/s/Hz, $C(x) \triangleq \log_2(1+x)$ is the (normalized) capacity of the Gaussian channel with SNR x at the receiver [20]. This outage definition, as well as the ones introduced later on, assume the use of Gaussian signaling and capacity-achieving coding with sufficiently long codewords. However, our analysis can be extended to include practical codes by computing the outage probabilities for the specific code considered. In (1), it is assumed that SU transmissions are treated as background Gaussian noise by the PU. This is a reasonable assumption in CRs in which the PU is oblivious to the presence of SUs. In general, we have $q_{pp}^{(\mathrm{A})}(R_p) \geq q_{pp}^{(\mathrm{I})}(R_p)$, where equality holds if and only if $\gamma_{sp} \equiv 0$ deterministically. We denote the expected PU throughput accrued in each time-slot, when the SU is idle and accesses the channel, as $T_p^{(\mathrm{I})}(R_p) = R_p[1 - q_{pp}^{(\mathrm{I})}(R_p)]$ and $T_p^{(\mathrm{A})}(R_p) = R_p[1 - q_{pp}^{(\mathrm{A})}(R_p)]$, respectively.

[1] We define $\mathbb{N}(n_0, n_1) = \{t \in \mathbb{N}, n_0 \leq t \leq n_1\}$ for $n_0 \leq n_1 \in \mathbb{N}$

[2] Herein, we denote the outage probability as $q_{xy}^{(Z)}$, where x and y are the source and the recipient of the message, respectively (PU if $x, y = p$, SU if $x, y = s$), and $Z \in \{\mathrm{A}, \mathrm{I}\}$ denotes the action of the SU (A if the SU is active and it accesses the channel, I if the SU remains idle). For example, $q_{ps}^{(\mathrm{A})}$ is the probability that the PU message is in outage at SUrx, when SUtx transmits.

A. Operation of the SU

Unlike the PU that uses a simple Type-I Hybrid ARQ mechanism, it is assumed that the SU uses "best effort" transmission. Moreover, the SU is provided with side-information about the PU, *e.g.*, ARQ deadline D, PU codebook and feedback information from PUrx (ACK/NACK messages). This is consistent with the common characterization of the PU as a legacy system, and of the SU as an opportunistic and cognitive system, which exploits the primary ARQ feedback to create a best-effort link with maximized throughput, while the flow control mechanisms are left to the upper layers. By overhearing the feedback information from PUrx, the SU can thus track the primary ARQ state t. Moreover, by leveraging the PU codebook, SUrx attempts, in any time-slot, to decode the PU message, which enables the following IC techniques at SUrx:

- *Forward IC* (FIC): by decoding the PU message, SUrx can perform IC in the current as well as in the following ARQ retransmissions, if these occur, to achieve a larger SU throughput;
- *Backward IC* (BIC): SUrx buffers the received signals corresponding to SU transmissions which undergo outage due to severe interference from the PU. These transmissions can later be recovered using IC on the buffered received signals, if the interfering PU message is successfully decoded by SUrx in a subsequent primary ARQ retransmission attempt.

We define the *SU buffer state* $b \in \mathbb{N}(0, B)$ as the number of received signals currently buffered at SUrx, where $B \in \mathbb{N}(0, D-1)$[3] denotes the buffer size. Moreover, we define the *PU message knowledge state* $\Phi \in \{\mathrm{K}, \mathrm{U}\}$, which denotes the knowledge at SUrx about the PU message currently handled by the PU. Namely, if $\Phi = \mathrm{K}$, then SUrx knows the PU message, thus enabling FIC/BIC; conversely ($\Phi = \mathrm{U}$), the PU message is unknown to SUrx.

Remark 1 (Feedback Information). Note that PUrx needs to report one feedback bit to inform PUtx (and the SU, which overhears the feedback) on the transmission outcome (ACK/NACK). On the other hand, two feedback bits need to be reported by SUrx to SUtx: one bit to inform SUtx as to whether the PU message has been successfully decoded, so that SUtx can track the PU message knowledge state Φ; and one bit to inform SUtx as to whether the received signal has been buffered, so that SUtx can track the SU buffer state b. Herein, we assume ideal (error-free) feedback channels, so that the SU can track (t, b, Φ), and the PU can track the ARQ state t. However, optimization is possible with imperfect observations as well [21]. □

We now further detail the operation of the SU for $\Phi \in \{\mathrm{K}, \mathrm{U}\}$.

1) PU message unknown to SUrx ($\Phi = \mathrm{U}$): When $\Phi = \mathrm{U}$ and the SU is idle, SUrx attempts to decode the PU message, so as to enable FIC/BIC. A decoding failure occurs if the rate of the PU message, R_p, exceeds the capacity of the channel PUtx→SUrx, with SNR γ_{ps}. We denote the corresponding outage probability as $q_{ps}^{(\mathrm{I})}(R_p) = \Pr(R_p > C(\gamma_{ps}))$.

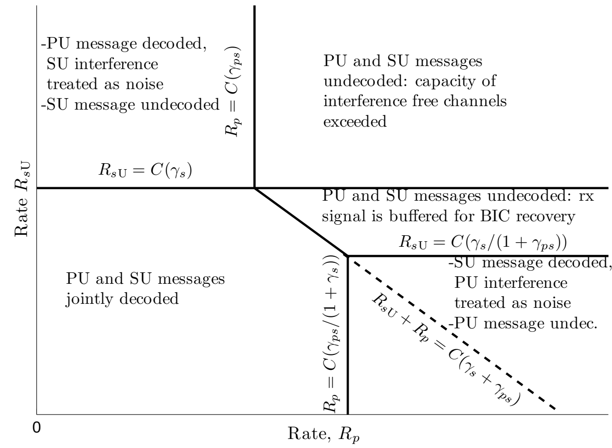


Fig. 2. Decodability regions for PU message (rate R_p) and SU message (rate R_{sU}) at SUrx, for a fixed SNR pair (γ_s, γ_{ps})

If the SU accesses the channel, SU transmissions are performed with rate R_{sU} (bits/s/Hz) and are interfered by the PU. SUrx thus attempts to decode both the SU and PU messages; moreover, if the decoding of the SU message fails due to severe interference from the PU, the received signal is buffered for future BIC recovery. Using standard information-theoretic results [20], with the help of Fig. 2, we define the following SNR regions associated with the decodability of the SU and PU messages at SUrx, where $\mathcal{A}^c$ denotes the complementary set of $\mathcal{A}$:[4]

$$\Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p) \triangleq \Big\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s), R_p \leq C(\gamma_{ps}), R_{s\mathrm{U}} + R_p \leq C(\gamma_s + \gamma_{ps}) \Big\} \quad (2)$$

$$\bigcup \left\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} > C(\gamma_s), R_p \leq C\left(\frac{\gamma_{ps}}{1+\gamma_s}\right) \right\} \quad (3)$$

$$\Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \triangleq \Big\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s), R_p \leq C(\gamma_{ps}), R_{s\mathrm{U}} + R_p \leq C(\gamma_s + \gamma_{ps}) \Big\} \quad (4)$$

$$\bigcup \left\{ (\gamma_s, \gamma_{ps}) : R_p > C(\gamma_{ps}), R_{s\mathrm{U}} \leq C\left(\frac{\gamma_{ps}}{1+\gamma_s}\right) \right\} \quad (5)$$

$$\Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p) \triangleq \Big\{ \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p) \cup \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \Big\}^c \bigcap \Big\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s) \Big\}. \quad (6)$$

The SNR regions (2) and (4) guarantee that the two rates R_p and R_{sU} are within the multiple access channel region formed by the two transmitters (PUtx and SUtx) and SUrx [20], so that both the SU and PU messages are correctly decoded via joint decoding techniques. On the other hand, in the SNR region (5) (respectively, (3)), only the SU (PU) message is successfully

[3] Note that $B \leq D-1$, since the same PU message is transmitted at most D times by PUtx. Once the ARQ deadline D is reached, a new PU transmission occurs, and the buffer is emptied.

[4] Herein, we assume optimal joint decoding techniques of the SU and PU messages. Using other techniques, *e.g.*, successive IC, the SNR regions may change accordingly, without providing any further insights in the following analysis.

decoded at SUrx by treating the interference from the PU (SU) as background noise. If the SNR pair falls outside the two regions (4) and (5) (respectively, (2) and (3)), then SUrx incurs a failure in decoding the SU (PU) message. Therefore, when $(\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)$, SUrx successfully decodes the SU message. The corresponding expected SU throughput is thus given by

$$T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p) \triangleq R_{s\mathrm{U}} \mathrm{Pr}\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)\right). \quad (7)$$

Similarly, when $(\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$, SUrx successfully decodes the PU message. We denote the corresponding outage probability as $q_{ps}^{(\mathrm{A})}(R_{s\mathrm{U}}, R_p) \triangleq \mathrm{Pr}\left((\gamma_s, \gamma_{ps}) \notin \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)\right)$. Note that $q_{ps}^{(\mathrm{A})}(R_{s\mathrm{U}}, R_p) > q_{ps}^{(\mathrm{I})}(R_p)$, since SU transmissions interfere with the decoding of the PU message.

Finally, in (6), the decoding of both the SU and PU messages fails, since the SNR pair (γ_s, γ_{ps}) falls outside both regions $\Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$ and $\Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)$. However, the rate $R_{s\mathrm{U}}$ is within the capacity region of the interference free channel ($R_{s\mathrm{U}} \leq C(\gamma_s)$), so that the SU message can be recovered via BIC, should the PU message become available in a future ARQ retransmission attempt. The received signal is thus buffered at SUrx. We denote the *buffering probability* as

$$\begin{aligned} p_{s,\mathrm{buf}}(R_{s\mathrm{U}}, R_p) &\triangleq \mathrm{Pr}\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p)\right) \\ &= \mathrm{Pr}\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, 0)\right) \\ &\quad - \mathrm{Pr}\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)\right) > 0, \end{aligned} \quad (8)$$

where the second equality follows from inspection of Fig. 2.

2) PU message known to SUrx ($\Phi = \mathrm{K}$): When $\Phi = \mathrm{K}$, SUrx performs FIC on the received signal, thus enabling interference free SU transmissions. The SU transmits with rate $R_{s\mathrm{K}}$, and the accrued throughput is given by $T_{s\mathrm{K}}(R_{s\mathrm{K}}) = R_{s\mathrm{K}} \mathrm{Pr}\left(R_{s\mathrm{K}} < C(\gamma_s)\right)$.

We now provide an example to illustrate the use of FIC/BIC at SUrx.

Example 1. Consider a sequence of 3 primary retransmission attempts in which the SU always accesses the channel. Initially, the PU message is unknown to SUrx, hence the PU message knowledge state is set to $\Phi = \mathrm{U}$ in the first time-slot, and the SU transmits with rate $R_{s\mathrm{U}}$. Assume that the SNR pair $(\gamma_s(1), \gamma_{ps}(1))$ falls in $\Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p)$. Then, neither the SU nor the PU messages are successfully decoded by SUrx, but the received signal is buffered for future BIC recovery. In the second time-slot, $(\gamma_s(2), \gamma_{ps}(2)) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \cap \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$, hence both the SU and PU messages are correctly decoded by SUrx, and the PU message knowledge state switches to $\Phi = \mathrm{K}$. At this point, SUrx performs BIC on the previously buffered received signal to recover the corresponding SU message. In the third time-slot, SUtx transmits with rate $R_{s\mathrm{K}}$, and decoding at SUrx takes place after cancellation of the interference from the PU via FIC. □

We now briefly elaborate on the choice of the transmission rate $R_{s\mathrm{K}}$. Since its value does not affect the outage performance at PUrx (1) and the evolution of the ARQ process, $R_{s\mathrm{K}}$ is chosen so as to maximize $T_{s\mathrm{K}}(R_{s\mathrm{K}})$. Therefore, from (8) we obtain

$$\begin{aligned} T_{s\mathrm{K}}(R_{s\mathrm{K}}) \geq & T_{s\mathrm{K}}(R_{s\mathrm{U}}) = T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p) \\ & + p_{s,\mathrm{buf}}(R_{s\mathrm{U}}, R_p) R_{s\mathrm{U}} > T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p). \end{aligned} \quad (9)$$

Conversely, the choice of the rate $R_{s\mathrm{U}}$ is not as straightforward, since its value reflects a trade-off between the potentially larger throughput accrued with a larger rate $R_{s\mathrm{U}}$ and the corresponding diminished capabilities for IC caused by the more difficult decoding of the PU message by SUrx.

In the following treatment, the rates $R_{s\mathrm{K}}$, $R_{s\mathrm{U}}$ and R_p are assumed to be fixed parameters of the system, and they are not considered part of the optimization (see Sec. VI for further elaboration in this regard). For the sake of notational convenience, we omit the dependence of the quantities defined above on them. Moreover, for clarity, we consider the case $B = D - 1$ in which SUrx can buffer up to $D - 1$ received signals. However, the following analysis can be extended to a generic value of B.

III. Policy Definition and Optimization Problem

We model the evolution of the network as a Markov Decision Process [11], [12]. Namely, we denote the state of the PU-SU system by the tuple (t, b, Φ), where $t \in \mathbb{N}(1, D)$ is the primary ARQ state, $b \in \mathbb{N}(0, B)$ is the SU buffer state and $\Phi \in \{\mathrm{U}, \mathrm{K}\}$ is the PU message knowledge state. (t, b, Φ) takes values in the state space $\mathcal{S} \equiv \mathcal{S}_{\mathrm{U}} \cup \mathcal{S}_{\mathrm{K}}$, where $\mathcal{S}_{\mathrm{K}} \equiv \{(t, 0, \mathrm{K}) : t \in \mathbb{N}(2, D)\}$ and $\mathcal{S}_{\mathrm{U}} \equiv \{(t, b, \mathrm{U}) : t \in \mathbb{N}(1, D), b \in \mathbb{N}(0, t-1)\}$ are the sets of states where the PU message is known and unknown to SUrx, respectively.

The SU follows a *stationary randomized access policy* $\mu \in \mathcal{U} \equiv \{\mu : \mathcal{S} \mapsto [0, 1]\}$, which determines the secondary access probability for each state $\mathbf{s} \in \mathcal{S}$. Note that, from [22], this choice is without loss of optimality for the specific problem at hand. Namely, in state $(t, b, \Phi) \in \mathcal{S}$, the SU is "active", *i.e.*, it accesses the channel, with probability $\mu(t, b, \Phi)$ and stays "idle" with probability $1 - \mu(t, b, \Phi)$. We denote the "active" and "idle" actions as A and I, respectively.

With these definitions at hand, we define the following average long-term metrics under μ: the SU throughput $\bar{T}_s(\mu)$, the SU power expenditure $\bar{P}_s(\mu)$ and the PU throughput $\bar{T}_p(\mu)$, given by

$$\begin{aligned} \bar{T}_s(\mu) = & \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} R_{s\Phi_n} \mathbf{1}\left(\{Q_n = \mathrm{A}\} \cap O_{s,n}^c\right) \middle| \mathbf{s}_0\right] \\ & + \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} R_{s\mathrm{U}} B_n \mathbf{1}(O_{ps,n}^c) \middle| \mathbf{s}_0\right], \end{aligned} \quad (10)$$

$$\bar{P}_s(\mu) = P_s \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} \mathbf{1}\left(\{Q_n = \mathrm{A}\}\right) \middle| \mathbf{s}_0\right], \quad (11)$$

$$\bar{T}_p(\mu) = \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} R_p \mathbf{1}\left(O_{p,n}^c\right) \middle| \mathbf{s}_0\right], \quad (12)$$

where n is the time-slot index, $\mathbf{s}_0 \in \mathcal{S}$ is the initial state in time-slot 0; $\Phi_n \in \{\mathrm{K}, \mathrm{U}\}$ is the PU message knowledge state and B_n is the SU buffer state in time-slot n; $Q_n \in \{\mathrm{A}, \mathrm{I}\}$ is the action of the SU, drawn according to the access policy

μ; $O_{s,n}$ and $O_{ps,n}$ denote the outage events at SUrx for the decoding of the SU and PU messages, so that $O_{s,n}^c$ and $O_{ps,n}^c$ denote successful decoding of the SU and PU messages by SUrx, respectively; $O_{p,n}$ denotes the outage event at PUrx, so that $O_{p,n}^c$ denotes successful decoding of the PU message by PUrx; and $\mathbf{1}(E)$ is the indicator function of the event E. Note that all the quantities defined above are independent of the initial state $\mathbf{s}_0$. In fact, starting from any $\mathbf{s}_0 \in \mathcal{S}$, the system reaches with probability 1 the positive recurrent state $(1, 0, \mathrm{U})$ (new PU transmission) within a finite number of time-slots, due to the ARQ deadline. Due to the Markov property, from this state on, the evolution of the process is independent of the initial transient behavior, which has no effect on the time averages defined in (10), (11) and (12).

In this work, we study the problem of maximizing the average long-term SU throughput subject to constraints on the average long-term PU throughput loss and SU power. Specifically,

$$\mu^* = \arg\max_{\mu} \bar{T}_s(\mu) \text{ s.t. } \bar{T}_p(\mu) \geq T_p^{(\mathrm{I})}(1-\epsilon_{\mathrm{PU}}), \quad \bar{P}_s(\mu) \leq \mathcal{P}_s^{(\mathrm{th})}, \tag{13}$$

where $\epsilon_{\mathrm{PU}} \in [0,1]$ and $\mathcal{P}_s^{(\mathrm{th})} \in [0, P_s]$ represent the (normalized) maximum tolerated PU throughput loss with respect to the case in which the SU is idle and the SU power constraint, respectively. This problem entails a trade-off in the operation of the SU. On the one hand, the SU is incentivized to transmit in order to increase its throughput and to optimize the buffer occupancy at SUrx (*i.e.*, failed SU transmissions which are potentially recovered via BIC). On the other hand, SU transmissions might jeopardize the correct decoding of the PU message at SUrx, thus impairing the use of FIC/BIC, and might violate the constraints in (13).

Under $\mu \in \mathcal{U}$, the state process is a stationary Markov chain, with steady state distribution π_μ [12], [23]. $\pi_\mu(\mathbf{s}), \mathbf{s} \in \mathcal{S}$, is the long-term fraction of the time-slots spent in state $\mathbf{s}$, *i.e.*, $\pi_\mu(\mathbf{s}) = \lim_{N\to+\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{Pr}_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$, where $\mathrm{Pr}_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$ is the n-step transition probability of the chain from state $\mathbf{s}_0$.[5] In state (t, b, U), the SU accesses the channel with probability $\mu(t, b, \mathrm{U})$, thus accruing the throughput $\mu(t, b, \mathrm{U})\, T_{s\mathrm{U}}$. Moreover, if SUrx successfully decodes the PU message (with probability $1 - q_{ps}^{(\mathrm{I})} - \mu(t, b, \mathrm{U})(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})})$), $bR_{s\mathrm{U}}$ bits are recovered by performing BIC on the buffered received signals, yielding an additional BIC throughput. Similarly, in state $(t, 0, \mathrm{K})$, the SU accrues the throughput $\mu(t, 0, \mathrm{K})\, T_{s\mathrm{K}}$. Then, we can rewrite (10) and (11) in terms of the steady state distribution and of the cost/reward in each state as

$$\bar{T}_s(\mu) = T_{s\mathrm{U}}\bar{W}_s(\mu) + \bar{F}_s(\mu) + \bar{B}_s(\mu), \;\; \bar{P}_s(\mu) = P_s \bar{W}_s(\mu), \tag{14}$$

where the *SU access rate* $\bar{W}_s(\mu)$, *i.e.*, the average long-term number of secondary channel accesses per time-slot, the *FIC throughput* $\bar{F}_s(\mu)$ and the *BIC throughput* $\bar{B}_s(\mu)$ are defined as

$$\begin{aligned}
\bar{W}_s(\mu) &\triangleq \textstyle\sum_{\mathbf{s}\in\mathcal{S}} \pi_\mu(\mathbf{s})\, \mu(\mathbf{s}), \\
\bar{F}_s(\mu) &\triangleq \textstyle\sum_{t=2}^{D} \pi_\mu(t, 0, \mathrm{K})\, \mu(t, 0, \mathrm{K})\, (T_{s\mathrm{K}} - T_{s\mathrm{U}}), \\
\bar{B}_s(\mu) &\triangleq \textstyle\sum_{t=1}^{D}\sum_{b=0}^{t-1} \pi_\mu(t, b, \mathrm{U})\, bR_{s\mathrm{U}} \\
&\quad \times \left[1 - q_{ps}^{(\mathrm{I})} - \mu(t, b, \mathrm{U})\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right)\right].
\end{aligned} \tag{15}$$

In (14), $T_{s\mathrm{U}}\bar{W}_s(\mu)$ is the SU throughput attained without FIC/BIC, while the terms $\bar{F}_s(\mu)$ and $\bar{B}_s(\mu)$ account for the throughput gains of FIC and BIC, respectively. Conversely, the PU accrues the throughput $T_p^{(\mathrm{I})}$ if the SU is idle and $T_p^{(\mathrm{A})}$ if the SU accesses the channel, so that (12) is given by

$$\bar{T}_p(\mu) = T_p^{(\mathrm{I})} - (T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\bar{W}_s(\mu). \tag{16}$$

The quantity $(T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\bar{W}_s(\mu)$ is referred to as the *PU throughput loss* induced by the secondary access policy μ [13]. The following result follows directly from (13), (14) and (16).

Lemma 1. The problem (13) is equivalent to

$$\mu^* = \arg\max_{\mu\in\mathcal{U}} \bar{T}_s(\mu) \tag{17}$$
$$\text{s.t. } \bar{W}_s(\mu) \leq \min\left\{\frac{(1-q_{pp}^{(\mathrm{I})})\epsilon_{\mathrm{PU}}}{q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})}}, \frac{\mathcal{P}_s^{(\mathrm{th})}}{P_s}\right\} \triangleq \epsilon_{\mathrm{W}}.$$

□

In the next section, we characterize the solution of (17). We will need the following definition.

Definition 1. Let μ be the policy such that secondary access takes place if and only if the PU message is known to SUrx, *i.e.*, $\mu(\mathbf{s}) = 1, \;\forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, $\mu(\mathbf{s}) = 0, \;\forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$. We denote the SU access rate achieved by such policy as $\epsilon_{\mathrm{th}} = \bar{W}(\mu)$. The system is in the *low SU access rate regime* if $\epsilon_{\mathrm{W}} \leq \epsilon_{\mathrm{th}}$ in (17). Otherwise, the system is in the *high SU access rate regime*. □

IV. Optimal Policy

In this section, we characterize in closed form the optimal policy in the low SU access rate regime, and we present an algorithm to derive the optimal policy in the high SU access rate regime.

A. Low SU Access Rate Regime

The next lemma shows that, in the low SU access rate regime, an optimal policy prescribes that secondary access only takes place in the states where the PU message is known to SUrx, with an equal probability in all such states. It follows that only FIC, and not BIC, is needed in this regime to attain optimal performance.

Lemma 2. In the low SU access rate regime $\epsilon_{\mathrm{W}} \leq \epsilon_{\mathrm{th}}$, an optimal policy is given by[6]

$$\mu^*(\mathbf{s}) = \frac{\epsilon_{\mathrm{W}}}{\epsilon_{\mathrm{th}}}, \;\forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}, \;\; \mu^*(\mathbf{s}) = 0, \;\forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}. \tag{18}$$

[5] Similarly to (10), (11) and (12), $\pi_\mu(\mathbf{s})$ is independent of the initial state $\mathbf{s}_0$, due to the recurrence of state $(1, 0, \mathrm{U})$.

[6] The optimal policy in the low SU access rate is not unique. In fact, any policy μ such that $\mu(\mathbf{s}) = 0, \;\forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$ and $\bar{W}_s(\mu) = \epsilon_{\mathrm{th}}$ is optimal, attaining the same throughput $\bar{T}_s(\mu) = T_{s\mathrm{K}}\epsilon_{\mathrm{th}}$ as (18).

Moreover, $\bar{T}_s(\mu^*) = T_{s\mathrm{K}}\epsilon_{\mathrm{W}}$, $\bar{P}_s(\mu^*) = P_s\epsilon_{\mathrm{W}}$, and $\bar{T}_p(\mu^*) = T_p^{(\mathrm{I})} - (T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\epsilon_{\mathrm{W}}$. □

Proof: For any policy $\mu \in \mathcal{U}$ obeying the SU access rate constraint $\bar{W}_s(\mu) \leq \epsilon_{\mathrm{W}}$, we have $\bar{T}_s(\mu) \leq \bar{W}_s(\mu)T_{s\mathrm{K}} \leq \epsilon_{\mathrm{W}}T_{s\mathrm{K}}$. The first inequality holds since $\bar{W}_s(\mu)T_{s\mathrm{K}}$ is the long-term throughput achievable when the PU message is known *a priori* at SUrx, which is an upper bound to the performance; the second from the SU access rate constraint. The upper bound $\epsilon_{\mathrm{W}}T_{s\mathrm{K}}$ is achieved by policy (18), as can be directly seen by substituting (18) in (14), (15). ■

Remark 2. Note that secondary accesses in states $\mathcal{S}_{\mathrm{U}}$, where the PU message is unknown to SUrx, would obtain a smaller throughput, namely at most $T_{s\mathrm{U}} + p_{s,\mathrm{buf}}R_{s\mathrm{U}} \leq T_{s\mathrm{K}}$, where $T_{s\mathrm{U}}$ is the "instantaneous" throughput and $p_{s,\mathrm{buf}}R_{s\mathrm{U}}$ is the BIC throughput, *possibly* recovered via BIC in a future ARQ retransmission. Therefore, SU accesses in states $\mathcal{S}_{\mathrm{K}}$ are more "cost effective". □

B. High SU Access Rate Regime

In this section, we study the high SU access rate regime in which $\epsilon_{\mathrm{W}} > \epsilon_{\mathrm{th}}$, thus complementing the analysis above for the regime where $\epsilon_{\mathrm{W}} \leq \epsilon_{\mathrm{th}}$. It will be seen that, if $\epsilon_{\mathrm{W}} > \epsilon_{\mathrm{th}}$, unlike in the low SU access rate regime, the SU should generally access the channel also in states $\mathcal{S}_{\mathrm{U}}$ where the PU message is unknown to SUrx in order to achieve the optimal performance. Therefore, both BIC and FIC are necessary to attain optimality. In this section, we derive the optimal policy. We first introduce some necessary definitions and notations.

Definition 2 (Secondary access efficiency)**.** We define the *secondary access efficiency* under policy $\mu \in \mathcal{U}$ in state $\mathbf{s} \in \mathcal{S}$ as

$$\eta_\mu(\mathbf{s}) = \frac{\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}}{\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}}. \tag{19}$$

□

The secondary access efficiency can be interpreted as follows. If the secondary access probability is increased in state $\mathbf{s} \in \mathcal{S}$ by a small amount δ, then the PU throughput loss is increased by an amount equal to $\delta(T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (from (16)), the SU power is increased by an amount equal to $\delta P_s\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (from (14)), and the SU throughput augments or diminishes by an amount equal to $\delta\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (depending on the sign of the derivative). Therefore, $\eta_\mu(\mathbf{s})$ yields the rate of increase (or decrease if $\eta_\mu(\mathbf{s}) < 0$) of the SU throughput per unit increase of the SU access rate, as induced by augmenting the secondary channel access probability in state $\mathbf{s}$. Equivalently, it measures how efficiently the SU can access the channel in state $\mathbf{s}$, in terms of maximizing the SU throughput gain while minimizing its negative impact on the PU throughput and on the SU power expenditure.

Remark 3. It is worth noting that the definition of $\eta_\mu(\mathbf{s})$ given in Def. 2 is not completely rigorous. In fact, under a generic policy μ, the Markov chain of the PU-SU system may not be irreducible [23], so that state $\mathbf{s}$ may not be accessible, hence $\pi_\mu(\mathbf{s}) = 0$ and $\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = \frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = 0$. One example is the idle policy $\mu(\mathbf{s}) = 0,\ \forall\mathbf{s}$: since the SU is always idle, the buffer at SUrx is always empty, hence states (t, b, U) with $b > 0$ are never accessed. To overcome this problem, a formal definition is given in App. B, by treating the Markov chain of the PU-SU system as the limit of an irreducible Markov chain. $\eta_\mu(\mathbf{s})$ is explicitly derived in Lemma 6 in App. B. □

We denote the indicator function of state $\mathbf{s}$ as $\delta_{\mathbf{s}} : \mathcal{S} \mapsto \{0, 1\}$, with $\delta_{\mathbf{s}}(\mathbf{s}) = 1$, $\delta_{\mathbf{s}}(\sigma) = 0,\ \forall\sigma \neq \mathbf{s}$. Moreover, we denote the policy at the ith iteration of the algorithm as $\mu^{(i)}$. We are now ready to describe the algorithm that obtains an optimal policy in the high SU access rate regime. An intuitive explanation of the algorithm can be found below.

Algorithm 1 (Derivation of the optimal policy)**.**

1) **Initialization**:
   - Let $\mu^{(0)}$ be the policy $\mu^{(0)}(\mathbf{s}) = 0,\ \forall\ \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$, $\mu^{(0)}(\mathbf{s}) = 1,\ \forall\ \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, and $i = 0$.
   - Let $\mathcal{S}_{\mathrm{idle}}^{(0)} \equiv \{\mathbf{s} \in \mathcal{S} : \mu^{(0)}(\mathbf{s}) = 0\} \equiv \mathcal{S}_{\mathrm{U}}$ be the set of states where the SU is idle.
2) **Stage** i:
   a) Compute $\eta_{\mu^{(i)}}(\mathbf{s}),\ \forall\ \mathbf{s} \in \mathcal{S}_{\mathrm{idle}}^{(i)}$ and let $\mathbf{s}^{(i)} \triangleq \arg\max_{\mathbf{s}\in\mathcal{S}_{idle}^{(i)}} \eta_{\mu^{(i)}}(\mathbf{s})$.
   b) If $\eta_{\mu^{(i)}}(\mathbf{s}^{(i)}) \leq 0$, go to step 3). Otherwise, let $\mu^{(i+1)} = \mu^{(i)} + \delta_{\mathbf{s}^{(i)}}$, $\mathcal{S}_{\mathrm{idle}}^{(i+1)} \equiv \mathcal{S}_{\mathrm{idle}}^{(i)} \setminus \{\mathbf{s}^{(i)}\}$.
   c) Set $i := i+1$. If $\mathcal{S}_{\mathrm{idle}}^{(i)} \equiv \emptyset$, go to step 3). Otherwise, repeat from step 2).
3) Let $N = i$, the sequence of states $(\mathbf{s}^{(0)}, \ldots, \mathbf{s}^{(N-1)})$ and of policies $(\mu^{(0)}, \ldots, \mu^{(N-1)})$.
4) **Optimal policy**: given ϵ_{W},
   a) If $\bar{W}_s(\mu^{(N-1)}) \leq \epsilon_{\mathrm{W}}$, then $\mu^* = \mu^{(N-1)}$.
   b) Otherwise, $\mu^* = \lambda\mu^{(j)} + (1-\lambda)\mu^{(j+1)}$, where $j \triangleq \max\left\{i : \bar{W}_s\left(\mu^{(i)}\right) \leq \epsilon_{\mathrm{W}}\right\}$ and $\lambda \in (0, 1]$ uniquely solves $\bar{W}_s(\lambda\mu^{(j)} + (1-\lambda)\mu^{(j+1)}) = \epsilon_{\mathrm{W}}$. □

The algorithm, starting from the optimal policy for the case $\epsilon_{\mathrm{W}} = \epsilon_{\mathrm{th}}$ (Lemma 2), ranks the states in the set $\mathcal{S}_{\mathrm{U}}$ in decreasing order of secondary access efficiency, and iteratively allocates the secondary access to the state with the highest efficiency, among the states where the SU is idle. The rationale of this step is that secondary access in the most efficient state yields the steepest increase of the SU throughput, per unit increase of the SU access rate or, equivalently, of the PU throughput loss and of the SU power expenditure. The optimality of Algorithm 1 is established in the following theorem.

Theorem 1. *Algorithm 1 returns an optimal policy for the optimization problem (17).* □

Proof: See App. C. ■

V. Special Case: Degenerate Cognitive Radio Network Scenario

We point out that Algorithm 1 determines the optimal policy for a generic set of system parameters. However, the resulting

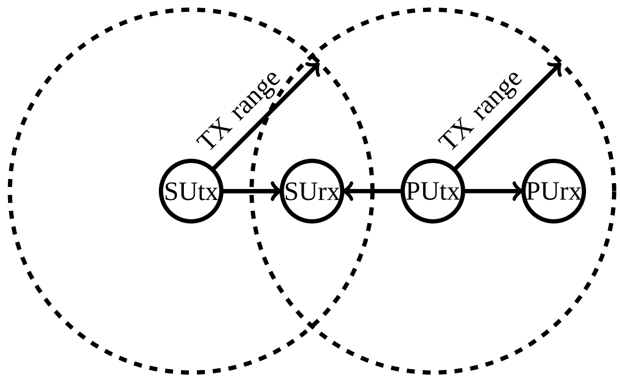


Fig. 3. Degenerate cognitive radio network

optimal policy does not always have a structure that is easily interpreted. In this section, we consider a special case of the general model discussed so far, a *degenerate cognitive radio network*, where the activity of the PU is unaffected by the transmissions of the SU, *i.e.*, the channel gain between the SU transmitter and the PU receiver is zero.

Consider the scenario depicted in Fig. 3, where PUrx is outside the transmission range of SUtx, whereas SUrx is inside the transmission range of both SUtx and SUrx. In this scenario, the interference produced by SU to PU is negligible. In contrast, the PU produces significant interference at the SU receiver. The SU thus potentially benefits by employing the BIC and FIC mechanisms. We denote this scenario as a *Degenerate cognitive radio network*, and we model it by assuming that the SNR of the interfering link SUtx→PUrx is deterministically equal to zero, *i.e.*, $\gamma_{sp} = 0$. From (1), we then have $q_{pp}^{(\mathrm{I})} = q_{pp}^{(\mathrm{A})} \triangleq q_{pp}$, *i.e.*, the outage performance of the PU is unaffected by the activity of the SU, and the primary ARQ process is independent of the secondary access policy. We define

$$\Delta_s \triangleq \frac{T_{s\mathrm{K}} - T_{s\mathrm{U}} - p_{s,\mathrm{buf}} R_{s\mathrm{U}}}{R_{s\mathrm{U}}}. \tag{20}$$

From (9), it follows that $\Delta_s \geq 0$, with equality if $R_{s\mathrm{U}} = R_{s\mathrm{K}}$. Therefore, $R_{s\mathrm{U}}\Delta_s$ is the marginal throughput gain accrued in the states where the PU message is known to SUrx, over the throughput accrued in the states where the PU message is unknown (instantaneous throughput $T_{s\mathrm{U}}$ plus BIC throughput $p_{s,\mathrm{buf}} R_{s\mathrm{U}}$, possibly recovered in a future ARQ retransmission). The following lemma proves that, if the marginal throughput gain Δ_s is "small", the secondary accesses in the high SU access rate regime in a degenerate cognitive radio network are allocated, in order, to the states in $\mathcal{S}_{\mathrm{K}}$ (Lemma 2), then to the idle states (t, b, U) in $\mathcal{S}_{\mathrm{U}}$, giving priority to states with low b and t over states with high b and t, respectively. An illustrative example of the optimal policy for this scenario is given in Fig. 4.

Lemma 3. In the degenerate cognitive radio network scenario with $q_{pp}^{(\mathrm{A})} = q_{pp}^{(\mathrm{I})} = q_{pp}$, if

$$\Delta_s < \frac{1 - q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}}, \tag{21}$$

the sequence of policies $(\mu^{(0)}, \dots, \mu^{(N-1)})$ returned by Algorithm 1 is such that, $\forall i \in \mathbb{N}(0, N-1)$,

$$\mu^{(i)}(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}, \tag{22}$$

$$\mu^{(i)}(t, b, \mathrm{U}) = \begin{cases} 1 & b < b^{(i)}(t) \\ 0 & b \geq b^{(i)}(t), \end{cases},\ \forall (t, b, \mathrm{U}) \in \mathcal{S}_{\mathrm{U}}, \tag{23}$$

where $b^{(i)}(t)$ is non-increasing in t and non-decreasing in i, with $b^{(0)}(t) = 0$ and $b^{(N-1)}(t) = \bar{b}_{\max}(t)$, *i.e.*,

$$\bar{b}_{\max}(t) = b^{(N-1)}(t) \geq \dots \geq b^{(i)}(t) \geq b^{(i-1)}(t) \geq \dots \geq b^{(0)}(t) = 0. \tag{24}$$

$$b^{(i)}(1) \geq b^{(i)}(2) \geq \dots \geq b^{(i)}(t-1) \geq b^{(i)}(t) \geq \dots \geq b^{(i)}(D), \tag{25}$$

where

$$\bar{b}_{\max}(t) = \left\lceil \frac{\begin{array}{c}\frac{T_{s\mathrm{U}}}{R_{s\mathrm{U}}}\left[1 - q_{pp}\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) A_0(t+1)\right] \\ + \left(\frac{1-q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}} - \Delta_s\right) \times \\ \times q_{pp}\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) A_0(t+1)\end{array}}{\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right)\left(1 - q_{pp}(1 - q_{ps}^{(\mathrm{I})}) A_0(t+1)\right)} \right\rceil - 1 \tag{26}$$

and we have defined

$$A_0(\tau) \triangleq \frac{1 - q_{pp}^{D-\tau+1} q_{ps}^{(\mathrm{I})(D-\tau+1)}}{1 - q_{pp} q_{ps}^{(\mathrm{I})}}, \tag{27}$$

$$A_1(\tau) \triangleq \frac{1 - q_{pp}^{D-\tau+1}}{1 - q_{pp}}. \tag{28}$$

Proof: See App. D. ■

Remark 4. Interestingly, this is the same result derived in our work [2] for $D = 2$. However, therein the result was shown to hold for general $q_{pp}^{(\mathrm{A})} \geq q_{pp}^{(\mathrm{I})}$ (not necessarily a degenerate cognitive radio network), whereas Lemma 3 holds for general D but only for a degenerate cognitive radio network scenario.

The lemma dictates that, in the degenerate cognitive radio network scenario, the SU should restrict its channel accesses to the states corresponding to a low primary ARQ index and small buffer occupancy at the SU receiver. Alternatively, the larger the ARQ index or the buffer occupancy, the smaller the incentive to access the channel. By doing so, the SU maximizes the buffer occupancy in the early HARQ retransmission attempts, and invests in the future BIC recovery. When the primary ARQ state t approaches the deadline D, the SU is incentivized to idle so as to help SUrx to decode the PU message, thus enabling the recovery of the failed SU transmissions from the buffered received signals via BIC, before the ARQ deadline D is reached and the buffer is depleted. Moreover, when the buffer state b grows, since $q_{ps}^{(\mathrm{A})} > q_{ps}^{(\mathrm{I})}$, the instantaneous reward accrued by staying idle ($(1 - q_{ps}^{(\mathrm{I})}) b R_{s\mathrm{U}}$) approaches and, at some point, becomes larger than the reward accrued by transmitting ($T_{s\mathrm{U}} + (1 - q_{ps}^{(\mathrm{A})}) b R_{s\mathrm{U}}$), hence the

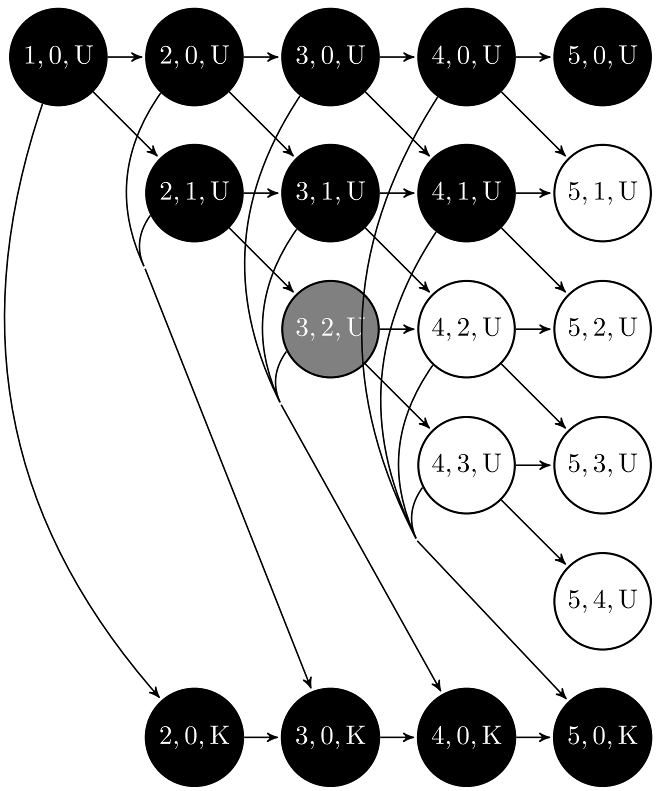


Fig. 4. Illustrative example of the structure of the optimal secondary access policy for the degenerate cognitive radio network; the SU is active in the black states, idle in the white ones, and randomly accesses the channel in the gray state; the arrows indicate the possible state transitions (transitions to state $(1,0,\mathrm{U})$ are omitted).

incentive to stay idle grows. On the other hand, if Δ_s is large, then the marginal throughput gain accrued in the states where the PU message is known to SUrx, over the throughput accrued in the states where the PU message is unknown, is large. The SU is thus incentivized to stay idle in the initial ARQ rounds, so as to help SUrx decode the PU message. Therefore, for large Δ_s, the optimal policy may not obey the structure of Lemma 3.

As a final remark, note that, in the degenerate cognitive radio network scenario, the only limitation to the activity of the SU is the secondary power expenditure $\bar{P}_s(\mu)$, since the primary throughput is unaffected. In the special case $\mathcal{P}_s^{(\mathrm{th})} = P_s$ in (13), neither the secondary power expenditure nor the primary throughput degradation limit the activity of the SU, hence the optimal policy solves the unconstrained maximization problem $\mu^* = \arg\max_\mu \bar{T}_s(\mu)$, whose solution follows as a corollary of Lemma 3.

Corollary 1. In the degenerate cognitive radio network scenario, the solution of the unconstrained optimization problem $\mu^* = \arg\max_\mu \bar{T}_s(\mu)$ yields

$$\mu^*(\mathbf{s}) = 1, \ \forall \mathbf{s} \in \mathcal{S}_\mathrm{K}, \tag{29}$$

$$\mu^*(t,b,\mathrm{U}) = \begin{cases} 1 & b < \bar{b}_{\max}(t) \\ 0 & b \geq \bar{b}_{\max}(t), \end{cases} \ , \ \forall (t,b,\mathrm{U}) \in \mathcal{S}_\mathrm{U}, \tag{30}$$

where $\bar{b}_{\max}(t)$ is defined in (26).

PU		
$R_p \simeq 2.52$	$q_{pp}^{(\mathrm{I})} \simeq 0.38$	$q_{pp}^{(\mathrm{A})} \simeq 0.68$
SU, $R_{s\mathrm{U}} = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p)$		
$R_{s\mathrm{U}} = 1.12$ $q_{ps}^{(\mathrm{I})} \simeq 0.61$ $R_{s\mathrm{K}} \simeq 1.91$	$T_{s\mathrm{U}} \simeq 0.59$ $q_{ps}^{(\mathrm{A})} \simeq 0.74$ $T_{s\mathrm{K}} \simeq 1.10$	$p_{s,\mathrm{buf}} = 0.26$
SU, $R_{s\mathrm{U}} = R_{s\mathrm{K}}$		
$R_{s\mathrm{U}} \simeq 1.91$ $q_{ps}^{(\mathrm{I})} \simeq 0.61$ $R_{s\mathrm{K}} \simeq 1.91$	$T_{s\mathrm{U}} \simeq 0.40$ $q_{ps}^{(\mathrm{A})} \simeq 0.88$ $T_{s\mathrm{K}} \simeq 1.10$	$p_{s,\mathrm{buf}} = 0.37$

TABLE I
PARAMETERS OF THE SU AND PU, FOR THE SNRS $\bar{\gamma}_s = 5$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$.

VI. Numerical Results

We consider a scenario with Rayleigh fading channels, *i.e.*, the SNR $\gamma_x,\ x \in \{s, p, sp, ps\}$, is an exponential random variable with mean $\mathbb{E}[\gamma_x] = \bar{\gamma}_x$. We consider the following parameters, unless otherwise stated. The average SNRs are set to $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{sp} = 2$. The ARQ deadline is $D = 5$. $R_{s\mathrm{K}}$ is chosen as $R_{s\mathrm{K}} = \arg\max_{R_s} T_{s\mathrm{K}}(R_s)$. The PU rate R_p is chosen as the maximizer of the instantaneous PU throughput under an idle SU, *i.e.*, $R_p = \arg\max_R T_p^{(\mathrm{I})}(R)$. For the rate $R_{s\mathrm{U}}$, we evaluate the two cases $R_{s\mathrm{U}} = R_{s\mathrm{U}}^*$ and $R_{s\mathrm{U}} = R_{s\mathrm{K}}$, where $R_{s\mathrm{U}}^* = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p)$. The former maximizes the instantaneous throughput under interference from the PU, thus neglecting the buffering capability at SUrx; therefore, the choice $R_{s\mathrm{U}} = R_{s\mathrm{U}}^*$ reflects a pessimistic expectation of the ability of SUrx to decode the PU message and to enable BIC. As to the latter, from (9) we have $R_{s\mathrm{U}} = R_{s\mathrm{K}} = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p) + p_{s,\mathrm{buf}}(R_s, R_p) R_{s\mathrm{K}}$, hence $R_{s\mathrm{U}} = R_{s\mathrm{K}}$ maximizes the sum of the instantaneous throughput and the future throughput possibly recovered via BIC, thus reflecting an optimistic expectation of the ability of SUrx to decode the PU message, which enables BIC. The PU throughput loss constraint is set to $\epsilon_{\mathrm{PU}} = 0.2$, and the constraint on the SU power is set to $\mathcal{P}_s^{(\mathrm{th})} = P_s$ (inactive). The resulting values of the system parameters are listed in Table I.

We consider the following schemes: "FIC/BIC", which employs both FIC and BIC; the optimal "FIC/BIC" policy is derived using Algorithm 1 and Lemma 2; "FIC only", which does not employ the buffering mechanism [1]; "no FIC/BIC", which employs neither BIC nor FIC. In this case, the SU message is decoded by leveraging the PU codebook structure [24]; however, possible knowledge of the PU message gained during the decoding operation is only used in the slot where the PU message is acquired, but is neglected in the past/future PU retransmissions. For "no FIC/BIC", the optimal policy consists in accessing the channel with a constant probability in all time-slots, independently of the underlying state, so as to attain the PU throughput loss constraint with equality. "PM known" refers to an ideal scenario where SUrx perfectly knows the current PU message in advance, and removes its interference;

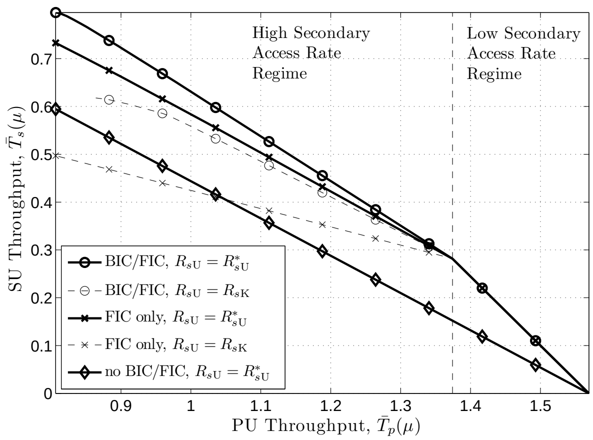


Fig. 5. SU throughput vs PU throughput. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. The other parameters are given in Table I.

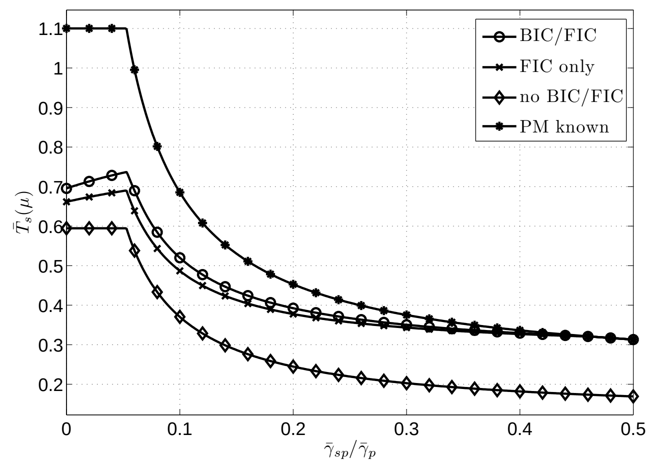


Fig. 6. SU throughput vs SNR ratio $\bar{\gamma}_{sp}/\bar{\gamma}_p$. PU throughput loss constraint $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_p = 10$. $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$.

specifically, SUtx transmits with rate $R_{s\mathrm{K}}$, thus accruing the throughput $T_{s\mathrm{K}}$ at each secondary access; "PM known" thus yields an upper bound to the performance of any other policy considered.

In Fig. 5, we plot the SU throughput versus the PU throughput, obtained by varying the SU access rate constraint ϵ_{W} in (17) from 0 to 1. As expected, the best performance is attained by "FIC/BIC", since the joint use of BIC and FIC enables IC at SUrx over the entire sequence of PU retransmissions. "FIC only" incurs a throughput penalty (except in the low SU access rate regime $\bar{T}_p(\mu) \geq 1.37$ where, from Lemma 2, "FIC/BIC" does not employ BIC), since the SU transmissions which undergo outage due to severe interference from the PU are simply dropped. "no FIC/BIC" incurs a further throughput loss, since possible knowledge about the PU message is not exploited to perform IC. Concerning the choice of the transmission rates, we note that the selection $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$ outperforms $R_{s\mathrm{U}} = R_{s\mathrm{K}}$ for the scenario considered. Note that, with $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$, the SU accrues a larger instantaneous throughput ($T_{s\mathrm{U}}$), but FIC and BIC are impaired, since both the buffering probability (8), $p_{s,\mathrm{buf}}$, and the probability that SUrx does not successfully decode the PU message, $q_{ps}^{(\mathrm{A})}$, diminish. Hence, in this case the instantaneous throughput maximization has a stronger impact on the performance than enabling FIC/BIC at SUrx.

In Fig. 6, we plot the SU throughput versus the SNR ratio $\bar{\gamma}_{sp}/\bar{\gamma}_p$, where $\bar{\gamma}_p = 5$ and $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$. Note that, for $\bar{\gamma}_{sp}/\bar{\gamma}_p \leq 0.5$, the SU throughput increases. In fact, in this regime the activity of the SU causes little harm to the PU, and the constraint on the PU throughput loss is inactive. The SU thus maximizes its own throughput. As $\bar{\gamma}_{sp}$ increases from 0 to $0.5\bar{\gamma}_p$, the activity of the SU induces more frequent primary ARQ retransmissions, hence there are more IC opportunities available and the SU throughput augments. On the other hand, as $\bar{\gamma}_{sp}$ grows beyond $0.5\bar{\gamma}_p$, the constraint on the PU throughput loss becomes active, secondary accesses become more and more harmful to the PU and take place more and more sparingly, hence the SU throughput degrades.

In Fig. 7, we plot the SU throughput versus the SNR ratio $\bar{\gamma}_{ps}/\bar{\gamma}_s$, where $\bar{\gamma}_s = 5$ and $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$, which is a function of $\bar{\gamma}_{ps}$. We notice that, when $\bar{\gamma}_{ps} = 0$, the upper bound is achieved with equality, since the SU operates under no interference from the PU. The upper bound is approached also for $\bar{\gamma}_{ps} \gg \bar{\gamma}_s$, corresponding to a strong interference regime where, with high probability, SUrx can successfully decode the PU message, remove its interference from the received signal, and then attempt to decode the SU message. The worst performance is attained when $\bar{\gamma}_{ps} \simeq \bar{\gamma}_s/2$. In fact, the interference from the PU is neither weak enough to be simply treated as noise, nor strong enough to be successfully decoded and then removed.

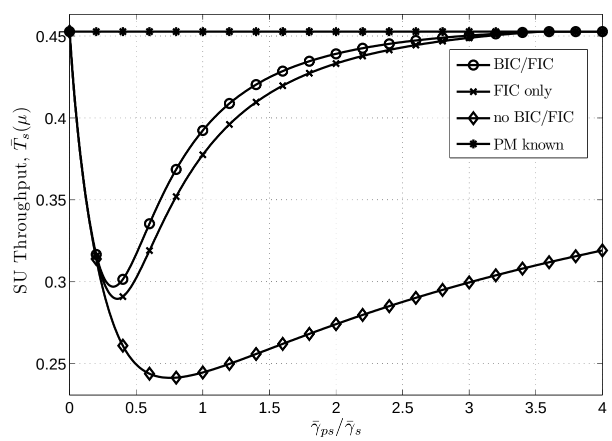


Fig. 7. SU throughput vs SNR ratio $\bar{\gamma}_{ps}/\bar{\gamma}_s$. PU throughput loss constraint $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$.

In Fig. 8, we plot the SU throughput versus the SU rate ratio $R_{s\mathrm{U}}/R_{s\mathrm{K}}$, where $R_{s\mathrm{K}} \simeq 1.91$ is kept fixed. Clearly, "no FIC/BIC" attains the best performance for $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$, which maximizes the throughput $T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p)$ achieved when neither FIC nor BIC are used. On the other hand, the performance of "FIC/BIC" is maximized for a slightly larger value of $R_{s\mathrm{U}}$. In fact, this value reflects the optimal trade-off between maximizing the throughput $T_{s\mathrm{U}}$ ($R_{s\mathrm{U}} \simeq 0.59 R_{s\mathrm{K}}$ in Fig. 9), maximizing the buffering probability, $p_{s,\mathrm{buf}}$ ($R_{s\mathrm{U}} \to 1$), and minimizing the probability that SUrx does not successfully

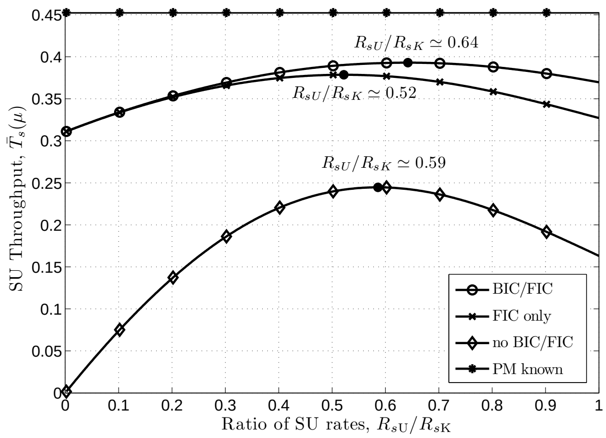


Fig. 8. SU throughput vs SU rate ratio $R_{s\mathrm{U}}/R_{s\mathrm{K}}$. $R_{s\mathrm{K}} \simeq 1.91$ is kept fixed. PU throughput loss constraint $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{ps} = 5$.

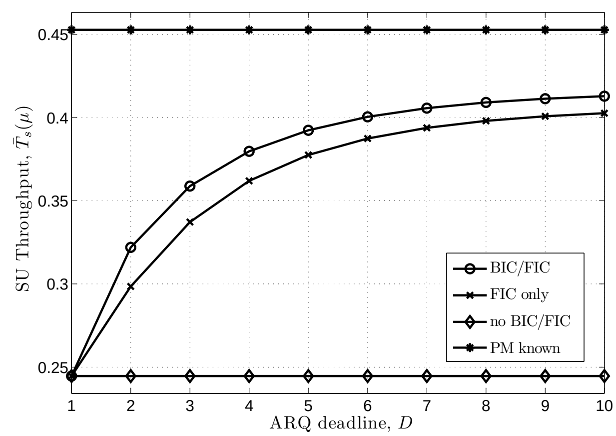


Fig. 10. SU throughput vs ARQ deadline D. PU throughput loss constraint $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. $R_{s\mathrm{U}} = R_{s\mathrm{U}}^*$.

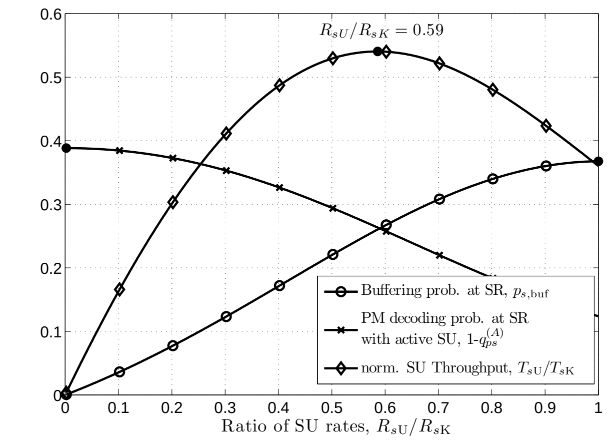


Fig. 9. Probabilities $p_{s,\mathrm{buf}}$, $1 - q_{ps}^{(\mathrm{A})}$ and normalized SU throughput $T_{s\mathrm{U}}$ vs the SU rate ratio $R_{s\mathrm{U}}/R_{s\mathrm{K}}$. $R_{s\mathrm{K}} \simeq 1.91$ is kept fixed. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$.

decode the PU message, $q_{ps}^{(\mathrm{A})}$ ($R_{s\mathrm{U}} \to 0$). Finally, "FIC only" is optimized by $R_{s\mathrm{U}} \simeq 0.52 R_{s\mathrm{K}} < R_{s\mathrm{U}}^*$. Since "FIC only" does not use BIC, this value reflects the optimal trade-off between maximizing $T_{s\mathrm{U}}$ and minimizing $q_{ps}^{(\mathrm{A})}$ ($R_{s\mathrm{U}} \to 0$).

In Fig. 10, we plot the SU throughput versus the ARQ deadline D. We notice that, when $D = 1$, all the IC mechanisms considered attain the same performance as "no FIC/BIC". In fact, this is a degenerate scenario where the PU does not employ ARQ, hence no redundancy is introduced in the primary transmission process. Interestingly, by employing FIC or BIC, the performance improves as D increases. In fact, the larger D, the more the redundancy introduced by the primary ARQ process, hence the more the opportunities for FIC/BIC at SUrx.

VII. Conclusion

In this work, we have investigated the idea of leveraging the redundancy introduced by the ARQ protocol implemented by a Primary User (PU) to perform Interference Cancellation (IC) at the receiver of a Secondary User (SU) pair: the SU receiver (SUrx), after decoding the PU message, exploits this knowledge to perform *Forward IC* (FIC) in the following ARQ retransmissions and *Backward IC* (BIC) in the previous ARQ retransmissions, corresponding to SU transmissions whose decoding failed due to severe interference from the PU. We have employed a stochastic optimization approach to optimize the SU access strategy which maximizes the average long-term SU throughput, under constraints on the average long-term PU throughput degradation and SU power expenditure. We have proved that the SU prioritizes its channel accesses in the states where SUrx knows the PU message, thus enabling FIC, and we have provided an algorithm to optimally allocate additional secondary access opportunities in the states where the PU message is unknown. Finally, we have shown numerically the throughput gain of the proposed schemes.

Appendix A

In this appendix, we compute $\bar{T}_s(\mu)$, $\bar{W}_s(\mu)$ and state properties of $\bar{W}_s(\mu)$.

Definition 3. We define $\mathbf{G}_\mu(t,b,\Phi)$, $\mathbf{V}_\mu(t,b,\Phi)$ and $\mathbf{D}_\mu(t,b,\Phi)$ as the average throughput, the average number of secondary channel accesses and the average number of time-slots, respectively, accrued starting from state (t,b,Φ) until the end of the primary ARQ cycle under policy μ (*i.e.*, until the recurrent state $(1,0,\mathrm{U})$ is reached). Starting from $\mathbf{X}_\mu(D+1,b,\Phi) = 0,\ \forall b, \forall \Phi \in \{\mathrm{U},\mathrm{K}\}$,[7] where $\mathbf{X}_\mu$ stands for $\mathbf{G}_\mu$, $\mathbf{V}_\mu$ or $\mathbf{D}_\mu$ (we write $\mathbf{X} \in \{\mathbf{G},\mathbf{V},\mathbf{D}\}$), these are defined recursively as, for $t \in \mathbb{N}(1,D)$, $b \in \mathbb{N}(0,t-1)$,

$$\begin{aligned}
\mathbf{X}_\mu(t,b,\mathrm{U}) &= x_\mu(t,b,\mathrm{U}) \\
&+\Pr_\mu(t+1,b,\mathrm{U}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,b,\mathrm{U}) \\
&+\Pr_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,b+1,\mathrm{U}) \\
&+\Pr_\mu(t+1,0,\mathrm{K}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,0,\mathrm{K}), \\
\mathbf{X}_\mu(t,0,\mathrm{K}) &= x_\mu(t,0,\mathrm{K}) \\
&+ \left[q_{pp}^{(\mathrm{I})} + \mu(t,0,\mathrm{K})(q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\right]\mathbf{X}_\mu(t+1,0,\mathrm{K}),
\end{aligned} \tag{31}$$

where $x_\mu(t,b,\Phi)$ is the cost/reward accrued in state (t,b,Φ) and $\Pr_\mu(\cdot|\cdot)$ is the one-step transition probability, which can be

[7] We introduce the fictitious state $(D+1,b,\Phi)$ for notational convenience.

TABLE II
TRANSITION PROBABILITIES. $X \in \{\mathrm{A}, \mathrm{I}\}$ DENOTES THE ACTION OF THE SU: ACTIVE (A) OR IDLE (I)

From \ To	$(1,0,\mathrm{U})$	$(t+1,b,\mathrm{U})$		$(t+1,b+1,\mathrm{U})$		$(t+1,0,\mathrm{K})$
	$X \in \{\mathrm{A}, \mathrm{I}\}$	A	I	A	I	$X \in \{\mathrm{A}, \mathrm{I}\}$
(t,b,U)	$1-q_{pp}^{(X)}$	$q_{pp}^{(\mathrm{A})}(q_{ps}^{(\mathrm{A})}-p_{s,\mathrm{buf}})$	$q_{pp}^{(\mathrm{I})}q_{ps}^{(\mathrm{I})}$	$q_{pp}^{(\mathrm{A})}p_{s,\mathrm{buf}}$	0	$q_{pp}^{(X)}(1-q_{ps}^{(X)})$
(D,b,U)	1	0		0		0
$(t,0,\mathrm{K})$	$1-q_{pp}^{(X)}$	0		0		$q_{pp}^{(X)}$
$(D,0,\mathrm{K})$	1	0		0		0

derived with the help of Table II by taking the expectation with respect to the actions *SU idle* (I, with probability $1-\mu(t,b,\Phi)$) and *SU active* (A, with probability $\mu(t,b,\Phi)$), yielding

$$\mathrm{Pr}_\mu(t+1,b,\mathrm{U}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}\left(q_{ps}^{(A)}-p_{s,\mathrm{buf}}\right) + (1-\mu(t,b,\mathrm{U}))q_{pp}^{(I)}q_{ps}^{(I)}, \tag{32}$$

$$\mathrm{Pr}_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}p_{s,\mathrm{buf}}, \tag{33}$$

$$\mathrm{Pr}_\mu(t+1,0,\mathrm{K}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}\left(1-q_{ps}^{(A)}\right) + (1-\mu(t,b,\mathrm{U}))q_{pp}^{(I)}\left(1-q_{ps}^{(I)}\right). \tag{34}$$

Namely, if $\mathbf{X} = \mathbf{G}$ (throughput), then $x_\mu(t,b,\Phi)$, $\Phi \in \{\mathrm{U},\mathrm{K}\}$, is the expected throughput accrued in state (t,b,Φ), and is given by

$$x_\mu(t,0,\mathrm{K}) = \mu(t,0,\mathrm{K})T_{s\mathrm{K}} \triangleq g_\mu(t,0,\mathrm{K}), \tag{35}$$

$$x_\mu(t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})T_{s\mathrm{U}} + \left[\mu(t,b,\mathrm{U})(1-q_{ps}^{(\mathrm{A})}) + (1-\mu(t,b,\mathrm{U}))(1-q_{ps}^{(\mathrm{I})})\right] bR_{s\mathrm{U}} \triangleq g_\mu(t,b,\mathrm{U}), \tag{36}$$

where the second term in (36) accounts for the successful recovery of the b SU messages from the buffered received signals via BIC, when the PU message is decoded by SUrx; if $\mathbf{X} = \mathbf{V}$ (secondary access), then $x_\mu(t,b,\Phi)$ is the SU access probability in state (t,b,Φ), *i.e.*,

$$x_\mu(t,b,\Phi) = \mu(t,b,\Phi) \triangleq v_\mu(t,b,\Phi); \tag{37}$$

finally, if $\mathbf{X} = \mathbf{D}$ (time-slots), then

$$x_\mu(t,b,\Phi) = 1 \triangleq d_\mu(t,b,\Phi), \tag{38}$$

corresponding to one time-slot. Moreover, we define, for $\mathbf{X} \in \{\mathbf{G},\mathbf{V},\mathbf{D}\}$,

$$\mathbf{X}'_\mu(\mathbf{s}) \triangleq \frac{\mathrm{d}\mathbf{X}'_\mu(\mathbf{s})}{\mathrm{d}\mu(\mathbf{s})}. \tag{39}$$

□

The number of visits to state $(1,0,\mathrm{U})$ up to time-slot n is a renewal process [25]. Each renewal interval (*i.e.*, the ARQ sequence in which the PU attempts to deliver a specific packet) has average duration $\mathbf{D}_\mu(1,0,\mathrm{U})$, over which the expected accrued SU throughput is $\mathbf{G}_\mu(1,0,\mathrm{U})$, and the expected number of secondary channel accesses is $\mathbf{V}_\mu(1,0,\mathrm{U})$. Then, the following lemma directly follows from the strong law of large numbers for renewal-reward processes [25].

Lemma 4. The average long-term SU throughput and access rate are given by $\bar{T}_s(\mu) = \frac{\mathbf{G}_\mu(1,0,\mathrm{U})}{\mathbf{D}_\mu(1,0,\mathrm{U})}$ and $\bar{W}_s(\mu) = \frac{\mathbf{V}_\mu(1,0,\mathrm{U})}{\mathbf{D}_\mu(1,0,\mathrm{U})}$, respectively. □

We have the following lemma.

Lemma 5. We have

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} \geq 0, \quad \forall \mathbf{s} \in \mathcal{S},\ \forall \mu \in \mathcal{U}. \tag{40}$$

The inequality is strict if and only if state $\mathbf{s}$ is accessible from $(1,0,\mathrm{U})$ under policy μ, *i.e.*, $\exists\ n > 0 : \mathrm{Pr}_\mu^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$. Moreover, for all $\mathbf{s} \in \mathcal{S}$ we have

$$\mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu) > 0. \tag{41}$$

□

Proof: If state $\mathbf{s}$ is not accessible from state $(1,0,\mathrm{U})$ under policy μ, then the steady state distribution satisfies $\pi_\mu(\mathbf{s}) = 0$, hence $\bar{W}_s(\mu)$ is unaffected by $\mu(\mathbf{s})$. Otherwise, from Lemma 4 we have that

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = \frac{\frac{\mathrm{d}\mathbf{V}_\mu(1,0,\mathrm{U})}{\mathrm{d}\mu(\mathbf{s})} - \frac{\mathrm{d}\mathbf{D}_\mu(1,0,\mathrm{U})}{\mathrm{d}\mu(\mathbf{s})}\bar{W}_s(\mu)}{\mathbf{D}_\mu(1,0,\mathrm{U})} \propto \mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu), \tag{42}$$

where $\propto$ represents equality up to a positive multiplicative factor, and the right hand side holds since, $\forall \mathbf{X} \in \{\mathbf{V},\mathbf{D}\}$ and $(t,b,\Phi) \in \mathcal{S}$, $\frac{\mathrm{d}\mathbf{X}_\mu(1,0,\mathrm{U})}{\mathrm{d}\mu(t,b,\Phi)} = \mathrm{Pr}_\mu^{(t)}(t,b,\Phi|1,0,\mathrm{U})\,\mathbf{X}'_\mu(t,b,\Phi)$.

If $\mathbf{s} \in \mathcal{S}_\mathrm{K}$, *i.e.*, $\mathbf{s} = (t,0,\mathrm{K})$, we have

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t,0,\mathrm{K})} \propto \mathbf{V}'_\mu(t,0,\mathrm{K}) - \mathbf{D}'_\mu(t,0,\mathrm{K})\bar{W}_s(\mu) \geq \mathbf{V}'_\mu(t,0,\mathrm{K}) - \mathbf{D}'_\mu(t,0,\mathrm{K}) \triangleq A_\mu(t), \tag{43}$$

where, from (31) we have used the fact that $\mathbf{D}'_\mu(t,0,\mathrm{K}) = +(q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\mathbf{D}_\mu(t+1,0,\mathrm{K}) \geq 0$ and $\bar{W}_s(\mu) \leq 1$.

We now prove by induction that $A_\mu(t) > 0, \forall\ t \in \mathbb{N}(1,T)$, so that (40) and (41) follow for $\mathbf{s} \in \mathcal{S}_\mathrm{K}$. From (31), for $t < D$, after algebraic manipulation we obtain

$$A_\mu(t) = 1 + (q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})[\mathbf{V}_\mu(t+1,0,\mathrm{K}) - \mathbf{D}_\mu(t+1,0,\mathrm{K})] = 1 - q_{pp}^{(\mathrm{A})} + \mathrm{Pr}_\mu(t+2,0,\mathrm{K}|t+1,0,\mathrm{K})A_\mu(t+1). \tag{44}$$

Since $A_\mu(D) = 1 > 0$, we obtain $A_\mu(t) > 0$ by induction.

If $\mathbf{s} \in \mathcal{S}_\mathrm{U}$, *i.e.*, $\mathbf{s} = (t,b,\mathrm{U})$, we have

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t,b,\mathrm{U})} \propto \mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U})\bar{W}_s(\mu). \tag{45}$$

We prove that $\mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U})\bar{W}_s(\mu) > 0$ in two steps, so that (40) and (41) follow for $\mathbf{s} \in \mathcal{S}_\mathrm{U}$. First, we prove that $C_\mu(t,b) \triangleq \mathbf{D}'_\mu(t,b,\mathrm{U}) \geq 0$. Then, since $\bar{W}_s(\mu) \leq 1$, we obtain

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t,b,0)} \propto \mathbf{V}'_\mu(t,b,\mathrm{U}) - C_\mu(t,b)\bar{W}_s(\mu) \geq \mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U}) \triangleq B_\mu(t,b). \quad (46)$$

Finally, we prove that $B_\mu(t,b) > 0$.

Proof of $C_\mu(t,b) \geq 0$: from (31), for $t < D$ we have

$$\begin{aligned} C_\mu(t,b) =&(q_{pp}^{(\mathrm{A})}(1-q_{ps}^{(\mathrm{A})}) - q_{pp}^{(\mathrm{I})}(1-q_{ps}^{(\mathrm{I})}))\mathbf{D}_\mu(t+1,0,\mathrm{K}) \\ &+ (q_{pp}^{(\mathrm{A})}(q_{ps}^{(\mathrm{A})} - p_{s,\mathrm{buf}}) - q_{pp}^{(\mathrm{I})}q_{ps}^{(\mathrm{I})})\mathbf{D}_\mu(t+1,b,\mathrm{U}) \\ &+ q_{pp}^{(\mathrm{A})}p_{s,\mathrm{buf}}\mathbf{D}_\mu(t+1,b+1,\mathrm{U}). \end{aligned} \quad (47)$$

Using the recursions (31) and rearranging the terms, we obtain the recursive expression

$$\begin{aligned} C_\mu(t,b) =& \Pr_\mu(t+2,b+2,\mathrm{U}|t+1,b+1,\mathrm{U})C_\mu(t+1,b+1) \\ &+ q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})} + \Pr_\mu(t+2,b,\mathrm{U}|t+1,b,\mathrm{U})C_\mu(t+1,b) \\ &+ \Big[(1-\mu(t+1,0,\mathrm{K}))q_{pp}^{(\mathrm{I})}(1-q_{ps}^{(\mathrm{I})}) \\ &+\mu(t+1,0,\mathrm{K})q_{pp}^{(\mathrm{A})}(1-q_{ps}^{(\mathrm{A})})\Big]\,(q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\mathbf{D}_\mu(t+2,0,\mathrm{K}). \end{aligned}$$

Since $C_\mu(D,b) = 0,\ \forall\ b \in \mathbb{N}(0,D-1)$, it follows by induction on t that $C_\mu(b,t) \geq 0$.

Proof of $B_\mu(t,b) > 0$: From (31), for $t < D$ we obtain the following recursive expression for $B_\mu(t,b)$, after algebraic manipulation,

$$\begin{aligned} B_\mu(t,b) =& 1 - q_{pp}^{(\mathrm{A})} + \Pr_\mu(t+2,b,\mathrm{U}|t+1,b,\mathrm{U})B_\mu(t+1,b) \\ &+ \Pr_\mu(t+2,b+2,\mathrm{U}|t+1,b+1,\mathrm{U})B_\mu(t+1,b+1) \\ &+ \Big[(1-\mu(t+1,0,\mathrm{K}))q_{pp}^{(\mathrm{I})}(1-q_{ps}^{(\mathrm{I})}) \\ &+\mu(t+1,0,\mathrm{K})q_{pp}^{(\mathrm{A})}(1-q_{ps}^{(\mathrm{A})})\Big]\,A_\mu(t+1), \end{aligned} \quad (48)$$

here $A_\mu(t)$ is defined in (43). The result follows by induction, since $B_\mu(D,b) = 1 > 0$ and $A_\mu(t+1) > 0$. ■

Appendix B

In this appendix, we give a rigorous definition of secondary access efficiency, thus complementing Def. 2. Moreover, in Lemma 6, we derive it. We recall that $\Pr_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$ is the n-step transition probability of the chain from $\mathbf{s}_0$ to $\mathbf{s}$.

Definition 4. Let $\tilde{\mu} \in \mathcal{U}$ be a policy such that $\exists n > 0 : \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$, and $\mu_\upsilon = (1-\upsilon)\mu + \upsilon\tilde{\mu}$, where $\upsilon \in (0,1]$, $\mu \in \mathcal{U}$. We define the *secondary access efficiency* under policy μ in state $\mathbf{s} \in \mathcal{S}$ as

$$\eta_\mu(\mathbf{s}) = \lim_{\upsilon\to0^+} \left.\frac{\frac{\mathrm{d}\bar{T}_s(\mu_\upsilon)}{\mathrm{d}\mu_\upsilon(\mathbf{s})}}{\frac{\mathrm{d}\bar{W}_s(\mu_\upsilon)}{\mathrm{d}\mu_\upsilon(\mathbf{s})}}\right|_{\mu_\upsilon}.$$

□

Remark 5. Notice that the condition $\exists\ n > 0 : \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$ guarantees that state $\mathbf{s}$ is accessible from state $(1,0,\mathrm{U})$ under policy μ_υ, for $\upsilon > 0$. Under this condition, $\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} > 0$ (Lemma 5 in App. A), hence the fraction within the limit is well defined for $\upsilon > 0$ and in the limit $\upsilon \to 0^+$. One such policy $\tilde{\mu}$ is $\tilde{\mu}(\mathbf{s}) = 0.5, \forall \mathbf{s} \in \mathcal{S}$. □

Using Lemma 4 and Def. 3 in App. A and Def. 4, $\eta_\mu(\mathbf{s})$ can be derived according to the following lemma.

Lemma 6. We have $\eta_\mu(\mathbf{s}) = \frac{\mathbf{G}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{T}_s(\mu)}{\mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu)}$. □

Remark 6. This is well defined, since $\mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu) > 0$ from Lemma 5 in App. A. □

Appendix C

Proof of Theorem 1: In the first part of the theorem, we prove that, by initializing Algorithm 1 with the idle policy $\mu^{(0)}$, $\mu^{(0)}(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}$, and with the set of idle states $\mathcal{S}_{\mathrm{idle}}^{(0)} \equiv \mathcal{S}$, we obtain an optimal policy. In the second part of the proof, we prove the optimality of the specific initialization of Algorithm 1 for the high SU access rate regime.

Let $\tilde{\mu}$ be a policy under which all states $\mathbf{s} \in \mathcal{S}$ are accessible from state $(1,0,\mathrm{U})$, *i.e.*, $\exists\ n > 0 :\ \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$. One such policy is $\tilde{\mu}(\mathbf{s}) = 0.5,\ \forall \mathbf{s} \in \mathcal{S}$. Consider a modified Markov Decision Process, parameterized by $\upsilon \in (0,1)$, obtained by applying the policy $(1-\upsilon)\mu + \upsilon\tilde{\mu}$ to the original system, where $\mu \in \mathcal{U}$. Since $\mu, \tilde{\mu} \in \mathcal{U}$ and $\upsilon \in (0,1)$, it follows that $(1-\upsilon)\mu + \upsilon\tilde{\mu} \in \mathcal{U}$. We define $\bar{T}_s(\mu,\upsilon) \triangleq \bar{T}_s((1-\upsilon)\mu + \upsilon\tilde{\mu})$ and $\bar{W}_s(\mu,\upsilon) \triangleq \bar{W}_s((1-\upsilon)\mu + \upsilon\tilde{\mu})$, and we study the problem

$$\mu^{*(\upsilon)} = \arg\max\nolimits_{\mu\in\mathcal{U}} \bar{T}_s(\mu,\upsilon) \text{ s.t. } \bar{W}_s(\mu,\upsilon) \leq \epsilon_\mathrm{W}, \quad (49)$$

where the parameter υ is small enough to guarantee a feasible problem, *i.e.*, $\exists\ \mu \in \mathcal{U} : \bar{W}_s(\mu,\upsilon) \leq \epsilon_\mathrm{W}$. (17) is obtained in the limit $\upsilon \to 0^+$. Notice that, $\forall\ \mu \in \mathcal{U}$, under policy $(1-\upsilon)\mu + \upsilon\tilde{\mu}$, all the states $\mathbf{s} \in \mathcal{S}$ are accessible from state $(1,0,\mathrm{U})$, and the Markov chain is irreducible. Hence, from Lemma 5 in App. A, $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s})$, $\forall \mathbf{s} \in \mathcal{S}$. This is an important assumption in the following proof.

Let $\mathcal{D} \subset \mathcal{U}$ be the set of all the deterministic policies, and $\mathcal{G}_\upsilon = \left\{\left(\bar{W}_s(\mu,\upsilon), \bar{T}_s(\mu,\upsilon)\right), \mu \in \mathcal{D}\right\}$. With the help of Fig. 11, for any $\mu \in \mathcal{U}$, we have that $\left(\bar{W}_s(\mu,\upsilon), \bar{T}_s(\mu,\upsilon)\right) \in \mathrm{conv}(\mathcal{G}_\upsilon)$, where $\mathrm{conv}(\mathcal{G}_\upsilon)$ is the convex hull of the set $\mathcal{G}_\upsilon$. In particular, for the optimal policy we have $\left(\bar{W}_s(\mu^{*(\upsilon)},\upsilon), \bar{T}_s(\mu^{*(\upsilon)},\upsilon)\right) \in \mathrm{bd}(\mathcal{G}_\upsilon)$, where $\mathrm{bd}(\mathcal{G}_\upsilon)$ denotes the boundary of $\mathrm{conv}(\mathcal{G}_\upsilon)$.

Algorithm 1 determines the sequence of vertices of the polyline $\mathrm{bd}(\mathcal{G}_\upsilon)$ in the limit $\upsilon \to 0^+$ (bold line in Fig. 11). For $\upsilon > 0$, starting from the leftmost vertex of $\mathrm{bd}(\mathcal{G}_\upsilon)$, achieved by the *idle* policy $\mu^{(0)}(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}$ (this follows from the fact that $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s})$, hence it is minimized by the idle policy), the algorithm determines iteratively the next vertex of $\mathrm{bd}(\mathcal{G}_\upsilon)$ as the maximizer of the slope

$$\mu^{(i+1)} = \underset{\mu\in\mathcal{D}:\bar{W}_s(\mu,\upsilon)>\bar{W}_s(\mu^{(i)},\upsilon)}{\arg\max} \frac{\bar{T}_s(\mu,\upsilon) - \bar{T}_s(\mu^{(i)},\upsilon)}{\bar{W}_s(\mu,\upsilon) - \bar{W}_s(\mu^{(i)},\upsilon)}. \quad (50)$$

Since (17) has one constraint, the optimal policy $\mu^{*(\upsilon)}$ is randomized in one state [22], and hence each segment on the boundary $\mathrm{bd}(\mathcal{G}_\upsilon)$ between pairs $(\bar{W}_s(\mu^{(i)},\upsilon), \bar{T}_s(\mu^{(i)},\upsilon))$

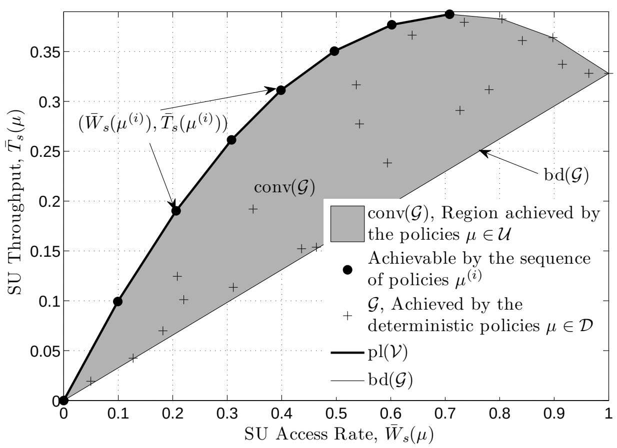


Fig. 11. Geometric interpretation of problem (49)

achievable with deterministic policies is attained by a policy that is randomized in only one state. It follows that $\mu^{(i)}$ and $\mu^{(i+1)}$ differ in only one state. Moreover, in (50) the maximization is over $\mu \in \mathcal{D}$ such that $\bar{W}_s(\mu,\upsilon) > \bar{W}_s(\mu^{(i)},\upsilon)$, *i.e.*, since $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s})$ and $\mu^{(i+1)}$ and $\mu^{(i)}$ differ in only one position, $\mu^{(i+1)}$ is obtained from $\mu^{(i)}$ by allocating one more secondary access to a state which is idle under $\mu^{(i)}$. In (50), the maximization is thus over $\left\{\mu^{(i)} + \delta_{\mathbf{s}} : \mathbf{s} \in \mathcal{S}_{\text{idle}}^{(i)}\right\}$, and, after algebraic manipulation, $\mu^{(i+1)}$ in (50) maximizes

$$\max_{\mathbf{s}\in\mathcal{S}_{\text{idle}}^{(i)}} \frac{\bar{T}_s(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon) - \bar{T}_s(\mu^{(i)},\upsilon)}{\bar{W}_s(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon) - \bar{W}_s(\mu^{(i)},\upsilon)} = \max_{\mathbf{s}\in\mathcal{S}_{\text{idle}}^{(i)}} \eta_{(1-\upsilon)\mu^{(i)}+\upsilon\tilde{\mu}}(\mathbf{s}).$$

Stage i of the algorithm is thus proved. If $\eta_{(1-\upsilon)\mu^{(i)}+\upsilon\tilde{\mu}}(\mathbf{s}) \leq 0$, we have $\bar{W}_s\left(\mu^{(i)} + \delta_{\mathbf{s}},\upsilon\right) > \bar{W}_s\left(\mu^{(i)},\upsilon\right)$ and $\bar{T}_s\left(\mu^{(i)} + \delta_{\mathbf{s}},\upsilon\right) \leq \bar{T}_s\left(\mu^{(i)},\upsilon\right)$. If this condition holds $\forall\ \mathbf{s} \in \mathcal{S}_{\text{idle}}^{(i)}$, any next vertex of the polyline $\text{bd}(\mathcal{G}_\upsilon)$ yields a decrease of the SU throughput and a larger SU access rate, hence a sub-optimal set of policies, and the algorithm stops.

By construction, the algorithm returns a sequence of policies $(\mu^{(i)}, i \in \mathbb{N}(0,N-1))$, characterized by strictly increasing values of the SU throughput and of the SU access rate. The optimal policy belongs to the polyline with vertices $\mathcal{V}_\upsilon \equiv \{(\bar{W}_s(\mu^{(i)},\upsilon), \bar{T}_s(\mu^{(i)},\upsilon)), i \in \mathbb{N}(0,N-1)\}$, denoted by $\text{pl}(\mathcal{V}_\upsilon)$ in Fig. 11. Then, (17) becomes equivalent to $T_s^{*(\upsilon)} = \max_{(W_s,T_s)\in\mathcal{V}_\upsilon} T_s$ s.t. $W_s \leq \epsilon_{\text{W}}$, whose solution is given in the last step of Algorithm 1. The result finally follows for $\upsilon \to 0^+$.

To conclude, we prove the initialization of Algorithm 1 for the high SU access rate. Let $(\mu^{(0)},\dots,\mu^{(N-1)})$ and $(\mathbf{s}^{(0)},\dots,\mathbf{s}^{(N-1)})$ be the sequence of deterministic policies and of states returned by Algorithm 1, obtained by initializing the algorithm as in the first part of the proof. Let $\mathcal{D}_0 \equiv \{\mu \in \mathcal{D} : \mu(t,0,0) = 0\ \forall\ t \in \mathbb{N}(1,T)\}$, $\tilde{\mathcal{D}}_0 \equiv \{\mu \in \mathcal{D}_0 : \mu(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\text{K}}\}$, and $N_0 \triangleq \max\{i \in \{0,\dots,N-1\} : \bar{W}_s(\mu^{(i)}) < \epsilon_{\text{th}}\}$. We prove that $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$, *i.e.*, $\mu^{(N_0+1)}(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\text{K}}$. From the definition of $\tilde{\mathcal{D}}_0$ and the construction of the algorithm, it follows that, for $i > N_0$, $\mu^{(i)}(\mathbf{s}) = 1,\ \forall\ \mathbf{s} \in \mathcal{S}_{\text{K}}$. Moreover, from Lemma 7, $\bar{W}_s(\mu^{(N_0+1)}) = \epsilon_{\text{th}}$. Hence, for the high SU access rate $\epsilon > \epsilon_{th}$, the optimal policy μ^* obeys $\mu^*(\mathbf{s}) = 1, \forall \mathbf{s} \in \mathcal{S}_{\text{K}}$. Then, letting $\mathcal{U}_1 \equiv \{\mu \in \mathcal{U} : \mu(\mathbf{s}) = 1, \forall \mathbf{s} \in \mathcal{S}_{\text{K}}\}$, the optimization problem (17) can be restricted to the set of randomized policies $\mu \in \mathcal{U}_1 \subset \mathcal{U}$ when $\epsilon > \epsilon_{th}$. Equivalently, secondary accesses taking place in $\mathcal{S}_{\text{U}}$ can be obtained by initializing the algorithm with $\mu^{(0)}(\mathbf{s}) = 0,\ \mathbf{s} \in \mathcal{S}_{\text{U}}$, $\mu^{(0)}(\mathbf{s}) = 1,\ \mathbf{s} \in \mathcal{S}_{\text{K}}$, $S_{\text{idle}}^{(0)} \equiv \mathcal{S}_{\text{U}}$. The initialization of Algorithm 1 is thus proved.

Proof of $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$: We prove by induction that $\mu^{(i)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0,\ \forall i \leq N_0$ and $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$. Assume that, for some $i \geq 0$, $\mu^{(j)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0,\ \forall j \leq i$. From Lemma 7, it follows that $N_0 \geq i$. This clearly holds for $i = 0$. We show that this implies that either $\mu^{(i+1)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$, hence $N_0 > i$, thus proving the induction step, or $\mu^{(i+1)} \in \tilde{\mathcal{D}}_0$, hence $N_0 = i$, thus proving the property. The result follows since $N_0 \leq 1 + |\mathcal{S}| < \infty$ (*i.e.*, $i = N_0$ is reached within a finite number of steps).

From Lemma 8, $\eta_{\mu^{(i)}}(\mathbf{s}) = T_{s\text{K}} > 0, \forall \mathbf{s} \in \mathcal{S}_{\text{K}} \cap \mathcal{S}_{\text{idle}}^{(i)}$ and $\eta_{\mu^{(i)}}(t,0,\text{U}) < T_{s\text{K}}, \forall t \in \mathbb{N}(1,D)$, hence, from the main iteration stage of the algorithm it follows that $\mu^{(i+1)} \in \mathcal{D}_0$. In particular, if $\mu^{(i+1)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$, then $N_0 > i$ from Lemma 7. On the other hand, if $\mu^{(i+1)} \in \tilde{\mathcal{D}}_0$, then, from Lemma 7, $N_0 = i$. The property is thus proved. ■

Lemma 7.
$\bar{W}_s(\mu) < \epsilon_{\text{th}}, \forall \mu \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$ and $\bar{W}_s(\mu) = \epsilon_{\text{th}}, \forall \mu \in \tilde{\mathcal{D}}_0$. □

Proof: Let $\mu \in \tilde{\mathcal{D}}_0$. Since the states (t,b,U) with $b > 0$ are not accessible from $(1,0,\text{U})$ under μ, the transmission probability $\mu(t,b,\text{U})$, $b > 0$, does not affect $\bar{W}_s(\mu)$. Then, from Def. 1, we have $\bar{W}_s(\mu) = \epsilon_{\text{th}}$.

Let $\mu \in \mathcal{D} \setminus \tilde{\mathcal{D}}_0$. Letting $\mathcal{S}_\mu = \{\mathbf{s} \in \mathcal{S}_{\text{K}} : \mu(\mathbf{s}) = 0\}$, we have that $\mu + \sum_{\mathbf{s}\in\mathcal{S}_\mu} \delta_{\mathbf{s}} \in \tilde{\mathcal{D}}_0$. Finally, since every $\mathbf{s} \in \mathcal{S}_\mu$ is accessible from $(1,0,\text{U})$ under μ, and $\mathcal{S}_\mu$ is non-empty, from Lemma 5 in App. A and the previous case, it follows that $\bar{W}_s(\mu) < \bar{W}_s(\mu + \sum_{\mathbf{s}\in\mathcal{S}_\mu} \delta_{\mathbf{s}}) = \epsilon_{\text{th}}$. ■

Lemma 8. Let $\mu \in \mathcal{U}$ such that $\mu(t,0,\text{U}) = 0\ \forall\ t \in \mathbb{N}(1,D)$. Then, $\eta_\mu(t,0,\text{U}) < T_{s\text{K}}$ and $\eta_\mu(t,0,\text{K}) = T_{s\text{K}},\ \forall t$. □

Proof: Let $\mu \in \mathcal{U}$ such that $\mu(t,0,\text{U}) = 0\ \forall\ t \in \mathbb{N}(1,D)$. It follows that the states (t,b,U) with $b > 0$ are not accessible, hence their steady state probability satisfies $\pi_\mu(t,b,\text{U}) = 0,\ \forall\ t,\ \forall\ b > 0$. It is then straightforward to show, by using the recursion (31), that $\mathbf{G}_\mu(t,0,\text{U}) = T_{s\text{K}}\mathbf{V}_\mu(t,0,\text{K})$, $\mathbf{G}_\mu(t,0,\text{K}) = T_{s\text{K}}\mathbf{V}_\mu(t,0,\text{K})$ and $\bar{T}_s(\mu) = T_{s\text{K}}\bar{W}_s(\mu)$. Then, using these expressions, the recursion (31) and Lemma 6, we obtain $\eta_\mu(t,0,\text{K}) = T_{s\text{K}}$ and

$$\eta_\mu(t,0,\text{U}) = T_{s\text{K}} - \frac{T_{s\text{K}}\mathbf{V}_\mu'(t,0,\text{U}) - \mathbf{G}_\mu'(t,0,\text{U})}{\mathbf{V}_\mu'(t,0,\text{U}) - \mathbf{D}_\mu'(t,0,\text{U})\bar{W}_s(\mu)}. \quad (51)$$

We now prove that $\eta_\mu(t,0,\text{U}) < T_{s\text{K}}$, which proves the lemma. Equivalently, using Lemma 5 in App. A and (31), we prove that

$$T_{s\text{K}}\mathbf{V}_\mu'(t,0,\text{U}) - \mathbf{G}_\mu'(t,0,\text{U}) = (T_{s\text{K}} - T_{s\text{U}}) + q_{pp}^{(\text{A})} p_{s,\text{buf}}[T_{s\text{K}}\mathbf{V}_\mu(t,1,\text{U}) - \mathbf{G}_\mu(t,1,\text{U})] > 0. \quad (52)$$

Letting

$$M_\mu(t,b) = b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) \quad (53)$$
$$+ q_{pp}^{(\mathrm{A})} p_{s,\mathrm{buf}}[T_{s\mathrm{K}}\mathbf{V}_\mu(t,b,\mathrm{U}) - \mathbf{G}_\mu(t,b,\mathrm{U})] > 0,\ \forall\ t,b \geq 1,$$

(52) is equivalent to $M_\mu(t,1) > 0$. We now prove by induction that $M_\mu(t,b) > 0,\ \forall\ t,b \geq 1$, yielding (52) as a special case when $b=1$. For $t = D+1$ we have $M_\mu(D+1,b) = b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) > 0$, since $T_{s\mathrm{K}} > T_{s\mathrm{U}}$ and $b \geq 1$. Now, let $t \leq D$ and assume $M_\mu(t+1,b) > 0$. Using (31), after algebraic manipulation we obtain

$$M_\mu(t,b) = b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) + q_{pp}^{(\mathrm{A})} p_{s,\mathrm{buf}} \mu(t,b,\mathrm{U})(T_{s\mathrm{K}} - T_{s\mathrm{U}})$$
$$- q_{pp}^{(\mathrm{A})} p_{s,\mathrm{buf}} \left[1 - \mu(t,b,\mathrm{U}) q_{ps}^{(\mathrm{A})} - (1-\mu(t,b,\mathrm{U})) q_{ps}^{(\mathrm{I})}\right] bR_{s\mathrm{U}}$$
$$+ \Pr_\mu(t+1,b,\mathrm{U}|t,b,\mathrm{U})[M_\mu(t+1,b) - b(T_{s\mathrm{K}} - T_{s\mathrm{U}})]$$
$$+ \Pr_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U}) M_\mu(t+1,b)$$
$$- \Pr_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U})(b+1)(T_{s\mathrm{K}} - T_{s\mathrm{U}}). \quad (54)$$

Finally, since $M_\mu(t+1,b) > 0$ by the induction hypothesis, using inequality (9) we obtain

$$M_\mu(t,b) > p_{s,\mathrm{buf}} bR_{s\mathrm{U}} \left(1 - q_{pp}^{(\mathrm{A})}\right)$$
$$+ p_{s,\mathrm{buf}} bR_{s\mathrm{U}} (1-\mu(t,b,\mathrm{U})) q_{ps}^{(\mathrm{I})} (q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})}) > 0, \quad (55)$$

which proves the induction step. The lemma is proved. ■

APPENDIX D

Proof of Lemma 3: Let $\mathcal{D} \subset \mathcal{U}$ be the set of all the deterministic (non-randomized) policies. Let

$$\tilde{\mathcal{D}} \equiv \{\mu \in \mathcal{D} : \mu(t,b,\mathrm{U}) = 1, \forall t, b < b(t);$$
$$\mu(t,b,\mathrm{U}) = 0, \forall t, b \geq b(t);\ \mu(\mathbf{s}) = 1, \mathbf{s} \in \mathcal{S}_\mathrm{K};$$
$$\exists\ b(\cdot) : b(t+1) \leq b(t)\ \forall t\}\ .$$

By inspection, we have that the sequences of policies (22) are such that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i \in \mathbb{N}(0,N-1)$. Therefore, the first part of the lemma states that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i \in \mathbb{N}(0,N-1)$. We prove this property by induction. Namely, we show that $\mu^{(i)} \in \tilde{\mathcal{D}} \Rightarrow \mu^{(i+1)} \in \tilde{\mathcal{D}}$. Then, since $\mu^{(0)} \in \tilde{\mathcal{D}}$ (initialization of Algorithm 1) it follows that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i$. Let $\mu^{(i)} \in \tilde{\mathcal{D}}$, *i.e.*, $\mu^{(i)}$ is given by (22) for some $b^{(i)}(t)$ non-increasing in t. The set of idle states is then given by

$$\mathcal{S}_{\mathrm{idle}}^{(i)} \equiv \left\{(t,b,\mathrm{U}) \in \mathcal{S}_\mathrm{U} : t \in \mathbb{N}(1,D), b \geq b^{(i)}(t)\right\}. \quad (56)$$

We then prove that, under the hypotheses of the lemma, $\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu^{(i)}}(t,b+1,\mathrm{U})$ and $\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_\mu(t+1,b,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. It follows that the SU access efficiency is maximized by the state in the idle set $\mathcal{S}_{\mathrm{idle}}^{(i)}$ with the lowest value of the primary ARQ state t, among the states with the same buffer occupancy b, and with the fewest number of buffered received signals b, among the states with the same primary ARQ state t. Therefore, in the main iteration stage of the algorithm, the SU access efficiency is maximized by $\mathbf{s}^{(i)} = \arg\max_{\mathbf{s} \in \mathcal{S}_{\mathrm{idle}}^{(i)}} \eta_{\mu^{(i)}}(\mathbf{s})$, where $\mathbf{s}^{(i)} = (t,b,\mathrm{U})$ is such that $\tau \geq t,\ \beta \geq b,\ \forall\ (\tau,\beta,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. By inspection, we have that $\mu^{(i+1)} = \mu^{(i)} + \delta_{\mathbf{s}^{(i)}} \in \tilde{\mathcal{D}}$, hence the induction step is proved.

We thus need to prove the induction step, *i.e.*, letting $\mu^{(i)} \in \tilde{\mathcal{D}}$, we show that

$$\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu^{(i)}}(t,b+1,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)},$$
$$\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_\mu(t+1,b,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}. \quad (57)$$

To this end, note that, in the degenerate cognitive radio network scenario, the primary ARQ process is not affected by the SU access scheme, hence, using the notation in App. A, $\mathbf{D}'_{\mu^{(i)}}(t,b,\mathrm{U}) = 0$. By the definition of SU access efficiency (6), we thus obtain

$$\eta_{\mu^{(i)}}(t,b,\mathrm{U}) = \frac{\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})}{\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})}, \quad (58)$$

where, using (31), (32-34), (36) and (37),

$$\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) = T_{s\mathrm{U}} + \left(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})}\right) bR_{s\mathrm{U}} \quad (59)$$
$$+ q_{pp}(q_{ps}^{(\mathrm{A})} - p_{s,\mathrm{buf}} - q_{ps}^{(\mathrm{I})})\mathbf{G}_{\mu^{(i)}}(t+1,b,\mathrm{U})$$
$$+ q_{pp} p_{s,\mathrm{buf}} \mathbf{G}_{\mu^{(i)}}(t+1,b+1,\mathrm{U})$$
$$+ q_{pp}(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})})\mathbf{G}_{\mu^{(i)}}(t+1,0,\mathrm{K}),$$
$$\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) = 1 + q_{pp}(q_{ps}^{(\mathrm{A})} - p_{s,\mathrm{buf}} - q_{ps}^{(\mathrm{I})})\mathbf{V}_{\mu^{(i)}}(t+1,b,\mathrm{U})$$
$$+ q_{pp} p_{s,\mathrm{buf}} \mathbf{V}_{\mu^{(i)}}(t+1,b+1,\mathrm{U})$$
$$+ q_{pp}(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})})\mathbf{V}_{\mu^{(i)}}(t+1,0,\mathrm{K}). \quad (60)$$

Using the fact that $\mu^{(i)}(\tau,\beta,\mathrm{U}) = 0,\ \forall \tau \geq t, \beta \geq b$, it can be proved that

$$\mathbf{V}_{\mu^{(i)}}(\tau,\beta,\mathrm{U}) = A_1(\tau) - A_0(\tau), \quad (61)$$
$$\mathbf{G}_{\mu^{(i)}}(\tau,\beta,\mathrm{U}) = (1 - q_{ps}^{(\mathrm{I})})\beta R_{s\mathrm{U}} A_0(\tau) \quad (62)$$
$$+ T_{s\mathrm{K}}(A_1(\tau) - A_0(\tau)),$$
$$\mathbf{V}_{\mu^{(i)}}(\tau,0,\mathrm{K}) = A_1(\tau), \quad (63)$$
$$\mathbf{G}_{\mu^{(i)}}(\tau,0,\mathrm{K}) = T_{s\mathrm{K}} A_1(\tau), \quad (64)$$

where $A_0(\cdot)$ and $A_1(\cdot)$ are defined in (27) and (28), respectively. The expressions (61-64) can be easily verified by induction, starting from $\tau = D+1$ backward. In fact, for $\tau = D+1$, we have $A_0(D+1) = A_1(D+1) = 0$, hence we obtain $\mathbf{V}_{\mu^{(i)}}(D+1,\beta,\mathrm{U}) = \mathbf{G}_{\mu^{(i)}}(D+1,\beta,\mathrm{U}) = \mathbf{V}_{\mu^{(i)}}(D+1,0,\mathrm{K}) = \mathbf{G}_{\mu^{(i)}}(D+1,0,\mathrm{K}) = 0$, which is consistent with Def. 3. The induction step can be proved by inspection, using the recursive expression (31) and the fact that $\mu(\tau,\beta,\mathrm{U}) = 0,\ \forall \tau \geq t, \beta \geq b$. Substituting the expressions (61-64) in (59) and (60), we obtain

$$\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) = T_{s\mathrm{U}} + q_{pp} p_{s,\mathrm{buf}} (1 - q_{ps}^{(\mathrm{I})}) R_{s\mathrm{U}} A_0(t+1)$$
$$+ \left(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})}\right) bR_{s\mathrm{U}} \left[1 - q_{pp}(1 - q_{ps}^{(\mathrm{I})}) A_0(t+1)\right]$$
$$+ q_{pp}(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})}) T_{s\mathrm{K}} A_0(t+1), \quad (65)$$
$$\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) = 1 - q_{pp}(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}) A_0(t+1). \quad (66)$$

Proof of $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu^{(i)}}(t,b+1,0)$

By substituting (65) and (66) in (58), and noticing that $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) = \mathbf{V}'_{\mu^{(i)}}(t,b+1,\mathrm{U})$ from (66) and $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) > 0$ (from Lemma 5 with $\mathbf{D}'_\mu(\mathbf{s}) = 0$), the condition $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu^{(i)}}(t,b+1,0)$ is equivalent to

$\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) > \mathbf{G}'_{\mu^{(i)}}(t,b+1,\mathrm{U})$, which is readily verified from (65), since

$$\begin{aligned}&\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) - \mathbf{G}'_{\mu^{(i)}}(t,b+1,\mathrm{U}) \\ &= \left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) R_{s\mathrm{U}} \left[1 - q_{pp}(1-q_{ps}^{(\mathrm{I})})A_0(t+1)\right] \\ &> \left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) \frac{1-q_{pp}}{1-q_{pp}q_{ps}^{(\mathrm{I})}} R_{s\mathrm{U}} > 0,\end{aligned} \tag{67}$$

where the first inequality follows from the fact that $A_0(t+1) < \frac{1}{1-q_{pp}q_{ps}^{(\mathrm{I})}}$, the second from $q_{ps}^{(\mathrm{I})} < q_{ps}^{(\mathrm{A})}$.

Proof of $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu}(t+1,b,0)$

Since $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) > 0$, the condition $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu}(t+1,b,0)$ is equivalent to

$$\begin{aligned}&\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) \left(\mathbf{V}'_{\mu^{(i)}}(t+1,b,\mathrm{U}) - \mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})\right) \\ &> \mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) \left(\mathbf{G}'_{\mu^{(i)}}(t+1,b,\mathrm{U}) - \mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})\right).\end{aligned} \tag{68}$$

Using (65) and (66), after algebraic manipulation we obtain the equivalent condition

$$\begin{aligned}&\left(1-q_{ps}^{(\mathrm{A})}\right) p_{s,\mathrm{buf}} + \left(1-q_{ps}^{(\mathrm{A})}\right)\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) b \\ &+ \left(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})}\right) \Delta_s > 0,\end{aligned} \tag{69}$$

where we have used the fact that $T_{s\mathrm{K}} = \Delta_s R_{s\mathrm{U}} + T_{s\mathrm{U}} + p_{s,\mathrm{buf}} R_{s\mathrm{U}}$. Since we require this condition to hold $\forall b \geq 0$ and the left hand expression is minimized by $b=0$, the condition (69) should be satisfied for $b=0$, yielding the equivalent condition $\Delta_s < \frac{1-q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})}-q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}}$, which is an hypothesis of the lemma.

It is thus proved that the sequence of policies returned by Algorithm 1 has the structure defined by (22), where $b^{(i)}(t)$ satisfies the inequality (24). Moreover, the inequality (25) holds since, by the algorithm construction, $\mu^{(i+1)}$ is obtained from $\mu^{(i)}$ by "activating" one additional state from the set of idle states $\mathcal{S}_{\mathrm{idle}}^{(i)}$.

The second part of the lemma states that $b^{(N-1)}(t) = \bar{b}_{\max}(t)$, where $\bar{b}_{\max}(t)$ is given by (26). This is a consequence of the fact that Algorithm 1 stops if the SU access efficiency becomes non-positive, *i.e.*, $\eta_{\mu^{(i)}}(\mathbf{s}) \leq 0$, $\forall \mathbf{s} \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. From (58), this condition is equivalent to $\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) \leq 0$, $\forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. By using (65) and by solving $\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) \leq 0$ with respect to b, the result follows. ■

REFERENCES

[1] N. Michelusi, O. Simeone, M. Levorato, P. Popovski, and M. Zorzi, "Optimal Cognitive Transmission Exploiting Redundancy in the Primary ARQ process," in *Information Theory and Applications Workshop*, Feb. 2011.

[2] N. Michelusi, P. Popovski, M. Levorato, O. Simeone, and M. Zorzi, "Cognitive transmissions under a primary ARQ process via Backward Interference Cancellation," in *49th Allerton Conference on Communication, Control, and Computing*, Sep. 2011, pp. 727–735.

[3] J. Mitola and G. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[4] FCC Spectrum Policy Task Force, "Report of the Spectrum Efficiency Working Group," FCC, Tech. Rep., available on http://www.fcc.gov/sptf/reports.html.

[5] J. Peha, "Approaches to spectrum sharing," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[6] ——, "Sharing Spectrum Through Spectrum Policy Reform and Cognitive Radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, Apr. 2009.

[7] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[8] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894 –914, May 2009.

[9] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[10] R. Comroe and D. Costello, "ARQ Schemes for Data Transmission in Mobile Radio Systems," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 472–481, July 1984.

[11] D. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.

[12] D. J. White, *Markov Decision Processes*. Wiley, 1993.

[13] M. Levorato, U. Mitra, and M. Zorzi, "Cognitive Interference Management in Retransmission-Based Wireless Networks," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3023–3046, May 2012.

[14] R. Tannious and A. Nosratinia, "Cognitive Radio Protocols Based on Exploiting Hybrid ARQ Retransmissions," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2833 –2841, Sep. 2010.

[15] M. Luby, "LT Codes," in *Proceedings of the 43rd Symposium on Foundations of Computer Science*, ser. FOCS'02. Washington, DC, USA: IEEE Computer Society, 2002.

[16] S. Gollakota and D. Katabi, "Zigzag decoding: combating hidden terminals in wireless networks," in *Proceedings of the ACM SIGCOMM conference on Data communication*. ACM, 2008, pp. 159–170.

[17] J. Li, W. Zhang, A. Nosratinia, and J. Yuan, "Opportunistic Spectrum Sharing Based on Exploiting ARQ Retransmission in Cognitive Radio Networks," in *IEEE Global Telecommunications Conference (Globecom)*, Dec. 2010, pp. 1–5.

[18] A. Jovicic and P. Viswanath, "Cognitive Radio: An Information-Theoretic Perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[19] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., New York, 2006.

[21] M. Levorato, S. Firouzabadi, and A. Goldsmith, "A Learning Framework for Cognitive Interference Networks with Partial and Noisy Observations," *IEEE Transactions on Wireless Communications*, 2012, accepted for publication.

[22] K. W. Ross, "Randomized and Past-Dependent Policies for Markov Decision Processes with Multiple Constraints," *Operations Research*, vol. 37, no. 3, pp. 474–477, 1989.

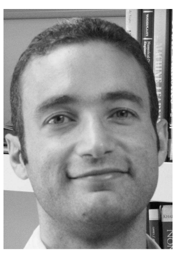
[23] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Springer, 1960.

[24] R. Narasimhan, "Individual Outage Rate Regions for Fading Multiple Access Channels," in *IEEE International Symposium on Information Theory, ISIT*, June 2007, pp. 24–29.

[25] R. Gallager, *Discrete Stochastic Processes*. Kluwer, Boston, 1996.

Nicolò Michelusi (S'09) received the B.S. (Electronics Engineering) and M.S. (Telecommunications Engineering) degrees summa cum laude from the University of Padova, Italy, in 2006 and 2009, respectively, and the M.S. degree in Telecommunications Engineering from Technical University of Denmark, Copenhagen, Denmark, in 2009, as part of the T.I.M.E. double degree program. Since January 2009, he is a Ph.D. student at University of Padova, Italy. In 2011, he was on leave at the University of Southern California, Los Angeles, United States, as a visiting Ph.D. student. His research interests include ultrawideband communications, wireless networks, cognitive radio networks, stochastic optimization, optimal control, energy harvesting for communications.

Petar Popovski (S'97-A'98-M'04-SM'10) received the Dipl.-Ing. in electrical engineering and Magister Ing. in communication engineering from Sts. Cyril and Methodius University, Skopje, Macedonia, in 1997 and 2000, respectively and Ph. D. from Aalborg University, Denmark, in 2004. He was Assistant Professor (2004-2009) and Associate Professor (2009-2012) at Aalborg University. From 2008 to 2009 he held part-time position as a wireless architect at Oticon A/S. Since 2012 he is a Professor at Aalborg University. He has more than 140 publications in journals, conference proceedings and books and has more than 25 patents and patent applications. He has received the Young Elite Researcher award and the SAPERE AUDE career grant from the Danish Council for Independent Research. He has received six best paper awards, including three from IEEE. Dr. Popovski serves on the editorial board of several journals, including IEEE Communications Letters (Senior Editor), IEEE Transactions on Communications and IEEE Transactions on Wireless Communications. His research interests are in the broad area of wireless communication and networking, information theory and protocol design.



Osvaldo Simeone (M'02) received the M.Sc. degree (with honors) and the Ph.D. degree in information engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively. He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark, where he is an Associate Professor. His current research interests concern the cross-layer analysis and design of wireless networks with emphasis on information-theoretic, signal processing, and queuing aspects. Specific topics of interest are: cognitive radio, cooperative communications, rate-distortion theory, ad hoc, sensor, mesh and hybrid networks, distributed estimation, and synchronization. Dr. Simeone is a co-recipient of Best Paper Awards of the IEEE SPAWC 2007 and IEEE WRECOM 2007. He currently serves as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS.

Marco Levorato (S'06, M'09) obtained both the BE (Electronics and Telecommunications Engineering) and the ME (Telecommunications Engineering) summa cum laude from the University of Ferrara, Italy, in 2002 and 2005, respectively. In 2009, he received a Ph.D. in Information Engineering from the University of Padova. During 2008 he was on leave at the University of Southern California, Los Angeles, United States. In 2009 he was a post doctorate researcher at the University of Padova. Since January 2010, he is a post doctorate researcher at Stanford and the University of Southern California (USC).

Michele Zorzi (S'89, M'95, SM'98, F'07) was born in Venice, Italy, on December 6th, 1966. He received the Laurea and the PhD degrees in Electrical Engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the Academic Year 1992/93, he was on leave at the University of California, San Diego (UCSD) as a visiting PhD student, working on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the University of Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty at the Information Engineering Department of the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, energy constrained communications protocols, broadband wireless access and underwater acoustic communications and networking.

Dr. Zorzi was the Editor-In-Chief of the IEEE WIRELESS COMMUNICATIONS MAGAZINE from 2003 to 2005 and the Editor-In-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2008 to 2011, and currently serves on the Editorial Board of the WILEY JOURNAL OF WIRELESS COMMUNICATIONS AND MOBILE COMPUTING. He was also guest editor for special issues in the IEEE PERSONAL COMMUNICATIONS MAGAZINE (Energy Management in Personal Communications Systems) IEEE WIRELESS COMMUNICATIONS MAGAZINE (Cognitive Wireless Networks) and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Multi-media Network Radios, and Underwater Wireless Communications Networks). He served as a Member-at-large of the Board of Governors of the IEEE Communications Society from 2009 to 2011.