Spectrum Leasing via Cooperative Opportunistic Routing Techniques

Davide Chiarotto, Student Member, IEEE, Osvaldo Simeone, Member, IEEE, and Michele Zorzi, Fellow, IEEE

Abstract—A licensed multihop network that coexists with a set of unlicensed nodes is considered. Coexistence is regulated via a spectrum leasing mechanism that is based on cooperation and opportunistic routing. Specifically, the primary network consists of a source and a destination communicating via a number of primary relay nodes. In each transmission block, the next hop is selected in an on-line fashion based on the channel conditions (and thus the decoding outcome) in the previous transmissions, according to the idea of opportunistic routing. The secondary nodes may serve as extra relays, and hence potential next hops, for the primary network, but only in exchange for spectrum leasing. Namely, in return for their forwarding of primary packets, secondary nodes are awarded spectral resources for transmission of their own traffic. Secondary nodes enforce Quality-of-Service requirements in terms of rate and reliability when deciding whether or not to cooperate. Four policies that exploit spectrum leasing via opportunistic routing in different ways are proposed. These policies are designed to span different operating points in the trade-off between gains in throughput and overall energy expenditure for the primary network. Analysis is carried out for networks with a linear geometry and quasistatic Rayleigh fading statistics by using Markov chain tools. Different multiplexing techniques are considered for multiplexing of the primary and secondary traffic at the secondary nodes, namely orthogonal multiplexing (such as time, frequency or orthogonal code division multiplexing) and superposition coding. The optimality in terms of both throughput and primary energy consumption of superposition coding over all possible multiplexing strategies, for the given routing techniques, is proved. Finally, numerical results demonstrate the advantages of the proposed spectrum leasing solution based on opportunistic routing and illustrate the trade-offs between primary throughput and energy consumption.

Index Terms—Cognitive radio networks, property-rights, spectrum leasing, cooperative transmission, opportunistic routing, end-to-end throughput, superposition coding.

I. INTRODUCTION

THE problem of scarce radio spectrum availability and the inefficiency of traditional fixed spectrum management

Manuscript received October 27, 2010; revised March 14, 2011; accepted May 12, 2011. The associate editor coordinating the review of this paper and approving it for publication was G. Abreu.

Part of this work was presented at the Asilomar Conference on Signals, Systems, and Computers, 2010. This work was partially supported by the U. S. National Science Foundation under Grant # CCF-0914899, and by the European Commission under the MEDIEVAL project (grant agreement no FP7-258053).

D. Chiarotto and M. Zorzi are with the Department of Information Engineering, University of Padova, Via G. Gradenigo 6/B, 35131 Padova (PD), Italy (e-mail: dchiarot, zorzi@dei.unipd.it), and with Consorzio Ferrara Ricerche, Via Saragat 1, 44100 Ferrara (FE), Italy.

O. Simeone is with the Center for Wireless Communications and Signal Processing Research, New Jersey Institute of Technology, Newark, New Jersey 07102-1982, USA (e-mail: osvaldo.simeone@njit.edu).

Digital Object Identifier 10.1109/TWC.2011.072011.101908

schemes call for new communications paradigms for spectrum sharing [1]. Spectrum leasing is one such paradigm in which licensed users are allowed to lease portions of the spectrum to unlicensed users. In a standard implementation, spectrum leasing (also called property-rights, [2]) would be effected at a system level with "spectrum servers" allocating resources to secondary users [2], [3]. Moreover, secondary users would be charged for their use of the spectral resources. Instead, references [4], [5] propose a novel approach in which spectrum leasing is performed locally and dynamically by primary devices, and remuneration from secondary to primary users takes place in the form of cooperation. In the approach of [4], [5], secondary nodes accept to cooperate only if granted enough spectrum with respect to their desired Quality-of-Service (QoS) requirements.

This work proposes to implement spectrum leasing via cooperation in a multihop scenario by means of opportunistic routing, which is a well-known technique that aims at increasing the throughput of multihop networks over fading channels by exploiting the channel diversity offered by the availability of multiple possible next hops. In particular, selection of the next hop is made in an opportunistic fashion based on the decoding outcomes of previous transmissions of the given packet, thanks to appropriate feedback from the decoders [6], [7]. An information-theoretic analysis of opportunistic routing in a linear multihop network over block-fading channels is presented in [8], showing better performance with respect to conventional strategies, which were studied in [9] for additive white Gaussian noise channels, and in [10], [11] for nonergodic and ergodic fading channels, respectively. Also related to this work are [12]-[14] and the references therein, where different techniques were proposed for multihop routing in secondary networks for cognitive radio scenarios operating according to the commons radio principle.

The main idea of the present paper is that secondary nodes may serve as potential hops for a primary network that routes packets based on opportunistic routing by leveraging the principle of spectrum leasing via cooperation. In particular, secondary nodes may be selected as next hops for the primary packet if this benefits the primary performance. However, in exchange for their cooperation, secondary nodes impose the condition that enough spectral resources should be leased to them to satisfy their QoS requirements, which are defined with respect to secondary single-hop communication with given rate and outage probability guarantee. The primary network, thanks to this approach, may gain on two fronts: (*i*) *Throughput*, due to the improved multiuser diversity in the selection of the next hop that is afforded by the availability



Fig. 1. A primary linear multihop network (gray circles) with k hops and a secondary network (white circles) aligned with respect to primary relay nodes. The destination of the secondary message is not necessarily shown in this picture.

of secondary nodes; (*ii*) *Primary energy consumption*, due to the fact that transmissions can be delegated to the secondary network. This work studies the trade-off between these two metrics by proposing protocols that work at different operating points of this trade-off.

Application scenarios in which the proposed framework applies are characterized by the coexistence of a primary, licensed, multihop network with secondary nodes that perform single-hop transmission with QoS constraints. For instance, primary and secondary users may belong to the same wireless network but have different priorities, perhaps corresponding to different subscription fees, where users with lower priority (secondary users) are limited to short-range (i.e., singlehop) transmissions. Primary users may be mobile terminals or infrastructure nodes such as femtocell access points. In the former case, energy is typically a relevant performance criterion, whereas in the latter case it may be more appropriate to only optimize throughput.

This paper proposes a number of spectrum leasing strategies based on opportunistic routing. Numerical results are provided to show the advantages of spectrum leasing over conventional transmission that involves only primary nodes, even when the secondary QoS requirements are significant. The analysis accounts for different possible multiplexing techniques at the secondary nodes, namely Orthogonal Multiplexing strategies (OM) (such as Time, Frequency or Orthogonal Code Division Multiplexing) and Superposition Coding (SC), showing the optimality of SC in terms of both throughput and primary energy consumption over all possible multiplexing strategies.

II. SYSTEM MODEL AND SECONDARY QOS REQUIREMENTS

In Fig. 1 we show a primary and a secondary network that coexist via spectrum leasing. The aim of the *primary source* P_0 is to communicate with the *primary destination* P_k , at a normalized distance of one, possibly taking advantage of multihop routing through two sets of additional nodes placed along two parallel linear geometries with vertical distance Δ_V . Both sets are composed of k - 1 nodes: the first one is formed by *primary nodes*, denoted by P_1, \ldots, P_{k-1} whose only role is that of forwarding information from P_0 to P_k ; the second set of nodes, instead, consists of *secondary nodes* S_1, \ldots, S_{k-1} that can access the channel only if spectrum is leased by the primary network (see Section II-B). Primary nodes have $\Delta_H = 1/k$ inter-node distance. Secondary nodes are aligned with primaries, and thus have the same internode distance. Generally, we will consider a *partial secondary*

deployment in which only one in every α secondary nodes is active so that the number of secondary nodes is $k/\alpha - 1$ with inter-node distance $\alpha \Delta_{\rm H}$. This work relies on geometrical simplifying assumptions with the objective of having both a solvable theoretical model and an insightful analysis of spectrum leasing via cooperative routing techniques. More general network topologies will be considered in future work.

All devices considered work in half-duplex mode and the transmission is organized in blocks of n (complex) channel uses each, where only one node is active (i.e., no spatial reuse is allowed). In the first block, the source P_0 transmits a packet of $n R_P$ bits, where R_P is the transmission rate of the original (primary) transmission in bits/s/Hz. In the following blocks, retransmissions take place, if necessary, according to a Type-I HARQ process (i.e., retransmissions are not combined at the destination). Retransmissions in each block may be performed by the source, or by the primary relays or secondary nodes, as long as the latter have correctly decoded in the previous block. After the packet is correctly delivered to the destination, the primary source transmits a new packet and the process repeats.

A. Signal Model and Secondary QoS Requirements

Considering a transmission between node $N_i \in \{P_0, \ldots, P_{k-1}, S_1, \ldots, S_{k-1}\}$ and node $N_j \in \{P_1, \ldots, P_k, S_1, \ldots, S_{k-1}\}$, we denote with

$$y_{N_iN_j}(b,t) = d_{N_iN_j}^{-\eta/2} h_{N_iN_j}(b) x_{N_i}(b,t) + z_{N_j}(b,t)$$
(1)

the discrete-time (complex) baseband sample received by node N_i during the b-th block, at channel use $t, t = 1, \ldots, n$. The path loss between the N_i -th transmitter and the N_j -th receiver with power path-loss exponent η is represented by $d_{N_iN_j}^{-\eta/2}$. The distance $d_{N_iN_j}$ can assume three forms (see also Fig. 1): (a) if both nodes N_i and N_j lie on the same line, i.e., for transmissions between primary relays or secondary nodes, $d_{N_iN_i}$ is given by $|j-i| \Delta_{\rm H}$; (b) if N_i is the source P_0 and the receiver is a primary or secondary relay, or N_i is a relay and N_j is the destination P_k , $d_{N_iN_j}$ is equal to $\Delta_{|j-i|}^{(T,D)}$, where $\Delta_a^{(T,D)} = \sqrt{(a\Delta_H)^2 + (\Delta_V/2)^2}$; (c) finally, if the transmission is between two relays, one in the primary and one in the secondary network, $d_{N_iN_j} = \Delta_{|j-i|}^{(R)}$ with $\Delta_a^{(R)} = \sqrt{(a\Delta_{\rm H})^2 + \Delta_{\rm V}^2}$. The channel coefficient between transmitter N_i and receiver N_j is represented by $h_{N_iN_j}(b)$, and assumed to be quasi-static Rayleigh fading, i.e., it is a complex Gaussian random variable with zero mean and unit power, assumed to be constant within each block, but to vary independently from block to block. The channel state information is not known to the transmitter, but only to the receiver. Network geometry, and thus distances $d_{N_iN_i}$, are known to all nodes. The term $x_{N_i}(b,t)$ represents the discrete-time (complex) baseband sample transmitted by the scheduled node N_i , with the per-symbol power constraint $\mathbb{E}[|x_{N_i}(b,t)|^2] \leq E_N$, $t = 1, \ldots, n$, where $E_{\rm N}$ is equal to $E_{\rm P}$ or $E_{\rm S}$ when the transmitter is a primary or a secondary node, respectively. Finally, we let $z_{N_i}(b, t)$ be the complex white Gaussian noise term with zero mean and power $\mathbb{E}[|z_{N_i}(b,t)|^2] = N_0$. We assume randomly generated Gaussian codebooks throughout. We define $\gamma_{\rm P}$ as the received signal-to-noise ratio (SNR) for primary users, which is given by the ratio between the

maximum average energy directly received by P_k from the source P_0 and the noise power N_0 , $\gamma_{\rm P} = E_{\rm P}/N_0$. Hence, for a transmission from a primary node that covers a distance d the average received SNR is $\gamma_{\rm P}d^{-\eta}$. For consistency, the SNR from a secondary node that covers a distance d is given by $\gamma_{\rm S}d^{-\eta}$, with $\gamma_{\rm S} = E_{\rm S}/N_0$.

B. Opportunistic Routing and Spectrum Leasing

The next hop decisions in the network are made by primary nodes based on the feedback received at the end of the previous block from all nodes (primary and secondary) that have successfully received the packet. Thanks to this information, the primary network can schedule transmissions in an opportunistic fashion based on the channel conditions, and thus the decoding outcomes, in previous blocks. Therefore, the choice of the next hop can only be based on the availability of the packet as a result of the previous transmission [6], [7]. Channel resources for feedback allocation will not be included in the analysis (as usually considered in the literature on opportunistic routing), assuming that the feedback information will be available to the primary nodes as required by the different protocols to be introduced below. We note that the amount of signaling present in our scenario is equivalent to that of classical opportunistic routing schemes. An analysis of the overhead involved is reported in [15], showing that opportunistic routing can retain significant benefits even with limited signaling.

As discussed above, secondary nodes may serve as relays for the current primary packet. However, secondary nodes follow the *spectrum leasing via cooperation* (relaying) principle stated in [4]: in fact, they do not cooperate for free, but they accept to serve as relays only if they are granted sufficient resource for their own traffic as well.

We consider Orthogonal Multiplexing (OM) and Superposition Coding (SC) to multiplex primary and secondary traffic. Specifically, in OM, the secondary node S_i multiplexes primary and secondary data by assigning a portion $0 \le \beta \le 1$ of the spectral resources (e.g., time or frequency) to forward the primary packet and the remaining part $1 - \beta$ to transmit its own data in the secondary network. SC, instead, is a physical layer technique in which a transmitter can simultaneously send independent messages to multiple receivers. Specifically, the secondary node transmits a signal obtained by superimposing the packet carrying its own data and the primary packet intended for the primary network. In particular, the secondary transmitter encodes and modulates both packets at the selected rates and scales the *power* of each modulated symbol to match the chosen power split, assigning a portion $0 \le \psi \le 1$ to the primary packet and $1 - \psi$ to its own packet. Finally, the complex baseband symbols (or waveforms with pass band representation) are added to obtain the transmitted signal [16]. Both the spectral resources fraction β for OM and the power fraction ψ for SC are selected by secondary nodes so as to satisfy their own QoS requirements in terms of rate and reliability.

To derive the spectral resources fraction β for OM and the power fraction ψ for SC that meet the QoS requirements of the secondary users, we assume that each secondary node wants

to transmit at rate R_S to a node, which is not shown in Fig.1, at distance d_S with outage probability ϵ_S . In Sections II-D and II-E derivations of β and ψ are detailed, respectively.

C. Outage Probabilities for Primary Transmission

Consider transmission from a primary node P_i . Assuming that the coding block is long enough, the probability that a packet transmitted by the primary node P_i is not decoded correctly by a node (primary, P_j , or secondary, S_j) at distance d^1 is given by [17]:

$$P_{\text{out},\mathbf{P}}(d) = \Pr\left\{\log_2\left(1+|h|^2\gamma_{\mathbf{P}}d^{-\eta}\right) \le \mathbf{R}_{\mathbf{P}}\right\}$$
$$=1-\exp\left(-\frac{2^{\mathbf{R}_{\mathbf{P}}}-1}{\gamma_{\mathbf{P}}d^{-\eta}}\right).$$
(2)

D. Outage Probabilities for Secondary Transmission: OM

In OM, the secondary node transmits the primary packet for a fraction β of the spectral resources and the secondary packet for the remaining part $1 - \beta$. In the following we evaluate the two corresponding outage probabilities. Let $P_{\text{out,SP}}^{(\text{OM})}(d)$ define the outage probability of a primary packet transmitted by a secondary node, at a distance d. Similarly to (2), this is given by:

$$P_{\text{out,SP}}^{(\text{OM})}(d) = \Pr\left\{\beta \log_2\left(1 + |h|^2 \gamma_{\text{S}} d^{-\eta}\right) \le \text{R}_{\text{P}}\right\}$$
$$= 1 - \exp\left(-\frac{2^{\text{R}_{\text{P}}/\beta} - 1}{\gamma_{\text{S}} d^{-\eta}}\right). \tag{3}$$

Notice that the rate for secondary transmissions of the primary packet needs to be increased to R_P / β to compensate for the fact that only a fraction of spectral resources β is used for primary data.

Similarly, the outage probability of a secondary packet transmitted by a secondary node is given by

$$P_{\text{out,SS}}^{(\text{OM})}(d) = 1 - \exp\left(-\frac{2^{\text{R}_{\text{S}}/(1-\beta)} - 1}{\gamma_{\text{S}}d^{-\eta}}\right).$$
 (4)

The choice of β depends, as discussed above, on the QoS requirements of the secondary nodes. Recalling that a fraction $1-\beta$ of the spectral resources is used for the secondary's own traffic and imposing the condition on the outage probability in (4) as $P_{\text{out,SS}}^{(0M)}(d_{\text{S}}) = \epsilon_{\text{S}}$, we obtain:

$$\beta = 1 - \frac{\mathrm{R}_{\mathrm{S}}}{\log_2 \left[1 - \log_e \left(1 - \epsilon_{\mathrm{S}}\right) \gamma_{\mathrm{S}} d_{\mathrm{S}}^{-\eta}\right]}.$$
 (5)

E. Outage Probabilities for Secondary Transmission: SC

With SC, the secondary node sends the sum of two (complex) codewords, one for the primary with power ψE_S and one for the secondary with power $(1 - \psi)E_S$. We consider a receiver that employs two decoders in parallel.² The first

¹In the rest of this work, for simplicity, we do not write explicitly the expressions of distance, $d_{N_iN_j}$, and channel coefficient, $h_{N_iN_j}$, between transmitter N_i and receiver N_j , but only d and h, with the understanding that the subscript N_iN_j is implied.

²We note that the SC receiver needs more decoding power with respect to the OM receiver, due to the successive interference cancellation algorithm used. This extra complexity could be factored in the problem formulation. However, since we do not focus on specific decoding schemes, it is not possible to specify exactly the number of operations. Instead, we could add a generic term to the energy expenditure to account for the larger power required by SC. This is fairly straightforward and will not be further pursued here.

decoder attempts to decode the desired packet by treating the undesired packet, which is superimposed, as additive Gaussian noise. The second decoder, instead, first attempts to estimate the undesired packet, cancels it from the received signal and then decodes the desired packet from the interference-free signal. The overall decoder successfully obtains the desired message if either of the two decoders discussed above decodes correctly (this can be checked via CRC, see for instance [18]). It is noted that this decoder is capacity-achieving for the Gaussian broadcast channel [16]. Further discussion on this issue can be found in Section III-C.

For the SC approach, the outage probability for a primary packet transmitted by a secondary node S_i to a node N_j (primary or secondary) at distance d can be found to be given by

$$P_{\text{out,SP}}^{(\text{sc)}}(d) = \Pr\left\{\log_{2}\left(1 + \frac{|h|^{2}\psi E_{\text{S}}d^{-\eta}}{N_{0} + |h|^{2}(1 - \psi)E_{\text{S}}d^{-\eta}}\right) \leq R_{\text{P}} \\ \cap \left[\left(\log_{2}\left(1 + \frac{|h|^{2}(1 - \psi)E_{\text{S}}d^{-\eta}}{N_{0} + |h|^{2}\psi E_{\text{S}}d^{-\eta}}\right) \leq R_{\text{S}}\right) \\ \cup \left(\log_{2}\left(1 + \frac{|h|^{2}\psi E_{\text{S}}d^{-\eta}}{N_{0}}\right) \leq R_{\text{P}}\right)\right]\right\}$$
(6)
$$= \Pr\left\{|h|^{2} \leq \min\left(\mathcal{H}_{\text{P}}^{(1)}, \mathcal{H}_{\text{P}}^{(2)}\right)\right\} = \\ = 1 - \exp\left[-\min\left(\mathcal{H}_{\text{P}}^{(1)}, \mathcal{H}_{\text{P}}^{(2)}\right)\right].$$
(7)

The first term in (6) represents the outage probability of the first decoder, in which the secondary packet is treated as noise. The remaining term is the outage probability of the successive decoding scheme, where the receiver first decodes the secondary packet and then the primary one. The overall outage probability (6) is the probability that both the first and the second decoder fail. As such, in equation (7), $\mathcal{H}_{\rm P}^{(1)}$ and $\mathcal{H}_{\rm P}^{(2)}$ represent the minimum values that the channel coefficient $|h|^2$ can assume without causing an outage for the two decoders:

$$\mathcal{H}_{P}^{(1)} = \begin{cases} \infty, & 0 \le \psi \le 1 - 2^{-R_{P}} \\ \frac{2^{R_{P}} - 1}{(1 - (1 - \psi)2^{R_{P}})\gamma_{S}d^{-\eta}}, & 1 - 2^{-R_{P}} < \psi \le 1 \end{cases}, \quad (8)$$

$$\mathcal{H}_{P}^{(2)} = \begin{cases} \max\left\{\frac{2^{R_{S}}-1}{(1-\psi 2^{R_{S}})\gamma_{S}d^{-\eta}}, \frac{2^{R_{P}}-1}{\psi\gamma_{S}d^{-\eta}}\right\}, & 0 < \psi < 2^{-R_{S}} \\ \infty, & \psi = 0 \text{ or } 2^{-R_{S}} \le \psi \le 1 \end{cases}$$
(9)

Notice from (9) that if the allocated power is too small, the channel gain threshold values for which there is no outage become infinite (i.e., outage occurs with probability one for all finite channel gains).

The outage probability that a secondary packet (superimposed with a primary packet) transmitted by a secondary node S_i is not decoded correctly by a secondary node S_j placed at distance d is given by

$$P_{\text{out,SS}}^{(\text{SC})}(d) = 1 - \exp\left[-\min\left(\mathcal{H}_{\text{S}}^{(1)}, \mathcal{H}_{\text{S}}^{(2)}\right)\right], \quad (10)$$

where $\mathcal{H}_{S}^{(1)}$ and $\mathcal{H}_{S}^{(2)}$ are

$$\mathcal{H}_{S}^{(1)} = \begin{cases} \frac{2^{R_{S}} - 1}{(1 - \psi 2^{R_{S}})\gamma_{S}d^{-\eta}}, & 0 \le \psi < 2^{-R_{S}} \\ \infty, & 2^{-R_{S}} \le \psi \le 1 \end{cases},$$
(11)

$$\mathcal{H}_{\rm S}^{(2)} = \begin{cases} \infty, & 0 \le \psi \le 1 - 2^{-\,{\rm R}_{\rm P}} \text{ and } \psi = 1\\ \max\left\{\frac{2^{{\rm R}_{\rm P}} - 1}{(1 - (1 - \psi)2^{{\rm R}_{\rm P}})\gamma_{\rm S}d^{-\eta}}, \frac{2^{{\rm R}_{\rm S}} - 1}{(1 - \psi)\gamma_{\rm S}d^{-\eta}}\right\}, & .\\ & 1 - 2^{-\,{\rm R}_{\rm P}} < \psi < 1 \end{cases}$$
(12)

As with OM, given the secondary QoS requirements ($d_{\rm S}$, $R_{\rm S}$, $\epsilon_{\rm S}$), one can obtain the resource allocation parameter ψ . While for OM this could be easily done in closed-form (5), for SC we had to resort to a numerical solution of the equation $P_{\rm out,SS}^{\rm (SC)}(d_{\rm S}) = \epsilon_{\rm S}$, for a given rate pair ($R_{\rm P}, R_{\rm S}$).

III. THROUGHPUT AND PRIMARY ENERGY ANALYSIS

The goal of this section is to define the performance metrics of interest, introduce four routing policies that exploit spectrum leasing via opportunistic routing in different ways and show the optimality of SC.

Let $T(k, R_P, Q)$ be the primary end-to-end throughput, defined as the average number of successfully transmitted bits per second per Hz, given the total number of hops k, the primary transmission rate R_P and the parameter Q, which represents the secondary QoS constraints (d_S, R_S, ϵ_S) or equivalently the parameter β for OM or ψ for SC. Using renewal theory, the throughput can be calculated as

$$T(k, \mathbf{R}_{\mathbf{P}}, \mathcal{Q}) = \frac{\mathbf{R}_{\mathbf{P}}}{\mathbb{E}[N]},$$
(13)

where N is the total number of blocks, including both primary and secondary transmissions, necessary to transmit a given packet correctly from the source P_0 to the destination P_k . We also define the primary energy $E(k, R_P, Q)$ as the average overall energy used by the primary network to deliver a packet successfully. When normalized with respect to the energy of a packet transmission, this quantity is the average number of primary transmissions necessary to correctly deliver a packet from the source P_0 to the destination P_k ,

$$E(k, \mathbf{R}_{\mathbf{P}}, \mathcal{Q}) = \mathbb{E}[N_{\mathbf{P}}], \tag{14}$$

where $N_{\rm P}$ represents the number of *primary* transmissions. The normalization of the primary energy (14) to the energy per packet $E_{\rm P}$ allows a more transparent performance comparison of different strategies across varying SNRs (and thus varying $E_{\rm P}$). The primary network selects one of the proposed routing strategies and corresponding parameters to be introduced in the next subsection, in order to provide different working points between the end-to-end throughput maximization and the primary energy minimization. This can be achieved through the definition of a cost function f(T, E), which is decreasing with increasing primary throughput T and increasing with increasing primary energy E. An example is

$$f(T, E) = -\mu_T T(k, \mathbf{R}_{\mathbf{P}}, \mathcal{Q}) + \mu_E E(k, \mathbf{R}_{\mathbf{P}}, \mathcal{Q}), \qquad (15)$$

where parameters $0 \le \mu_T$, $\mu_E \le 1$ set the relative importance between throughput and energy. In this case, if the primary network is interested in maximizing the throughput with little regard for energy expenditure, we should select $\mu_T \gg \mu_E$, and vice versa if the primary energy is at a premium. We stress that the primary network has full control of the secondary network in that it dictates how routing should be done, and hence which nodes should transmit and how. The only requirement that the secondary nodes put forth is that their QoS Q be satisfied in case they are selected as next hops.

A. Proposed Policies

We now detail the four proposed transmission policies for the primary packets. We remark that all four policies are based on a Type-I HARQ (extension to more complex forms of HARQ is left as future work). All policies are implemented using both receiver techniques introduced above (i.e., OM and SC). Note that the policy descriptions below apply identically to both cases, as they only differ at the PHY level. Finally, we recall that the proposed routing strategies leverage the concept of spectrum leasing via opportunistic routing, which enables the choice of secondary next hops in an adaptive way depending on the current channel conditions. This allows the primary network to accrue performance benefits in terms of f(T, E), while at the same time letting secondary nodes transmit as well.

1) Policy 1: only Primary (only-P): the only-P policy does not exploit spectrum leasing and is introduced here for reference. Only the primary nodes are involved in transmissions. A basic opportunistic routing strategy is assumed: in each block the transmitter is selected as the primary node that has decoded the previous transmission and is the closest to the destination. Since we assume Type-I HARQ, the current transmitter retransmits the packet until at least one of the downstream nodes has successfully decoded.

2) Policy 2: only Secondary (only-S): this policy aims at reducing to a minimum the primary transmissions, and thus $E(k, R_P, Q)$. So, in a somewhat dual fashion to only-P, the only-S policy forces the source to send the information only through secondary nodes (i.e., without exploiting any primary relay), allowing primary (re)transmissions only from P_0 . An opportunistic routing scheme is used on the secondary network, where transmission is granted to the secondary node that has decoded the previous transmission and that is the closest to the destination. So, only-S has the same topology of only-P, but a different exploitation of the relays, due to both a different portion of the shared resource used to transmit the primary packet (β or ψ) and a different transmission power.

3) Policy 3: Primary to Secondary (P-to-S): the only-S policy minimizes the primary transmissions thanks to spectrum leasing, but may suffer from a poor throughput as, once the packet has entered the secondary network, the multiuser diversity arising from the presence of primary nodes, and the higher primary transmission power, are not leveraged. The *P-to-S* policy, proposed here, and *P-and-S*, to be discussed below, attempt, to different extents, to offer a better trade-off between primary throughput and energy, managing the multiuser diversity via a parameter m.

Unlike *only-S*, the idea of *P-to-S* is to use primary relays unless a secondary node in a "sufficiently good" position, as dictated by m, has decoded. From that point on, the packet is handled by the secondary network as in *only-S*. Specifically, at each block in which a primary node is the transmitter, it first determines the type of relay closest to P_k that has successfully decoded. If the latter node is a secondary, it is selected for the next hop. If it is a primary, in order to save primary energy, the node is selected only if the best secondary node is at least m hops behind. That is, the next transmitter is selected as either the primary node at hand or the closest secondary node as long as the latter is within a window of m hops from the



Fig. 2. An illustration of the *P*-to-S policy. Spectrum leasing is performed from node P_1 to S_3 , with backward window with parameter m = 2.

position of the primary node toward P_0 . This window will be referred to as *backward window* and we generally have $0 \le m \le k - 2$. An example of the idea is illustrated in Fig. 2, where P_1 is the current transmitter, P_4 is the most advanced decoding node (a black cross indicates a node which has not successfully decoded) and m = 2. Before selecting node P_4 as the next hop, the primary network checks whether any secondary relay within the *backward window* $\{S_4, S_3, S_2\}$ has decoded the packet. If this is the case, the transmitter picks the most advanced such node as the next hop. In this particular example, relay S_3 is selected, because node S_4 has not successfully decoded.

4) Policy 4: Primary and Secondary (P-and-S): in the P-to-S policy, when a packet enters the secondary network, it cannot return to the primary one, except for the final destination P_k . This is again done in an attempt to save the primary energy, but limits the multiuser diversity and the resource available to the secondary transmitter, causing a throughput reduction. The proposed *P-and-S* policy removes this constraint to favor throughput maximization with respect to other policies that use spectrum leasing, provided that the secondary QoS constraints are not too strict (i.e., (d_S, R_S, ϵ_S) are chosen in order to enable the cooperation between primary and secondary users). The policy is again described by m.

Let us start with m = 0. The idea here is simply to select in each block the node that is the closest to the destination among those that have decoded, irrespective of whether such node is primary or secondary. This strategy clearly privileges primary throughput, since it exploits all the transmission opportunities afforded by the network. In order to obtain a more controllable trade-off between throughput and energy, we generalize this policy by letting m > 0 and operating as follows. Let m > 0. The policy extends *P-to-S* allowing transmissions from secondary back to primary relays, but with a constraint on the minimum progress given by the so called *forward window*. In particular, if the transmitter is a primary node, the strategy works as for the *P*-to-S policy. However, if the transmitter is a secondary, we enable the selection also of primary nodes, as long as the primary node to be selected is at least m hops ahead of the most advanced secondary decoding node. Thus, a primary relay can receive the packet from a secondary node only if it is outside the forward window, which is of size m hops³ and starts from the most advanced decoding secondary node towards the destination. In Fig. 3 we illustrate a possible scenario, where S_1 is the current transmitter and m = 2. Node S_2 is the most advanced secondary decoding relay, whereas

³In principle, one could choose two different sizes for the forward and backward windows, but this is not further investigated here.



Fig. 3. An illustration of the *P*-and-S policy. The transmission granted is handed between secondary relays S_1 and S_2 .

the best decoding primary relay is P_3 . However, nodes P_2 and P_3 cannot be selected for the next hop because they are inside the *forward window* (a gray dashed cross indicates this fact). Therefore, the *P*-and-S policy selects node S_2 in this case.

B. Evaluating Primary Throughput and Energy

In order to evaluate the performance metrics throughput (13) and average primary energy (14) for the protocols discussed above, we use the theory of Markov chains. We model the network with a chain of 2k states, one for each node. State P_0 refers to a situation where the current packet is at the source P_0 , the primary states P_i and secondary states S_i , $i = 1, \ldots, k-1$, are similarly defined, and P_k represents the state where the destination has successfully decoded. Recalling that we assume Type-I HARQ, the current transmitter retransmits the packet until at least one of the nodes admitted by the specific policy has successfully decoded. So, the transition matrix is organized in four blocks as

$$\Phi = \begin{bmatrix} \Phi_{P,P} & \Phi_{P,S} \\ \hline \Phi_{S,P} & \Phi_{S,S} \end{bmatrix},$$
(16)

where the states are ordered as $P_0, P_1, \ldots, P_k, S_1, \ldots, S_{k-1}$, and $\Phi_{A,B}$, $A, B \in \{P, S\}$ are the submatrices that collect all the transition probabilities from nodes of type A (Primary, P, or Secondary, S) to nodes of type B. In general, in matrix $\Phi_{A,B}$ the term $\Phi_{A,B}(i, j)$ represents the probability that, given the current state A_i (i.e., the transmitter is node A_i , with $i = 0, \ldots, k-1$ if A = P and $i = 1, \ldots, k-1$ if A = S), the next state is B_j , with $j = i, i+1, \ldots, k$ when B = P and $j = 1, \ldots, k-1$ if B = S.

In matrix (16), the first k states and the last k - 1 states are transient, whereas the k + 1-th state, corresponding to the packet being received at the destination, is absorbing. Depending on the routing policy adopted, the transition probabilities will assume different expressions and will be detailed in the Appendix. The average primary energy and throughput are derived as detailed in the Lemma below.

Lemma 1. The end-to-end throughput (13) and the primary energy (14) for fixed primary transmission rate \mathbb{R}_P are given by $T(k, \mathbb{R}_P, Q) = \mathbb{R}_P/v_{P_0}$ and $E(k, \mathbb{R}_P, Q) =$ w_{P_0} , where v_{P_0} and w_{P_0} are the first elements of vectors $\mathbf{v} = [\mathbf{v}_P, \mathbf{v}_S] = [v_{P_0}, \dots, v_{P_{k-1}}, v_{S_1}, \dots, v_{S_{k-1}}]$ and $\mathbf{w} =$ $[\mathbf{w}_P, \mathbf{w}_S] = [w_{P_0}, \dots, w_{P_{k-1}}, w_{S_1}, \dots, w_{S_{k-1}}]$, which are evaluated as $\mathbf{v} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}$ and $\mathbf{w} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{r}$, where $\mathbf{1}$ is a $(2k-1) \times 1$ vector with all entries equal to 1 and \mathbf{r} is the reward vector $\mathbf{r} = [\mathbf{r}_P, \mathbf{r}_S] = [r_{P_0}, \dots, r_{P_{k-1}}, r_{S_1}, \dots, r_{S_{k-1}}]$ where \mathbf{r}_P is a $k \times 1$ vector with all ones, \mathbf{r}_S is a $(k-1) \times 1$ vector with all zero elements and \mathbf{I} is the $(2k-1) \times (2k-1)$ *identity matrix. Finally, matrix* \mathbf{Q} *is obtained from* $\boldsymbol{\Phi}$ *by removing the* (k + 1)*-th row and the* (k + 1)*-th column.*

Proof: The lemma follows from Markov chain theory [19, Ch. 3]. Specifically, both equations $\mathbf{v} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}$ and $\mathbf{w} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{r}$ follow from the standard first-step analysis [19] and represent the matrix formulation of the recursive equations $v_{A_i} = 1 + \sum_{B_j \neq P_k} \Phi_{A,B}(i,j)v_{B_j}$ and $w_{A_i} = r_{A_i} + \sum_{B_j \neq P_k} \Phi_{A,B}(i,j)v_{B_j}$, respectively, with $A_i \neq P_k$, $A, B \in \{P, S\}$ and $r_{P_i} = 1$ or $r_{S_i} = 0$.

C. Optimality of the SC approach

In [20] the authors consider the outage capacity of a twouser quasi-static fading broadcast channel when the transmitter has no information about the instantaneous state of the channel. It is proved that SC with Gaussian codewords is optimum. The following Proposition uses this result to prove that, in our system, the considered SC scheme is optimal for both throughput and primary energy consumption, that is, it is the best among all possible multiplexing schemes to be employed by the secondary nodes.

Proposition 1 (Optimality of SC). Fix primary rate \mathbb{R}_P , secondary QoS requirements $(d_S, \mathbb{R}_S, \epsilon_S)$ and any of the proposed routing strategies. The following holds: (i) Any throughput $T(k, \mathbb{R}_P, \mathcal{Q})$ that can be attained by any multiplexing scheme of primary and secondary codewords (not necessarily randomly generated according to a Gaussian codebook) at the secondary nodes can also be achieved by the SC scheme with Gaussian codewords studied in Section II-E; (ii) The primary energy $E(k, \mathbb{R}_P, \mathcal{Q})$ used by the SC scheme is no larger than the amount of energy expended by any other multiplexing scheme.

Proof: Consider any multiplexing scheme secondary node S_i . In our model, at the let $\mathcal{M} \subset \{P_1, \ldots, P_k, S_1, \ldots, S_{k-1}\}$ denote the set with cardinality $|\mathcal{M}| = M$ of primary and secondary receivers that, depending on the specific routing strategy, try to decode the primary message and let $N^{(2)}$ be the intended receiver of the secondary message. Fix the outage probability of such scheme at all nodes. We want to show that for the given outage probabilities (and the given transmission powers), SC is able to support rates (R_P, R_S) for the primary and secondary packets, respectively, as large as any other multiplexing scheme at secondary node S_i . This would prove that SC is also throughput and primary energy optimal. In fact, given any transmission rates (R_P, R_S) , recalling that the outage probability is non-decreasing in the transmission rate, the above would imply that the outage probabilities at all nodes with SC are always as low as with any other scheme. Therefore, primary throughput and energy are not degraded (and potentially improved) with SC and so is the secondary QoS.

We now prove that SC achieves rates as large as any other scheme at S_i for given outage probabilities. We start by observing that, since the probability of outage at each receiver only depends on the corresponding fading channel coefficient, there are M thresholds $k_{N^{(1)}}$ such that if and only if $|h_{S_iN^{(1)}}|^2 < k_{N^{(1)}}$ and for $N^{(1)} \in \mathcal{M}$, transmission from



Fig. 4. Primary energy and end-to-end throughput as a function of the primary transmission rate R_P for a network with the same number of primary and secondary relays (k - 1) (full secondary deployment) for k = 12 hops, SNR $\gamma = -3$ dB, m = 1 and R_S = 1 bits/s/Hz.

 S_i to $N^{(1)}$ is in outage, and there exists a threshold k_S such that if and only if $|h_{S_iN^{(2)}}|^2 < k_S$, the secondary packet is received in outage [20]. Moreover, it is clear that if decoding is successful at the threshold channel values, it should also be successful for larger channel values. So, fix the channel gains to be equal to their respective thresholds. Rates (R_P, R_S) will be correctly decoded if and only if they are inside the capacity region of the Gaussian (non-fading) broadcast channel with these channel coefficients. In particular, one can focus on the two-user broadcast channel formed by the secondary receiver and the primary channel in \mathcal{M} with the worst overall channel (since decoding of the primary packet at this node also implies decoding at the better nodes). The desired result follows from the fact that Gaussian SC can achieve any rate pair in the capacity region of the Gaussian broadcast channel [16].

IV. NUMERICAL RESULTS

In this section we first provide some numerical evidence about the superiority of SC over OM, which was proved in Proposition 1. Then, focusing on the SC scheme, we elaborate on the advantages of spectrum leasing and on the design of the proposed schemes. Throughout this section we fix the following parameters: number of hops k = 12, path loss $\eta = 3$, geometry of the network $\Delta_{\rm V} = \Delta_{\rm H} = 1/k$ and transmit power of secondary users $E_{\rm S} = E_{\rm P}$ (then also the received SNRs are equal, $\gamma_{\rm P} = \gamma_{\rm S} = \gamma$, see Section II-A).⁴ We consider two secondary deployments: (*i*) Full ($\alpha = 1$) and (*ii*) Partial ($\alpha > 1$). As for the secondary QoS requirements Q, we assume that each secondary node wants to transmit its own traffic at rate $R_{\rm S}$ to a node with SNR equal to γ at distance $d_{\rm S} = 1/10$ with outage probability $\epsilon_{\rm S} = 0.1$.

A comment on the calculation of the secondary transmission parameters based on the QoS requirements Q is in order. For OM, from (5), the spectral resources fraction β is equal to 0.83, for $R_S = 1$ bits/s/Hz. As for SC, the value of ψ that satisfies the secondary QoS constraint $P_{out,SS}^{(sC)}(d_S) = \epsilon_S$ in



Fig. 5. End-to-end throughput and overall primary energy as a function of the primary transmission rate R_P for a network with the same number of primary and secondary relays (k - 1) (full secondary deployment) for k = 12 hops, SNR $\gamma = -3$ dB, m = 1 and $R_S = 1$ bits/s/Hz. Each line is obtained by varying the primary transmission rate R_P as {1.8, 2.4, 2.7, 2.9, 3.2, 3.4, 3.6, 3.9, 4.1, 4.8, 5.2} bits/s/Hz.

(10) is not necessarily unique, but depends on the particular decoder employed, namely treating interference as noise $(\mathcal{H}_{S}^{(1)})$ in (12)) or adopting a successive decoding scheme $(\mathcal{H}_{S}^{(2)})$ in (12)). We select the solution that maximizes the primary rate, i.e., the highest feasible ψ .

1) Comparison of SC and OM: Figs. 4 and 5 evaluate the gain of the SC scheme over OM for P-to-S and P-and-S policies, by varying the primary transmission rate R_P in a full secondary deployment for $\gamma = -3$ dB, m = 1 and $R_s = 1$ bits/s/Hz. For the rest of this section, the normalized primary energy $E(k, \mathbf{R}_{\mathbf{P}}, \mathcal{Q})$ defined in (14) is expressed in dB, i.e. $10 \log_{10} E(k, R_P, Q)$. Fig. 4 confirms the optimality of the SC scheme for both P-to-S and P-and-S, regardless of R_P. In fact, as already proved in Proposition 1, the better outage provided by SC results in a better throughput and in a primary energy saving. It is also seen that, for each policy, there exists a different rate that maximizes the throughput. In order to reduce the primary energy consumption, at the cost of a reduced throughput, one can decrease the transmission rate R_P for all policies except *P-to-S*. In fact, only a higher R_P causes a narrower coverage range of the primary transmission so that it is more likely that a primary packet enters the secondary network due to the backward window. In fact, the latter forces the network to choose a secondary node as the next hop when the distance between the transmitter and the best primary relay is comparable with the distance between the transmitter and the best secondary relay.⁵ Thus, when R_P decreases, the coverage area is larger and it is more likely to find a primary node that satisfies the backward window (especially when m is low, as in the case of Figs. 4 and 5), even if there are a good number of secondary users that have correctly received the primary packet. This behavior is also present in *P*-and-S, but is well balanced by the forward window, which gives to the primary packet the possibility to return

⁴We keep these values fixed to better focus on what we believe are the main issues of the proposed scenario.

⁵With "best" we mean the relay, primary or secondary, that is closest to the destination among those that correctly received the primary packet.



Fig. 6. End-to-end throughput and overall primary energy plotted varying the SNR, γ , for a network with full secondary deployment, for k = 12 hops, primary transmission rate $R_P = 2.9$ bits/s/Hz, m = 1 and $R_S = 1$ bits/s/Hz. Each line is obtained by varying γ as $\{-20, -15, \pm 10, \pm 8, \pm 5, \pm 3, 0\}$ dB.



Fig. 7. End-to-end throughput and overall primary energy plotted varying the SNR, γ , for a network with partial secondary deployment for $\alpha = 3$, k = 12 hops, primary transmission rate $R_P = 2.9$ bits/s/Hz, m = 1, $R_S = 1$ bits/s/Hz and γ that assumes the following values $\{-20, -15, \pm 10, \pm 8, \pm 5, \pm 3, 0\}$ dB.

to the primary network, increasing the end-to-end throughput at the price of a higher primary energy consumption. It is finally noted that the energy gain of spectrum leasing over only-P (no spectrum leasing) is substantial irrespective of the choice of R_P. Moreover, spectrum leasing combined with the SC scheme outperforms only-P in terms of throughput for $R_P \leq 4.1$ bits/s/Hz. In addition, Fig. 5 clearly shows the trade-offs available between primary throughput and energy, which can be formalized through a cost function f(T, E) (e.g., see Eq. (15)). Each curve is obtained by evaluating the pair end-to-end throughput T and primary energy E of a given scheme (i.e., SC or OM) and policy for different values of R_P , while keeping all the other parameters fixed. Therefore, the best heuristic policy (among those described in Section III-A) selected by the primary network for a given priority between throughput and energy (i.e., fixing μ_T and μ_E) is the one that minimizes the cost function f(T, E), for any pair (T, E).

2) Design and Advantages of Spectrum Leasing: In the rest of this section we numerically evaluate the impact of secondary relays on the primary network only, focusing on the optimal scheme SC to numerically study the spectrum leasing features. We first study the trade-off between end-toend throughput and the overall primary energy consumption defined in (14) as a function of the SNR γ for two different secondary node deployments (full in Fig. 6 and partial, with $\alpha = 3$, in Fig. 7), fixing the remaining protocol parameters to m = 1, $R_P = 2.9$ bits/s/Hz and $R_S = 1$ bits/s/Hz. Each curve is obtained by evaluating the pair end-to-end throughput and normalized primary energy of a given policy for different γ (ranging from -20 dB to 10 dB), thus different ψ , and keeping all the other parameters fixed. In these figures (and also Fig. 8) we use the same approach as in Fig. 5 to better highlight the available trade-off between primary throughput and energy, by considering the minimization of a cost function f(T, E) (e.g., see Eq. (15)) performed by the primary network to select the best routing strategy. From Fig. 6, it is seen that, with the given parameters, spectrum leasing policies with full

secondary deployment are more energy efficient than only-P, especially as the SNR decreases, due to the larger benefits afforded by opportunistic routing. However, the throughput may not always be better than only-P, since with spectrum leasing more transmissions may be necessary to deliver a primary packet when the secondary QoS requirements are sufficiently strict. In such cases, if throughput is the main performance criterion of interest, then spectrum leasing should not be used. Such gains, while still substantial, decrease with a partial secondary deployment as shown in Fig. 7 for $\alpha = 3$. In both cases, however, when γ is low, the secondary QoS requirements are satisfied only with ψ close to 0, where the throughput of the spectrum leasing policies is almost 0, due to the low power assigned to the primary packet. Moreover, when the SNR decreases, the throughput of only-S and P-to-S is affected by the partial deployment due to the longer secondary hops. In this case, P-and-S is to be preferred as it is able to keep the same level of throughput of Fig. 6 (though with larger primary energy). Better performance can be obtained by optimizing the *window* size m, as discussed next.

Fig. 8 shows end-to-end throughput and primary energy by varying m for full and partial secondary deployment with $\alpha = 4$ and for parameters $\gamma = -3$ dB, $R_P = 3.4$ bits/s/Hz and $R_S = 1$ bits/s/Hz. Similar to the discussion above, *P-and-S* outperforms *only-S* and *P-to-S* from a throughput point of view, especially in the partial secondary deployment scenario.⁶ Moreover, it is clear that m allows to trade off energy and throughput. For *P-to-S* and *P-and-S*, increasing m ($m \ge 4$) trades throughput for a decreased primary energy consumption, due to the larger number of secondary transmissions admitted. When m is sufficiently low ($m \le 4$), the throughput increases differently in *P-to-S* and *P-and-S*. For *P-to-S*, which employs only the *backward window* and blocks secondary transmissions to primary relays, the

⁶In Fig. 8, the performance of the *only-S* policy in the full secondary relay scenario is very close to that of *P-to-S* for high m, and is not visible in the graph. However, this behavior confirms the strict relationship between these two policies, especially for high m.



Fig. 8. End-to-end throughput and overall primary energy shown varying m for a network with full and partial secondary relay deployment for $\alpha = 4$, k = 12 hops, transmission rate $R_P = 3.4$ bits/s/Hz, $\gamma = -3$ dB and $R_S = 1$ bits/s/Hz. The lines are obtained by varying m from 0 to 10.



Fig. 9. Primary energy and end-to-end throughput as a function of R_S for a network with full secondary deployment, k = 12 hops, transmission rate $R_P = 2.9$ bits/s/Hz, $\gamma = -3$ dB and m = 1.

throughput and primary energy are larger due to the lower number of secondary nodes available to lease the spectrum. In *P-and-S* this limit is overcome by removing the block from the secondary transmissions and by introducing the *forward window*. Thus, due to the capability of exploiting more path diversity, *P-and-S* is able to obtain larger throughput (for larger energy consumption) than *P-to-S*.

Finally, in Figs. 9 and 10 we consider the impact of R_S on the four policies for the SC scheme, for $\gamma = -3$ dB, $R_P = 2.9$ bits/s/Hz and m = 1 and for full and partial secondary deployment with $\alpha = 4$, respectively. We note in Fig. 9 that increasing the secondary QoS requirements (i.e., increasing R_S) leads to a decreased throughput without affecting the primary energy for all policies, except *P*-and-*S*. Indeed, in all policies except *P*-and-*S*, modifying R_S does not change the number of primary transmissions, but only the portion of the spectrum leased to the secondary node that is used to serve primary traffic. Instead, for *P*-and-*S*, a higher R_S leads to both a decreased throughput and an increased primary energy, due to the larger number of secondary transmissions towards the primary network. In fact, the number of relays



Fig. 10. Primary energy and end-to-end throughput as a function of R_S for a network with partial secondary deployment for $\alpha = 4$, k = 12 hops, transmission rate $R_P = 2.9$ bits/s/Hz, $\gamma = -3$ dB and m = 1.

that are potentially reachable at each transmission from a secondary relay increases with R_S . So, when R_S is high, the number of secondary transmissions increases, and therefore the possibility of returning to the primary network increases. If this happens, the next hop will be covered by a primary transmission, which affects the primary energy expenditure. Moreover, with partial secondary deployment for $\alpha = 4$ (see Fig. 10), *P*-and-S confirms to be able to best adapt to the lack of secondary nodes, though at the price of an increased primary energy consumption.

V. CONCLUDING REMARKS

This paper has proposed a novel approach to regulate the coexistence of primary and secondary nodes in multihop networks based on spectrum leasing and opportunistic routing. In particular, it is proposed that primary nodes may, in a local and dynamic fashion, select secondary nodes as next hops for primary traffic by allowing the latter to exploit the spectral resources for secondary data with QoS guarantees. This approach is an implementation of the previously proposed idea of spectrum leasing via cooperation. We have designed different routing strategies based on this principle that provide different trade-offs between gains in terms of primary throughput and energy. Moreover, we have shown that secondary nodes can optimally multiplex primary and secondary traffic using superposition coding. Our results demonstrate the effectiveness of the proposed paradigm. Future work should evaluate the performance of the proposed protocols in networks of arbitrary topology, where it is expected that the performance gains of the proposed spectrum leasing approach may be even more significant due to the generally larger number of secondary nodes at comparable distance that the primary can choose from.

APPENDIX

In this Appendix, we define the transition probabilities of the policies described in Section III needed to calculate the matrix (16). In order to keep the expressions simple, we will use the following notation: (1) $P_{\text{out},\text{TP}}(a) = P_{\text{out},\text{TS}}(a) =$ $P_{\text{out},\text{P}}(\Delta_a^{\text{(TD)}})$ for transmissions from source (T) to primary (P) or secondary (S) relays; (2) $P_{out,TD}(k) = P_{out,P}(k\Delta_{H})$ for transmissions from source (T) to destination (D); (3) $P_{out,PP}(a) = P_{out,P}(a\Delta_{H}), P_{out,SS}(a) = P_{out,SP}^{(A)}(a\Delta_{H})$ for transmissions between primary relays (PP) and between secondary relays (SS), with $A \in \{OM, SC\}$ as in (3) and (6); (4) $P_{out,PS}(a) = P_{out,P}(\Delta_{a}^{(R)}), P_{out,SP}(a) = P_{out,SP}^{(A)}(\Delta_{a}^{(R)})$ for transmissions between the two sets of relays, primary to secondary (PS) and vice versa (SP), with $A \in \{OM, SC\}$; (5) $P_{out,PD}(a) = P_{out,P}(\Delta_{a}^{(T,D)}), P_{out,SD}(a) = P_{out,SP}^{(A)}(\Delta_{a}^{(T,D)})$ for primary (P) or secondary (S) transmissions to destination (D), with $A \in \{OM, SC\}$.

1) only-P: The only non-zero submatrix in only-P (no spectrum leasing) is $\Phi_{P,P}$, that describes the transition probabilities between primary nodes. We have:

$$\mathbf{\Phi}_{\mathbf{P},\mathbf{P}} = \begin{bmatrix} \Phi_{\mathbf{P},\mathbf{P}}(0,0) & \dots & \Phi_{\mathbf{P},\mathbf{P}}(0,k) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\mathbf{P},\mathbf{P}}(k,k) \end{bmatrix};$$
(17)

$$\begin{split} \Phi_{\mathrm{P,P}}(0,0) &= \prod_{\ell=1}^{k-1} P_{\mathrm{out,TP}}(\ell) P_{\mathrm{out,TD}}(k); \\ \Phi_{\mathrm{P,P}}(0,k) &= 1 - P_{\mathrm{out,TD}}(k); \\ \Phi_{\mathrm{P,P}}(0,j) &= (1 - P_{\mathrm{out,TP}}(j)) \prod_{\ell=j+1}^{k-1} P_{\mathrm{out,TP}}(\ell) P_{\mathrm{out,TD}}(k), \\ & j = 1, \dots, k-1; \\ \Phi_{\mathrm{P,P}}(i,j) &= \prod_{\ell=j+1}^{k-1} P_{\mathrm{out,PP}}(\ell-i) P_{\mathrm{out,PD}}(k-i) \times \\ &\times (1 - P_{\mathrm{out,PP}}(j-i)), i = 1, \dots, k-1, j = i, \dots, k-1; \\ \Phi_{\mathrm{P,P}}(i,k) &= 1 - P_{\mathrm{out,PD}}(k-i), \quad i = 1, \dots, k-1; \\ \Phi_{\mathrm{P,P}}(k,k) &= 1; \quad \Phi_{\mathrm{P,P}}(i,j) = 0, \quad \text{otherwise.} \end{split}$$

The other submatrices are zero, i.e., $\Phi_{P,S} = \mathbf{0}_{[k+1,k-1]}$, $\Phi_{S,P} = \mathbf{0}_{[k-1,k+1]}$ and $\Phi_{S,S} = \mathbf{0}_{[k-1,k-1]}$, where $\mathbf{0}_{[c,d]}$ is a zero matrix with *c* rows and *d* columns.

2) only-S: In only-S the only primary transmissions allowed are from the source, which leads to submatrices $\Phi_{P,P}$ and $\Phi_{P,S}$:

$$\boldsymbol{\Phi}_{\mathrm{P},\mathrm{P}} = \begin{bmatrix} \Phi_{\mathrm{P},\mathrm{P}}(0,0) & 0 & \dots & 0 & \Phi_{\mathrm{P},\mathrm{P}}(0,k) \\ \hline 0 & 0_{[k-1,k+1]} & \\ \hline 0 & \dots & 0 & \Phi_{\mathrm{P},\mathrm{P}}(k,k) \end{bmatrix}; \quad (18)$$

$$\Phi_{P,P}(0,0) = P_{\text{out,TD}}(k) \prod_{q=1}^{k-1} P_{\text{out,TS}}(q);$$

$$\Phi_{P,P}(0,k) = 1 - P_{\text{out,TD}}(k); \quad \Phi_{P,P}(k,k) = 1.$$

$$\Phi_{P,S} = \left[\frac{\Phi_{P,S}(0,1) \dots \Phi_{P,S}(0,k-1)}{\mathbf{0}_{[k,k-1]}} \right]; \quad (19)$$

$$\Phi_{P,S}(0,j) = P_{\text{out,TD}}(k) \left(1 - P_{\text{out,TS}}(j)\right) \prod_{q=j+1}^{k-1} P_{\text{out,TS}}(q), \\ j = 1, \dots, k-1.$$

Finally, submatrices $\Phi_{S,P}$ and $\Phi_{S,S}$ reflect the fact that secondary transmissions can reach only other secondary relays or the destination:

$$\mathbf{\Phi}_{\mathbf{S},\mathbf{P}} = \begin{bmatrix} \mathbf{0}_{[k-1,k]} & \vdots \\ \mathbf{0}_{\mathbf{F},\mathbf{P}}(k-1,k) & \vdots \end{bmatrix}; \quad (20)$$

$$\Phi_{S,P}(i,k) = 1 - P_{\text{out},SD}(k-i), \quad i = 1, \dots, k-1.$$

$$\Phi_{S,S} = \begin{bmatrix} \Phi_{S,S}(1,1) & \dots & \Phi_{S,S}(1,k-1) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{S,S}(k-1,k-1) \end{bmatrix}; \quad (21)$$

 $\Phi_{\text{S,S}}(i,j) = (1 - P_{\text{out,SS}}(j-i)) P_{\text{out,SD}}(k-i) \times \\ \times \prod_{q=j+1}^{k-1} P_{\text{out,SS}}(q-i), i = 1, \dots, k-1, j = i, \dots, k-1.$

 $\Phi_{A,B}(i,j) = 0$ with $A, B \in \{P, S\}$, in all other cases.

3) *P-to-S:* Submatrix $\Phi_{P,P}$ assumes the same structure of the *only-P* (no spectrum leasing) policy, but the transition probabilities have to consider the presence of the unlicensed network. We have:

$$\mathbf{\Phi}_{\mathbf{P},\mathbf{P}} = \begin{bmatrix} \Phi_{\mathbf{P}\mathbf{P}}(0,0) & \dots & \Phi_{\mathbf{P}\mathbf{P}}(0,k) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\mathbf{P}\mathbf{P}}(k,k) \end{bmatrix};$$
(22)

$$\begin{split} & \Phi_{\text{P,P}}(0,0) = \prod_{\ell=1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k) \prod_{q=1}^{k-1} P_{\text{out,TS}}(q); \\ & \Phi_{\text{P,P}}(0,k) = 1 - P_{\text{out,TD}}(k); \\ & \Phi_{\text{P,P}}(0,j) = (1 - P_{\text{out,TP}}(j)) \prod_{\ell=j+1}^{k-1} P_{\text{out,TP}}(\ell) P_{\text{out,TD}}(k) \times \\ & \times \prod_{q=j-m}^{k-1} \left[1 + \mathbf{1}_{\{q>0\}} \left(P_{\text{out,TS}}(q) - 1 \right) \right], j = 1, \dots, k-1; \\ & \Phi_{\text{P,P}}(i,j) = \prod_{\ell=j+1}^{k-1} P_{\text{out,PP}}(\ell-i) P_{\text{out,PD}}(k-i) \times \\ & \times \prod_{q=j-m}^{k-1} \left[1 + \mathbf{1}_{\{q>0\}} \left(P_{\text{out,PS}}(|q-i|) - 1 \right) \right] \times \\ & \times (1 - P_{\text{out,PP}}(j-i)), i = 1, \dots, k-1; j = i, \dots, k-1; \\ & \Phi_{\text{P,P}}(i,k) = 1 - P_{\text{out,PD}}(k-i), i = 1, \dots, k-1; \\ & \Phi_{\text{P,P}}(k,k) = 1. \end{split}$$

In the following submatrix the effect of spectrum leasing and of the *backward window* with parameter m are taken into account to express the transition between primary and secondary relays:

$$\mathbf{\Phi}_{\rm P,S} = \begin{bmatrix} \Phi_{\rm PS}(0,1) & \dots & \Phi_{\rm PS}(0,k-1) \\ \vdots & & \vdots \\ \Phi_{\rm P,S}(k-1,1) & \dots & \Phi_{\rm P,S}(k-1,k-1) \\ 0 & \dots & 0 \end{bmatrix};$$
(23)

$$\begin{split} \Phi_{\text{P,S}}(0,j) &= (1 - P_{\text{out,TS}}(j)) \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,TP}}(\ell) \times \\ &\times P_{\text{out,TD}}(k) \prod_{q=j+1}^{k-1} P_{\text{out,TS}}(q), j = 1, \dots, k-1; \\ \Phi_{\text{P,S}}(i,j) &= (1 - P_{\text{out,PS}}(j-i)) \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,PP}}(\ell-i) \times \\ &\times P_{\text{out,PD}}(k-i) \prod_{q=j+1}^{k-1} P_{\text{out,PS}}(q-i), \\ &i = 1, \dots, k-1, j = i, \dots, k-1; \\ \Phi_{\text{P,S}}(i,j) &= \mathbf{1}_{\{(i-j) \leq m\}} (1 - P_{\text{out,PS}}(i-j)) P_{\text{out,PD}}(k-i) \times \\ &\times \prod_{\ell=(j+m)+1}^{k-1} P_{\text{out,PP}}(\ell-i) \prod_{q=j+1}^{k-1} P_{\text{out,PS}}(|i-q|), \\ &i = 2, \dots, k-1, j = 1, \dots, i-1. \end{split}$$

and $\Phi_{A,B}(i, j) = 0$ with $A, B \in \{P, S\}$ in all other cases. $\Phi_{S,P}$ and $\Phi_{S,S}$ are equal to the submatrices of the *only-S* policy in (20) and (21), respectively.

4) *P-and-S:* Submatrices $\Phi_{P,P}$ and $\Phi_{P,S}$ are equal to those of the *P-to-S* policy, so they are not reported. However, in this policy the behavior of the secondary network is different, so $\Phi_{S,P}$ and $\Phi_{S,S}$ are derived considering the presence of primary relays, as limited by the *forward window*. We have:

$$\mathbf{\Phi}_{\mathbf{S},\mathbf{P}} = \begin{bmatrix} 0 & 0 & \Phi_{\mathbf{S},\mathbf{P}}(1,2) & \dots & \Phi_{\mathbf{S},\mathbf{P}}(1,k) \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \Phi_{\mathbf{S},\mathbf{P}}(k-1,k) \end{bmatrix}; \quad (24)$$

$$\begin{split} \Phi_{\text{S,P}}(i,j) &= \mathbf{1}_{\{m \leq (j-i)\}} \left(1 - P_{\text{out,SP}}(j-i)\right) P_{\text{out,SD}}(k-i) \times \\ &\times \prod_{\ell=j+1}^{k-1} P_{\text{out,SP}}(\ell-i) \prod_{q=j}^{k-1} P_{\text{out,SS}}(q-i) \times \\ &\times \left[1 + \mathbf{1}_{\{m \geq 2\}} \left(-1 + \prod_{t=1}^{m-1} P_{\text{out,SS}}(j-t-i)\right)\right], \\ &\quad i = 1, \dots, k-2, j = i+1, \dots, k-1; \\ \Phi_{\text{S,P}}(i,k) &= 1 - P_{\text{out,SD}}(k-i), \quad i = 1, \dots, k-1. \\ &\quad \Phi_{\text{S,S}} = \begin{bmatrix} \Phi_{\text{S,S}}(1,1) & \cdots & \Phi_{\text{S,S}}(1,k-1) \\ 0 & \ddots & \vdots \\ 0 & 0 & \Phi_{\text{S,S}}(k-1,k-1) \end{bmatrix}; \quad (25) \\ &\quad \Phi_{\text{S,S}}(i,j) = (1 - P_{\text{out,SS}}(j-i)) P_{\text{out,SD}}(k-i) \times \\ &\times \prod_{\ell=m+j}^{k-1} P_{\text{out,SP}}(\ell-i) \prod_{q=j+1}^{k-1} P_{\text{out,SS}}(q-i), \end{split}$$

 $i = 1, \dots, k-1, \quad j = i, \dots, k-1.$

 $\Phi_{S,B}(i,j) = 0$ with $B \in \{P, S\}$, in all other cases.

5) Partial secondary deployment: The previous expressions can be derived also for a secondary deployment with $\alpha > 1$. In this case matrix Φ (16) and vectors v, 1, w and r in Lemma 1 have to be reduced in accordance to the number of active secondary relays. This change has effects on the calculation of the transition probabilities derived above. In particular, it is sufficient to set the values of $P_{\text{out},\text{TS}}(a)$ and $P_{\text{out},\text{SS}}(a)$ equal to 1 if the remainder of the integer division $\text{rem}(a, \alpha) \neq 0$, otherwise they remain unchanged. For transmission from primary (P_i) to secondary relay (S_j) , the probability $P_{\text{out},\text{PS}}(|j-i|)$ has to be set to 1 only if $\text{rem}(j, \alpha) \neq 0$. The other probabilities remain unchanged.

REFERENCES

- Spectrum Efficiency Working Group, "Report of the Spectrum Efficiency Working Group," FCC, Tech. Rep., Nov. 2002.
- [2] J. M. Peha, "Approaches to spectrum sharing," IEEE Commun. Mag., vol. 43, no. 2, pp. 10–12, Feb. 2005.
- [3] J. Acharya and R. Yates, "Dynamic spectrum allocation for uplink users with heterogeneous utilities," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1405–1413, Mar. 2009.
- [4] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, Jan. 2008.
- [5] I. Stanojev, O. Simeone, U. Spagnolini, Y. Bar-Ness, and R. Pickholtz, "Cooperative ARQ via auction-based spectrum leasing," *IEEE Trans. Commun.*, vol. 58, no. 6, June 2010.
- [6] M. Zorzi and R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Trans. Mobile Comput.*, vol. 2, no. 4, pp. 337–348, Oct. 2003.
- [7] S. Biswas and R. Morris, "Opportunistic routing in multi-hop wireless networks," ACM SIGCOMM Comput. Commun. Rev., vol. 34, no. 1, pp. 69–74, Jan. 2004.
- [8] D. Chiarotto, O. Simeone, and M. Zorzi, "Throughput and energy efficiency of opportunistic routing with type-I HARQ in linear multihop networks," in *Proc. IEEE GlobeCom*, Dec. 2010.
- [9] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. E. Fuja, "Bandwidth-and power-efficient routing in linear wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2624–2633, June 2006.
- [10] O. Oyman and S. Sandhu, "A Shannon-Theoretic perspective on fading multihop networks," in *Proc. CISS*, Mar. 2006, pp. 525–530.
- [11] A. Bader and E. Ekici, "Performance optimization of interferencelimited multihop networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 5, pp. 1147–1160, Oct. 2008.
- [12] M. Cesana, F. Cuomo, and E. Ekici, "Routing in cognitive radio networks: challenges and solutions," Ad Hoc Netw. (2010), doi:10.1016/j.adhoc.2010.06.009.
- [13] A. Abbagnale and F. Cuomo, "Gymkhana: a connectivity-based routing scheme for cognitive radio ad hoc networks," in *Proc. IEEE INFOCOM*, Mar. 2010.
- [14] L. Ding, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Crosslayer routing and dynamic spectrum allocation in cognitive radio ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1969–1979, May 2010.
- [15] M. Naghshvar and T. Javidi, "Opportunistic routing with congestion diversity and tunable overhead," in *Proc. ISCCSP*, Mar. 2010.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edition. John Wiley & Sons, Inc., 2006.
- [17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
 [18] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols
- [18] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, July 2001.

- [19] H. Taylor and S. Karlin, An Introduction to Stochastic Modeling, 3rd edition. Academic Press, 1998.
- [20] M. Uppal, A. Høst-Madsen, and Z. Xiong, "Outage capacity of the broadcast channel in the low power regime," in *Proc. IEEE ISIT*, June 2010, pp. 609–613.



Davide Chiarotto (S'09) received the M.Sc. degree (with honors) and the Ph.D. degree in Telecommunications Engineering from University of Padova, Italy, in 2007 and 2011, respectively. During the first half of 2010 he was on leave at the Center for Wireless Communications and Signal Processing Research (CWCSPR), at the New Jersey Institute of Technology (NJIT), Newark, New Jersey, working on information-theoretic aspects of cognitive radio techniques in ad hoc networks. His current research interests include cognitive radio, cooperative com-

munications, cellular networks, resource management, cross-layer protocol design for wireless networks and PHY-aware MAC protocols, with a particular focus on MIMO ad hoc networks.



Osvaldo Simeone (M'02) received the M.Sc. degree (with honors) and the Ph.D. degree in information engineering from the 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 Assistant Professor. His current research interests concern the cross-layer analysis and design of wireless networks with emphasis on informationtheoretic, signal processing, and queuing aspects.

Specific topics of interest are: cooperative communications, unlicensed spectrum access, ad hoc, sensor, mesh and hybrid networks, distributed estimation, and synchronization.

Dr. Simeone is the co-recipient of the best paper award of the IEEE SPAWC 2007 and the IEEE WRECOM 2007. He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2007–2010 and he currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.



Michele Zorzi (F'07) was born in Venice, Italy, in 1966. He received his Laurea degree and Ph.D. in electrical engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During academic year 1992/93, he was on leave at the University of California, San Diego (UCSD) attending graduate courses and doing research 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, and in 2003 joined the Department of Information Engineering of the University of Padova, Italy, where he is currently a Professor. 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, cognitive networks, and underwater communications and networking.

He was Editor-in-Chief of *IEEE Wireless Communications* from 2003 to 2005, is currently Editor-in-Chief of IEEE TRANSACTIONS ON COMMUNI-CATIONS, and serves on the Editorial Board of the *Wiley Journal of Wireless Communications and Mobile Computing*. He was also guest editor for special issues in *IEEE Personal Communications* and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is a Member-at-Large of the Board of Governors of the IEEE Communications Society.