An Amplify-and-Forward Scheme for Spectrum Sharing in Cognitive Radio Channels

Francesco Verde, *Senior Member, IEEE*, Anna Scaglione, *Fellow, IEEE*,
Donatella Darsena, *Member, IEEE*, and Giacinto Gelli

***Abstract*—In this paper, we propose a cognitive radio scheme that allows a secondary user (SU) to transmit over the same time-frequency slot of a primary user (PU), even when the PU is active. In our scheme, the SU amplifies and forwards the signal of the PU, by using as scaling factor the value of its information symbol to be transmitted towards the secondary receiver. The information-theoretic limits of the proposed protocol are investigated in terms of ergodic channel capacities of both the PU and SU links. It is shown that: 1) under certain operating conditions, the SU can superimpose its information symbols on the PU signal, without violating the cognitive radio principle of protecting the PU transmission; and 2) when the primary link is busy, the SU offers the PU its own transmitting power in exchange for a low-capacity communication channel, which improves the packet delay performance of the SU. In this barter, the tempting incentive for the PU consists of a noticeable improvement of its achievable rate at the price of a slight increase in the computational complexity of the primary receiver.**

***Index Terms*—Amplify-and-forward relaying, cognitive radio, ergodic channel capacity, wireless communications.**

I. Introduction

MEASUREMENT studies [1] have recently confirmed that the licensed radio spectrum is relatively underutilized: as a consequence, numerous *cognitive radio schemes* [2] have been proposed, wherein secondary users (SUs) can temporarily share a portion of licensed spectrum, provided that they generate a minimal amount of interference to the licensed primary users (PUs). Common spectrum sharing strategies [3] belong to two categories: (i) the SUs are allowed to transmit also when the PUs are transmitting; (ii) the SUs use the licensed spectrum only when the PUs are not transmitting. In this paper, we consider a dynamic spectrum sharing scheme belonging to the first family.

Most of the literature on spectrum sharing for cognitive radios (see, e.g., [4]–[8]) relies on the idea of minimizing the interference caused by the SU. In our scheme, instead, we allow the SU to superimpose its transmission to the PU one, albeit in a "symbiotic" form, in order to possibly improve the primary link quality rather than degrading it. Specifically, when the PU is inactive, the SU uses the primary channel in a conventional manner. When, however, the PU is active, the SU is still allowed to transmit its data, by employing an amplify-and-forward (AF) relaying strategy, which has been widely used in cooperative systems [9]–[16]: specifically, the PU signal received by the SU is multiplied by the information symbols of the SU and retransmitted.

It should be observed that our approach differs from "cognitive relaying" [17]–[19], wherein relays with cognitive functionalities forward source data by using the spectrum white space(s) they have previously detected; indeed, in our scheme the SU is allowed to transmit also when the primary channel is busy. Moreover, our approach is also different from analog network coding (ANC) [20]–[23], since, in ANC schemes, the nodes use the wireless channel only to relay third-party information, without transmitting their own information. Our method also shares some resemblance with the concept of hiding information onto another signal without significantly distorting it [24], [25], a paradigm adopted in several applications, such as copyright protection for digital media, watermarking, fingerprinting, stenography, and data embedding. However, the proposed approach differs from information hiding mainly because superimposition of the SU symbols on the PU signal not only preserves the information of the PU, but also improves its performance (due to relaying) under suitable conditions.

The fundamental limits of the cognitive radio approach can be studied by modeling the cognitive radio channel as a classical interference channel [26]–[29]. It has been shown in [29] that the capacity-achieving strategy for the SU is to perform superposition coding of its codeword (generated by Costa precoding [30]), as well as of the codeword of the PU. Such a strategy requires the SU to know the PU codeword before it is transmitted (this is referred to as *noncausal knowledge* hereinafter). Even though such a constraint can be relaxed if the PU and SU are in close proximity of each other, the SU code selection would require instantaneous decoding of the PU message while it is transmitted on the air. In contrast, our proposed scheme does not require noncausal knowledge.

The main analytical contribution of the paper is to provide upper and lower bounds on the ergodic channel capacity [31] of the proposed scheme, evaluated by assuming a block-fading

Manuscript received February 1, 2014; revised June 21, 2014, November 8, 2014, and March 18, 2015; accepted May 22, 2015. This work was supported in part by the Italian National Project Servizi per lInfrastruttura di Rete Wireless Oltre il 3G (SIRIO). The associate editor coordinating the review of this paper and approving it for publication was G. Yue.

F. Verde and G. Gelli are with the Department of Electrical Engineering and Information Technology, University Federico II, Naples I-80125, Italy (e-mail: f.verde@unina.it; gelli@unina.it).

A. Scaglione is with the School of Electrical Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-9309 USA (e-mail: Anna.Scaglione@asu.edu).

D. Darsena is with the Department of Engineering, Parthenope University, Naples I-80143, Italy (e-mail: darsena@uniparthenope.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2440359

channel model and considering different assumptions on the amount of side information available at the receivers.[1] The ergodic capacity serves as a useful upper bound on the performance of any communication system and it can be achieved if the length of the codebook is long enough to reflect the ergodic nature of fading (i.e., the transmission duration of the codeword is much greater than the channel coherence time) [33]. It is shown that the concurrent transmission of the SU might improve the capacity of the PU link, provided that certain non-restrictive conditions are fulfilled: in this case, the SU earns an unlicensed channel with low transmission rates, which allow to reduce its average delay per symbol.

The paper is organized as follows. The model of the proposed cognitive radio network is introduced in Section II. Upper and lower bounds on the ergodic channel capacities of the PU and SU links are calculated in Sections III and IV, respectively, along with Monte Carlo numerical results aimed at assessing the ultimate achievable performances of both the PU and SU. Conclusions are drawn in Section V.

A. Notations and Preliminaries

The fields of complex, real, and nonnegative integer numbers are denoted with $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{N}$, respectively; matrices [vectors] are denoted with upper [lower] case boldface letters (e.g., $\mathbf{A}$ or $\mathbf{a}$); the field of $m \times n$ complex matrices is denoted as $\mathbb{C}^{m\times n}$, with $\mathbb{C}^m$ used as a shorthand for $\mathbb{C}^{m\times 1}$; the superscripts T and H denote the transpose and the conjugate transpose of a matrix, respectively; the symbol $*$ stands for (linear) convolution; $j \triangleq \sqrt{-1}$ denotes the imaginary unit; $\delta(n)$ is the Kronecker delta; $|a|$ and $\measuredangle a$ denote the magnitude and the phase of $a \in \mathbb{C}$, respectively; $\text{int}(x)$ gives the integer part of $x \in \mathbb{R}$; $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ denotes the identity matrix; $\|\mathbf{a}\|$ is the Euclidean norm of $\mathbf{a} \in \mathbb{C}^n$; matrix $\mathbf{A} = \text{diag}(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{C}^{n\times n}$ is diagonal; $\det(\mathbf{A})$ and $\text{tr}(\mathbf{A})$ are the determinant and trace of matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$, respectively; the operator $\mathbb{E}_{\mathbf{x}}[\cdot]$ denotes ensemble averaging with respect to the random vector $\mathbf{x} \in \mathbb{C}^n$ (the subscript is omitted when it can be deduced from the context) and $\mathbb{E}_{\mathbf{x}|\mathbf{y}}[\cdot]$ is the conditional mean with respect to $\mathbf{x}$ given the random vector $\mathbf{y} \in \mathbb{C}^m$; let $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{y} \in \mathbb{C}^n$, and $\mathbf{z} \in \mathbb{C}^p$ be random vectors, $p(\mathbf{x})$ is the probability density function (pdf) of $\mathbf{x}$, $p(\mathbf{x}|\mathbf{y})$ is the conditional pdf of $\mathbf{x}$, given $\mathbf{y}$, $\mathsf{I}(\mathbf{x};\mathbf{y})$ denotes the mutual information [31] between $\mathbf{x}$ and $\mathbf{y}$, $\mathsf{I}(\mathbf{x};\mathbf{y}|\mathbf{z})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}$, given $\mathbf{z}$, $\mathsf{h}(\mathbf{x})$ denote the differential entropy [31] of $\mathbf{x}$, and $\mathsf{h}(\mathbf{x}|\mathbf{y})$ is the conditional differential entropy of $\mathbf{x}$, given $\mathbf{y}$; a circularly symmetric complex Gaussian random vector $\mathbf{x} \in \mathbb{C}^n$ with mean $\boldsymbol{\mu} \in \mathbb{C}^n$ and covariance matrix $\Sigma \in \mathbb{C}^{n\times n}$ is denoted as $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \Sigma)$; finally, we define the following function (see, e.g., [34])

$$
\begin{aligned}
f(A) &\triangleq \int_0^{+\infty} \exp(-u)\ln(1+Au)\mathrm{d}u \\
&= -\exp\left(\frac{1}{A}\right)\mathrm{Ei}\left(-\frac{1}{A}\right) \\
&\approx \begin{cases} A, & \text{for} 0 < A \ll 1; \\ \ln(1+A) - \gamma, & \text{for} A \gg 1. \end{cases}
\end{aligned} \tag{1}
$$

[1]Preliminary results of such an analysis are reported in [32].

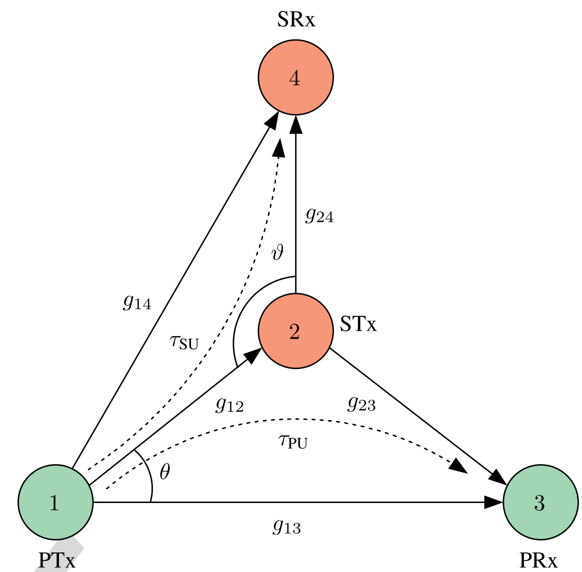

Fig. 1. The considered wireless network model: in green, the PU transmitting/receiving nodes, in red the SU transmitting/receiving nodes.

where, for $x < 0$,

$$
\mathrm{Ei}(x) \triangleq \int_{-\infty}^{x} \frac{\exp(u)}{u}\mathrm{d}u = \gamma + \ln(-x) + \sum_{k=1}^{+\infty} \frac{x^k}{k!k}
$$

denotes the exponential integral function and

$$
\gamma \triangleq \lim_{n\to\infty}\left(n^{-1}\sum_{k=1}^{n} k^{-1} - \ln n\right) \approx 0.57721
$$

is the Euler-Mascheroni constant.

II. System Model and Proposed Cooperative Protocol

The cognitive radio network includes (Fig. 1) a primary transmitter/receiver pair (PTx/PRx) represented by nodes 1 and 3, and a secondary transmitter/receiver pair (STx/SRx) represented by nodes 2 and 4. A single channel is licensed to the primary user (PU), who uses it in a bursty manner, by alternating between busy (ON) and idle (OFF) intervals. During the ON intervals, the PU transmits a sequence $x_{\text{PU}}(\cdot)$ of independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex symbols, with variance $\sigma^2_{\text{PU}} = \mathsf{P}_{\text{PU}}$ and signaling interval T_{PU}. Such a sequence is arranged in consecutive frames of M symbols, whose duration is comparable with the coherence time of the channel. With regard to the secondary user (SU), the STx attempts to send towards the SRx a sequence $x_{\text{SU}}(\cdot)$ of i.i.d. zero-mean circularly symmetric complex symbols with variance σ^2_{SU}, statistically independent of $x_{\text{PU}}(\cdot)$. The power budget of the SU is given by P_{SU}. We assume that the SU knows the occurrence of the PU ON/OFF intervals: such an information is provided by the PU or estimated by the SU itself (e.g., by sensing). The symbol rate $1/\mathsf{T}_{\text{SU}}$ and the variance σ^2_{SU} of the SU depend on the status of the PU channel, as explained in the sequel. The channel corresponding to the $i \to \ell$

link is assumed to be frequency-flat and quasi-stationary: it is modeled by the fading coefficient $g_{i\ell} \sim \mathcal{CN}(0, \sigma_{i\ell}^2)$, which is constant within one frame, but is allowed to vary independently from frame to frame. Fading coefficients of different links are statistically independent among themselves and of the symbol sequences. CSI is not available at the transmitters, whereas it can be acquired at the receivers by training [35].

Depending on PU activity, the STx selects one of two transmission modes: *white-space* (idle PU channel) or *dirty-space* (busy PU channel) one. In white-space mode, the STx gains exclusive use of the channel and employs a conventional point-to-point technique to transmit with signaling interval $\mathsf{T}_{\mathrm{SU}} \equiv \mathsf{T}_{\mathrm{SU,white}}$. In dirty-space mode, since the signal transmitted by the PU is also overheard by the STx, the latter acts as a full-duplex AF relay and transmits one symbol per PU frame, i.e., $\mathsf{T}_{\mathrm{SU}} \equiv \mathsf{T}_{\mathrm{SU,dirty}} \triangleq M\mathsf{T}_{\mathrm{PU}}$, as explained soon after.

The T_{PU}-spaced baseband equivalent signal received at the STx during the mth PU symbol period of a frame of length M is expressed as

$$r_2(m) = g_{12} x_{\mathrm{PU}}(m) + v_2(m), \quad m \in \{0, 1, \ldots, M-1\} \tag{2}$$

where $v_2(\cdot)$ denotes AWGN at the STx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_2}^2)$ random variables (RVs), statistically independent of g_{12} and $x_{\mathrm{PU}}(\cdot)$. Let x_{SU} denote the SU symbol to be transmitted during the considered PU frame, the signal (2) is scaled at the STx by the factor x_{SU} and forwarded to the SRx. The variance of x_{SU} is adjusted by the STx according to the average power constraint $\mathbb{E}[|x_{\mathrm{SU}} r_2(m)|^2] = \mathsf{P}_{\mathrm{SU}}$, i.e.,

$$\sigma_{\mathrm{SU}}^2 \equiv \sigma_{\mathrm{SU,dirty}}^2 \triangleq \frac{\mathsf{P}_{\mathrm{SU}}}{\sigma_{12}^2 \mathsf{P}_{\mathrm{PU}} + \sigma_{n_2}^2}. \tag{3}$$

Further details on the signal model in white space-mode are omitted, since it is a simple block flat-fading Rayleigh channel affected by additive white Gaussian noise (AWGN) [36], [37]. In the next subsections, we derive and discuss the received signal models at the PRx and SRx in dirty-space mode.

A. Received Signal at the PRx in the Dirty-Space Mode

Preliminarily (see also Fig. 1), let $t_{i\ell}$ denotes the propagation delays for the $i \to \ell$ link: since only relative delays are important [38], the delays t_{13} and t_{14} of the PTx → PRx and PTx → SRx links are conventionally set to zero. Thus, the cumulative PTx → STx → PRx and PTx → STx → SRx delays can be expressed as $t_{\mathrm{PU}} = t_{12} + t_{\mathrm{p}} + t_{23}$ and $t_{\mathrm{SU}} = t_{12} + t_{\mathrm{p}} + t_{24}$, respectively, with t_{p} denoting the processing time at the STx, which depends on various parameters [39], such as the frame length M and the hardware/memory characteristics of the STx node. We assume that M is chosen such that both t_{PU} and t_{SU} are much smaller than the frame duration $M\,\mathsf{T}_{\mathrm{PU}}$.

The T_{PU}-spaced baseband equivalent received signal at the PRx can be written as

$$y_3(m) = g_{13} x_{\mathrm{PU}}(m) + g_{23} x_{\mathrm{SU}} r_2(m - \tau_{\mathrm{PU}}) + v_3(m) \tag{4}$$

for $m \in \{0, 1, \ldots, M-1\}$, where $\tau_{\mathrm{PU}} \triangleq \mathrm{int}(t_{\mathrm{PU}}/\mathsf{T}_{\mathrm{PU}})$ is the integer[2] delay and $v_3(\cdot)$ denotes AWGN at the PRx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_3}^2)$ RVs, statistically independent of g_{13}, $x_{\mathrm{PU}}(\cdot)$, g_{23}, x_{SU}, and $r_2(\cdot)$. Setting $y_3(m) \equiv y_{\mathrm{PU}}(m)$, eq. (4) becomes

$$y_{\mathrm{PU}}(m) = g_{\mathrm{PU}}(m) * x_{\mathrm{PU}}(m) + v_{\mathrm{PU}}(m) \tag{5}$$

where $g_{\mathrm{PU}}(m) \triangleq g_{13}\delta(m) + g_{12} g_{23} x_{\mathrm{SU}} \delta(m - \tau_{\mathrm{PU}})$ and $v_{\mathrm{PU}}(m) \triangleq g_{23} x_{\mathrm{SU}} v_2(m - \tau_{\mathrm{PU}}) + v_3(m)$ represent the impulse response of the overall PU relay channel towards the PRx and the equivalent noise term at the PRx, respectively. It can be seen from (5) that, when $\tau_{\mathrm{PU}} > 0$, the PU experiences ISI through a *two-ray frequency-selective channel* $g_{\mathrm{PU}}(m)$, with the second channel tap gain incorporating the contribution of the SU transmitted symbol x_{SU}. The delay τ_{PU} depends on the sum of the delays t_{12} and t_{23} (see Fig. 1), the processing time t_{p} at the STx, and the PU symbol period T_{PU}: it is noteworthy that τ_{PU} is greater than zero when either $t_{12} + t_{23} \geq \mathsf{T}_{\mathrm{PU}}$ or $t_{\mathrm{p}} \geq \mathsf{T}_{\mathrm{PU}}$—conditions that are likely to be fulfilled if the PU transmits at high baud rates (T_{PU} small).

The "composite" channel impulse response $g_{\mathrm{PU}}(m)$ can be directly estimated at the PRx using the training symbols transmitted by the PTx and, thus, knowledge of x_{SU} by the PU is not required. Such a channel estimate can then be utilized to coherently recover the PU data symbols [36].

B. Received Signal at the SRx in the Dirty-Space Mode

The T_{PU}-spaced baseband equivalent received signal at the SRx is given by

$$y_4(m) = g_{24} x_{\mathrm{SU}} r_2(m - \tau_{\mathrm{SU}}) + g_{14} x_{\mathrm{PU}}(m) + v_4(m) \tag{6}$$

for $m \in \{0, 1, \ldots, M-1\}$, where $\tau_{\mathrm{SU}} \triangleq \mathrm{int}(t_{\mathrm{SU}}/\mathsf{T}_{\mathrm{PU}})$ is the integer[3] delay and the sequence $v_4(\cdot)$ denotes AWGN at the SRx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_4}^2)$ RVs, statistically independent of g_{14}, $x_{\mathrm{PU}}(\cdot)$, g_{24}, x_{SU}, and $r_2(\cdot)$. Setting $y_4(m) \equiv y_{\mathrm{SU}}(m)$, eq. (6) becomes

$$y_{\mathrm{SU}}(m) = g_{\mathrm{SU}}(m) x_{\mathrm{SU}} + v_{\mathrm{SU}}(m) \tag{7}$$

where $g_{\mathrm{SU}}(m) \triangleq g_{24} r_2(m - \tau_{\mathrm{SU}}) = g_{24} g_{12} x_{\mathrm{PU}}(m - \tau_{\mathrm{SU}}) + g_{24} v_2(m - \tau_{\mathrm{SU}})$ and $v_{\mathrm{SU}}(m) \triangleq g_{14} x_{\mathrm{PU}}(m) + v_4(m)$ represent the time-varying fading gain towards the SRx and the equivalent noise term at the SRx, respectively. It results from (7) that the SU sees a *fast flat-fading channel* $g_{\mathrm{SU}}(m)$, which also depends on the PU symbol $x_{\mathrm{PU}}(m)$. Since the SU transmits only one symbol per frame in dirty-space mode, it may be not able to reliably estimate the time-varying channel $g_{\mathrm{SU}}(m)$ by using its own training sequence, for $m \in \{0, 1, \ldots, M-1\}$, and, thus, coherent detection of x_{SU} might be difficult to implement in the dirty-space mode. Without any additional knowledge, recovery of x_{SU} can be accomplished at the SRx by resorting to noncoherent or generalized maximum-likelihood (ML)

[2] Any residual fractional delay $t_{\mathrm{PU}} - \tau_{\mathrm{PU}} \mathsf{T}_{\mathrm{PU}}$ can be absorbed into the channel coefficient g_{23}.

[3] Any residual fractional delay $t_{\mathrm{SU}} - \tau_{\mathrm{SU}} \mathsf{T}_{\mathrm{PU}}$ can be absorbed into the channel coefficient g_{24}.

detection rules [40]–[42], which require at most knowledge of the second-order statistics of $g_{\rm SU}(m)$. In Section IV, the performance of the SU is evaluated under the assumption that the SRx has the additional knowledge of the training symbols of the PU. If the signal-to-noise ratio (SNR) at the STx is sufficiently large, the term $g_{24}v_2(m-\tau_{\rm SU})$ in $g_{\rm SU}(m)$ can be neglected and the SRx can acquire the channel parameters $g_{24}g_{12}$ and g_{14}, by jointly exploiting the training sequences sent by the PTx and STx. In this case, the SU symbols can be estimated at the SRx by resorting to partially-coherent detection algorithms [43].

III. Performance Analysis of the Primary User

Herein, we derive the conditions assuring that the achievable long-term rate of the PU link is not worsened when the SU is transmitting. To this aim, it is assumed that the training sequence for the PU link is long enough to acquire the relevant CSI with negligible error at the PRx. As a benchmark, we first study the performance limit of the PU link in the case of direct PTx $\to$ PRx transmission, i.e., when the SU is silent. In this case, the model for the received signal at the PRx can be simply obtained by setting $x_{\rm SU}=0$ in (4), thus yielding $y_{\rm PU}(m)=g_{13}x_{\rm PU}(m)+v_3(m)$, which shows that the direct PU transmission sees a block flat-fading Rayleigh channel with AWGN. The ergodic capacity (nats/symbol) [33], [37] of the direct PU link with CSI at the receiver (CSIR) is given by

$$\mathsf{C}_{\rm PU,direct}=\mathbb{E}\left[\ln\left(1+|g_{13}|^2\frac{\mathsf{P}_{\rm PU}}{\sigma_{n_3}^2}\right)\right]=f(\mathsf{ASNR}_{\rm PU,direct}) \quad (8)$$

where we have used (1) and the fact that $|g_{13}|^2$ has an exponential distribution with mean σ_{13}^2, whereas $\mathsf{ASNR}_{\rm PU,direct}\triangleq(\sigma_{13}^2\mathsf{P}_{\rm PU})/\sigma_{n_3}^2$ is the average (over the channel) SNR of the PU at the PRx (node 3) when $x_{\rm SU}=0$. As expected, $\mathsf{C}_{\rm PU,direct}$ is a monotonically increasing function of $\mathsf{ASNR}_{\rm PU,direct}$: for high ASNR values, it results from (1) that $\mathsf{C}_{\rm PU,direct}\approx\ln(1+\mathsf{ASNR}_{\rm PU,direct})-\gamma$; in the low-SNR regime, one has $\mathsf{C}_{\rm PU,direct}\approx\mathsf{ASNR}_{\rm PU,direct}$, that is, the capacity increases linearly with $\mathsf{ASNR}_{\rm PU,direct}$.

Let us now consider the case wherein the STx is active, i.e., $x_{\rm SU}\neq 0$; in this case, as discussed in Section II-A, the PU transmission experiences block frequency-selective fading [see eq. (5)]. Evaluation of the ergodic capacity of a single-user frequency-selective channel with CSIR can be carried out [37] by decomposing the channel into an equivalent number of independent frequency-flat (i.e., memoryless) subchannels, whose input-output relationships are given by

$$\tilde{y}_{\rm PU}(\ell)=G_{\rm PU}(\ell)\tilde{x}_{\rm PU}(\ell)+\tilde{v}_{\rm PU}(\ell) \quad (9)$$

for $\ell\in\{0,1,\ldots,M-1\}$, where $\tilde{y}_{\rm PU}(\cdot)$, $\tilde{x}_{\rm PU}(\cdot)$, and $\tilde{v}_{\rm PU}(\cdot)$ are the M-point discrete Fourier transform (DFT) of $y_{\rm PU}(\cdot)$, $x_{\rm PU}(\cdot)$, and $v_{\rm PU}(\cdot)$, respectively, and

$$G_{\rm PU}(\ell)=g_{13}+g_{12}g_{23}x_{\rm SU}e^{-j\frac{2\pi}{M}\ell\tau_{\rm PU}} \quad (10)$$

is the M-point DFT of the channel $g_{\rm PU}(m)$ in (5). In compact form, let the M-dimensional vectors $\mathbf{x}_{\rm PU}$, $\tilde{\mathbf{y}}_{\rm PU}$, $\tilde{\mathbf{v}}_{\rm PU}$ gather the corresponding scalar quantities, one has

$$\tilde{\mathbf{y}}_{\rm PU}\triangleq\widetilde{\mathbf{G}}_{\rm PU}\tilde{\mathbf{x}}_{\rm PU}+\tilde{\mathbf{v}}_{\rm PU} \quad (11)$$

where $\widetilde{\mathbf{G}}_{\rm PU}\triangleq\mathrm{diag}[G_{\rm PU}(0),G_{\rm PU}(1),\ldots,G_{\rm PU}(M-1)]\in\mathbb{C}^{M\times M}$ and $\tilde{\mathbf{x}}_{\rm PU}$ is subject to the power constraint $\mathbb{E}[\|\tilde{\mathbf{x}}_{\rm PU}\|^2]=M\mathsf{P}_{\rm PU}$. Under the assumption of exact CSIR, the PRx is assumed to have perfect knowledge of the channel vector $\mathbf{g}_{\rm PU}\triangleq[g_{13},g_{12}g_{23}x_{\rm SU}]^{\rm T}\in\mathbb{C}^2$, and the ergodic channel capacity (in nats/symbol) of the PU can be calculated [37] as

$$\mathsf{C}_{\rm PU}\triangleq\lim_{M\to+\infty}\mathsf{C}_{\rm PU}(M) \quad (12)$$

where $\mathsf{C}_{\rm PU}(M)$ is obtained by taking the supremum of the average mutual information $\mathsf{I}(\tilde{\mathbf{x}}_{\rm PU};\tilde{\mathbf{y}}_{\rm PU}|\mathbf{g}_{\rm PU})/M$ over all possible distributions of the vector $\tilde{\mathbf{x}}_{\rm PU}$ that satisfy the power constraint $\mathbb{E}[\|\tilde{\mathbf{x}}_{\rm PU}\|^2]=M\mathsf{P}_{\rm PU}$. Since the subchannels (9) are memoryless and the entries of $\tilde{\mathbf{x}}_{\rm PU}$ are statistically independent, one has

$$\mathsf{I}(\tilde{\mathbf{x}}_{\rm PU};\tilde{\mathbf{y}}_{\rm PU}|\mathbf{g}_{\rm PU})=\sum_{\ell=0}^{M-1}\mathsf{I}\left(\tilde{x}_{\rm PU}(\ell);\tilde{y}_{\rm PU}(\ell)|\mathbf{g}_{\rm PU}\right). \quad (13)$$

Therefore, calculation of $\mathsf{C}_{\rm PU}$ boils down to evaluating $\mathsf{I}(\tilde{x}_{\rm PU}(\ell);\tilde{y}_{\rm PU}(\ell)|\mathbf{g}_{\rm PU})$. It is worth noting that, for a given $\mathbf{g}_{\rm PU}$, the equivalent noise term $\tilde{v}_{\rm PU}(\ell)$ in (9) is a non-Gaussian RV, which complicates exact evaluation of (13). For this reason, we provide in Theorem 1 upper and lower bounds on $\mathsf{C}_{\rm PU}$.

Theorem 1 (Upper and Lower Bounds on the PU Capacity): The ergodic channel capacity of the PU can be lower- and upper-bounded as follows:

$$\begin{aligned}\mathbb{E}\left\{f\left[\Gamma_{3,\rm lower}\left(|g_{23}|^2|x_{\rm SU}|^2\right)\right]\right\}&\triangleq\mathsf{C}_{\rm PU,lower}\leq\mathsf{C}_{\rm PU}\\ &\leq\mathsf{C}_{\rm PU,upper}\triangleq\mathbb{E}\left\{f\left[\Gamma_{3,\rm upper}\left(|g_{23}|^2|x_{\rm SU}|^2\right)\right]\right\}\end{aligned} \quad (14)$$

with

$$\Gamma_{3,\rm upper}\left(|g_{23}|^2|x_{\rm SU}|^2\right)\triangleq\mathsf{ASNR}_{\rm PU,direct}\frac{1+|g_{23}|^2|x_{\rm SU}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}}{1+|g_{23}|^2|x_{\rm SU}|^2\frac{\sigma_{n_2}^2}{\sigma_{n_3}^2}} \quad (15)$$

$$\Gamma_{3,\rm lower}(|g_{23}|^2|x_{\rm SU}|^2)\triangleq\mathsf{ASNR}_{\rm PU,direct}\frac{1+|g_{23}|^2|x_{\rm SU}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}}{1+\sigma_{23}^2\sigma_{\rm SU,dirty}^2\frac{\sigma_{n_2}^2}{\sigma_{n_3}^2}}. \quad (16)$$

Proof: See Appendices A and B. □

The upper bound in (14) is obtained in Appendix A by assuming that the PRx has the additional perfect knowledge of g_{12}.[4] Instead, the lower bound in (14) is derived in Appendix B by replacing $\tilde{x}_{\rm PU}(\ell)$ and $\tilde{v}_{\rm PU}(\ell)$ in (9) with $\tilde{x}_{\rm PU,G}(\ell)\sim\mathcal{CN}(0,\mathsf{P}_{\rm PU})$ and $\tilde{v}_{\rm PU,G}(\cdot)\sim\mathcal{CN}\left(0,\sigma_{n_3}^2+\sigma_{n_2}^2\sigma_{23}^2\sigma_{\rm SU,dirty}^2\right)$,

[4]To acquire g_{12} in practice, the PRx would require additional help from the STx in the form of channel-state feedback.

respectively. Since $g_{23} \sim \mathcal{CN}(0, \sigma_{23}^2)$, the averages needed to evaluate $\mathsf{C}_{\mathrm{PU,lower}}$ and $\mathsf{C}_{\mathrm{PU,upper}}$ can be computed numerically by assuming a prior distribution for the symbol x_{SU} transmitted by the SU. Finally, observe that both $\mathsf{C}_{\mathrm{PU,upper}}$ and $\mathsf{C}_{\mathrm{PU,lower}}$ do not depend on the delay τ_{PU}, provided that $\tau_{\mathrm{PU}} \ll M$.

At this point, we can draw some interesting conclusions. To this end, we refer for simplicity to the path-loss model $\sigma_{i\ell}^2 = d_{i\ell}^{-\eta}$, where $d_{i\ell}$ is the distance between nodes i and ℓ and η denotes the path-loss exponent. The upper and lower bounds in (14) depend on $\Gamma_{3,\mathrm{upper}}(|g_{23}|^2|x_{\mathrm{SU}}|^2)$ and $\Gamma_{3,\mathrm{lower}}(|g_{23}|^2|x_{\mathrm{SU}}|^2)$, respectively. The right-hand sides (RHSs) in (15) and (16) can be intuitively explained as follows. On one side, the SU transmission is beneficial since it increases the frequency diversity of the PU; mathematically, the gain in frequency diversity comes in (15) and (16) from the multiplication of $\mathsf{ASNR}_{\mathrm{PU,direct}}$ by the factor $1 + |g_{23}|^2|x_{\mathrm{SU}}|^2(\sigma_{12}^2/\sigma_{13}^2)$. On the other hand, the AF relaying carried out by the SU is detrimental, due to the noise propagation phenomenon from the STx to the PRx; mathematically, the adverse effect of noise propagation is represented in (15) by the division of $\mathsf{ASNR}_{\mathrm{PU,direct}}$ by the factor $1 + |g_{23}|^2|x_{\mathrm{SU}}|^2(\sigma_{n_2}^2/\sigma_{n_3}^2)$ and in (16) by the division of $\mathsf{ASNR}_{\mathrm{PU,direct}}$ by the factor $1 + \sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2(\sigma_{n_2}^2/\sigma_{n_3}^2)$.

Let us consider the case when $\sigma_{12}^2/\sigma_{13}^2 \leq \sigma_{n_2}^2/\sigma_{n_3}^2$. One has from (15) that $\Gamma_{3,\mathrm{upper}}(|g_{23}|^2|x_{\mathrm{SU}}|^2) \leq \mathsf{ASNR}_{\mathrm{PU,direct}}$ for each realization of $g_{23}x_{\mathrm{SU}}$. Since $f(A)$ is a monotonically increasing function of $A \geq 0$, one obtains from (8) and (14) that $\mathsf{C}_{\mathrm{PU,upper}} \leq \mathsf{C}_{\mathrm{PU,direct}}$. According to the path-loss model, *the capacity of the PU may degrade when*

$$d_{12} \geq d_{13} \sqrt[\eta]{\sigma_{n_3}^2/\sigma_{n_2}^2}$$

since the SU prevalently forwards noise. In particular, when $\sigma_{n_3}^2 \geq \sigma_{n_2}^2$, that is, the PRx is noisier than the STx, the cognitive radio principle of protecting the PU might be violated if $d_{12} \geq d_{13}$, that is, the PTx is farther from the STx than from the PRx or it is equidistant from them. Since the SU can determine whether $\sigma_{12}^2/\sigma_{13}^2 > \sigma_{n_2}^2/\sigma_{n_3}^2$ or, equivalently, $d_{12} < d_{13} \sqrt[\eta]{\sigma_{n_3}^2/\sigma_{n_2}^2}$ (we call it the *symbiotic region*), in the following we restrict attention to this case.

By virtue of Theorem 1, the variance $\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2$ of the RV $|g_{23}|^2|x_{\mathrm{SU}}|^2$ plays a crucial role in determining the performance of the PU. Remembering (3), we observe that

$$\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2 = \frac{\sigma_{23}^2\mathsf{P}_{\mathrm{SU}}}{\sigma_{12}^2\mathsf{P}_{\mathrm{PU}} + \sigma_{n_2}^2} < \frac{\sigma_{23}^2}{\sigma_{12}^2}\frac{\mathsf{P}_{\mathrm{SU}}}{\mathsf{P}_{\mathrm{PU}}} = \left(\frac{d_{23}}{d_{12}}\right)^{-\eta}\frac{\mathsf{P}_{\mathrm{SU}}}{\mathsf{P}_{\mathrm{PU}}} \tag{17}$$

where, as a consequence of the Carnot's cosine law, we can also write

$$\frac{d_{23}}{d_{12}} = \sqrt{1 + \frac{d_{13}^2}{d_{12}^2} - 2\,\frac{d_{13}}{d_{12}}\cos(\theta)} \geq \left|\frac{d_{13}}{d_{12}} - 1\right| \tag{18}$$

with θ denoting the angle contained between sides of lengths d_{12} and d_{13} (see Fig. 1). The minimum value of d_{23}/d_{12} [corresponding to the equality in (18)] is achieved when $\cos(\theta) = 1$. It can be inferred from (15) and (16) that, when $\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2 \ll \epsilon$, with $\epsilon > 0$ sufficiently small, the benefits in frequency diversity prevail over the losses caused by noise propagation. In such a case, Chebychev's inequality [44] implies that $\mathsf{P}(|g_{23}|^2|x_{\mathrm{SU}}|^2 \geq \epsilon) \leq \sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2/\epsilon \ll 1$ and, consequently, the RV $|g_{23}|^2|x_{\mathrm{SU}}|^2$ takes on values significantly smaller than one, with high probability. Hence, when $\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2 \ll 1$, $\Gamma_{3,\mathrm{upper}}(|g_{23}|^2|x_{\mathrm{SU}}|^2)$ in (15) and $\Gamma_{3,\mathrm{lower}}(|g_{23}|^2|x_{\mathrm{SU}}|^2)$ in (16) can be approximated as

$$\begin{aligned}\Gamma_{3,\mathrm{upper}}\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right) &\approx \Gamma_{3,\mathrm{lower}}\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right)\\ &\approx \Gamma_3\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right) \triangleq \mathsf{ASNR}_{\mathrm{PU,direct}}\left(1+|g_{23}|^2|x_{\mathrm{SU}}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}\right)\end{aligned} \tag{19}$$

which leads to

$$\mathsf{C}_{\mathrm{PU}} \approx \mathsf{C}_{\mathrm{PU,lower}} \approx \mathsf{C}_{\mathrm{PU,upper}} \approx \mathbb{E}\left\{f\left[\Gamma_3\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right)\right]\right\}. \tag{20}$$

Comparing (8) with (20), since $\Gamma_3(|g_{23}|^2|x_{\mathrm{SU}}|^2) > \mathsf{ASNR}_{\mathrm{PU,direct}}$ for each realization of $g_{23}x_{\mathrm{SU}}$ and $f(A)$ is a monotonically increasing function of $A \geq 0$, one readily obtains that $\mathsf{C}_{\mathrm{PU}} > \mathsf{C}_{\mathrm{PU,direct}}$. Thus, *the capacity of the PU improves as a result of the SU transmission when* $\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2 \ll 1$, no matter what the distributions of g_{23} and x_{SU} are.[5] In the symbiotic region $\sigma_{12}^2/\sigma_{13}^2 > \sigma_{n_2}^2/\sigma_{n_3}^2$, $\Gamma_3(|g_{23}|^2|x_{\mathrm{SU}}|^2)$ is lower bounded as

$$\begin{aligned}&\Gamma_3\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right)\\ &\qquad > \mathsf{ASNR}_{\mathrm{PU,direct}}\left[1 + |g_{23}|^2|x_{\mathrm{SU}}|^2\left(\sigma_{n_2}^2/\sigma_{n_3}^2\right)\right].\end{aligned}$$

By virtue of (17), (18), condition $\sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2 \ll 1$ is met if

$$|d_{13}/d_{12} - 1| \gg \sqrt[\eta]{\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}}.$$

Since the symbiotic region is equivalently characterized by the inequality $d_{13}/d_{12} > \sqrt[\eta]{\sigma_{n_2}^2/\sigma_{n_3}^2}$, this imposes that, when $\sigma_{n_2}^2 \geq \sigma_{n_3}^2$, the STx has to be sufficiently closer to the PTx than to the PRx.

In order to assess the capacity improvement of the PU, it is noteworthy that, when $\mathsf{ASNR}_{\mathrm{PU,direct}} \gg 1$, it also results from (19) that $\Gamma_3(|g_{23}|^2|x_{\mathrm{SU}}|^2) > \mathsf{ASNR}_{\mathrm{PU,direct}} \gg 1$ for each realization of $g_{23}x_{\mathrm{SU}}$ and, hence, eq. (20) admits, in the high-SNR region, the approximated expression [see eq. (1)]

$$\begin{aligned}\mathsf{C}_{\mathrm{PU}} &\approx \mathbb{E}\left[\ln\left(1 + \Gamma_3\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right)\right)\right] - \gamma\\ &\approx \mathbb{E}\left[\ln\left(\Gamma_3\left(|g_{23}|^2|x_{\mathrm{SU}}|^2\right)\right)\right] - \gamma\\ &= \mathsf{C}_{\mathrm{PU,direct}} + \mathbb{E}\left[\ln\left(1 + \sigma_{23}^2\sigma_{\mathrm{SU,dirty}}^2|\overline{g}_{23}|^2|\overline{x}_{\mathrm{SU}}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}\right)\right]\end{aligned} \tag{21}$$

[5] Such an improvement is achieved if the PU is willing to equalize a two-ray frequency selective channel rather than a flat-fading one. While this behavior is atypical for conventional cognitive radio scheme where the PU is unaware of the SU transmissions, some advanced cognitive radio systems have been proposed where the PUs either aid or react to the SUs [3].

where $\overline{g}_{23} \triangleq g_{23}/\sigma_{23}$ and $\overline{x}_{\rm SU} \triangleq x_{\rm SU}/\sigma_{\rm SU}$ are normalized versions of g_{23} and $x_{\rm SU}$, respectively, with $\sigma^2_{\rm SU,dirty}$ given by (3), and we have used the fact that $\mathsf{C}_{\rm PU,direct} \approx \ln(\mathsf{ASNR}_{\rm PU,direct}) - \gamma$ for $\mathsf{ASNR}_{\rm PU,direct} \gg 1$. Since $|\overline{g}_{23}|^2$ is a unit-mean exponential RV statistically independent of $\overline{x}_{\rm SU}$, by evaluating the expectation in (21) with respect to the distribution of $\overline{g}_{23}$, and using (1), one has, for $\mathsf{ASNR}_{\rm PU,direct} \gg 1$, that the *capacity gain* of the PU is given by

$$\Delta\mathsf{C}_{\rm PU} \triangleq \mathsf{C}_{\rm PU} - \mathsf{C}_{\rm PU,direct} \approx \mathbb{E}\left[f\left(\frac{\sigma^2_{23} \mathsf{P}_{\rm SU} |\overline{x}_{\rm SU}|^2}{\sigma^2_{13} \mathsf{P}_{\rm PU}} \right) \right] \quad (22)$$

where we have observed from (17) that, in the high-SNR regime, $\sigma^2_{23}\sigma^2_{\rm SU,dirty}\sigma^2_{12}/\sigma^2_{13} \approx (\sigma^2_{23}\mathsf{P}_{\rm SU})/(\sigma^2_{13}\mathsf{P}_{\rm PU})$. We observe that, if the SU adopts a constant-modulus constellation with average energy $\sigma^2_{\rm SU,dirty}$, the capacity gain of the PU in the high-SNR regime is obtained in closed-form by setting $|\overline{x}_{\rm SU}|^2 = 1$ and removing the average in (22). Interestingly, the capacity gain $\Delta\mathsf{C}_{\rm PU}$ becomes significant in the high-SNR region if $\mathsf{P}_{\rm SU} \gg \mathsf{P}_{\rm PU}$: this is in contrast with conventional cognitive radio approaches, for which concurrent transmission of the SU is allowed only if its transmission power is subject to a strict constraint, such that the interference at the PRx is within the interference temperature limit [2]. Intuitively, such a behavior is a consequence of the fact that increasing the variance of $x_{\rm SU}$ [see eq. (3)] leads to a raise in the variance of the second tap of the impulse response $g_{\rm PU}(m)$ [see eq. (5)].

To corroborate the information-theoretic findings, we report herein some results of numerical simulations. Specifically, we plot the maximum $(\Delta\mathsf{C}_{\rm PU})_{\max} \triangleq \mathsf{C}_{\rm PU,upper} - \mathsf{C}_{\rm PU,direct}$ and minimum capacity gain $(\Delta\mathsf{C}_{\rm PU})_{\min} \triangleq \mathsf{C}_{\rm PU,lower} - \mathsf{C}_{\rm PU,direct}$ of the PU (reported in bits/symbol and referred to as "ub" and "lb," respectively), where the ensemble averages in (14) were evaluated by carrying out 10^4 Monte Carlo trials. Obviously, it results that $(\Delta\mathsf{C}_{\rm PU})_{\min} \le \Delta\mathsf{C}_{\rm PU} \le (\Delta\mathsf{C}_{\rm PU})_{\max}$. With reference to Fig. 1, we normalized the distance between the PTx and the PRx, as well as the transmitting power of the PU, by setting $d_{13} = 1$ and $\mathsf{P}_{\rm PU} = 1$, respectively. Moreover, we chose $\theta = \pi/3$ and $\eta = 2$. The SU symbol $x_{\rm SU}$ was drawn from a QPSK constellation having average energy $\sigma^2_{\rm SU,dirty}$ given by (3) and we chose $\sigma^2_{n_2} = \sigma^2_{n_3}$.[6]

1) Fig. 2 : It reports the curves of PU capacity gain as a function of $\mathsf{SNR}_{\rm PU} \triangleq \mathsf{P}_{\rm PU}/\sigma^2_{n_2}$ for different values of the distance d_{12} between the PTx and the STx (the distance d_{23} between the STx and the PRx is calculated according to the Carnot's cosine law), with $\mathsf{P}_{\rm SU} = 1$. It is seen that the PU can harvest a noticeable capacity gain from the concurrent transmission of the SU, which rapidly increases as d_{12} decreases. Moreover, when the PTx and the STx are sufficiently close to each other, the upper and lower bounds in (14) tend to coincide, thus yielding an accurate approximation of the PU capacity.

2) Fig. 3 : The curves of the PU capacity gain are depicted as a function of $\mathsf{SNR}_{\rm PU}$ for different values of the SU transmitting power $\mathsf{P}_{\rm SU}$, with $d_{12} = 0.5$. As analytically predicted, results show that, the larger the transmitting power of the SU, the greater the capacity improvement of the PU will be. Moreover, as the difference between $\mathsf{P}_{\rm SU}$ and $\mathsf{P}_{\rm PU}$ increases, the upper and lower bounds in (14) tend to slightly space out, allowing anyhow to accurately predict the achievable PU rate.

[6]Such an assumption is reasonable when nodes 2 and 3 (approximatively) have the same noise figure, i.e., they are equipped with similar hardware components and operate in the same environment.

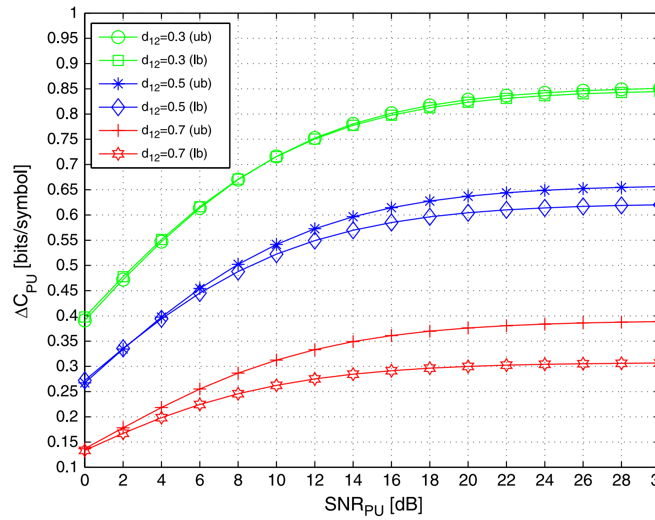

Fig. 2. $(\Delta\mathsf{C}_{\rm PU})_{\max}$, $(\Delta\mathsf{C}_{\rm PU})_{\min}$ versus $\mathsf{SNR}_{\rm PU}$ for different values of d_{12} ($d_{13} = 1$, $\mathsf{P}_{\rm PU} = \mathsf{P}_{\rm SU} = 1$, $\sigma^2_{n_2} = \sigma^2_{n_3}$).

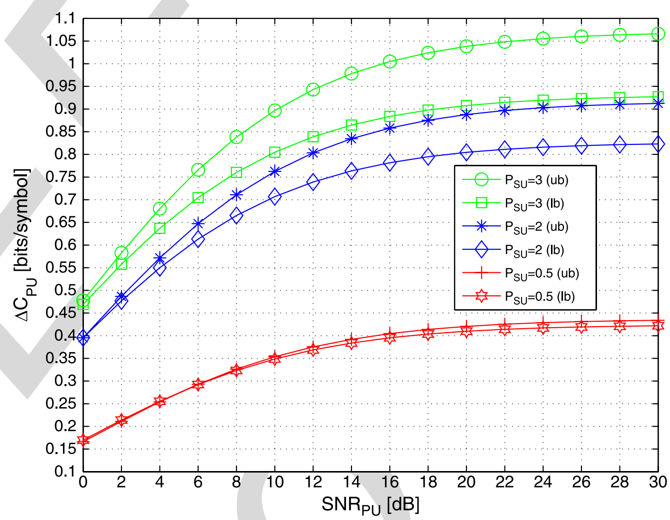

Fig. 3. $(\Delta\mathsf{C}_{\rm PU})_{\max}$, $(\Delta\mathsf{C}_{\rm PU})_{\min}$ versus $\mathsf{SNR}_{\rm PU}$ for different values of $\mathsf{P}_{\rm SU}$ ($d_{13} = 1$, $d_{12} = 0.5$, $\mathsf{P}_{\rm PU} = 1$, $\sigma^2_{n_2} = \sigma^2_{n_3}$).

IV. Performance Analysis of the Secondary User

In white-space mode (PU OFF intervals), the STx directly transmits to the SRx with average power per symbol $\sigma^2_{\rm SU} \equiv \sigma^2_{\rm SU,white} \triangleq \mathsf{P}_{\rm SU}$, by exclusively using the available PU channel. In this case, the ergodic channel capacity $\mathsf{C}_{\rm SU,white}$ with CSIR can be obtained from (8) by replacing σ^2_{13}, $\mathsf{P}_{\rm PU}$, and $\sigma^2_{n_3}$ with σ^2_{24}, $\mathsf{P}_{\rm SU}$, and $\sigma^2_{n_4}$, respectively, thus obtaining

$$\mathsf{C}_{\rm SU,white} = f(\mathsf{ASNR}_{\rm SU,white}) \qquad \text{(nats/symbol)} \qquad (23)$$

where $\mathsf{ASNR}_{\rm SU,white} \triangleq (\sigma^2_{24}\mathsf{P}_{\rm SU})/\sigma^2_{n_4}$ is the ASNR at the SRx in white-space mode.

Let us evaluate the achievable throughput of the SU in dirty-space mode (PU ON intervals). As discussed in Section II-B, the SU sees a fast flat-fading channel [see eq. (7)] in dirty-space mode. According to the results of the capacity analysis of the PU, we assume hereinafter that the ASNR at the STx is sufficiently large, i.e., $\sigma_{12}^2 \mathsf{P}_{\text{PU}} \gg \sigma_{n_2}^2$, which is a reasonable assumption for SNR values of practical interest when the STx is not too far from the PTx. In this case, the noise term in (2) can be neglected, allowing one to simplify the fading gain as $g_{\text{SU}}(m) \approx g_{24} g_{12} x_{\text{PU}}(m - \tau_{\text{SU}})$. Therefore, let $\mathbf{y}_{\text{SU}} \triangleq [y_{\text{SU}}(0), y_{\text{SU}}(1), \ldots, y_{\text{SU}}(M-1)]^{\text{T}} \in \mathbb{C}^M$ be the vector of samples that the SU observes over a frame, accounting for (7), the block $\mathbf{y}_{\text{SU}}$ can be expressed as

$$\mathbf{y}_{\text{SU}} = \left(g_{24} g_{12} \mathbf{J} \check{\mathbf{x}}_{\text{PU}}\right) x_{\text{SU}} + g_{14} \mathbf{x}_{\text{PU}} + \mathbf{v}_4 \tag{24}$$

where the matrix $\mathbf{J} \in \mathbb{C}^{M \times (M+\tau_{\text{SU}})}$ is obtained from the identity matrix $\mathbf{I}_{M+\tau_{\text{SU}}}$ by picking its first M rows, $\check{\mathbf{x}}_{\text{PU}} \triangleq [x_{\text{PU}}(-\tau_{\text{SU}}), \ldots, x_{\text{PU}}(-1), \mathbf{x}_{\text{PU}}^T]^{\text{T}} \in \mathbb{C}^{M+\tau_{\text{SU}}}$, and $\mathbf{v}_4 \triangleq [v_4(0), v_4(1), \ldots, v_4(M-1)]^{\text{T}} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_{n_4}^2 \mathbf{I}_M)$. Note that $\mathbf{y}_{\text{SU}}$ also depends on the block $\check{\mathbf{x}}_{\text{PU}}$ of PU symbols.

Regarding the CSIR of the SU, since the SU "composite" channel $g_{\text{SU}}(m)$, $m \in \{0, 1, \ldots, M-1\}$, rapidly changes due to the dependence on the PU symbols, the accuracy with which *full* CSI can be retrieved at the SRx during a predefined frame may be unsatisfactory. However, *partial* CSI can be reliably acquired at the SRx. Indeed, the SRx can attain an accurate estimate of the vector $\mathbf{g}_{\text{SU}} \triangleq [g_{24} g_{12}, g_{14}]^{\text{T}} \in \mathbb{C}^2$ by assuming that, besides having knowledge of the training symbol transmitted by the STx, the SRx additionally knows the training signal sent by the PTx.[7] Under this assumption, the channel $\mathbf{g}_{\text{SU}}$ can be estimated at the SRx by resorting to standard estimators [45], such as the ML or the Bayesian linear minimum mean-square error ones. Therefore, the capacity $\mathsf{C}_{\text{SU,dirty}}$ (in nats/symbol) of the SU link in dirty-space mode turns out to be the supremum of $\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}} | \mathbf{g}_{\text{SU}})/M$ over all distributions on x_{SU} satisfying (3). Upper and lower bounds on $\mathsf{C}_{\text{SU,dirty}}$ are given by Theorem 2.

Theorem 2 (Upper and Lower Bounds on the SU Capacity): The ergodic channel capacity of the SU can be lower- and upper-bounded as shown in (25) at the bottom of the page where

$$\begin{aligned}\mathbf{R}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) &\triangleq \Omega(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \mathbf{I}_M + \Upsilon^*(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \mathbf{B}^{\tau_{\text{SU}}} \\ &\quad + \Upsilon(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \mathbf{F}^{\tau_{\text{SU}}} \\ \boldsymbol{\Sigma}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) &\triangleq \left(|g_{24}|^2 |g_{12}|^2 x_{\text{SU}} \mathbf{I}_M + g_{24}^* g_{12}^* g_{14} \mathbf{B}^{\tau_{\text{SU}}}\right) \mathsf{P}_{\text{PU}}\end{aligned}$$

with $\Omega(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \triangleq (|g_{24}|^2 |g_{12}|^2 |x_{\text{SU}}|^2 + |g_{14}|^2) \mathsf{P}_{\text{PU}} + \sigma_{n_4}^2$, $\Upsilon(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \triangleq g_{24} g_{12} x_{\text{SU}} g_{14}^* \mathsf{P}_{\text{PU}}$, the matrices $\mathbf{B} \in \mathbb{R}^{M \times M}$ and $\mathbf{F} \in \mathbb{R}^{M \times M}$ being backward- and forward-shift matrices, respectively, and the inner ensemble average in the lower bound is taken over $x_{\text{SU}} \sim \mathcal{CN}(0, \sigma_{\text{SU,dirty}}^2)$.

Proof: See Appendices C and D. □

Some remarks are now in order about the bounds in (25). The upper bound $\mathsf{C}_{\text{SU,dirty,upper}}$ is obtained in Appendix C by assuming that the SRx has perfect knowledge of the pair $(\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}})$. Remembering that $g_{i\ell} \sim \mathcal{CN}(0, \sigma_{i\ell}^2)$, the capacity $\mathsf{C}_{\text{SU,dirty,upper}}$ can be evaluated numerically by assuming a prior distribution for the symbol sequence $x_{\text{PU}}(\cdot)$ transmitted by the PU. In addition to the Gaussian assumption for the SU symbol, the lower bound $\mathsf{C}_{\text{SU,dirty,lower}}$ is derived in Appendix D by assuming that the PU symbols are circularly symmetric complex Gaussian RVs. When the PU symbols are drawn from discrete symbol constellations, it can be argued that $\mathsf{C}_{\text{SU,dirty,lower}}$ is still valid as an approximated lower bound [46].

The upper bound in (25) can be (approximately) achieved if the SU is capable of decoding the PU symbols with arbitrarily small error probability. In order for this to be true, the information rate R_{PU} (in nats/symbol) of the PU must be smaller than the ergodic channel capacity $\mathsf{C}_{\text{PU}\to\text{SU}}$ (in nats/symbol) of the overall link between the PTx and the SRx, i.e., $\mathsf{R}_{\text{PU}} < \mathsf{C}_{\text{PU}\to\text{SU}}$. Due to the complete symmetry between the PRx and SRx with respect to the PTx, we can obtain an upper bound on $\mathsf{C}_{\text{PU}\to\text{SU}}$ exactly as we have done in Appendix A to get $\mathsf{C}_{\text{PU,upper}}$. Hence, similarly to (14), the following inequality holds $\mathsf{C}_{\text{PU}\to\text{SU}} \le \mathbb{E}\{f[\Gamma_{4,\text{upper}}(g_{24} x_{\text{SU}})]\}$, with

$$\Gamma_{4,\text{upper}}\left(|g_{24}|^2 |x_{\text{SU}}|^2\right) \triangleq \mathsf{ASNR}_{\text{PU}\to\text{SU}} \cdot \frac{1 + |g_{24}|^2 |x_{\text{SU}}|^2 \left(\sigma_{12}^2/\sigma_{14}^2\right)}{1 + |g_{24}|^2 |x_{\text{SU}}|^2 \left(\sigma_{n_2}^2/\sigma_{n_4}^2\right)}$$

where $\mathsf{ASNR}_{\text{PU}\to\text{SU}} \triangleq (\sigma_{14}^2 \mathsf{P}_{\text{PU}})/\sigma_{n_4}^2$ is the ASNR at the SRx when $x_{\text{SU}} = 0$. Such an upper bound can be approximated by using the same arguments that led to (19), (20), and (21).

At this point, we discuss some special cases/approximations of the bounds in (25), which provide insights on the ergodic channel capacity of the SU. Let us first consider the lower bound $\mathsf{C}_{\text{SU,dirty,lower}}$. The matrix $\mathbf{R}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}})$ in (25) is an Hermitian Toeplitz matrix and, for $\tau_{\text{SU}} \neq 0$, the behavior of its inverse can be characterized by using asymptotic (i.e., for $M \to +\infty$) arguments [47]. A substantial simplification occurs when $\tau_{\text{SU}} = 0$:[8] indeed, in this case, both $\mathbf{R}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}})$ and

[7]Such an assumption is reasonable when the SRx is sufficiently close to both the PTx and STx.

[8]This is a reasonable assumption when the PTx, STx, and SRx are sufficiently close in space, i.e., $d_{12} + d_{24} < c \mathsf{T}_{\text{PU}}$, and the processing time at the STx is smaller than the symbol period of the PU, i.e., $t_{\text{p}} < \mathsf{T}_{\text{PU}}$. Results of numerical simulations (not reported here in the interest of saving space) show that the impact of τ_{SU} on the worst-case ergodic channel capacity of the SU is negligible if $\tau_{\text{SU}} \ll M$.

$$\begin{aligned}&\frac{1}{M} \mathbb{E}_{\mathbf{g}_{\text{SU}}} \left\{ \ln \left[1 + \sigma_{\text{SU,dirty}}^2 \mathbb{E}_{x_{\text{SU}} | \mathbf{g}_{\text{SU}}} \left\{ \text{tr} \left[\mathbf{R}_{\text{SU}}^{-1}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \boldsymbol{\Sigma}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \mathbf{R}_{\text{SU}}^{-1}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \boldsymbol{\Sigma}_{\text{SU}}^{\text{H}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) \right] \right\} \right] \right\} \\ &\qquad \triangleq \mathsf{C}_{\text{SU,dirty,lower}} \le \mathsf{C}_{\text{SU,dirty}} \le \mathsf{C}_{\text{SU,dirty,upper}} \triangleq \frac{1}{M} \mathbb{E} \left\{ \ln \left[1 + \frac{\sigma_{\text{SU,dirty}}^2 |g_{24}|^2 |g_{12}|^2}{\sigma_{n_4}^2} \sum_{m=0}^{M-1} |x_{\text{PU}}(m - \tau_{\text{SU}})|^2 \right] \right\}\end{aligned} \tag{25}$$

$\boldsymbol{\Sigma}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}})$ turn out to be scaled identity matrices, thus yielding (26), shown at the bottom of the page.

Lastly, we consider the upper bound $\mathsf{C}_{\text{SU,dirty,upper}}$. We observe that $\sum_{m=0}^{M-1} |x_{\text{PU}}(m-\tau_{\text{SU}})|^2/M$ converges almost surely to $\sigma_{\text{PU}}^2 = \mathsf{P}_{\text{PU}}$ for $M \to +\infty$ by the strong law of large number [44]. If M is sufficiently large, we can use the approximation $\sum_{m=0}^{M-1} |x_{\text{PU}}(m-\tau_{\text{SU}})|^2 \approx M\mathsf{P}_{\text{PU}}$.[9] Thus,

$$\begin{aligned}\mathsf{C}_{\text{SU,dirty,upper}} &\approx \frac{1}{M}\mathbb{E}\left[\ln\left(1+M\frac{\mathsf{P}_{\text{SU}}}{\sigma_{n_4}^2}|g_{24}|^2|\overline{g}_{12}|^2\right)\right]\\ &= \frac{1}{M}\mathbb{E}\left[f\left(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2\right)\right] \quad (27)\end{aligned}$$

where we have used the approximation $\sigma_{\text{SU,dirty}}^2 \approx \mathsf{P}_{\text{SU}}/(\sigma_{12}^2\mathsf{P}_{\text{PU}})$, whereas $\overline{g}_{12} \triangleq g_{12}/\sigma_{12}$ denotes the normalized version of the channel coefficient characterizing the PTx $\to$ STx link. The equality in (27) is obtained by observing that $|g_{24}|^2$ is an exponential RV with mean σ_{24}^2, evaluating the expectation with respect to the distribution of g_{24}, and using (1). Low- and high-SNR approximations of (27) can be obtained by using the fact that $|\overline{g}_{12}|^2$ is a unit-mean exponential RV. Specifically, let $\epsilon > 0$ be a sufficiently small real number, one has $P(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2 \geq \epsilon) = \exp[-\epsilon/(M\mathsf{ASNR}_{\text{SU,white}})]$. This shows that, for $M\mathsf{ASNR}_{\text{SU,white}} \ll \epsilon$, the RV $M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2$ takes on values significantly smaller than one, with high probability. In this case, using (1), the following approximation of (27) holds

$$\mathsf{C}_{\text{SU,dirty,upper}} \approx \frac{1}{M}\mathbb{E}\left[M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2\right] = \mathsf{ASNR}_{\text{SU,white}}$$

with $\mathsf{ASNR}_{\text{SU,white}} \approx \mathsf{C}_{\text{SU,white}}$ for $\mathsf{ASNR}_{\text{SU,white}} \ll 1$ [see (1) and (23)]. As intuitively expected, if the SRx has perfect knowledge of $\mathbf{g}_{\text{SU}}$ and $\check{\mathbf{x}}_{\text{PU}}$, the performance limit of the SU in the low-SNR regime is nearly equal to that when the STx and SRx communicate directly using a dedicated exclusive channel. On the other hand, let $K > 0$ be a sufficiently large real number, one has

$$\begin{aligned}&P\left(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2 \leq K\right)\\ &\qquad = 1-\exp\left[-K/(M\mathsf{ASNR}_{\text{SU,white}})\right]. \quad (28)\end{aligned}$$

This shows that, for $M\mathsf{ASNR}_{\text{SU,white}} \gg K$, the RV $M\ \mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2$ takes on values significantly greater than one, with high probability. In this case, using (1) over and over again, we can approximate (27) as

$$\begin{aligned}\mathsf{C}_{\text{SU,dirty,upper}} &\approx \frac{1}{M}\left\{\mathbb{E}\left[\ln(1+M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2)\right]-\gamma\right\}\\ &= \frac{1}{M}\left[f(M\mathsf{ASNR}_{\text{SU,white}})-\gamma\right]\\ &\approx \frac{1}{M}\left[\ln(1+M\mathsf{ASNR}_{\text{SU,white}})-2\gamma\right]\\ &\approx \frac{1}{M}\left[\ln(M\mathsf{ASNR}_{\text{SU,white}})-2\gamma\right]\\ &= \frac{1}{M}\left[\ln(M)+\mathsf{C}_{\text{SU,white}}-\gamma\right] \quad (29)\end{aligned}$$

where we have observed that $\ln(\mathsf{ASNR}_{\text{SU,white}})-\gamma \approx \mathsf{C}_{\text{SU,white}}$ for $\mathsf{ASNR}_{\text{SU,white}} \gg 1$ [see eqs. (1) and (23)]. In the high-SNR regime, the achievable best-case capacity of the SU in the dirty-space mode becomes vanishingly small for $M \to +\infty$. This result is essentially due to the fact that the SU transmits only one symbol per PU frame, in order to assure primary channel stationarity during a frame.

In wireless networks, due to the fluctuation of the instantaneous capacity of fading channels, a buffer is typically used at the transmitter to adapt the source data traffic flow to the channel transmission capability. We will show by means of simple models that the use of portions of frequency band that are being used by the PU leads to a significant advantage for the SU in terms of average symbol delay (i.e., waiting time in queue plus transmission time), even when it has a low physical-layer rate in dirty-space mode. To this end, let us assume that SU data arrive at the buffer in symbols each carrying Q_{SU} bits. We evaluate the average delay D_{SU} for the transmission of a symbol with the proposed protocol and compare it with the average symbol delay $\mathsf{D}_{\text{SU,white}}$ when the SU interrupts its transmission during the ON intervals of the PU, i.e., it transmits in white-space mode only. The service system for the proposed protocol can be modeled as an M/G/1 queue [48], for which the mean of the symbol transmission delay is given by

$$\mathsf{D}_{\text{SU}} = \mathbb{E}[\mathsf{X}_{\text{SU}}] + \frac{\lambda_{\text{SU}}\mathbb{E}\left[\mathsf{X}_{\text{SU}}^2\right]}{2\left(1-\lambda_{\text{SU}}\mathbb{E}[\mathsf{X}_{\text{SU}}]\right)} \quad (30)$$

where X_{SU} is the symbol service time of the proposed protocol and λ_{SU} (in symbols for second) is the average arrival rate. On the other hand, when the SU uses all its available power to transmit data only when the PU is inactive, we model the service system as an M/G/1 queue with vacations [48] and, in this case, the mean of the symbol transmission delay can be found as

$$\mathsf{D}_{\text{SU,white}} = \mathbb{E}[\mathsf{X}_{\text{SU,white}}] + \frac{\lambda_{\text{SU}}\mathbb{E}\left[\mathsf{X}_{\text{SU,white}}^2\right]}{2\left(1-\lambda_{\text{SU}}\mathbb{E}[\mathsf{X}_{\text{SU,white}}]\right)} + \frac{\mathbb{E}\left[\mathsf{T}_{\text{ON}}^2\right]}{2\mathbb{E}[\mathsf{T}_{\text{ON}}]} \quad (31)$$

[9] Such an approximation turns out to be an equality if the PU adopts a constant-modulus constellation with average energy P_{PU}.

$$\mathsf{C}_{\text{SU,dirty,lower}} = \frac{1}{M}\mathbb{E}_{\mathbf{g}_{\text{SU}}}\left\{\ln\left[1+M\mathsf{P}_{\text{PU}}^2\sigma_{\text{SU,dirty}}^2\mathbb{E}_{x_{\text{SU}}|\mathbf{g}_{\text{SU}}}\left\{\frac{\left||g_{24}|^2|g_{12}|^2x_{\text{SU}}+g_{24}^*g_{12}^*g_{14}\right|^2}{\left[\Omega(x_{\text{SU}},\mathbf{g}_{\text{SU}})+\Upsilon^*(x_{\text{SU}},\mathbf{g}_{\text{SU}})+\Upsilon(x_{\text{SU}},\mathbf{g}_{\text{SU}})\right]^2}\right\}\right]\right\}. \quad (26)$$

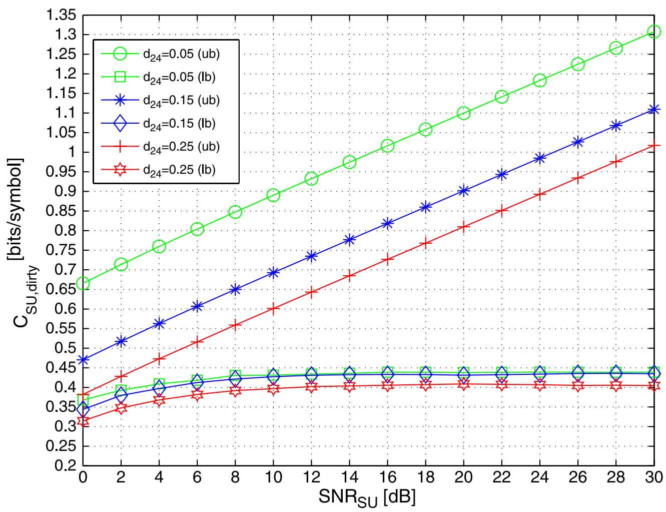

Fig. 4. $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ versus $\mathsf{SNR}_{\mathrm{SU}}$ for different values of d_{24} ($d_{12} = 0.5$, $\mathsf{P}_{\mathrm{PU}} = \mathsf{P}_{\mathrm{SU}} = 1$).

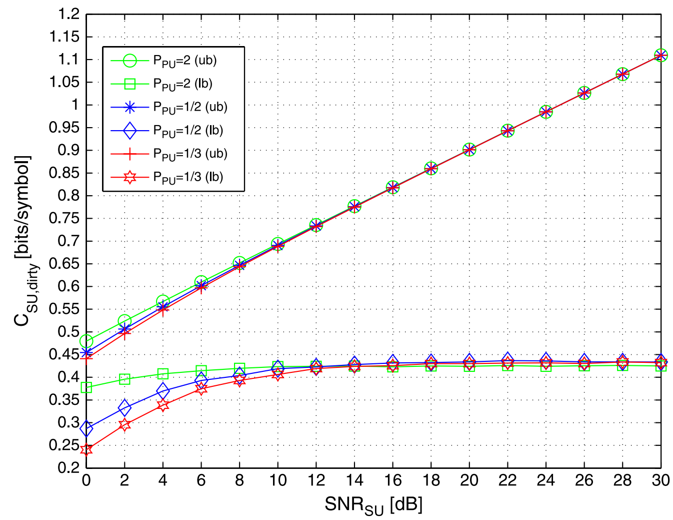

Fig. 5. $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ versus $\mathsf{SNR}_{\mathrm{SU}}$ for different values of P_{PU} ($d_{12} = 0.5$, $d_{24} = 0.15$, $\mathsf{P}_{\mathrm{SU}} = 1$).

where $\mathsf{X}_{\mathrm{SU,white}}$ is the symbol service time in white-space mode and T_{ON} is the duration of the vacations (i.e., PU ON intervals) taken by the SU transmitter.

To evaluate the ensemble averages involved in (30) and (31), since the PU transmission protocol works on a frame-by-frame basis and a PU frame lasts $\mathsf{T}_{\mathrm{SU,dirty}}$ seconds, we assume that $\mathsf{T}_{\mathrm{ON}} = \mathsf{I}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,dirty}}$, where I_{SU} is a geometric random variable with success probability p_{ON} and range $\{0, 1, 2, \ldots\}$. As a consequence, it results [44] that $\mathbb{E}[\mathsf{I}_{\mathrm{SU}}] = (1 - \mathsf{p}_{\mathrm{ON}})/\mathsf{p}_{\mathrm{ON}}$ and $\mathbb{E}[\mathsf{I}_{\mathrm{SU}}^2] = [(1 - \mathsf{p}_{\mathrm{ON}})(2 - \mathsf{p}_{\mathrm{ON}})]/\mathsf{p}_{\mathrm{ON}}^2$. Regarding the first and second moments of X_{SU} and $\mathsf{X}_{\mathrm{SU,white}}$, we assume that the SU can transmit at rates near the information-theoretic limits in both white- and dirty-space modes and, hence, we use the ergodic capacity to measure the data transmission capability of the wireless link.[10] Specifically, for the proposed protocol, we assume that the packet service time X_{SU} is a discrete binary random variable assuming the values

$$\mathsf{x}_{\mathrm{SU,dirty}} \triangleq (\mathsf{Q}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,dirty}}) \big/ \left[\log_2(e)\mathsf{C}_{\mathrm{SU,dirty}}\right]$$

$$\mathsf{x}_{\mathrm{SU,white}} \triangleq (\mathsf{Q}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,white}}) \big/ \left[\log_2(e)\mathsf{C}_{\mathrm{SU,white}}\right]$$

with probabilities P_{dirty} and $1 - P_{\mathrm{dirty}}$, respectively, where P_{dirty} is the probability that the PU channel is busy. In white-space mode, we assume that the service time $\mathsf{X}_{\mathrm{SU,white}}$ is deterministic and it is given by $\mathsf{x}_{\mathrm{SU,white}}$.

To support the performance analysis of the SU, we report the results of numerical simulations. Specifically, we plot in Figs. 4 and 5 the upper and lower bounds on $\mathsf{C}_{\mathrm{SU,dirty}}$ (reported in bits/symbol and referred to as “ub” and “lb,” respectively), whereas the *worst-case* value $(\Delta\mathsf{D}_{\mathrm{SU}})_{\min}$ of the difference $\Delta\mathsf{D}_{\mathrm{SU}} \triangleq \mathsf{D}_{\mathrm{SU,white}} - \mathsf{D}_{\mathrm{SU}}$ is reported in Fig. 6 by replacing $\mathsf{C}_{\mathrm{SU,dirty}}$ with $\mathsf{C}_{\mathrm{SU,dirty,lower}}$. The ensemble averages in (25)

[10]Depending on the dynamics of the fading process, a long coding delay (i.e., the amount of time required to encode/decode packets) may be required to approach ergodic capacity [33]. However, since the SU uses the same encoding/decoding strategy in both white- and dirty-space modes, such a coding delay does not affect the comparison between D_{SU} and $\mathsf{D}_{\mathrm{SU,white}}$.

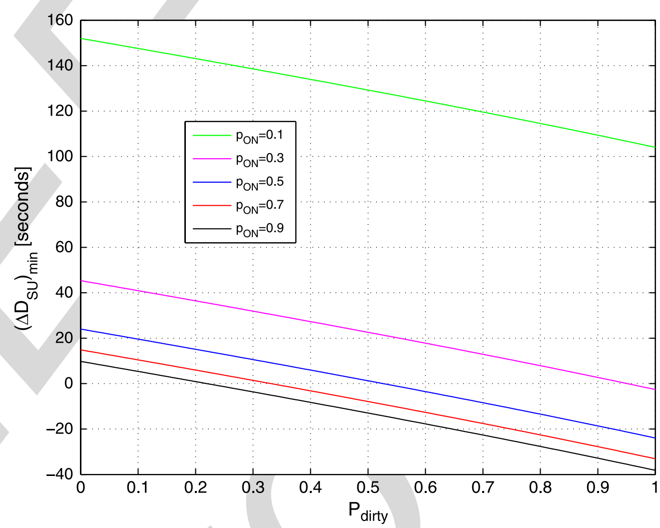

Fig. 6. $(\Delta\mathsf{D}_{\mathrm{SU}})_{\min}$ versus P_{dirty} for different values of p_{ON} ($\mathsf{SNR}_{\mathrm{SU}} =$ 20 dB, $d_{24}/d_{12} = 0.3$, $\mathsf{P}_{\mathrm{PU}} = 1$, $\lambda_{\mathrm{SU}} = 0.01$).

are evaluated through 10^4 Monte Carlo trials. With reference to Fig. 1, the distance between the PTx and the STx is fixed to $d_{12} = 0.5$ and $\vartheta = \pi/3$, whereas the transmitting power of the SU is normalized by setting $\mathsf{P}_{\mathrm{SU}} = 1$; moreover, we chose $M = 16$, $\tau_{\mathrm{SU}} = 2$, $\mathsf{Q}_{\mathrm{SU}} = 1$ (i.e., binary modulation), $\mathsf{T}_{\mathrm{SU,white}} = \mathsf{T}_{\mathrm{PU}} = 1$, and set the path-loss exponent equal to $\eta = 2$. The PU symbols $x_{\mathrm{PU}}(\cdot)$ were generated as circularly symmetric complex Gaussian RVs with average energy P_{PU} and, similarly, the SU symbol x_{SU} was modeled as a circularly symmetric complex Gaussian RV with variance $\sigma^2_{\mathrm{SU,dirty}}$ and $\sigma^2_{\mathrm{SU,white}}$ in dirty- and white-space mode, respectively.

3) Fig. 4 : It reports the upper and lower bounds on the SU capacity as a function of $\mathsf{SNR}_{\mathrm{SU}} \triangleq \mathsf{P}_{\mathrm{SU}}/\sigma^2_{n_4}$ for different values of the distance d_{24} between the STx and the SRx (the distance d_{14} between the PTx and the SRx is calculated according to the Carnot’s cosine law), with $\mathsf{P}_{\mathrm{PU}} = 1$. It is seen that the best performance significantly increases as the SRx brings nearer to the STx. For instance, when the distance d_{24} between the STx and the SRx is about one-third of the distance d_{14} between the PTx and the STx, the achievable rate is greater than or equal to one bit per symbol for $\mathsf{SNR}_{\mathrm{SU}} \geq 24$ dB. On the other hand, the worst performance is less dependent on d_{24}. There is a significant gap between the best and the worst performances of the SU, thus evidencing that decoding of the PU symbols at the SRx is important to achieve reasonable rates. In particular, contrary to $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, the worst-case capacity $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ in (25) does not grow without bound as $\sigma^2_{n_4} \to 0$, thus exhibiting a marked floor in the high-SNR region. Such a behavior is due to the fact that the $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ is derived under the assumption that the PU symbols are unknown at the SRx (they are modeled as Gaussian RVs).

4) Fig. 5 : The upper and lower bounds of the SU capacity are depicted as a function of $\mathsf{SNR}_{\mathrm{SU}}$ for different values of the PU transmitting power P_{PU}, with $d_{24}/d_{12} = 0.3$. Differently from the PU case (see Fig. 3), increasing the power ratio $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ has negligible impact on the SU capacity, except for very small values of $\mathsf{SNR}_{\mathrm{SU}}$. In accordance with (29), the upper bound $\mathsf{C}_{\mathrm{SU,dirty,upper}}$ does not depend on P_{PU} at all for moderate-to-high SNR values, since in this case the SRx has perfect knowledge of the PU symbols; on the other hand, the lower bound $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ exhibits a weak dependence on P_{PU}.

5) Fig. 6 : It reports the worst-case difference $(\Delta\mathsf{D}_{\mathrm{SU}})_{\min}$ between the average delays (31) and (30) as a function of the probability P_{dirty} that the PU channel is busy, for different values of the success probability p_{ON}, which lead to different values of $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}]/\mathsf{T}_{\mathrm{SU,dirty}} = (1-\mathsf{p}_{\mathrm{ON}})/\mathsf{p}_{\mathrm{ON}}$. Results of Fig. 6 are obtained by setting $\mathsf{P}_{\mathrm{PU}} = 1$, $\mathsf{SNR}_{\mathrm{SU}} = 20$ dB, $d_{24}/d_{12} = 0.3$, and $\lambda_{\mathrm{SU}} = 0.01$ symbol/seconds. It is seen that transmitting at a low data rate in dirty-space mode allows to significantly reduce the average delay incurred by the SU data, for each value of P_{dirty}, when $\mathsf{p}_{\mathrm{ON}} < 0.3$, i.e., the mean duration $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}]$ of the PU ON intervals is roughly 2.4 times greater than the PU frame duration $\mathsf{T}_{\mathrm{SU,dirty}}$. On the other hand, for $\mathsf{p}_{\mathrm{ON}} \geq 0.3$, the proposed protocol ensures a delay improvement for low-to-moderate values of P_{dirty}: for instance, when $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}] = \mathsf{T}_{\mathrm{SU,dirty}}$, that is, $\mathsf{p}_{\mathrm{ON}} = 0.5$, an improvement can be observed when the PU channel is occupied for about less than 50% of time.

V. Conclusion

We proposed an AF scheme that allows a SU to concurrently transmit in the same frequency band of a PU not only when the PU is inactive, but also when the PU channel is busy. This can be obtained without requiring any noncausal knowledge of the PU information symbols. The main results of our performance analyses in terms of both PU and SU ergodic channel capacities can be summarized as follows. The cognitive radio principle of protecting the PU is not only guaranteed, but even a performance improvement can be gained by the PU in terms of ergodic channel capacity. Such a performance gain depends on the distance ratio d_{12}/d_{13}, the path-loss exponent η, and the power ratios $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ and $\sigma^2_{n_2}/\sigma^2_{n_3}$. Regarding the achievable rate of the SU, the secondary link can support more than one bit per symbol for moderate-to-high SNR values if the SRx is able to decode the PU data. Moreover, a variation of the power ratio $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ does not lead to appreciable effects for the SU. Notwithstanding the transmission of a single SU symbol per frame gives low information rates in dirty-space mode, the delay performance of the SU improves noticeably.

It is noteworthy that several PUs typically multiplex the frame resources in time or in frequency. Therefore, the SU might be active in parallel over all the channels allocated for the PUs, potentially attaining larger transmission rates without adding interference. Modification of the proposed protocol and evaluation of the corresponding capacity performance is left as a future development.

Appendix A
Upper Bound on the PU Ergodic Channel Capacity

An upper bound on C_{PU} can be obtained by assuming that the PRx additionally has perfect knowledge of the fading coefficient g_{12} characterizing the PTx $\to$ STx link. Indeed, let $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12})$ denote the conditional mutual information between $\tilde{x}_{\mathrm{PU}}(\ell)$ and $\tilde{y}_{\mathrm{PU}}(\ell)$, given $\mathbf{g}_{\mathrm{PU}}$ and g_{12}, by using the chain rule for mutual information [31], it can be proven that

$$\begin{aligned}\mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12}\right) &= \mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}\right) \\ &+ \underbrace{\mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}\right)}_{\geq 0} \geq \mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}\right) \qquad (32)\end{aligned}$$

where the equality holds if and only if $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}) = 0$, i.e., when $\tilde{x}_{\mathrm{PU}}(\ell)$ and g_{12} are conditionally independent given $\tilde{y}_{\mathrm{PU}}(\ell)$ and $\mathbf{g}_{\mathrm{PU}}$.[11]

For given values of $\mathbf{g}_{\mathrm{PU}}$ and g_{12}, the equivalent noise term $v_{\mathrm{PU}}(\ell)$ in (5) is a circular symmetric zero-mean complex Gaussian RV with variance $\sigma^2_{n_3} + \sigma^2_{n_2}|g_{23}|^2|x_{\mathrm{SU}}|^2$. Consequently, the vector $\tilde{\mathbf{v}}_{\mathrm{PU}}$ in (11) is composed of i.i.d. $\mathcal{CN}(0, \sigma^2_{n_3} + \sigma^2_{n_2}|g_{23}|^2|x_{\mathrm{SU}}|^2)$ RVs and, thus, the subchannels (9) are also Gaussian. In this case, the supremum of $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12})$ over all distributions on $\tilde{x}_{\mathrm{PU}}(\ell)$ that satisfy the power constraint $\sigma^2_{\mathrm{PU}} = \mathsf{P}_{\mathrm{PU}}$ is attained [31] when $\tilde{x}_{\mathrm{PU}}(\ell) \equiv \tilde{x}_{\mathrm{PU,G}}(\ell) \sim \mathcal{CN}(0, \mathsf{P}_{\mathrm{PU}})$, regardless of the operating SNR. With this choice, one obtains

$$\begin{aligned}&\mathsf{I}\left(\tilde{x}_{\mathrm{PU,G}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12}\right) \\ &\qquad = \mathbb{E}_{\mathbf{g}_{\mathrm{PU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\mathrm{PU}}\,|G_{\mathrm{PU}}(\ell)|^2}{\sigma^2_{n_3} + \sigma^2_{n_2}|g_{23}|^2|x_{\mathrm{SU}}|^2}\right]\right\} \qquad (33)\end{aligned}$$

and, accounting for (32), the ergodic capacity (nats/symbol) of the parallel fading channel (9) is upper bounded by

$$\mathsf{C}_{\mathrm{PU}}(M) \leq \frac{1}{M}\sum_{\ell=0}^{M-1}\mathbb{E}_{\mathbf{g}_{\mathrm{PU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\mathrm{PU}}|G_{\mathrm{PU}}(\ell)|^2}{\sigma^2_{n_3} + \sigma^2_{n_2}|g_{23}|^2|x_{\mathrm{SU}}|^2}\right]\right\}. \qquad (34)$$

[11] The fact that $\tilde{x}_{\mathrm{PU}}(\ell)$ and g_{12} are statistically independent neither implies nor is implied by $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}) = 0$.

To evaluate the expectation in (34), it is useful to observe that, conditioned on $g_{23}x_{\rm SU}$, one has $G_{\rm PU}(\ell)|g_{23}x_{\rm SU} \sim \mathcal{CN}(0, \sigma_{13}^2 + \sigma_{12}^2|g_{23}|^2|x_{\rm SU}|^2)$, $\forall \ell \in \{0, 1, \ldots, M-1\}$, whose squared magnitude is exponentially distributed with mean $\sigma_{13}^2 + \sigma_{12}^2|g_{23}|^2|x_{\rm SU}|^2$. Thus, using (1), it follows that:

$$\mathbb{E}_{\mathbf{g}_{\rm PU}|g_{23}x_{\rm SU}}\left\{\ln\left[1+\frac{\mathsf{P}_{\rm PU}\,|G_{\rm PU}(\ell)|^2}{\sigma_{n_3}^2+\sigma_{n_2}^2|g_{23}|^2|x_{\rm SU}|^2}\right]\right\} = f\left[\Gamma_{3,\rm upper}\left(|g_{23}|^2|x_{\rm SU}|^2\right)\right] \quad (35)$$

where $\Gamma_{3,\rm upper}(|g_{23}|^2|x_{\rm SU}|^2)$ is a transformation of the RV $g_{23}x_{\rm SU}$ defined in (15). Therefore, the terms of the sum in (34) do not depend on ℓ and, thus, capacity (34) does not depend on M. Hence, by applying the conditional expectation rule in (34) and using (35), one obtains the upper bound in (14) on the ergodic channel capacity of the PU [see eq. (12)].

Appendix B
Lower Bound on the PU Ergodic Channel Capacity

To find a lower bound on $\mathsf{C}_{\rm PU}$, we observe that the Gaussian distribution might not be the one maximizing $\mathsf{I}(\tilde{x}_{\rm PU}(\ell); \tilde{y}_{\rm PU}(\ell)|\mathbf{g}_{\rm PU})$ [see eqs. (12), (9)] and, thus, we choose $\tilde{x}_{\rm PU}(\ell) \equiv \tilde{x}_{\rm PU,G}(\ell) \sim \mathcal{CN}(0, \mathsf{P}_{\rm PU})$. Moreover, let $v_{\rm PU,G}(\cdot)$ be a sequence of i.i.d. circularly symmetric complex Gaussian RVs having the same mean and variance as $v_{\rm PU}(\cdot)$ in (5), i.e., $v_{\rm PU,G}(\cdot) \sim \mathcal{CN}\left(0, \sigma_{n_3}^2 + \sigma_{n_2}^2\sigma_{23}^2\sigma_{\rm SU,dirty}^2\right)$, independent of $\tilde{x}_{\rm PU,G}(\cdot)$. Let us also consider the frequency-domain vector

$$\tilde{\mathbf{v}}_{\rm PU,G} = \left[\tilde{v}_{\rm PU,G}(0), \tilde{v}_{\rm PU,G}(1), \ldots, \tilde{v}_{\rm PU,G}(M-1)\right]^{\rm T} \triangleq \mathbf{W}_{\rm DFT}\mathbf{v}_{\rm PU,G} \in \mathbb{C}^M$$

with $\mathbf{v}_{\rm PU,G} \triangleq [v_{\rm PU,G}(0), v_{\rm PU,G}(1), \ldots, v_{\rm PU,G}(M-1)]^{\rm T} \in \mathbb{C}^M$. It is readily seen that the entries of $\tilde{\mathbf{v}}_{\rm PU,G}$ are i.i.d. $\mathcal{CN}\left(0, \sigma_{n_3}^2 + \sigma_{n_2}^2\sigma_{23}^2\sigma_{\rm SU,dirty}^2\right)$ RVs. By replacing $\tilde{x}_{\rm PU}(\ell)$ and $\tilde{v}_{\rm PU}(\ell)$ in (9) with $\tilde{x}_{\rm PU,G}(\ell)$ and $\tilde{v}_{\rm PU,G}(\ell)$, respectively, we get the subchannels with both Gaussian input and Gaussian noise:

$$\tilde{y}_{\rm PU,G}(\ell) = G_{\rm PU}(\ell)\tilde{x}_{\rm PU,G}(\ell) + \tilde{v}_{\rm PU,G}(\ell)$$

for $\ell \in \{0, 1, \ldots, M-1\}$. Additionally, let

$$\tilde{y}_{\rm PU,NG}(\ell) = G_{\rm PU}(\ell)\tilde{x}_{\rm PU,G}(\ell) + \tilde{v}_{\rm PU}(\ell)$$

be the corresponding subchannels with Gaussian input and non-Gaussian noise, $\ell \in \{0, 1, \ldots, M-1\}$. Since conditional mutual information can be equivalently expressed as difference between conditional differential entropies [31], one obtains

$$\begin{aligned}\mathsf{I}\left(\tilde{x}_{\rm PU,G}(\ell); \tilde{y}_{\rm PU,NG}(\ell)|\mathbf{g}_{\rm PU}\right) - \mathsf{I}\left(\tilde{x}_{\rm PU,G}(\ell); \tilde{y}_{\rm PU,G}(\ell)|\mathbf{g}_{\rm PU}\right) &= \mathsf{h}\left(\tilde{x}_{\rm PU,G}(\ell)|\tilde{y}_{\rm PU,G}(\ell), \mathbf{g}_{\rm PU}\right)\\ &\quad - \mathsf{h}\left(\tilde{x}_{\rm PU,G}(\ell)|\tilde{y}_{\rm PU,NG}(\ell), \mathbf{g}_{\rm PU}\right) \geq 0 \end{aligned} \quad (36)$$

where the inequality holds for each realization of $\mathbf{g}_{\rm PU}$ and whatever is the probability distribution of $\tilde{v}_{\rm PU}(\ell)$.[12] Consequently, we have $\mathsf{I}(\tilde{x}_{\rm PU,G}(\ell); \tilde{y}_{\rm PU,NG}(\ell)|\mathbf{g}_{\rm PU}) \geq \mathsf{I}(\tilde{x}_{\rm PU,G}(\ell); \tilde{y}_{\rm PU,G}(\ell)|\mathbf{g}_{\rm PU})$ and, by doing calculations similar to those of Appendix A [see, in particular, eqs. (33) and (35)], we get the lower bound (in nats/symbol) on the ergodic channel capacity of the PU [see eq. (12)]

$$\begin{aligned}\mathsf{C}_{\rm PU} \geq \mathsf{C}_{\rm PU,lower} &\triangleq \lim_{M\to+\infty}\frac{1}{M}\sum_{\ell=0}^{M-1}\mathsf{I}\left(\tilde{x}_{\rm PU,G}(\ell); \tilde{y}_{\rm PU,G}(\ell)|\mathbf{g}_{\rm PU}\right)\\ &= \lim_{M\to+\infty}\frac{1}{M}\sum_{\ell=0}^{M-1}\mathbb{E}_{\mathbf{g}_{\rm PU}}\left\{\ln\left[1+\frac{\mathsf{P}_{\rm PU}|G_{\rm PU}(\ell)|^2}{\sigma_{n_3}^2+\sigma_{n_2}^2\sigma_{23}^2\sigma_{\rm SU,dirty}^2}\right]\right\}\\ &= \mathbb{E}\left\{f\left[\Gamma_{3,\rm lower}\left(|g_{23}|^2|x_{\rm SU}|^2\right)\right]\right\}\end{aligned}$$

where $\Gamma_{3,\rm lower}(|g_{23}|^2|x_{\rm SU}|^2)$ is a transformation of the RV $g_{23}x_{\rm SU}$ defined in (16).

Appendix C
Upper Bound on the SU Ergodic Channel Capacity

An upper bound on $\mathsf{C}_{\rm SU,dirty}$ can be obtained by assuming that the SRx has perfect knowledge of both the realization of the channel vector $\mathbf{g}_{\rm SU}$ and the PU symbol block $\check{\mathbf{x}}_{\rm PU}$. In this case, the channel output consists of the triplet $(\mathbf{y}_{\rm SU}, \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})$ and, thus, the mutual information between channel input and output (in nats/PU frame) is represented by $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}, \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})$. Owing to the statistical independence among $x_{\rm SU}$, $\mathbf{g}_{\rm SU}$, and $\check{\mathbf{x}}_{\rm PU}$, which implies that $\mathsf{I}(x_{\rm SU}; \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) = 0$, application of the chain rule [31] for mutual information allows one to write that $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}, \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) = \mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})$, i.e., it is equal to the conditional mutual information between $x_{\rm SU}$ and $\mathbf{y}_{\rm SU}$, given $\mathbf{g}_{\rm SU}$ and $\check{\mathbf{x}}_{\rm PU}$. Similarly to (32), it results that

$$\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) = \mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}) + \underbrace{\mathsf{I}(x_{\rm SU}; \check{\mathbf{x}}_{\rm PU}|\mathbf{y}_{\rm SU}, \mathbf{g}_{\rm SU})}_{\geq 0} \geq \mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}). \quad (37)$$

Consequently, an upper bound $\mathsf{C}_{\rm SU,dirty,upper}$ (in nats/symbol) on the ergodic channel capacity of the SU link is obtained by taking the supremum of $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})/M$ over all possible distributions of the symbol $x_{\rm SU}$ that satisfy the power constraint (3). Since conditional mutual information can be equivalently expressed as difference between conditional differential entropies [31], one has

$$\begin{aligned}\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) &= \mathsf{h}(\mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})\\ &\quad - \mathsf{h}(\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) = \mathsf{h}(\mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})\\ &\quad - \mathsf{h}(\mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}\,|x_{\rm SU}, \mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})\\ &= \mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})\end{aligned} \quad (38)$$

[12]The inequality is a consequence of two facts: (i) given $G_{\rm PU}(\ell)$, the RVs $\tilde{x}_{\rm PU,G}(\ell)$, and $G_{\rm PU}(\ell)\tilde{x}_{\rm PU,G}(\ell) + \tilde{v}_{\rm PU}(\ell)$ are jointly circularly symmetric complex Gaussian since $\tilde{x}_{\rm PU,G}(\ell)$ and $\tilde{v}_{\rm PU}(\ell)$ are independent of each other and each one is circularly symmetric complex Gaussian; (ii) jointly Gaussian RVs maximize conditional differential entropy [31], [49].

$$\mathsf{MSE}(\mathbf{g}_{\mathrm{SU}}) \geq J^{-1}(\mathbf{g}_{\mathrm{SU}}) \triangleq \left\{ \mathbb{E}_{\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}} \left[\left| \frac{\partial}{\partial x_{\mathrm{SU}}^*} \ln p(\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}) \right|^2 \right] \right\}^{-1} \tag{41}$$

$$J(\mathbf{g}_{\mathrm{SU}}) = \mathbb{E}_{x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}} \left\{ \mathbb{E}_{\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}} \left[\left| \frac{\partial}{\partial x_{\mathrm{SU}}^*} \ln p(\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \right|^2 \middle| x_{\mathrm{SU}} \right] \right\} + \mathbb{E}_{x_{\mathrm{SU}}} \left[\left| \frac{\partial}{\partial x_{\mathrm{SU}}^*} \ln p(x_{\mathrm{SU}}) \right|^2 \right] \tag{44}$$

where we have used the fact that subtracting a constant does not change differential entropy [31]. Strictly speaking, since SRx knows g_{14} and $\check{\mathbf{x}}_{\mathrm{PU}}$, it can decode x_{SU} by subtracting $g_{14}\mathbf{x}_{\mathrm{PU}}$ from (24), which amounts to a repetition coding [37] transmission scheme over a fast flat-fading channel with i.i.d. $\mathcal{CN}(0, \sigma_{n_4}^2)$ noise samples. According to (24), conditioned on $\mathbf{g}_{\mathrm{SU}}$ and $\check{\mathbf{x}}_{\mathrm{PU}}$, a sufficient statistic for detecting x_{SU} from $\mathbf{y}_{\mathrm{SU}} - g_{14}\mathbf{x}_{\mathrm{PU}}$ is given by the scalar

$$\begin{aligned} \check{y}_{\mathrm{SU}} &\triangleq (g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\mathrm{PU}})^{\mathrm{H}}(\mathbf{y}_{\mathrm{SU}} - g_{14}\mathbf{x}_{\mathrm{PU}}) \\ &= \|g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\mathrm{PU}}\|^2 x_{\mathrm{SU}} + (g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\mathrm{PU}})^{\mathrm{H}}\mathbf{v}_4 \end{aligned} \tag{39}$$

which is interpreted as an AWGN channel with SNR equal to $(\sigma_{\mathrm{SU,dirty}}^2 |g_{24}|^2 |g_{12}|^2 \sum_{m=0}^{M-1} |x_{\mathrm{PU}}(m-\tau_{\mathrm{SU}})|^2)/\sigma_{n_4}^2$. Since sufficient statistics preserve mutual information [31], one has $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}} - g_{14}\mathbf{x}_{\mathrm{PU}}|\mathbf{g}_{\mathrm{SU}}, \check{\mathbf{x}}_{\mathrm{PU}}) = \mathsf{I}(x_{\mathrm{SU}}; \check{y}_{\mathrm{SU}}\,|\mathbf{g}_{\mathrm{SU}}, \check{\mathbf{x}}_{\mathrm{PU}})$. The supremum of $\mathsf{I}(x_{\mathrm{SU}}; \check{y}_{\mathrm{SU}}\,|\mathbf{g}_{\mathrm{SU}}, \check{\mathbf{x}}_{\mathrm{PU}})$ over all distributions on x_{SU} satisfying (3) is attained [31] when $x_{\mathrm{SU}} \equiv x_{\mathrm{SU,G}} \sim \mathcal{CN}\left(0, \sigma_{\mathrm{SU,dirty}}^2\right)$, thus leading to the upper bound in (25).

APPENDIX D
LOWER BOUND ON THE SU ERGODIC CHANNEL CAPACITY

A general lower bound on $\mathsf{C}_{\mathrm{SU,dirty}}$ can be derived by linking the mutual information $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ to the conditional symbol estimation error of the SU. Specifically, the SRx uses the observation vector $\mathbf{y}_{\mathrm{SU}}$ in (24) to produce a reliable estimate $\hat{x}_{\mathrm{SU}}$ of the symbol x_{SU}. By virtue of the data processing theorem [31], any function of the channel output $\mathbf{y}_{\mathrm{SU}}$ cannot increase the information about x_{SU}, i.e., $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}) \geq \mathsf{I}(x_{\mathrm{SU}}; \hat{x}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$, and using a technique similar to that used in [50], the following general lower bound on $\mathsf{I}(x_{\mathrm{SU}}; \hat{x}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ can be obtained

$$\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}) \geq \mathsf{h}(x_{\mathrm{SU}}) - \mathbb{E}_{\mathbf{g}_{\mathrm{SU}}}\left\{\ln\left[\pi e \mathsf{MSE}(\mathbf{g}_{\mathrm{SU}})\right]\right\} \tag{40}$$

where, according to the maximum-entropy theorem for complex RVs [51], the inequality holds with equality if and only if $x_{\mathrm{SU}} - \hat{x}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}} \sim \mathcal{CN}[0, \mathsf{MSE}(\mathbf{g}_{\mathrm{SU}})]$, with $\mathsf{MSE}(\mathbf{g}_{\mathrm{SU}}) \triangleq \mathbb{E}_{\mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}}[|x_{\mathrm{SU}} - \hat{x}_{\mathrm{SU}}|^2]$ denoting the mean-square error (MSE) of the symbol estimate, given $\mathbf{g}_{\mathrm{SU}}$. Interestingly, eq. (40) provides a lower bound on $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ that depends on the conditional mean-square value of the symbol estimation error and applies for *any* symbol estimation strategy.

By modeling x_{SU} as a random variable with a given *a priori* pdf $p(x_{\mathrm{SU}})$, whose particular realization has to be estimated, a lower bound on $\mathsf{MSE}(\mathbf{g}_{\mathrm{SU}})$ is given by the complex counterpart of the Bayesian Cramér-Rao inequality [46], [52] that is reported in (41), shown at the top of the page, where $p(\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ is the conditional joint pdf of $\mathbf{y}_{\mathrm{SU}}$ and x_{SU}, given $\mathbf{g}_{\mathrm{SU}}$, whereas $J(\mathbf{g}_{\mathrm{SU}})$ is referred to as the Bayesian Fisher information.[13] The lower bound (41) is valid for any $p(\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ satisfying some regularity conditions in addition to a "weak unbiasedness" condition [52], which are typically fulfilled by Gaussian distributions [46]. The lower bound (40) is valid for any $\mathsf{MSE}(\mathbf{g}_{\mathrm{SU}})$ and, therefore, it also holds when $\mathsf{MSE}(\mathbf{g}_{\mathrm{SU}})$ is replaced with its minimum value $J^{-1}(\mathbf{g}_{\mathrm{SU}})$ given by (41), thus having

$$\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}) \geq \mathsf{h}(x_{\mathrm{SU}}) - \ln(\pi e) + \mathbb{E}_{\mathbf{g}_{\mathrm{SU}}}\left\{\ln\left[J(\mathbf{g}_{\mathrm{SU}})\right]\right\} \tag{42}$$

which is actually a lower bound if an efficient symbol estimator exists, i.e., it attains the Bayesian CRB.

To facilitate the derivation of the Fisher information $J(\mathbf{g}_{\mathrm{SU}})$ in a closed form, the PU symbols are modelled as $\check{\mathbf{x}}_{\mathrm{PU}} \equiv \check{\mathbf{x}}_{\mathrm{PU,G}} \sim \mathcal{CN}(\mathbf{0}_M, \mathsf{P}_{\mathrm{PU}}\mathbf{I}_{M+\tau_{\mathrm{SU}}})$. In such a case, it is verified that, given x_{SU} and $\mathbf{g}_{\mathrm{SU}}$, one has $\mathbf{y}_{\mathrm{SU}} \sim \mathcal{CN}[\mathbf{0}_M, \mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})]$ [see eq. (24)], where

$$\begin{aligned} \mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) &\triangleq \mathbb{E}\left[\mathbf{y}_{\mathrm{SU}}\mathbf{y}_{\mathrm{SU}}^{\mathrm{H}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}\right] = \Omega(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})\mathbf{I}_M \\ &\quad + \Upsilon^*(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})\mathbf{B}^{\tau_{\mathrm{SU}}} + \Upsilon(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})\mathbf{F}^{\tau_{\mathrm{SU}}} \end{aligned} \tag{43}$$

is a tridiagonal Hermitian Toeplitz matrix, with $\Omega(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \triangleq (|g_{24}|^2|g_{12}|^2|x_{\mathrm{SU}}|^2 + |g_{14}|^2)\mathsf{P}_{\mathrm{PU}} + \sigma_{n_4}^2$, $\Upsilon(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \triangleq g_{24}\, g_{12}\, x_{\mathrm{SU}} g_{14}^* \mathsf{P}_{\mathrm{PU}}$, $\mathbf{B} \in \mathbb{R}^{M\times M}$ and $\mathbf{F} \in \mathbb{R}^{M\times M}$ being backward-shift and forward-shift matrices [53], respectively.[14] By applying the conditional expectation rule, using the fact that $p(\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}) = p(\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})p(x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$, remembering that x_{SU} and $\mathbf{g}_{\mathrm{SU}}$ are statistically independent, and accounting for the regularity conditions [46], [52], one obtains (44), shown at the top of the page, where, since $\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}} \sim \mathcal{CN}[\mathbf{0}_M, \mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})]$, it follows that (see, e.g., [45])

$$\begin{aligned} &\mathbb{E}_{x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}} \left\{ \mathbb{E}_{\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}} \left[\left| \frac{\partial}{\partial x_{\mathrm{SU}}^*} \ln p(\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \right|^2 |x_{\mathrm{SU}} \right] \right\} \\ &\quad = \mathbb{E}_{x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}} \left\{ \mathrm{tr}\left[\mathbf{R}_{\mathrm{SU}}^{-1}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \boldsymbol{\Sigma}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \right.\right. \\ &\qquad \left.\left. \cdot\, \mathbf{R}_{\mathrm{SU}}^{-1}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \boldsymbol{\Sigma}_{\mathrm{SU}}^{\mathrm{H}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \right] \right\} \end{aligned} \tag{45}$$

[13] Inequality (41) also holds for biased estimators, in contrast to the standard Cramér-Rao bound (CRB) where the Fisher information is obtained from (41) by replacing $p(\mathbf{y}_{\mathrm{SU}}, x_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}})$ with $p(\mathbf{y}_{\mathrm{SU}}|x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})$ and carrying out the ensemble average with respect to $\mathbf{y}_{\mathrm{SU}}|\mathbf{g}_{\mathrm{SU}}$.

[14] It results that $\mathbf{B}^{\mathrm{T}} = \mathbf{F}$ by construction [53].

with [see eq. (43)]

$$\mathbf{\Sigma}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \triangleq \partial/\partial x_{\mathrm{SU}}^{*} \mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) = \left(|g_{24}|^2 |g_{12}|^2 x_{\mathrm{SU}} \mathbf{I}_M + g_{24}^{*} g_{12}^{*} g_{14} \mathbf{B}^{\tau_{\mathrm{SU}}}\right) \mathsf{P}_{\mathrm{PU}} \tag{46}$$

and we have recalled that, for arbitrary matrices $\mathbf{A} \in \mathbb{C}^{n \times n}$ and $\mathbf{B} \in \mathbb{C}^{n \times n}$, it results [45] $\mathbb{E}[\mathbf{x}^{\mathrm{H}} \mathbf{A}\, \mathbf{x}] = \mathrm{tr}(\mathbf{A}\, \Sigma)$ and, for $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_n, \Sigma)$, $\mathbb{E}[\mathbf{x}^{\mathrm{H}} \mathbf{A}\, \mathbf{x}\, \mathbf{x}^{\mathrm{H}} \mathbf{B} \mathbf{x}] = \mathrm{tr}(\mathbf{A}\Sigma)\mathrm{tr}(\mathbf{B}\Sigma) + \mathrm{tr}(\mathbf{A}\Sigma\mathbf{B}\Sigma)$. Accounting for (42), (44), and (45), the lower bound in (25) comes from the two facts: (i) if $x_{\mathrm{SU}} \equiv x_{\mathrm{SU,G}} \sim \mathcal{CN}(0, \sigma_{\mathrm{SU,dirty}}^2)$, then $\mathsf{h}(x_{\mathrm{SU}}) = \ln(\pi e) + \ln(\sigma_{\mathrm{SU,dirty}}^2)$ [51] and $\mathbb{E}_{x_{\mathrm{SU}}}[|\partial/\partial x_{\mathrm{SU}}^{*} \ln p(x_{\mathrm{SU}})|^2] = 1/\sigma_{\mathrm{SU,dirty}}^2$; (ii) $\mathsf{C}_{\mathrm{SU,dirty}} \geq \mathsf{I}(x_{\mathrm{SU,G}}; \mathbf{y}_{\mathrm{SU}} | \mathbf{g}_{\mathrm{SU}})/M$ since the Gaussian distribution may not maximize $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}} | \mathbf{g}_{\mathrm{SU}})$.

REFERENCES

[1] M. McHenry, E. Livsics, T. Nguyen, and N. Majumdar, "XG dynamic spectrum access field test results [topics in radio communications]," *IEEE Commun. Mag.*, vol. 45, no. 6, pp. 51–57, Jun. 2007.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[3] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. New York, NY, USA: Cambridge Univ. Press, 2013.

[4] Y. Han, A. Pandharipande, and S. H. Ting, "Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4945–4950, Oct. 2009.

[5] R. Manna, R. H. Y. Louie, Y. Li, and B. Vucetic, "Cooperative amplify-and-forward relaying in cognitive radio networks," in *Proc. IEEE CROWNCOM*, Cannes, France, Jun. 2010, pp. 1–5.

[6] D. Hu and S. Mao, "Cooperative relay in cognitive radio networks: Decode-and-forward or amplify-and-forward," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.

[7] Q. Wang, D.-W. Yue, and F. C. M. Lau, "Outage performance and cooperative diversity under amplify and forward relaying in cognitive radio networks," *Wireless Pers. Commun.*, vol. 69, no. 2, pp. 891–914, Apr. 2012.

[8] Y. Han, S. H. Ting, and A. Pandharipande, "Cooperative spectrum sharing protocol with selective relaying system," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 62–67, Jan. 2012.

[9] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I & II," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.

[10] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[11] F. Verde and A. Scaglione, "Randomized space-time block coding for distributed amplify-and-forward cooperative relays," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 3030–3033.

[12] F. Verde, "Design of randomized space-time block codes for amplify-and-forward cooperative relaying," in *Proc. 5th ISCCSP*, Rome, Italy, May 2012, pp. 1–5.

[13] F. Verde, "Performance analysis of randomized space-time block codes for amplify-and-forward cooperative relaying," *IET Commun.*, vol. 67, no. 17, pp. 1883–1898, Nov. 2013.

[14] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Optimization of coherent amplify-and-forward cooperative transmissions in spatially-multiplexed MIMO-OFDM systems," in *Proc. EUCNC*, Bologna, Italy, Jun. 2014, pp. 1–5.

[15] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Error probability analysis of a MIMO system with noncoherent relaying and MMSE reception," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1198–1201, Jul. 2014.

[16] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Performance analysis of amplify-and-forward multiple-relay MIMO systems with ZF reception," *IEEE Trans. Veh. Technol.*, to be published. DOI 10.1109/TVT.2014.2349740.

[17] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Cognitive multiple access via cooperation: Protocol design and performance analysis," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3677–3696, Oct. 2007.

[18] T. Fujii and Y. Suzuki, "Ad-hoc cognitive radio-development to frequency sharing system by using multi-hop network," in *Proc. IEEE Int. Symp. DySPAN*, Baltimore, MD, USA, Nov. 2005, pp. 589–592.

[19] J. Jia, J. Zhang, and Q. Zhang, "Cooperative relay for cognitive radio networks," in *Proc. IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2304–2312.

[20] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. IEEE SIGCOMM*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[21] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[22] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[23] I. Marić, A. Goldsmith, and M. Médard, "Analog network coding in the high-SNR regime," in *Proc. IEEE WiNC*, Boston, MA, USA, Jun. 2010, pp. 1–6.

[24] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, Jun. 2002.

[25] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 563–593, Mar. 2003.

[26] W. Yu, A. Sutivong, D. Julian, T. M. Cover, and M. Chiang, "Writing on colored paper," in *Proc. IEEE ISIT*, Washington, DC, USA, Jun. 2001, p. 302.

[27] A. Khisti, U. Erez, and G. Wornell, "Writing on many pieces of dirty paper at once: The binary case," in *Proc. IEEE ISIT*, Chicago, IL, USA, Jun. 2004, p. 533.

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

[29] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[30] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[31] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[32] F. Verde, A. Scaglione, D. Darsena, and G. Gelli, "An amplify-and-forward scheme for cognitive radios," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 2743–2747.

[33] E. Biglieri, J. G. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2916–2692, Oct. 1998.

[34] L. Ozarow, S. Shamai, and A. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[35] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[36] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2001.

[37] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[38] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3413–3431, Oct. 2007.

[39] R. Ramaswamy, N. Weng, and T. Wolf, "Characterizing network processing delay," in *Proc. IEEE GLOBECOM*, Dallas, TX, USA, Dec. 2004, pp. 1629–1634.

[40] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[41] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receiver in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2999, Sep. 2001.

[42] D. S. Papailiopoulos, G. A. Elkheir, and G. N. Karystinos, "Maximum-likelihood noncoherent PAM detection," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1152–1159, Mar. 2013.

[43] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*. New York, NY, USA: Wiley, 2000.

[44] A. Papoulis, *Probability, Random Variables, and Stochastics Processes*, 3rd ed. Singapore: McGraw-Hill, 1991.

[45] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[46] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[47] R. M. Gray, "Toeplitz and circulant matrices: A review," *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[48] D. Bertsekas and R. G. Gallager, *Data Networks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1992.

[49] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Inf. Control*, vol. 37, no. 1, pp. 34–39, 1978.

[50] A. Vosoughi and A. Scaglione, "On the effect of receiver estimation error upon channel mutual information," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 459–472, Feb. 2006.

[51] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[52] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part I*. New York, NY, USA: Wiley, 1968.

[53] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

Francesco Verde (M'10–SM'14) was born in Santa Maria Capua Vetere, Italy, on June 12, 1974. He received the Dr. Eng. degree (*summa cum laude*) in electronic engineering from the Second University of Napoli, Italy, in 1998, and the Ph.D. degree in information engineering from the University of Napoli Federico II, in 2002. Since December 2002, he has been with the University of Napoli Federico II. He first served as an Assistant Professor of signal theory and mobile communications and, since December 2011, he has served as an Associate Professor of telecommunications with the Department of Electrical Engineering and Information Technology. His research activities include cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems, multicarrier modulation systems, space-time processing for cooperative and cognitive communications systems, and software-defined networks. He has served as an Associate Editor for the IEEE Signal Processing Letters since November 2014. He was an Associate Editor of the IEEE Transactions on Signal Processing from September 2010 to October 2014 and a Guest Editor of the *EURASIP Journal on Advances in Signal Processing* in 2010.

Anna Scaglione (M'99–F'11) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rome, "La Sapienza", Rome, Italy, in 1995 and 1999, respectively. She is a Professor with the Department of Electrical and Computer Engineering, Arizona State University, Tempe, AZ, USA. She was previously with the University of California at Davis, Davis, CA, USA, during 2008–2014, Cornell University, Ithaca, NY, USA, between 2001 and 2007, and the University of New Mexico, Albuquerque, NM, USA, between 2000 and 2001. Her expertise is in the broad area of signal processing for communication systems and networks. Her current research focuses on studying and enabling decentralized learning and signal processing in networks of sensors social networks, and on demand side management and reliable energy delivery. She was Editor-in-Chief of the IEEE Signal Processing Letters, and served in many capacities for communication and signal processing journals and conferences. She was on the Board of Governors of the Signal Processing Society from 2012 to 2014. She is a corecipient of the 2013 IEEE Donald G. Fink Prize Paper Award, and the 2013 SPS society Young Author Best Paper award (with her student) first author of the paper that received the 2000 IEEE Signal Processing Transactions Best Paper Award. She has also received the NSF Career Award in 2002 and is corecipient of the SmartGridcomm 2014 Best Student Paper Award and the Ellersick Best Paper Award (MILCOM 2005).

Donatella Darsena (M'06) received the Dr. Eng. degree (*summa cum laude*) in telecommunications engineering in 2001, and the Ph.D. degree in electronic and telecommunications engineering in 2005, both from the University of Napoli Federico II, Italy.

From 2001 to 2002, she was an Engineer in the Telecommunications, Peripherals and Automotive Group, STMicroelectronics, Milano, Italy. Since 2005, she has been an Assistant Professor with the Department of Engineering, University of Napoli Parthenope, Italy. She has also held teaching positions at the University of Napoli Federico II, Italy. Her research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, her current interests are focused on equalization, channel identification, narrowband-interference suppression for multicarrier systems, space-time processing for cooperative communications systems and cognitive communications systems, and software-defined networks.

Giacinto Gelli was born in Napoli, Italy, on July 29, 1964. He received the Dr. Eng. degree (*summa cum laude*) in electronic engineering in 1990, and the Ph.D. degree in computer science and electronic engineering in 1994, both from the University of Napoli Federico II.

From 1994 to 1998, he was an Assistant Professor with the Department of Information Engineering, Second University of Napoli. Since 1998, he has been with the Department of Electrical Engineering and Information Technology, University of Napoli Federico II, first as an Associate Professor, and since November 2006 as a Full Professor of Telecommunications. He also held teaching positions at the University Parthenope of Napoli, Italy.

His research interests are in the broad area of signal and array processing for communications, with current emphasis on multicarrier modulation systems and space-time techniques for cooperative and cognitive communications systems.

An Amplify-and-Forward Scheme for Spectrum Sharing in Cognitive Radio Channels

Francesco Verde, *Senior Member, IEEE*, Anna Scaglione, *Fellow, IEEE*,
Donatella Darsena, *Member, IEEE*, and Giacinto Gelli

***Abstract*—In this paper, we propose a cognitive radio scheme that allows a secondary user (SU) to transmit over the same time-frequency slot of a primary user (PU), even when the PU is active. In our scheme, the SU amplifies and forwards the signal of the PU, by using as scaling factor the value of its information symbol to be transmitted towards the secondary receiver. The information-theoretic limits of the proposed protocol are investigated in terms of ergodic channel capacities of both the PU and SU links. It is shown that: 1) under certain operating conditions, the SU can superimpose its information symbols on the PU signal, without violating the cognitive radio principle of protecting the PU transmission; and 2) when the primary link is busy, the SU offers the PU its own transmitting power in exchange for a low-capacity communication channel, which improves the packet delay performance of the SU. In this barter, the tempting incentive for the PU consists of a noticeable improvement of its achievable rate at the price of a slight increase in the computational complexity of the primary receiver.**

***Index Terms*—Amplify-and-forward relaying, cognitive radio, ergodic channel capacity, wireless communications.**

I. INTRODUCTION

MEASUREMENT studies [1] have recently confirmed that the licensed radio spectrum is relatively underutilized: as a consequence, numerous *cognitive radio schemes* [2] have been proposed, wherein secondary users (SUs) can temporarily share a portion of licensed spectrum, provided that they generate a minimal amount of interference to the licensed primary users (PUs). Common spectrum sharing strategies [3] belong to two categories: (i) the SUs are allowed to transmit also when the PUs are transmitting; (ii) the SUs use the licensed spectrum only when the PUs are not transmitting. In this paper, we consider a dynamic spectrum sharing scheme belonging to the first family.

Most of the literature on spectrum sharing for cognitive radios (see, e.g., [4]–[8]) relies on the idea of minimizing the interference caused by the SU. In our scheme, instead, we allow the SU to superimpose its transmission to the PU one, albeit in a “symbiotic” form, in order to possibly improve the primary link quality rather than degrading it. Specifically, when the PU is inactive, the SU uses the primary channel in a conventional manner. When, however, the PU is active, the SU is still allowed to transmit its data, by employing an amplify-and-forward (AF) relaying strategy, which has been widely used in cooperative systems [9]–[16]: specifically, the PU signal received by the SU is multiplied by the information symbols of the SU and retransmitted.

It should be observed that our approach differs from “cognitive relaying” [17]–[19], wherein relays with cognitive functionalities forward source data by using the spectrum white space(s) they have previously detected; indeed, in our scheme the SU is allowed to transmit also when the primary channel is busy. Moreover, our approach is also different from analog network coding (ANC) [20]–[23], since, in ANC schemes, the nodes use the wireless channel only to relay third-party information, without transmitting their own information. Our method also shares some resemblance with the concept of hiding information onto another signal without significantly distorting it [24], [25], a paradigm adopted in several applications, such as copyright protection for digital media, watermarking, fingerprinting, stenography, and data embedding. However, the proposed approach differs from information hiding mainly because superimposition of the SU symbols on the PU signal not only preserves the information of the PU, but also improves its performance (due to relaying) under suitable conditions.

The fundamental limits of the cognitive radio approach can be studied by modeling the cognitive radio channel as a classical interference channel [26]–[29]. It has been shown in [29] that the capacity-achieving strategy for the SU is to perform superposition coding of its codeword (generated by Costa precoding [30]), as well as of the codeword of the PU. Such a strategy requires the SU to know the PU codeword before it is transmitted (this is referred to as *noncausal knowledge* hereinafter). Even though such a constraint can be relaxed if the PU and SU are in close proximity of each other, the SU code selection would require instantaneous decoding of the PU message while it is transmitted on the air. In contrast, our proposed scheme does not require noncausal knowledge.

The main analytical contribution of the paper is to provide upper and lower bounds on the ergodic channel capacity [31] of the proposed scheme, evaluated by assuming a block-fading

Manuscript received February 1, 2014; revised June 21, 2014, November 8, 2014, and March 18, 2015; accepted May 22, 2015. This work was supported in part by the Italian National Project Servizi per lInfrastruttura di Rete Wireless Oltre il 3G (SIRIO). The associate editor coordinating the review of this paper and approving it for publication was G. Yue.

F. Verde and G. Gelli are with the Department of Electrical Engineering and Information Technology, University Federico II, Naples I-80125, Italy (e-mail: f.verde@unina.it; gelli@unina.it).

A. Scaglione is with the School of Electrical Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-9309 USA (e-mail: Anna.Scaglione@asu.edu).

D. Darsena is with the Department of Engineering, Parthenope University, Naples I-80143, Italy (e-mail: darsena@uniparthenope.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2440359

channel model and considering different assumptions on the amount of side information available at the receivers.[1] The ergodic capacity serves as a useful upper bound on the performance of any communication system and it can be achieved if the length of the codebook is long enough to reflect the ergodic nature of fading (i.e., the transmission duration of the codeword is much greater than the channel coherence time) [33]. It is shown that the concurrent transmission of the SU might improve the capacity of the PU link, provided that certain non-restrictive conditions are fulfilled: in this case, the SU earns an unlicensed channel with low transmission rates, which allow to reduce its average delay per symbol.

The paper is organized as follows. The model of the proposed cognitive radio network is introduced in Section II. Upper and lower bounds on the ergodic channel capacities of the PU and SU links are calculated in Sections III and IV, respectively, along with Monte Carlo numerical results aimed at assessing the ultimate achievable performances of both the PU and SU. Conclusions are drawn in Section V.

A. Notations and Preliminaries

The fields of complex, real, and nonnegative integer numbers are denoted with $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{N}$, respectively; matrices [vectors] are denoted with upper [lower] case boldface letters (e.g., $\mathbf{A}$ or $\mathbf{a}$); the field of $m \times n$ complex matrices is denoted as $\mathbb{C}^{m\times n}$, with $\mathbb{C}^m$ used as a shorthand for $\mathbb{C}^{m\times 1}$; the superscripts T and H denote the transpose and the conjugate transpose of a matrix, respectively; the symbol $*$ stands for (linear) convolution; $j \triangleq \sqrt{-1}$ denotes the imaginary unit; $\delta(n)$ is the Kronecker delta; $|a|$ and $\measuredangle a$ denote the magnitude and the phase of $a \in \mathbb{C}$, respectively; $\text{int}(x)$ gives the integer part of $x \in \mathbb{R}$; $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ denotes the identity matrix; $\|\mathbf{a}\|$ is the Euclidean norm of $\mathbf{a} \in \mathbb{C}^n$; matrix $\mathbf{A} = \text{diag}(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{C}^{n\times n}$ is diagonal; $\det(\mathbf{A})$ and $\text{tr}(\mathbf{A})$ are the determinant and trace of matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$, respectively; the operator $\mathbb{E}_{\mathbf{x}}[\cdot]$ denotes ensemble averaging with respect to the random vector $\mathbf{x} \in \mathbb{C}^n$ (the subscript is omitted when it can be deduced from the context) and $\mathbb{E}_{\mathbf{x}|\mathbf{y}}[\cdot]$ is the conditional mean with respect to $\mathbf{x}$ given the random vector $\mathbf{y} \in \mathbb{C}^m$; let $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{y} \in \mathbb{C}^n$, and $\mathbf{z} \in \mathbb{C}^p$ be random vectors, $p(\mathbf{x})$ is the probability density function (pdf) of $\mathbf{x}$, $p(\mathbf{x}|\mathbf{y})$ is the conditional pdf of $\mathbf{x}$, given $\mathbf{y}$, $\mathsf{I}(\mathbf{x};\mathbf{y})$ denotes the mutual information [31] between $\mathbf{x}$ and $\mathbf{y}$, $\mathsf{I}(\mathbf{x};\mathbf{y}|\mathbf{z})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}$, given $\mathbf{z}$, $\mathsf{h}(\mathbf{x})$ denote the differential entropy [31] of $\mathbf{x}$, and $\mathsf{h}(\mathbf{x}|\mathbf{y})$ is the conditional differential entropy of $\mathbf{x}$, given $\mathbf{y}$; a circularly symmetric complex Gaussian random vector $\mathbf{x} \in \mathbb{C}^n$ with mean $\boldsymbol{\mu} \in \mathbb{C}^n$ and covariance matrix $\Sigma \in \mathbb{C}^{n\times n}$ is denoted as $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \Sigma)$; finally, we define the following function (see, e.g., [34])

$$
\begin{aligned}
f(A) &\triangleq \int_0^{+\infty} \exp(-u)\ln(1+Au)\mathrm{d}u \\
&= -\exp\left(\frac{1}{A}\right)\text{Ei}\left(-\frac{1}{A}\right) \\
&\approx \begin{cases} A, & \text{for} 0 < A \ll 1; \\ \ln(1+A)-\gamma, & \text{for} A \gg 1. \end{cases}
\end{aligned} \tag{1}
$$

[1]Preliminary results of such an analysis are reported in [32].

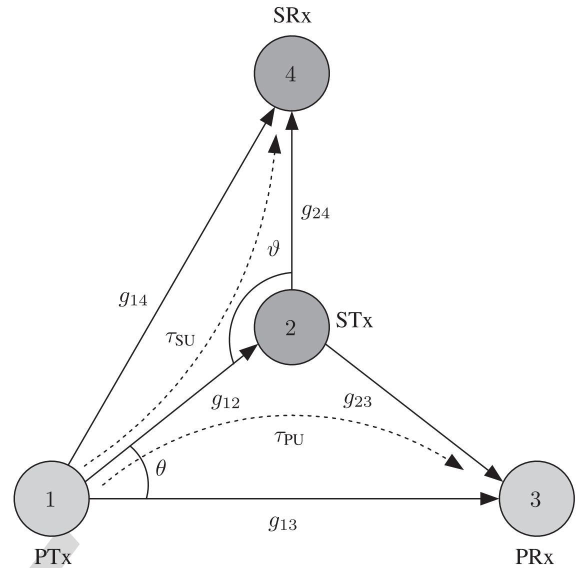

Fig. 1. The considered wireless network model: in green, the PU transmitting/receiving nodes, in red the SU transmitting/receiving nodes.

where, for $x < 0$,

$$
\text{Ei}(x) \triangleq \int_{-\infty}^{x} \frac{\exp(u)}{u}\mathrm{d}u = \gamma + \ln(-x) + \sum_{k=1}^{+\infty} \frac{x^k}{k!k}
$$

denotes the exponential integral function and

$$
\gamma \triangleq \lim_{n\to\infty}\left(n^{-1}\sum_{k=1}^{n} k^{-1} - \ln n\right) \approx 0.57721
$$

is the Euler-Mascheroni constant.

II. System Model and Proposed Cooperative Protocol

The cognitive radio network includes (Fig. 1) a primary transmitter/receiver pair (PTx/PRx) represented by nodes 1 and 3, and a secondary transmitter/receiver pair (STx/SRx) represented by nodes 2 and 4. A single channel is licensed to the primary user (PU), who uses it in a bursty manner, by alternating between busy (ON) and idle (OFF) intervals. During the ON intervals, the PU transmits a sequence $x_{\text{PU}}(\cdot)$ of independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex symbols, with variance $\sigma^2_{\text{PU}} = \mathsf{P}_{\text{PU}}$ and signaling interval T_{PU}. Such a sequence is arranged in consecutive frames of M symbols, whose duration is comparable with the coherence time of the channel. With regard to the secondary user (SU), the STx attempts to send towards the SRx a sequence $x_{\text{SU}}(\cdot)$ of i.i.d. zero-mean circularly symmetric complex symbols with variance σ^2_{SU}, statistically independent of $x_{\text{PU}}(\cdot)$. The power budget of the SU is given by P_{SU}. We assume that the SU knows the occurrence of the PU ON/OFF intervals: such an information is provided by the PU or estimated by the SU itself (e.g., by sensing). The symbol rate $1/\mathsf{T}_{\text{SU}}$ and the variance σ^2_{SU} of the SU depend on the status of the PU channel, as explained in the sequel. The channel corresponding to the $i \to \ell$

link is assumed to be frequency-flat and quasi-stationary: it is modeled by the fading coefficient $g_{i\ell} \sim \mathcal{CN}(0, \sigma_{i\ell}^2)$, which is constant within one frame, but is allowed to vary independently from frame to frame. Fading coefficients of different links are statistically independent among themselves and of the symbol sequences. CSI is not available at the transmitters, whereas it can be acquired at the receivers by training [35].

Depending on PU activity, the STx selects one of two transmission modes: *white-space* (idle PU channel) or *dirty-space* (busy PU channel) one. In white-space mode, the STx gains exclusive use of the channel and employs a conventional point-to-point technique to transmit with signaling interval $\mathsf{T}_{\text{SU}} \equiv \mathsf{T}_{\text{SU,white}}$. In dirty-space mode, since the signal transmitted by the PU is also overheard by the STx, the latter acts as a full-duplex AF relay and transmits one symbol per PU frame, i.e., $\mathsf{T}_{\text{SU}} \equiv \mathsf{T}_{\text{SU,dirty}} \triangleq M\mathsf{T}_{\text{PU}}$, as explained soon after.

The T_{PU}-spaced baseband equivalent signal received at the STx during the mth PU symbol period of a frame of length M is expressed as

$$r_2(m) = g_{12} x_{\text{PU}}(m) + v_2(m), \quad m \in \{0, 1, \ldots, M-1\} \quad (2)$$

where $v_2(\cdot)$ denotes AWGN at the STx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_2}^2)$ random variables (RVs), statistically independent of g_{12} and $x_{\text{PU}}(\cdot)$. Let x_{SU} denote the SU symbol to be transmitted during the considered PU frame, the signal (2) is scaled at the STx by the factor x_{SU} and forwarded to the SRx. The variance of x_{SU} is adjusted by the STx according to the average power constraint $\mathbb{E}[|x_{\text{SU}} r_2(m)|^2] = \mathsf{P}_{\text{SU}}$, i.e.,

$$\sigma_{\text{SU}}^2 \equiv \sigma_{\text{SU,dirty}}^2 \triangleq \frac{\mathsf{P}_{\text{SU}}}{\sigma_{12}^2 \mathsf{P}_{\text{PU}} + \sigma_{n_2}^2}. \quad (3)$$

Further details on the signal model in white space-mode are omitted, since it is a simple block flat-fading Rayleigh channel affected by additive white Gaussian noise (AWGN) [36], [37]. In the next subsections, we derive and discuss the received signal models at the PRx and SRx in dirty-space mode.

A. Received Signal at the PRx in the Dirty-Space Mode

Preliminarily (see also Fig. 1), let $t_{i\ell}$ denotes the propagation delays for the $i \to \ell$ link: since only relative delays are important [38], the delays t_{13} and t_{14} of the PTx → PRx and PTx → SRx links are conventionally set to zero. Thus, the cumulative PTx → STx → PRx and PTx → STx → SRx delays can be expressed as $t_{\text{PU}} = t_{12} + t_{\text{p}} + t_{23}$ and $t_{\text{SU}} = t_{12} + t_{\text{p}} + t_{24}$, respectively, with t_{p} denoting the processing time at the STx, which depends on various parameters [39], such as the frame length M and the hardware/memory characteristics of the STx node. We assume that M is chosen such that both t_{PU} and t_{SU} are much smaller than the frame duration $M\,\mathsf{T}_{\text{PU}}$.

The T_{PU}-spaced baseband equivalent received signal at the PRx can be written as

$$y_3(m) = g_{13} x_{\text{PU}}(m) + g_{23} x_{\text{SU}} r_2(m - \tau_{\text{PU}}) + v_3(m) \quad (4)$$

for $m \in \{0, 1, \ldots, M-1\}$, where $\tau_{\text{PU}} \triangleq \text{int}(t_{\text{PU}}/\mathsf{T}_{\text{PU}})$ is the integer[2] delay and $v_3(\cdot)$ denotes AWGN at the PRx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_3}^2)$ RVs, statistically independent of g_{13}, $x_{\text{PU}}(\cdot)$, g_{23}, x_{SU}, and $r_2(\cdot)$. Setting $y_3(m) \equiv y_{\text{PU}}(m)$, eq. (4) becomes

$$y_{\text{PU}}(m) = g_{\text{PU}}(m) * x_{\text{PU}}(m) + v_{\text{PU}}(m) \quad (5)$$

where $g_{\text{PU}}(m) \triangleq g_{13}\delta(m) + g_{12} g_{23} x_{\text{SU}} \delta(m - \tau_{\text{PU}})$ and $v_{\text{PU}}(m) \triangleq g_{23} x_{\text{SU}} v_2(m - \tau_{\text{PU}}) + v_3(m)$ represent the impulse response of the overall PU relay channel towards the PRx and the equivalent noise term at the PRx, respectively. It can be seen from (5) that, when $\tau_{\text{PU}} > 0$, the PU experiences ISI through a *two-ray frequency-selective channel* $g_{\text{PU}}(m)$, with the second channel tap gain incorporating the contribution of the SU transmitted symbol x_{SU}. The delay τ_{PU} depends on the sum of the delays t_{12} and t_{23} (see Fig. 1), the processing time t_{p} at the STx, and the PU symbol period T_{PU}: it is noteworthy that τ_{PU} is greater than zero when either $t_{12} + t_{23} \geq \mathsf{T}_{\text{PU}}$ or $t_{\text{p}} \geq \mathsf{T}_{\text{PU}}$—conditions that are likely to be fulfilled if the PU transmits at high baud rates (T_{PU} small).

The "composite" channel impulse response $g_{\text{PU}}(m)$ can be directly estimated at the PRx using the training symbols transmitted by the PTx and, thus, knowledge of x_{SU} by the PU is not required. Such a channel estimate can then be utilized to coherently recover the PU data symbols [36].

B. Received Signal at the SRx in the Dirty-Space Mode

The T_{PU}-spaced baseband equivalent received signal at the SRx is given by

$$y_4(m) = g_{24} x_{\text{SU}} r_2(m - \tau_{\text{SU}}) + g_{14} x_{\text{PU}}(m) + v_4(m) \quad (6)$$

for $m \in \{0, 1, \ldots, M-1\}$, where $\tau_{\text{SU}} \triangleq \text{int}(t_{\text{SU}}/\mathsf{T}_{\text{PU}})$ is the integer[3] delay and the sequence $v_4(\cdot)$ denotes AWGN at the SRx, modeled as a sequence of i.i.d. $\mathcal{CN}(0, \sigma_{n_4}^2)$ RVs, statistically independent of g_{14}, $x_{\text{PU}}(\cdot)$, g_{24}, x_{SU}, and $r_2(\cdot)$. Setting $y_4(m) \equiv y_{\text{SU}}(m)$, eq. (6) becomes

$$y_{\text{SU}}(m) = g_{\text{SU}}(m) x_{\text{SU}} + v_{\text{SU}}(m) \quad (7)$$

where $g_{\text{SU}}(m) \triangleq g_{24} r_2(m - \tau_{\text{SU}}) = g_{24} g_{12} x_{\text{PU}}(m - \tau_{\text{SU}}) + g_{24} v_2(m - \tau_{\text{SU}})$ and $v_{\text{SU}}(m) \triangleq g_{14} x_{\text{PU}}(m) + v_4(m)$ represent the time-varying fading gain towards the SRx and the equivalent noise term at the SRx, respectively. It results from (7) that the SU sees a *fast flat-fading channel* $g_{\text{SU}}(m)$, which also depends on the PU symbol $x_{\text{PU}}(m)$. Since the SU transmits only one symbol per frame in dirty-space mode, it may be not able to reliably estimate the time-varying channel $g_{\text{SU}}(m)$ by using its own training sequence, for $m \in \{0, 1, \ldots, M-1\}$, and, thus, coherent detection of x_{SU} might be difficult to implement in the dirty-space mode. Without any additional knowledge, recovery of x_{SU} can be accomplished at the SRx by resorting to noncoherent or generalized maximum-likelihood (ML)

[2] Any residual fractional delay $t_{\text{PU}} - \tau_{\text{PU}} \mathsf{T}_{\text{PU}}$ can be absorbed into the channel coefficient g_{23}.

[3] Any residual fractional delay $t_{\text{SU}} - \tau_{\text{SU}} \mathsf{T}_{\text{SU}}$ can be absorbed into the channel coefficient g_{24}.

detection rules [40]–[42], which require at most knowledge of the second-order statistics of $g_{SU}(m)$. In Section IV, the performance of the SU is evaluated under the assumption that the SRx has the additional knowledge of the training symbols of the PU. If the signal-to-noise ratio (SNR) at the STx is sufficiently large, the term $g_{24}v_2(m-\tau_{SU})$ in $g_{SU}(m)$ can be neglected and the SRx can acquire the channel parameters $g_{24}g_{12}$ and g_{14}, by jointly exploiting the training sequences sent by the PTx and STx. In this case, the SU symbols can be estimated at the SRx by resorting to partially-coherent detection algorithms [43].

III. Performance Analysis of the Primary User

Herein, we derive the conditions assuring that the achievable long-term rate of the PU link is not worsened when the SU is transmitting. To this aim, it is assumed that the training sequence for the PU link is long enough to acquire the relevant CSI with negligible error at the PRx. As a benchmark, we first study the performance limit of the PU link in the case of direct PTx → PRx transmission, i.e., when the SU is silent. In this case, the model for the received signal at the PRx can be simply obtained by setting $x_{SU}=0$ in (4), thus yielding $y_{PU}(m)=g_{13}x_{PU}(m)+v_3(m)$, which shows that the direct PU transmission sees a block flat-fading Rayleigh channel with AWGN. The ergodic capacity (nats/symbol) [33], [37] of the direct PU link with CSI at the receiver (CSIR) is given by

$$\mathsf{C}_{\text{PU,direct}}=\mathbb{E}\left[\ln\left(1+|g_{13}|^2\frac{\mathsf{P}_{\text{PU}}}{\sigma_{n_3}^2}\right)\right]=f(\mathsf{ASNR}_{\text{PU,direct}}) \quad (8)$$

where we have used (1) and the fact that $|g_{13}|^2$ has an exponential distribution with mean σ_{13}^2, whereas $\mathsf{ASNR}_{\text{PU,direct}} \triangleq (\sigma_{13}^2\mathsf{P}_{\text{PU}})/\sigma_{n_3}^2$ is the average (over the channel) SNR of the PU at the PRx (node 3) when $x_{SU}=0$. As expected, $\mathsf{C}_{\text{PU,direct}}$ is a monotonically increasing function of $\mathsf{ASNR}_{\text{PU,direct}}$: for high ASNR values, it results from (1) that $\mathsf{C}_{\text{PU,direct}} \approx \ln(1+\mathsf{ASNR}_{\text{PU,direct}})-\gamma$; in the low-SNR regime, one has $\mathsf{C}_{\text{PU,direct}} \approx \mathsf{ASNR}_{\text{PU,direct}}$, that is, the capacity increases linearly with $\mathsf{ASNR}_{\text{PU,direct}}$.

Let us now consider the case wherein the STx is active, i.e., $x_{SU}\neq 0$; in this case, as discussed in Section II-A, the PU transmission experiences block frequency-selective fading [see eq. (5)]. Evaluation of the ergodic capacity of a single-user frequency-selective channel with CSIR can be carried out [37] by decomposing the channel into an equivalent number of independent frequency-flat (i.e., memoryless) subchannels, whose input-output relationships are given by

$$\tilde{y}_{PU}(\ell)=G_{PU}(\ell)\tilde{x}_{PU}(\ell)+\tilde{v}_{PU}(\ell) \quad (9)$$

for $\ell \in \{0,1,\ldots,M-1\}$, where $\tilde{y}_{PU}(\cdot)$, $\tilde{x}_{PU}(\cdot)$, and $\tilde{v}_{PU}(\cdot)$ are the M-point discrete Fourier transform (DFT) of $y_{PU}(\cdot)$, $x_{PU}(\cdot)$, and $v_{PU}(\cdot)$, respectively, and

$$G_{PU}(\ell)=g_{13}+g_{12}g_{23}x_{SU}e^{-j\frac{2\pi}{M}\ell\tau_{PU}} \quad (10)$$

is the M-point DFT of the channel $g_{PU}(m)$ in (5). In compact form, let the M-dimensional vectors $\mathbf{x}_{PU}$, $\tilde{\mathbf{y}}_{PU}$, $\tilde{\mathbf{v}}_{PU}$ gather the corresponding scalar quantities, one has

$$\tilde{\mathbf{y}}_{PU} \triangleq \widetilde{\mathbf{G}}_{PU}\tilde{\mathbf{x}}_{PU}+\tilde{\mathbf{v}}_{PU} \quad (11)$$

where $\widetilde{\mathbf{G}}_{PU}\triangleq\text{diag}[G_{PU}(0),G_{PU}(1),\ldots,G_{PU}(M-1)]\in\mathbb{C}^{M\times M}$ and $\tilde{\mathbf{x}}_{PU}$ is subject to the power constraint $\mathbb{E}[\|\tilde{\mathbf{x}}_{PU}\|^2]=M\mathsf{P}_{PU}$. Under the assumption of exact CSIR, the PRx is assumed to have perfect knowledge of the channel vector $\mathbf{g}_{PU}\triangleq[g_{13},g_{12}g_{23}x_{SU}]^{\mathrm{T}}\in\mathbb{C}^2$, and the ergodic channel capacity (in nats/symbol) of the PU can be calculated [37] as

$$\mathsf{C}_{PU}\triangleq\lim_{M\to+\infty}\mathsf{C}_{PU}(M) \quad (12)$$

where $\mathsf{C}_{PU}(M)$ is obtained by taking the supremum of the average mutual information $\mathsf{I}(\tilde{\mathbf{x}}_{PU};\tilde{\mathbf{y}}_{PU}|\mathbf{g}_{PU})/M$ over all possible distributions of the vector $\tilde{\mathbf{x}}_{PU}$ that satisfy the power constraint $\mathbb{E}[\|\tilde{\mathbf{x}}_{PU}\|^2]=M\mathsf{P}_{PU}$. Since the subchannels (9) are memoryless and the entries of $\tilde{\mathbf{x}}_{PU}$ are statistically independent, one has

$$\mathsf{I}(\tilde{\mathbf{x}}_{PU};\tilde{\mathbf{y}}_{PU}|\mathbf{g}_{PU})=\sum_{\ell=0}^{M-1}\mathsf{I}\left(\tilde{x}_{PU}(\ell);\tilde{y}_{PU}(\ell)|\mathbf{g}_{PU}\right). \quad (13)$$

Therefore, calculation of C_{PU} boils down to evaluating $\mathsf{I}(\tilde{x}_{PU}(\ell);\tilde{y}_{PU}(\ell)|\mathbf{g}_{PU})$. It is worth noting that, for a given $\mathbf{g}_{PU}$, the equivalent noise term $\tilde{v}_{PU}(\ell)$ in (9) is a non-Gaussian RV, which complicates exact evaluation of (13). For this reason, we provide in Theorem 1 upper and lower bounds on C_{PU}.

Theorem 1 (Upper and Lower Bounds on the PU Capacity): The ergodic channel capacity of the PU can be lower- and upper-bounded as follows:

$$\begin{aligned}\mathbb{E}\left\{f\left[\Gamma_{3,\text{lower}}\left(|g_{23}|^2|x_{SU}|^2\right)\right]\right\}&\triangleq\mathsf{C}_{\text{PU,lower}}\leq\mathsf{C}_{PU}\\ &\leq\mathsf{C}_{\text{PU,upper}}\triangleq\mathbb{E}\left\{f\left[\Gamma_{3,\text{upper}}\left(|g_{23}|^2|x_{SU}|^2\right)\right]\right\}\end{aligned} \quad (14)$$

with

$$\Gamma_{3,\text{upper}}\left(|g_{23}|^2|x_{SU}|^2\right)\triangleq\mathsf{ASNR}_{\text{PU,direct}}\frac{1+|g_{23}|^2|x_{SU}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}}{1+|g_{23}|^2|x_{SU}|^2\frac{\sigma_{n_2}^2}{\sigma_{n_3}^2}} \quad (15)$$

$$\Gamma_{3,\text{lower}}(|g_{23}|^2|x_{SU}|^2)\triangleq\mathsf{ASNR}_{\text{PU,direct}}\frac{1+|g_{23}|^2|x_{SU}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}}{1+\sigma_{23}^2\sigma_{\text{SU,dirty}}^2\frac{\sigma_{n_2}^2}{\sigma_{n_3}^2}}. \quad (16)$$

Proof: See Appendices A and B. □

The upper bound in (14) is obtained in Appendix A by assuming that the PRx has the additional perfect knowledge of g_{12}.[4] Instead, the lower bound in (14) is derived in Appendix B by replacing $\tilde{x}_{PU}(\ell)$ and $\tilde{v}_{PU}(\ell)$ in (9) with $\tilde{x}_{PU,G}(\ell)\sim\mathcal{CN}(0,\mathsf{P}_{PU})$ and $\tilde{v}_{PU,G}(\cdot)\sim\mathcal{CN}\left(0,\sigma_{n_3}^2+\sigma_{n_2}^2\sigma_{23}^2\sigma_{\text{SU,dirty}}^2\right)$,

[4] To acquire g_{12} in practice, the PRx would require additional help from the STx in the form of channel-state feedback.

respectively. Since $g_{23} \sim \mathcal{CN}(0, \sigma_{23}^2)$, the averages needed to evaluate $\mathsf{C}_{\text{PU,lower}}$ and $\mathsf{C}_{\text{PU,upper}}$ can be computed numerically by assuming a prior distribution for the symbol x_{SU} transmitted by the SU. Finally, observe that both $\mathsf{C}_{\text{PU,upper}}$ and $\mathsf{C}_{\text{PU,lower}}$ do not depend on the delay τ_{PU}, provided that $\tau_{\text{PU}} \ll M$.

At this point, we can draw some interesting conclusions. To this end, we refer for simplicity to the path-loss model $\sigma_{i\ell}^2 = d_{i\ell}^{-\eta}$, where $d_{i\ell}$ is the distance between nodes i and ℓ and η denotes the path-loss exponent. The upper and lower bounds in (14) depend on $\Gamma_{3,\text{upper}}(|g_{23}|^2|x_{\text{SU}}|^2)$ and $\Gamma_{3,\text{lower}}(|g_{23}|^2|x_{\text{SU}}|^2)$, respectively. The right-hand sides (RHSs) in (15) and (16) can be intuitively explained as follows. On one side, the SU transmission is beneficial since it increases the frequency diversity of the PU; mathematically, the gain in frequency diversity comes in (15) and (16) from the multiplication of $\mathsf{ASNR}_{\text{PU,direct}}$ by the factor $1 + |g_{23}|^2|x_{\text{SU}}|^2(\sigma_{12}^2/\sigma_{13}^2)$. On the other hand, the AF relaying carried out by the SU is detrimental, due to the noise propagation phenomenon from the STx to the PRx; mathematically, the adverse effect of noise propagation is represented in (15) by the division of $\mathsf{ASNR}_{\text{PU,direct}}$ by the factor $1 + |g_{23}|^2|x_{\text{SU}}|^2(\sigma_{n_2}^2/\sigma_{n_3}^2)$ and in (16) by the division of $\mathsf{ASNR}_{\text{PU,direct}}$ by the factor $1 + \sigma_{23}^2\sigma_{\text{SU,dirty}}^2(\sigma_{n_2}^2/\sigma_{n_3}^2)$.

Let us consider the case when $\sigma_{12}^2/\sigma_{13}^2 \leq \sigma_{n_2}^2/\sigma_{n_3}^2$. One has from (15) that $\Gamma_{3,\text{upper}}(|g_{23}|^2|x_{\text{SU}}|^2) \leq \mathsf{ASNR}_{\text{PU,direct}}$ for each realization of $g_{23}x_{\text{SU}}$. Since $f(A)$ is a monotonically increasing function of $A \geq 0$, one obtains from (8) and (14) that $\mathsf{C}_{\text{PU,upper}} \leq \mathsf{C}_{\text{PU,direct}}$. According to the path-loss model, *the capacity of the PU may degrade when*

$$d_{12} \geq d_{13}\sqrt[\eta]{\sigma_{n_3}^2/\sigma_{n_2}^2}$$

since the SU prevalently forwards noise. In particular, when $\sigma_{n_3}^2 \geq \sigma_{n_2}^2$, that is, the PRx is noisier than the STx, the cognitive radio principle of protecting the PU might be violated if $d_{12} \geq d_{13}$, that is, the PTx is farther from the STx than from the PRx or it is equidistant from them. Since the SU can determine whether $\sigma_{12}^2/\sigma_{13}^2 > \sigma_{n_2}^2/\sigma_{n_3}^2$ or, equivalently, $d_{12} < d_{13}\sqrt[\eta]{\sigma_{n_3}^2/\sigma_{n_2}^2}$ (we call it the *symbiotic region*), in the following we restrict attention to this case.

By virtue of Theorem 1, the variance $\sigma_{23}^2\sigma_{\text{SU,dirty}}^2$ of the RV $|g_{23}|^2|x_{\text{SU}}|^2$ plays a crucial role in determining the performance of the PU. Remembering (3), we observe that

$$\sigma_{23}^2\sigma_{\text{SU,dirty}}^2 = \frac{\sigma_{23}^2\mathsf{P}_{\text{SU}}}{\sigma_{12}^2\mathsf{P}_{\text{PU}} + \sigma_{n_2}^2} < \frac{\sigma_{23}^2}{\sigma_{12}^2}\frac{\mathsf{P}_{\text{SU}}}{\mathsf{P}_{\text{PU}}} = \left(\frac{d_{23}}{d_{12}}\right)^{-\eta}\frac{\mathsf{P}_{\text{SU}}}{\mathsf{P}_{\text{PU}}} \tag{17}$$

where, as a consequence of the Carnot's cosine law, we can also write

$$\frac{d_{23}}{d_{12}} = \sqrt{1 + \frac{d_{13}^2}{d_{12}^2} - 2\,\frac{d_{13}}{d_{12}}\cos(\theta)} \geq \left|\frac{d_{13}}{d_{12}} - 1\right| \tag{18}$$

with θ denoting the angle contained between sides of lengths d_{12} and d_{13} (see Fig. 1). The minimum value of d_{23}/d_{12} [corresponding to the equality in (18)] is achieved when $\cos(\theta) = 1$. It can be inferred from (15) and (16) that, when $\sigma_{23}^2\sigma_{\text{SU,dirty}}^2 \ll \epsilon$, with $\epsilon > 0$ sufficiently small, the benefits in frequency diversity prevail over the losses caused by noise propagation. In such a case, Chebychev's inequality [44] implies that $\mathsf{P}(|g_{23}|^2|x_{\text{SU}}|^2 \geq \epsilon) \leq \sigma_{23}^2\sigma_{\text{SU,dirty}}^2/\epsilon \ll 1$ and, consequently, the RV $|g_{23}|^2|x_{\text{SU}}|^2$ takes on values significantly smaller than one, with high probability. Hence, when $\sigma_{23}^2\sigma_{\text{SU,dirty}}^2 \ll 1$, $\Gamma_{3,\text{upper}}(|g_{23}|^2|x_{\text{SU}}|^2)$ in (15) and $\Gamma_{3,\text{lower}}(|g_{23}|^2|x_{\text{SU}}|^2)$ in (16) can be approximated as

$$\begin{aligned}\Gamma_{3,\text{upper}}\left(|g_{23}|^2|x_{\text{SU}}|^2\right) &\approx \Gamma_{3,\text{lower}}\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\\ &\approx \Gamma_3\left(|g_{23}|^2|x_{\text{SU}}|^2\right) \triangleq \mathsf{ASNR}_{\text{PU,direct}}\left(1+|g_{23}|^2|x_{\text{SU}}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}\right)\end{aligned} \tag{19}$$

which leads to

$$\mathsf{C}_{\text{PU}} \approx \mathsf{C}_{\text{PU,lower}} \approx \mathsf{C}_{\text{PU,upper}} \approx \mathbb{E}\left\{f\left[\Gamma_3\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\right]\right\}. \tag{20}$$

Comparing (8) with (20), since $\Gamma_3(|g_{23}|^2|x_{\text{SU}}|^2) > \mathsf{ASNR}_{\text{PU,direct}}$ for each realization of $g_{23}x_{\text{SU}}$ and $f(A)$ is a monotonically increasing function of $A \geq 0$, one readily obtains that $\mathsf{C}_{\text{PU}} > \mathsf{C}_{\text{PU,direct}}$. Thus, *the capacity of the PU improves as a result of the SU transmission when* $\sigma_{23}^2\sigma_{\text{SU,dirty}}^2 \ll 1$, no matter what the distributions of g_{23} and x_{SU} are.[5] In the symbiotic region $\sigma_{12}^2/\sigma_{13}^2 > \sigma_{n_2}^2/\sigma_{n_3}^2$, $\Gamma_3(|g_{23}|^2|x_{\text{SU}}|^2)$ is lower bounded as

$$\begin{aligned}&\Gamma_3\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\\ &\qquad > \mathsf{ASNR}_{\text{PU,direct}}\left[1 + |g_{23}|^2|x_{\text{SU}}|^2\left(\sigma_{n_2}^2/\sigma_{n_3}^2\right)\right].\end{aligned}$$

By virtue of (17), (18), condition $\sigma_{23}^2\sigma_{\text{SU,dirty}}^2 \ll 1$ is met if

$$|d_{13}/d_{12} - 1| \gg \sqrt[\eta]{\mathsf{P}_{\text{SU}}/\mathsf{P}_{\text{PU}}}.$$

Since the symbiotic region is equivalently characterized by the inequality $d_{13}/d_{12} > \sqrt[\eta]{\sigma_{n_2}^2/\sigma_{n_3}^2}$, this imposes that, when $\sigma_{n_2}^2 \geq \sigma_{n_3}^2$, the STx has to be sufficiently closer to the PTx than to the PRx.

In order to assess the capacity improvement of the PU, it is noteworthy that, when $\mathsf{ASNR}_{\text{PU,direct}} \gg 1$, it also results from (19) that $\Gamma_3(|g_{23}|^2|x_{\text{SU}}|^2) > \mathsf{ASNR}_{\text{PU,direct}} \gg 1$ for each realization of $g_{23}x_{\text{SU}}$ and, hence, eq. (20) admits, in the high-SNR region, the approximated expression [see eq. (1)]

$$\begin{aligned}\mathsf{C}_{\text{PU}} &\approx \mathbb{E}\left[\ln\left(1 + \Gamma_3\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\right)\right] - \gamma\\ &\approx \mathbb{E}\left[\ln\left(\Gamma_3\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\right)\right] - \gamma\\ &= \mathsf{C}_{\text{PU,direct}} + \mathbb{E}\left[\ln\left(1+\sigma_{23}^2\sigma_{\text{SU,dirty}}^2|\overline{g}_{23}|^2|\overline{x}_{\text{SU}}|^2\frac{\sigma_{12}^2}{\sigma_{13}^2}\right)\right]\end{aligned} \tag{21}$$

[5] Such an improvement is achieved if the PU is willing to equalize a two-ray frequency selective channel rather than a flat-fading one. While this behavior is atypical for conventional cognitive radio scheme where the PU is unaware of the SU transmissions, some advanced cognitive radio systems have been proposed where the PUs either aid or react to the SUs [3].

where $\bar{g}_{23} \triangleq g_{23}/\sigma_{23}$ and $\bar{x}_{\text{SU}} \triangleq x_{\text{SU}}/\sigma_{\text{SU}}$ are normalized versions of g_{23} and x_{SU}, respectively, with $\sigma^2_{\text{SU,dirty}}$ given by (3), and we have used the fact that $\mathsf{C}_{\text{PU,direct}} \approx \ln(\mathsf{ASNR}_{\text{PU,direct}}) - \gamma$ for $\mathsf{ASNR}_{\text{PU,direct}} \gg 1$. Since $|\bar{g}_{23}|^2$ is a unit-mean exponential RV statistically independent of $\bar{x}_{\text{SU}}$, by evaluating the expectation in (21) with respect to the distribution of $\bar{g}_{23}$, and using (1), one has, for $\mathsf{ASNR}_{\text{PU,direct}} \gg 1$, that the *capacity gain* of the PU is given by

$$\Delta\mathsf{C}_{\text{PU}} \triangleq \mathsf{C}_{\text{PU}} - \mathsf{C}_{\text{PU,direct}} \approx \mathbb{E}\left[f\left(\frac{\sigma^2_{23}\mathsf{P}_{\text{SU}}|\bar{x}_{\text{SU}}|^2}{\sigma^2_{13}\mathsf{P}_{\text{PU}}} \right) \right] \quad (22)$$

where we have observed from (17) that, in the high-SNR regime, $\sigma^2_{23}\sigma^2_{\text{SU,dirty}}\sigma^2_{12}/\sigma^2_{13} \approx (\sigma^2_{23}\mathsf{P}_{\text{SU}})/(\sigma^2_{13}\mathsf{P}_{\text{PU}})$. We observe that, if the SU adopts a constant-modulus constellation with average energy $\sigma^2_{\text{SU,dirty}}$, the capacity gain of the PU in the high-SNR regime is obtained in closed-form by setting $|\bar{x}_{\text{SU}}|^2 = 1$ and removing the average in (22). Interestingly, the capacity gain $\Delta\mathsf{C}_{\text{PU}}$ becomes significant in the high-SNR region if $\mathsf{P}_{\text{SU}} \gg \mathsf{P}_{\text{PU}}$: this is in contrast with conventional cognitive radio approaches, for which concurrent transmission of the SU is allowed only if its transmission power is subject to a strict constraint, such that the interference at the PRx is within the interference temperature limit [2]. Intuitively, such a behavior is a consequence of the fact that increasing the variance of x_{SU} [see eq. (3)] leads to a raise in the variance of the second tap of the impulse response $g_{\text{PU}}(m)$ [see eq. (5)].

To corroborate the information-theoretic findings, we report herein some results of numerical simulations. Specifically, we plot the maximum $(\Delta\mathsf{C}_{\text{PU}})_{\max} \triangleq \mathsf{C}_{\text{PU,upper}} - \mathsf{C}_{\text{PU,direct}}$ and minimum capacity gain $(\Delta\mathsf{C}_{\text{PU}})_{\min} \triangleq \mathsf{C}_{\text{PU,lower}} - \mathsf{C}_{\text{PU,direct}}$ of the PU (reported in bits/symbol and referred to as "ub" and "lb," respectively), where the ensemble averages in (14) were evaluated by carrying out 10^4 Monte Carlo trials. Obviously, it results that $(\Delta\mathsf{C}_{\text{PU}})_{\min} \leq \Delta\mathsf{C}_{\text{PU}} \leq (\Delta\mathsf{C}_{\text{PU}})_{\max}$. With reference to Fig. 1, we normalized the distance between the PTx and the PRx, as well as the transmitting power of the PU, by setting $d_{13} = 1$ and $\mathsf{P}_{\text{PU}} = 1$, respectively. Moreover, we chose $\theta = \pi/3$ and $\eta = 2$. The SU symbol x_{SU} was drawn from a QPSK constellation having average energy $\sigma^2_{\text{SU,dirty}}$ given by (3) and we chose $\sigma^2_{n_2} = \sigma^2_{n_3}$.[6]

1) Fig. 2 : It reports the curves of PU capacity gain as a function of $\mathsf{SNR}_{\text{PU}} \triangleq \mathsf{P}_{\text{PU}}/\sigma^2_{n_2}$ for different values of the distance d_{12} between the PTx and the STx (the distance d_{23} between the STx and the PRx is calculated according to the Carnot's cosine law), with $\mathsf{P}_{\text{SU}} = 1$. It is seen that the PU can harvest a noticeable capacity gain from the concurrent transmission of the SU, which rapidly increases as d_{12} decreases. Moreover, when the PTx and the STx are sufficiently close to each other, the upper and lower bounds in (14) tend to coincide, thus yielding an accurate approximation of the PU capacity.

2) Fig. 3 : The curves of the PU capacity gain are depicted as a function of SNR_{PU} for different values of the SU transmitting power P_{SU}, with $d_{12} = 0.5$. As analytically predicted, results show that, the larger the transmitting power of the SU, the greater the capacity improvement of the PU will be. Moreover, as the difference between P_{SU} and P_{PU} increases, the upper and lower bounds in (14) tend to slightly space out, allowing anyhow to accurately predict the achievable PU rate.

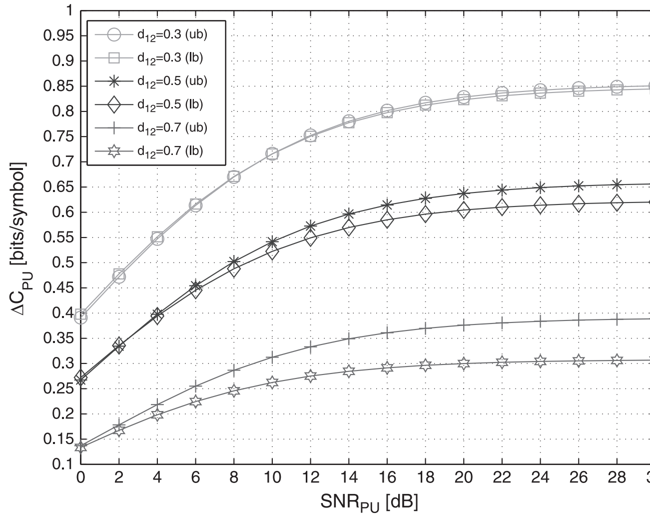

Fig. 2. $(\Delta\mathsf{C}_{\text{PU}})_{\max}$, $(\Delta\mathsf{C}_{\text{PU}})_{\min}$ versus SNR_{PU} for different values of d_{12} ($d_{13} = 1$, $\mathsf{P}_{\text{PU}} = \mathsf{P}_{\text{SU}} = 1$, $\sigma^2_{n_2} = \sigma^2_{n_3}$).

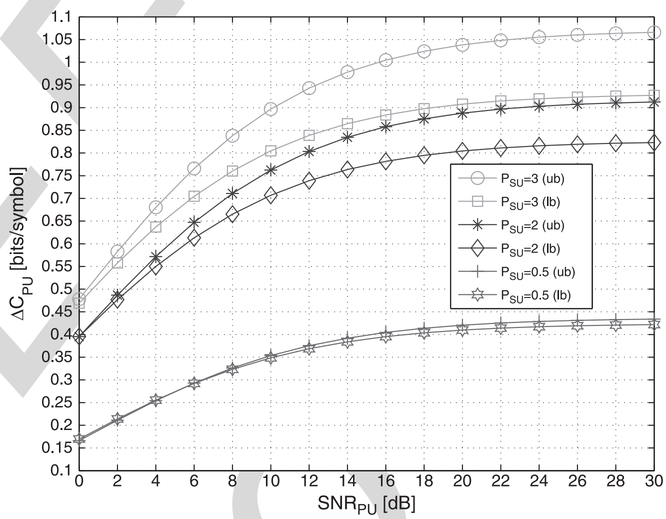

Fig. 3. $(\Delta\mathsf{C}_{\text{PU}})_{\max}$, $(\Delta\mathsf{C}_{\text{PU}})_{\min}$ versus SNR_{PU} for different values of P_{SU} ($d_{13} = 1$, $d_{12} = 0.5$, $\mathsf{P}_{\text{PU}} = 1$, $\sigma^2_{n_2} = \sigma^2_{n_3}$).

IV. Performance Analysis of the Secondary User

In white-space mode (PU OFF intervals), the STx directly transmits to the SRx with average power per symbol $\sigma^2_{\text{SU}} \equiv \sigma^2_{\text{SU,white}} \triangleq \mathsf{P}_{\text{SU}}$, by exclusively using the available PU channel. In this case, the ergodic channel capacity $\mathsf{C}_{\text{SU,white}}$ with CSIR can be obtained from (8) by replacing σ^2_{13}, P_{PU}, and $\sigma^2_{n_3}$ with σ^2_{24}, P_{SU}, and $\sigma^2_{n_4}$, respectively, thus obtaining

$$\mathsf{C}_{\text{SU,white}} = f(\mathsf{ASNR}_{\text{SU,white}}) \qquad \text{(nats/symbol)} \qquad (23)$$

where $\mathsf{ASNR}_{\text{SU,white}} \triangleq (\sigma^2_{24}\mathsf{P}_{\text{SU}})/\sigma^2_{n_4}$ is the ASNR at the SRx in white-space mode.

[6] Such an assumption is reasonable when nodes 2 and 3 (approximatively) have the same noise figure, i.e., they are equipped with similar hardware components and operate in the same environment.

Let us evaluate the achievable throughput of the SU in dirty-space mode (PU ON intervals). As discussed in Section II-B, the SU sees a fast flat-fading channel [see eq. (7)] in dirty-space mode. According to the results of the capacity analysis of the PU, we assume hereinafter that the ASNR at the STx is sufficiently large, i.e., $\sigma_{12}^2 \mathsf{P}_{\mathrm{PU}} \gg \sigma_{n_2}^2$, which is a reasonable assumption for SNR values of practical interest when the STx is not too far from the PTx. In this case, the noise term in (2) can be neglected, allowing one to simplify the fading gain as $g_{\mathrm{SU}}(m) \approx g_{24} g_{12} x_{\mathrm{PU}}(m-\tau_{\mathrm{SU}})$. Therefore, let $\mathbf{y}_{\mathrm{SU}} \triangleq [y_{\mathrm{SU}}(0), y_{\mathrm{SU}}(1), \ldots, y_{\mathrm{SU}}(M-1)]^{\mathrm{T}} \in \mathbb{C}^M$ be the vector of samples that the SU observes over a frame, accounting for (7), the block $\mathbf{y}_{\mathrm{SU}}$ can be expressed as

$$\mathbf{y}_{\mathrm{SU}} = \left(g_{24} g_{12} \mathbf{J} \check{\mathbf{x}}_{\mathrm{PU}}\right) x_{\mathrm{SU}} + g_{14} \mathbf{x}_{\mathrm{PU}} + \mathbf{v}_4 \tag{24}$$

where the matrix $\mathbf{J} \in \mathbb{C}^{M \times (M+\tau_{\mathrm{SU}})}$ is obtained from the identity matrix $\mathbf{I}_{M+\tau_{\mathrm{SU}}}$ by picking its first M rows, $\check{\mathbf{x}}_{\mathrm{PU}} \triangleq [x_{\mathrm{PU}}(-\tau_{\mathrm{SU}}), \ldots, x_{\mathrm{PU}}(-1), \mathbf{x}_{\mathrm{PU}}^T]^{\mathrm{T}} \in \mathbb{C}^{M+\tau_{\mathrm{SU}}}$, and $\mathbf{v}_4 \triangleq [v_4(0), v_4(1), \ldots, v_4(M-1)]^{\mathrm{T}} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_{n_4}^2 \mathbf{I}_M)$. Note that $\mathbf{y}_{\mathrm{SU}}$ also depends on the block $\check{\mathbf{x}}_{\mathrm{PU}}$ of PU symbols.

Regarding the CSIR of the SU, since the SU "composite" channel $g_{\mathrm{SU}}(m)$, $m \in \{0, 1, \ldots, M-1\}$, rapidly changes due to the dependence on the PU symbols, the accuracy with which *full* CSI can be retrieved at the SRx during a predefined frame may be unsatisfactory. However, *partial* CSI can be reliably acquired at the SRx. Indeed, the SRx can attain an accurate estimate of the vector $\mathbf{g}_{\mathrm{SU}} \triangleq [g_{24} g_{12}, g_{14}]^{\mathrm{T}} \in \mathbb{C}^2$ by assuming that, besides having knowledge of the training symbol transmitted by the STx, the SRx additionally knows the training signal sent by the PTx.[7] Under this assumption, the channel $\mathbf{g}_{\mathrm{SU}}$ can be estimated at the SRx by resorting to standard estimators [45], such as the ML or the Bayesian linear minimum mean-square error ones. Therefore, the capacity $\mathsf{C}_{\mathrm{SU,dirty}}$ (in nats/symbol) of the SU link in dirty-space mode turns out to be the supremum of $\mathsf{I}(x_{\mathrm{SU}}; \mathbf{y}_{\mathrm{SU}} | \mathbf{g}_{\mathrm{SU}})/M$ over all distributions on x_{SU} satisfying (3). Upper and lower bounds on $\mathsf{C}_{\mathrm{SU,dirty}}$ are given by Theorem 2.

Theorem 2 (Upper and Lower Bounds on the SU Capacity): The ergodic channel capacity of the SU can be lower- and upper-bounded as shown in (25) at the bottom of the page where

$$\begin{aligned}\mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) &\triangleq \Omega(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \mathbf{I}_M + \Upsilon^*(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \mathbf{B}^{\tau_{\mathrm{SU}}} \\ &\quad + \Upsilon(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \mathbf{F}^{\tau_{\mathrm{SU}}} \\ \boldsymbol{\Sigma}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) &\triangleq \left(|g_{24}|^2 |g_{12}|^2 x_{\mathrm{SU}} \mathbf{I}_M + g_{24}^* g_{12}^* g_{14} \mathbf{B}^{\tau_{\mathrm{SU}}}\right) \mathsf{P}_{\mathrm{PU}}\end{aligned}$$

with $\Omega(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \triangleq (|g_{24}|^2 |g_{12}|^2 |x_{\mathrm{SU}}|^2 + |g_{14}|^2) \mathsf{P}_{\mathrm{PU}} + \sigma_{n_4}^2$, $\Upsilon(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \triangleq g_{24} g_{12} x_{\mathrm{SU}} g_{14}^* \mathsf{P}_{\mathrm{PU}}$, the matrices $\mathbf{B} \in \mathbb{R}^{M \times M}$ and $\mathbf{F} \in \mathbb{R}^{M \times M}$ being backward- and forward-shift matrices, respectively, and the inner ensemble average in the lower bound is taken over $x_{\mathrm{SU}} \sim \mathcal{CN}(0, \sigma_{\mathrm{SU,dirty}}^2)$.

Proof: See Appendices C and D. □

Some remarks are now in order about the bounds in (25). The upper bound $\mathsf{C}_{\mathrm{SU,dirty,upper}}$ is obtained in Appendix C by assuming that the SRx has perfect knowledge of the pair $(\mathbf{g}_{\mathrm{SU}}, \check{\mathbf{x}}_{\mathrm{PU}})$. Remembering that $g_{i\ell} \sim \mathcal{CN}(0, \sigma_{i\ell}^2)$, the capacity $\mathsf{C}_{\mathrm{SU,dirty,upper}}$ can be evaluated numerically by assuming a prior distribution for the symbol sequence $x_{\mathrm{PU}}(\cdot)$ transmitted by the PU. In addition to the Gaussian assumption for the SU symbol, the lower bound $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ is derived in Appendix D by assuming that the PU symbols are circularly symmetric complex Gaussian RVs. When the PU symbols are drawn from discrete symbol constellations, it can be argued that $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ is still valid as an approximated lower bound [46].

The upper bound in (25) can be (approximately) achieved if the SU is capable of decoding the PU symbols with arbitrarily small error probability. In order for this to be true, the information rate R_{PU} (in nats/symbol) of the PU must be smaller than the ergodic channel capacity $\mathsf{C}_{\mathrm{PU}\to\mathrm{SU}}$ (in nats/symbol) of the overall link between the PTx and the SRx, i.e., $\mathsf{R}_{\mathrm{PU}} < \mathsf{C}_{\mathrm{PU}\to\mathrm{SU}}$. Due to the complete symmetry between the PRx and SRx with respect to the PTx, we can obtain an upper bound on $\mathsf{C}_{\mathrm{PU}\to\mathrm{SU}}$ exactly as we have done in Appendix A to get $\mathsf{C}_{\mathrm{PU,upper}}$. Hence, similarly to (14), the following inequality holds $\mathsf{C}_{\mathrm{PU}\to\mathrm{SU}} \leq \mathbb{E}\{f[\Gamma_{4,\mathrm{upper}}(g_{24} x_{\mathrm{SU}})]\}$, with

$$\Gamma_{4,\mathrm{upper}}\left(|g_{24}|^2 |x_{\mathrm{SU}}|^2\right) \triangleq \mathsf{ASNR}_{\mathrm{PU}\to\mathrm{SU}} \cdot \frac{1 + |g_{24}|^2 |x_{\mathrm{SU}}|^2 \left(\sigma_{12}^2/\sigma_{14}^2\right)}{1 + |g_{24}|^2 |x_{\mathrm{SU}}|^2 \left(\sigma_{n_2}^2/\sigma_{n_4}^2\right)}$$

where $\mathsf{ASNR}_{\mathrm{PU}\to\mathrm{SU}} \triangleq (\sigma_{14}^2 \mathsf{P}_{\mathrm{PU}})/\sigma_{n_4}^2$ is the ASNR at the SRx when $x_{\mathrm{SU}} = 0$. Such an upper bound can be approximated by using the same arguments that led to (19), (20), and (21).

At this point, we discuss some special cases/approximations of the bounds in (25), which provide insights on the ergodic channel capacity of the SU. Let us first consider the lower bound $\mathsf{C}_{\mathrm{SU,dirty,lower}}$. The matrix $\mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})$ in (25) is an Hermitian Toeplitz matrix and, for $\tau_{\mathrm{SU}} \neq 0$, the behavior of its inverse can be characterized by using asymptotic (i.e., for $M \to +\infty$) arguments [47]. A substantial simplification occurs when $\tau_{\mathrm{SU}} = 0$:[8] indeed, in this case, both $\mathbf{R}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}})$ and

[7]Such an assumption is reasonable when the SRx is sufficiently close to both the PTx and STx.

[8]This is a reasonable assumption when the PTx, STx, and SRx are sufficiently close in space, i.e., $d_{12} + d_{24} < c \mathsf{T}_{\mathrm{PU}}$, and the processing time at the STx is smaller than the symbol period of the PU, i.e., $t_{\mathrm{p}} < \mathsf{T}_{\mathrm{PU}}$. Results of numerical simulations (not reported here in the interest of saving space) show that the impact of τ_{SU} on the worst-case ergodic channel capacity of the SU is negligible if $\tau_{\mathrm{SU}} \ll M$.

$$\begin{aligned}&\frac{1}{M} \mathbb{E}_{\mathbf{g}_{\mathrm{SU}}} \left\{ \ln \left[1 + \sigma_{\mathrm{SU,dirty}}^2 \mathbb{E}_{x_{\mathrm{SU}} | \mathbf{g}_{\mathrm{SU}}} \left\{ \mathrm{tr} \left[\mathbf{R}_{\mathrm{SU}}^{-1}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \boldsymbol{\Sigma}_{\mathrm{SU}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \mathbf{R}_{\mathrm{SU}}^{-1}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \boldsymbol{\Sigma}_{\mathrm{SU}}^{\mathrm{H}}(x_{\mathrm{SU}}, \mathbf{g}_{\mathrm{SU}}) \right] \right\} \right] \right\} \\ &\quad \triangleq \mathsf{C}_{\mathrm{SU,dirty,lower}} \leq \mathsf{C}_{\mathrm{SU,dirty}} \leq \mathsf{C}_{\mathrm{SU,dirty,upper}} \triangleq \frac{1}{M} \mathbb{E} \left\{ \ln \left[1 + \frac{\sigma_{\mathrm{SU,dirty}}^2 |g_{24}|^2 |g_{12}|^2}{\sigma_{n_4}^2} \sum_{m=0}^{M-1} |x_{\mathrm{PU}}(m - \tau_{\mathrm{SU}})|^2 \right] \right\}\end{aligned} \tag{25}$$

$\boldsymbol{\Sigma}_{\text{SU}}(x_{\text{SU}}, \mathbf{g}_{\text{SU}})$ turn out to be scaled identity matrices, thus yielding (26), shown at the bottom of the page.

Lastly, we consider the upper bound $\mathsf{C}_{\text{SU,dirty,upper}}$. We observe that $\sum_{m=0}^{M-1} |x_{\text{PU}}(m-\tau_{\text{SU}})|^2/M$ converges almost surely to $\sigma^2_{\text{PU}} = \mathsf{P}_{\text{PU}}$ for $M \to +\infty$ by the strong law of large number [44]. If M is sufficiently large, we can use the approximation $\sum_{m=0}^{M-1} |x_{\text{PU}}(m-\tau_{\text{SU}})|^2 \approx M\mathsf{P}_{\text{PU}}$.[9] Thus,

$$\mathsf{C}_{\text{SU,dirty,upper}} \approx \frac{1}{M}\mathbb{E}\left[\ln\left(1 + M\frac{\mathsf{P}_{\text{SU}}}{\sigma^2_{n_4}}|g_{24}|^2|\overline{g}_{12}|^2\right)\right]$$
$$= \frac{1}{M}\mathbb{E}\left[f\left(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2\right)\right] \quad (27)$$

where we have used the approximation $\sigma^2_{\text{SU,dirty}} \approx \mathsf{P}_{\text{SU}}/(\sigma^2_{12}\mathsf{P}_{\text{PU}})$, whereas $\overline{g}_{12} \triangleq g_{12}/\sigma_{12}$ denotes the normalized version of the channel coefficient characterizing the PTx $\to$ STx link. The equality in (27) is obtained by observing that $|g_{24}|^2$ is an exponential RV with mean σ^2_{24}, evaluating the expectation with respect to the distribution of g_{24}, and using (1). Low- and high-SNR approximations of (27) can be obtained by using the fact that $|\overline{g}_{12}|^2$ is a unit-mean exponential RV. Specifically, let $\epsilon > 0$ be a sufficiently small real number, one has $P(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2 \geq \epsilon) = \exp[-\epsilon/(M\mathsf{ASNR}_{\text{SU,white}})]$. This shows that, for $M\mathsf{ASNR}_{\text{SU,white}} \ll \epsilon$, the RV $M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2$ takes on values significantly smaller than one, with high probability. In this case, using (1), the following approximation of (27) holds

$$\mathsf{C}_{\text{SU,dirty,upper}} \approx \frac{1}{M}\mathbb{E}\left[M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2\right] = \mathsf{ASNR}_{\text{SU,white}}$$

with $\mathsf{ASNR}_{\text{SU,white}} \approx \mathsf{C}_{\text{SU,white}}$ for $\mathsf{ASNR}_{\text{SU,white}} \ll 1$ [see (1) and (23)]. As intuitively expected, if the SRx has perfect knowledge of $\mathbf{g}_{\text{SU}}$ and $\check{\mathbf{x}}_{\text{PU}}$, the performance limit of the SU in the low-SNR regime is nearly equal to that when the STx and SRx communicate directly using a dedicated exclusive channel. On the other hand, let $K > 0$ be a sufficiently large real number, one has

$$P\left(M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2 \leq K\right)$$
$$= 1 - \exp\left[-K/(M\mathsf{ASNR}_{\text{SU,white}})\right]. \quad (28)$$

This shows that, for $M\mathsf{ASNR}_{\text{SU,white}} \gg K$, the RV $M\ \mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2$ takes on values significantly greater than one, with high probability. In this case, using (1) over and over again, we can approximate (27) as

$$\mathsf{C}_{\text{SU,dirty,upper}} \approx \frac{1}{M}\left\{\mathbb{E}\left[\ln(1 + M\mathsf{ASNR}_{\text{SU,white}}|\overline{g}_{12}|^2)\right] - \gamma\right\}$$
$$= \frac{1}{M}\left[f(M\mathsf{ASNR}_{\text{SU,white}}) - \gamma\right]$$
$$\approx \frac{1}{M}\left[\ln(1 + M\mathsf{ASNR}_{\text{SU,white}}) - 2\gamma\right]$$
$$\approx \frac{1}{M}\left[\ln(M\mathsf{ASNR}_{\text{SU,white}}) - 2\gamma\right]$$
$$= \frac{1}{M}\left[\ln(M) + \mathsf{C}_{\text{SU,white}} - \gamma\right] \quad (29)$$

where we have observed that $\ln(\mathsf{ASNR}_{\text{SU,white}}) - \gamma \approx \mathsf{C}_{\text{SU,white}}$ for $\mathsf{ASNR}_{\text{SU,white}} \gg 1$ [see eqs. (1) and (23)]. In the high-SNR regime, the achievable best-case capacity of the SU in the dirty-space mode becomes vanishingly small for $M \to +\infty$. This result is essentially due to the fact that the SU transmits only one symbol per PU frame, in order to assure primary channel stationarity during a frame.

In wireless networks, due to the fluctuation of the instantaneous capacity of fading channels, a buffer is typically used at the transmitter to adapt the source data traffic flow to the channel transmission capability. We will show by means of simple models that the use of portions of frequency band that are being used by the PU leads to a significant advantage for the SU in terms of average symbol delay (i.e., waiting time in queue plus transmission time), even when it has a low physical-layer rate in dirty-space mode. To this end, let us assume that SU data arrive at the buffer in symbols each carrying Q_{SU} bits. We evaluate the average delay D_{SU} for the transmission of a symbol with the proposed protocol and compare it with the average symbol delay $\mathsf{D}_{\text{SU,white}}$ when the SU interrupts its transmission during the ON intervals of the PU, i.e., it transmits in white-space mode only. The service system for the proposed protocol can be modeled as an M/G/1 queue [48], for which the mean of the symbol transmission delay is given by

$$\mathsf{D}_{\text{SU}} = \mathbb{E}[\mathsf{X}_{\text{SU}}] + \frac{\lambda_{\text{SU}}\mathbb{E}\left[\mathsf{X}^2_{\text{SU}}\right]}{2\left(1 - \lambda_{\text{SU}}\mathbb{E}[\mathsf{X}_{\text{SU}}]\right)} \quad (30)$$

where X_{SU} is the symbol service time of the proposed protocol and λ_{SU} (in symbols for second) is the average arrival rate. On the other hand, when the SU uses all its available power to transmit data only when the PU is inactive, we model the service system as an M/G/1 queue with vacations [48] and, in this case, the mean of the symbol transmission delay can be found as

$$\mathsf{D}_{\text{SU,white}} = \mathbb{E}[\mathsf{X}_{\text{SU,white}}] + \frac{\lambda_{\text{SU}}\mathbb{E}\left[\mathsf{X}^2_{\text{SU,white}}\right]}{2\left(1 - \lambda_{\text{SU}}\mathbb{E}[\mathsf{X}_{\text{SU,white}}]\right)} + \frac{\mathbb{E}\left[\mathsf{T}^2_{\text{ON}}\right]}{2\mathbb{E}[\mathsf{T}_{\text{ON}}]} \quad (31)$$

[9]Such an approximation turns out to be an equality if the PU adopts a constant-modulus constellation with average energy P_{PU}.

$$\mathsf{C}_{\text{SU,dirty,lower}} = \frac{1}{M}\mathbb{E}_{\mathbf{g}_{\text{SU}}}\left\{\ln\left[1 + M\mathsf{P}^2_{\text{PU}}\sigma^2_{\text{SU,dirty}}\mathbb{E}_{x_{\text{SU}}|\mathbf{g}_{\text{SU}}}\left\{\frac{\left||g_{24}|^2|g_{12}|^2 x_{\text{SU}} + g^*_{24}g^*_{12}g_{14}\right|^2}{\left[\Omega(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) + \Upsilon^*(x_{\text{SU}}, \mathbf{g}_{\text{SU}}) + \Upsilon(x_{\text{SU}}, \mathbf{g}_{\text{SU}})\right]^2}\right\}\right]\right\}. \quad (26)$$

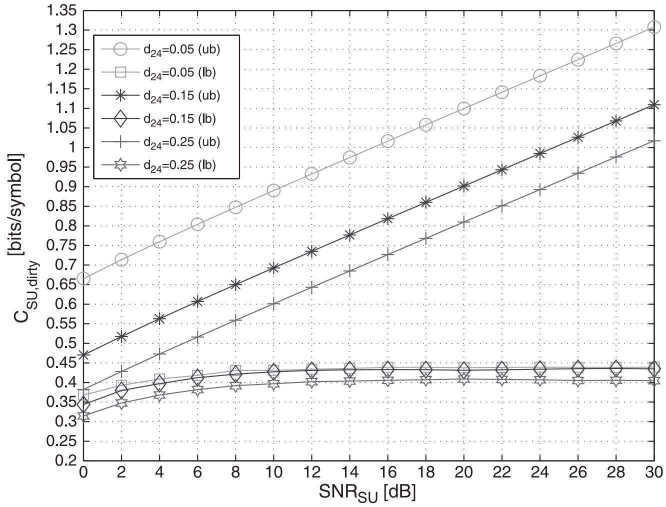

Fig. 4. $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ versus $\mathsf{SNR}_{\mathrm{SU}}$ for different values of d_{24} ($d_{12} = 0.5$, $\mathsf{P}_{\mathrm{PU}} = \mathsf{P}_{\mathrm{SU}} = 1$).

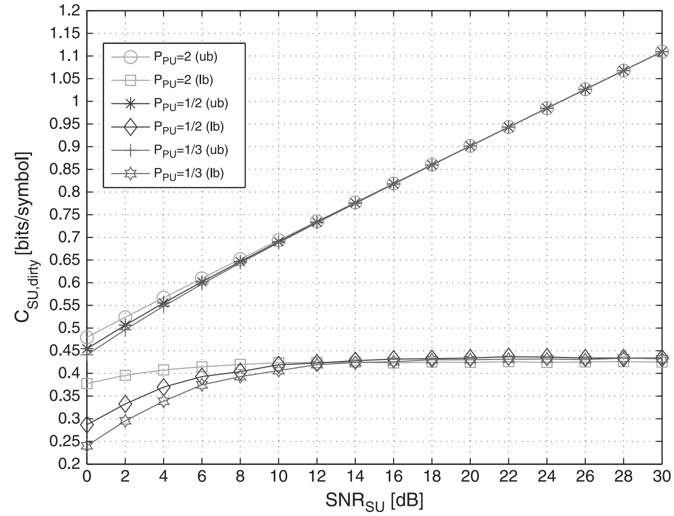

Fig. 5. $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ versus $\mathsf{SNR}_{\mathrm{SU}}$ for different values of P_{PU} ($d_{12} = 0.5$, $d_{24} = 0.15$, $\mathsf{P}_{\mathrm{SU}} = 1$).

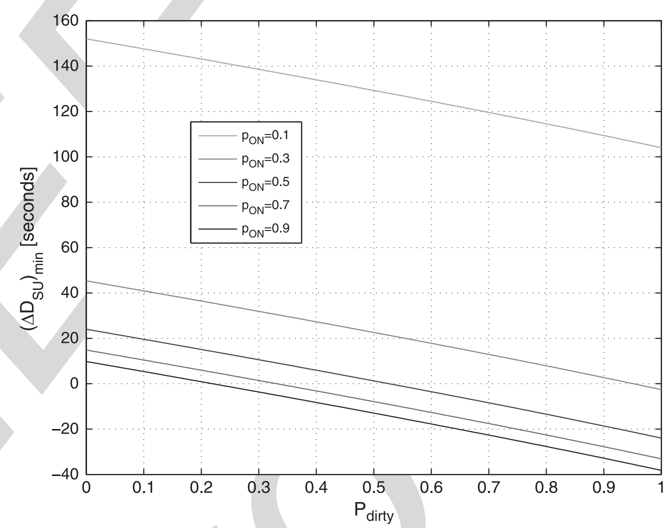

Fig. 6. $(\Delta\mathsf{D}_{\mathrm{SU}})_{\min}$ versus P_{dirty} for different values of p_{ON} ($\mathsf{SNR}_{\mathrm{SU}} = 20$ dB, $d_{24}/d_{12} = 0.3$, $\mathsf{P}_{\mathrm{PU}} = 1$, $\lambda_{\mathrm{SU}} = 0.01$).

where $\mathsf{X}_{\mathrm{SU,white}}$ is the symbol service time in white-space mode and T_{ON} is the duration of the vacations (i.e., PU ON intervals) taken by the SU transmitter.

To evaluate the ensemble averages involved in (30) and (31), since the PU transmission protocol works on a frame-by-frame basis and a PU frame lasts $\mathsf{T}_{\mathrm{SU,dirty}}$ seconds, we assume that $\mathsf{T}_{\mathrm{ON}} = \mathsf{I}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,dirty}}$, where I_{SU} is a geometric random variable with success probability p_{ON} and range $\{0, 1, 2, \ldots\}$. As a consequence, it results [44] that $\mathbb{E}[\mathsf{I}_{\mathrm{SU}}] = (1 - \mathsf{p}_{\mathrm{ON}})/\mathsf{p}_{\mathrm{ON}}$ and $\mathbb{E}[\mathsf{I}_{\mathrm{SU}}^2] = [(1 - \mathsf{p}_{\mathrm{ON}})(2 - \mathsf{p}_{\mathrm{ON}})]/\mathsf{p}_{\mathrm{ON}}^2$. Regarding the first and second moments of X_{SU} and $\mathsf{X}_{\mathrm{SU,white}}$, we assume that the SU can transmit at rates near the information-theoretic limits in both white- and dirty-space modes and, hence, we use the ergodic capacity to measure the data transmission capability of the wireless link.[10] Specifically, for the proposed protocol, we assume that the packet service time X_{SU} is a discrete binary random variable assuming the values

$$\mathsf{x}_{\mathrm{SU,dirty}} \triangleq (\mathsf{Q}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,dirty}}) \Big/ \left[\log_2(e)\mathsf{C}_{\mathrm{SU,dirty}}\right]$$

$$\mathsf{x}_{\mathrm{SU,white}} \triangleq (\mathsf{Q}_{\mathrm{SU}}\mathsf{T}_{\mathrm{SU,white}}) \Big/ \left[\log_2(e)\mathsf{C}_{\mathrm{SU,white}}\right]$$

with probabilities P_{dirty} and $1 - P_{\mathrm{dirty}}$, respectively, where P_{dirty} is the probability that the PU channel is busy. In white-space mode, we assume that the service time $\mathsf{X}_{\mathrm{SU,white}}$ is deterministic and it is given by $\mathsf{x}_{\mathrm{SU,white}}$.

To support the performance analysis of the SU, we report the results of numerical simulations. Specifically, we plot in Figs. 4 and 5 the upper and lower bounds on $\mathsf{C}_{\mathrm{SU,dirty}}$ (reported in bits/symbol and referred to as "ub" and "lb," respectively), whereas the *worst-case* value $(\Delta\mathsf{D}_{\mathrm{SU}})_{\min}$ of the difference $\Delta\mathsf{D}_{\mathrm{SU}} \triangleq \mathsf{D}_{\mathrm{SU,white}} - \mathsf{D}_{\mathrm{SU}}$ is reported in Fig. 6 by replacing $\mathsf{C}_{\mathrm{SU,dirty}}$ with $\mathsf{C}_{\mathrm{SU,dirty,lower}}$. The ensemble averages in (25) are evaluated through 10^4 Monte Carlo trials. With reference to Fig. 1, the distance between the PTx and the STx is fixed to $d_{12} = 0.5$ and $\vartheta = \pi/3$, whereas the transmitting power of the SU is normalized by setting $\mathsf{P}_{\mathrm{SU}} = 1$; moreover, we chose $M = 16$, $\tau_{\mathrm{SU}} = 2$, $\mathsf{Q}_{\mathrm{SU}} = 1$ (i.e., binary modulation), $\mathsf{T}_{\mathrm{SU,white}} = \mathsf{T}_{\mathrm{PU}} = 1$, and set the path-loss exponent equal to $\eta = 2$. The PU symbols $x_{\mathrm{PU}}(\cdot)$ were generated as circularly symmetric complex Gaussian RVs with average energy P_{PU} and, similarly, the SU symbol x_{SU} was modeled as a circularly symmetric complex Gaussian RV with variance $\sigma^2_{\mathrm{SU,dirty}}$ and $\sigma^2_{\mathrm{SU,white}}$ in dirty- and white-space mode, respectively.

3) Fig. 4 : It reports the upper and lower bounds on the SU capacity as a function of $\mathsf{SNR}_{\mathrm{SU}} \triangleq \mathsf{P}_{\mathrm{SU}}/\sigma^2_{n_4}$ for different values of the distance d_{24} between the STx and the SRx (the distance d_{14} between the PTx and the SRx is calculated according to the Carnot's cosine law), with $\mathsf{P}_{\mathrm{PU}} = 1$. It is seen that the best performance significantly increases as the SRx

[10]Depending on the dynamics of the fading process, a long coding delay (i.e., the amount of time required to encode/decode packets) may be required to approach ergodic capacity [33]. However, since the SU uses the same encoding/decoding strategy in both white- and dirty-space modes, such a coding delay does not affect the comparison between D_{SU} and $\mathsf{D}_{\mathrm{SU,white}}$.

brings nearer to the STx. For instance, when the distance d_{24} between the STx and the SRx is about one-third of the distance d_{14} between the PTx and the STx, the achievable rate is greater than or equal to one bit per symbol for $\mathsf{SNR}_{\mathrm{SU}} \geq 24$ dB. On the other hand, the worst performance is less dependent on d_{24}. There is a significant gap between the best and the worst performances of the SU, thus evidencing that decoding of the PU symbols at the SRx is important to achieve reasonable rates. In particular, contrary to $\mathsf{C}_{\mathrm{SU,dirty,upper}}$, the worst-case capacity $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ in (25) does not grow without bound as $\sigma_{n_4}^2 \to 0$, thus exhibiting a marked floor in the high-SNR region. Such a behavior is due to the fact that the $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ is derived under the assumption that the PU symbols are unknown at the SRx (they are modeled as Gaussian RVs).

4) Fig. 5 : The upper and lower bounds of the SU capacity are depicted as a function of $\mathsf{SNR}_{\mathrm{SU}}$ for different values of the PU transmitting power P_{PU}, with $d_{24}/d_{12} = 0.3$. Differently from the PU case (see Fig. 3), increasing the power ratio $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ has negligible impact on the SU capacity, except for very small values of $\mathsf{SNR}_{\mathrm{SU}}$. In accordance with (29), the upper bound $\mathsf{C}_{\mathrm{SU,dirty,upper}}$ does not depend on P_{PU} at all for moderate-to-high SNR values, since in this case the SRx has perfect knowledge of the PU symbols; on the other hand, the lower bound $\mathsf{C}_{\mathrm{SU,dirty,lower}}$ exhibits a weak dependence on P_{PU}.

5) Fig. 6 : It reports the worst-case difference $(\Delta \mathsf{D}_{\mathrm{SU}})_{\min}$ between the average delays (31) and (30) as a function of the probability P_{dirty} that the PU channel is busy, for different values of the success probability p_{ON}, which lead to different values of $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}]/\mathsf{T}_{\mathrm{SU,dirty}} = (1-\mathsf{p}_{\mathrm{ON}})/\mathsf{p}_{\mathrm{ON}}$. Results of Fig. 6 are obtained by setting $\mathsf{P}_{\mathrm{PU}} = 1$, $\mathsf{SNR}_{\mathrm{SU}} = 20$ dB, $d_{24}/d_{12} = 0.3$, and $\lambda_{\mathrm{SU}} = 0.01$ symbol/seconds. It is seen that transmitting at a low data rate in dirty-space mode allows to significantly reduce the average delay incurred by the SU data, for each value of P_{dirty}, when $\mathsf{p}_{\mathrm{ON}} < 0.3$, i.e., the mean duration $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}]$ of the PU ON intervals is roughly 2.4 times greater than the PU frame duration $\mathsf{T}_{\mathrm{SU,dirty}}$. On the other hand, for $\mathsf{p}_{\mathrm{ON}} \geq 0.3$, the proposed protocol ensures a delay improvement for low-to-moderate values of P_{dirty}: for instance, when $\mathbb{E}[\mathsf{T}_{\mathrm{ON}}] = \mathsf{T}_{\mathrm{SU,dirty}}$, that is, $\mathsf{p}_{\mathrm{ON}} = 0.5$, an improvement can be observed when the PU channel is occupied for about less than 50% of time.

V. Conclusion

We proposed an AF scheme that allows a SU to concurrently transmit in the same frequency band of a PU not only when the PU is inactive, but also when the PU channel is busy. This can be obtained without requiring any noncausal knowledge of the PU information symbols. The main results of our performance analyses in terms of both PU and SU ergodic channel capacities can be summarized as follows. The cognitive radio principle of protecting the PU is not only guaranteed, but even a performance improvement can be gained by the PU in terms of ergodic channel capacity. Such a performance gain depends on the distance ratio d_{12}/d_{13}, the path-loss exponent η, and the power ratios $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ and $\sigma_{n_2}^2/\sigma_{n_3}^2$. Regarding the achievable rate of the SU, the secondary link can support more than one bit per symbol for moderate-to-high SNR values if the SRx is able to decode the PU data. Moreover, a variation of the power ratio $\mathsf{P}_{\mathrm{SU}}/\mathsf{P}_{\mathrm{PU}}$ does not lead to appreciable effects for the SU. Notwithstanding the transmission of a single SU symbol per frame gives low information rates in dirty-space mode, the delay performance of the SU improves noticeably.

It is noteworthy that several PUs typically multiplex the frame resources in time or in frequency. Therefore, the SU might be active in parallel over all the channels allocated for the PUs, potentially attaining larger transmission rates without adding interference. Modification of the proposed protocol and evaluation of the corresponding capacity performance is left as a future development.

Appendix A
Upper Bound on the PU Ergodic Channel Capacity

An upper bound on C_{PU} can be obtained by assuming that the PRx additionally has perfect knowledge of the fading coefficient g_{12} characterizing the PTx $\to$ STx link. Indeed, let $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12})$ denote the conditional mutual information between $\tilde{x}_{\mathrm{PU}}(\ell)$ and $\tilde{y}_{\mathrm{PU}}(\ell)$, given $\mathbf{g}_{\mathrm{PU}}$ and g_{12}, by using the chain rule for mutual information [31], it can be proven that

$$\mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12}\right) = \mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}\right) + \underbrace{\mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}\right)}_{\geq 0} \geq \mathsf{I}\left(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}\right) \quad (32)$$

where the equality holds if and only if $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}) = 0$, i.e., when $\tilde{x}_{\mathrm{PU}}(\ell)$ and g_{12} are conditionally independent given $\tilde{y}_{\mathrm{PU}}(\ell)$ and $\mathbf{g}_{\mathrm{PU}}$.[11]

For given values of $\mathbf{g}_{\mathrm{PU}}$ and g_{12}, the equivalent noise term $v_{\mathrm{PU}}(\ell)$ in (5) is a circular symmetric zero-mean complex Gaussian RV with variance $\sigma_{n_3}^2 + \sigma_{n_2}^2 |g_{23}|^2 |x_{\mathrm{SU}}|^2$. Consequently, the vector $\tilde{\mathbf{v}}_{\mathrm{PU}}$ in (11) is composed of i.i.d. $\mathcal{CN}(0, \sigma_{n_3}^2 + \sigma_{n_2}^2 |g_{23}|^2 |x_{\mathrm{SU}}|^2)$ RVs and, thus, the subchannels (9) are also Gaussian. In this case, the supremum of $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12})$ over all distributions on $\tilde{x}_{\mathrm{PU}}(\ell)$ that satisfy the power constraint $\sigma_{\mathrm{PU}}^2 = \mathsf{P}_{\mathrm{PU}}$ is attained [31] when $\tilde{x}_{\mathrm{PU}}(\ell) \equiv \tilde{x}_{\mathrm{PU,G}}(\ell) \sim \mathcal{CN}(0, \mathsf{P}_{\mathrm{PU}})$, regardless of the operating SNR. With this choice, one obtains

$$\mathsf{I}\left(\tilde{x}_{\mathrm{PU,G}}(\ell); \tilde{y}_{\mathrm{PU}}(\ell)|\mathbf{g}_{\mathrm{PU}}, g_{12}\right) = \mathbb{E}_{\mathbf{g}_{\mathrm{PU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\mathrm{PU}}\, |G_{\mathrm{PU}}(\ell)|^2}{\sigma_{n_3}^2 + \sigma_{n_2}^2 |g_{23}|^2 |x_{\mathrm{SU}}|^2}\right]\right\} \quad (33)$$

and, accounting for (32), the ergodic capacity (nats/symbol) of the parallel fading channel (9) is upper bounded by

$$\mathsf{C}_{\mathrm{PU}}(M) \leq \frac{1}{M}\sum_{\ell=0}^{M-1} \mathbb{E}_{\mathbf{g}_{\mathrm{PU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\mathrm{PU}}|G_{\mathrm{PU}}(\ell)|^2}{\sigma_{n_3}^2 + \sigma_{n_2}^2 |g_{23}|^2 |x_{\mathrm{SU}}|^2}\right]\right\}. \quad (34)$$

[11] The fact that $\tilde{x}_{\mathrm{PU}}(\ell)$ and g_{12} are statistically independent neither implies nor is implied by $\mathsf{I}(\tilde{x}_{\mathrm{PU}}(\ell); g_{12}|\tilde{y}_{\mathrm{PU}}(\ell), \mathbf{g}_{\mathrm{PU}}) = 0$.

To evaluate the expectation in (34), it is useful to observe that, conditioned on $g_{23}x_{\text{SU}}$, one has $G_{\text{PU}}(\ell)|g_{23}x_{\text{SU}} \sim \mathcal{CN}(0, \sigma_{13}^2 + \sigma_{12}^2|g_{23}|^2|x_{\text{SU}}|^2)$, $\forall \ell \in \{0, 1, \ldots, M-1\}$, whose squared magnitude is exponentially distributed with mean $\sigma_{13}^2 + \sigma_{12}^2|g_{23}|^2|x_{\text{SU}}|^2$. Thus, using (1), it follows that:

$$\mathbb{E}_{\mathbf{g}_{\text{PU}}|g_{23}x_{\text{SU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\text{PU}}\,|G_{\text{PU}}(\ell)|^2}{\sigma_{n_3}^2 + \sigma_{n_2}^2|g_{23}|^2|x_{\text{SU}}|^2}\right]\right\} = f\left[\Gamma_{3,\text{upper}}\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\right] \quad (35)$$

where $\Gamma_{3,\text{upper}}(|g_{23}|^2|x_{\text{SU}}|^2)$ is a transformation of the RV $g_{23}x_{\text{SU}}$ defined in (15). Therefore, the terms of the sum in (34) do not depend on ℓ and, thus, capacity (34) does not depend on M. Hence, by applying the conditional expectation rule in (34) and using (35), one obtains the upper bound in (14) on the ergodic channel capacity of the PU [see eq. (12)].

APPENDIX B
LOWER BOUND ON THE PU ERGODIC CHANNEL CAPACITY

To find a lower bound on C_{PU}, we observe that the Gaussian distribution might not be the one maximizing $\mathsf{I}(\tilde{x}_{\text{PU}}(\ell); \tilde{y}_{\text{PU}}(\ell)|\mathbf{g}_{\text{PU}})$ [see eqs. (12), (9)] and, thus, we choose $\tilde{x}_{\text{PU}}(\ell) \equiv \tilde{x}_{\text{PU,G}}(\ell) \sim \mathcal{CN}(0, \mathsf{P}_{\text{PU}})$. Moreover, let $v_{\text{PU,G}}(\cdot)$ be a sequence of i.i.d. circularly symmetric complex Gaussian RVs having the same mean and variance as $v_{\text{PU}}(\cdot)$ in (5), i.e., $v_{\text{PU,G}}(\cdot) \sim \mathcal{CN}\left(0, \sigma_{n_3}^2 + \sigma_{n_2}^2\sigma_{23}^2\sigma_{\text{SU,dirty}}^2\right)$, independent of $\tilde{x}_{\text{PU,G}}(\cdot)$. Let us also consider the frequency-domain vector

$$\tilde{\mathbf{v}}_{\text{PU,G}} = \left[\tilde{v}_{\text{PU,G}}(0), \tilde{v}_{\text{PU,G}}(1), \ldots, \tilde{v}_{\text{PU,G}}(M-1)\right]^{\text{T}} \triangleq \mathbf{W}_{\text{DFT}}\mathbf{v}_{\text{PU,G}} \in \mathbb{C}^M$$

with $\mathbf{v}_{\text{PU,G}} \triangleq [v_{\text{PU,G}}(0), v_{\text{PU,G}}(1), \ldots, v_{\text{PU,G}}(M-1)]^{\text{T}} \in \mathbb{C}^M$. It is readily seen that the entries of $\tilde{\mathbf{v}}_{\text{PU,G}}$ are i.i.d. $\mathcal{CN}\left(0, \sigma_{n_3}^2 + \sigma_{n_2}^2\sigma_{23}^2\sigma_{\text{SU,dirty}}^2\right)$ RVs. By replacing $\tilde{x}_{\text{PU}}(\ell)$ and $\tilde{v}_{\text{PU}}(\ell)$ in (9) with $\tilde{x}_{\text{PU,G}}(\ell)$ and $\tilde{v}_{\text{PU,G}}(\ell)$, respectively, we get the subchannels with both Gaussian input and Gaussian noise:

$$\tilde{y}_{\text{PU,G}}(\ell) = G_{\text{PU}}(\ell)\tilde{x}_{\text{PU,G}}(\ell) + \tilde{v}_{\text{PU,G}}(\ell)$$

for $\ell \in \{0, 1, \ldots, M-1\}$. Additionally, let

$$\tilde{y}_{\text{PU,NG}}(\ell) = G_{\text{PU}}(\ell)\tilde{x}_{\text{PU,G}}(\ell) + \tilde{v}_{\text{PU}}(\ell)$$

be the corresponding subchannels with Gaussian input and non-Gaussian noise, $\ell \in \{0, 1, \ldots, M-1\}$. Since conditional mutual information can be equivalently expressed as difference between conditional differential entropies [31], one obtains

$$\begin{aligned}
&\mathsf{I}\left(\tilde{x}_{\text{PU,G}}(\ell); \tilde{y}_{\text{PU,NG}}(\ell)|\mathbf{g}_{\text{PU}}\right) - \mathsf{I}\left(\tilde{x}_{\text{PU,G}}(\ell); \tilde{y}_{\text{PU,G}}(\ell)|\mathbf{g}_{\text{PU}}\right) \\
&\quad = \mathsf{h}\left(\tilde{x}_{\text{PU,G}}(\ell)|\tilde{y}_{\text{PU,G}}(\ell), \mathbf{g}_{\text{PU}}\right) \\
&\qquad - \mathsf{h}\left(\tilde{x}_{\text{PU,G}}(\ell)|\tilde{y}_{\text{PU,NG}}(\ell), \mathbf{g}_{\text{PU}}\right) \geq 0 \qquad (36)
\end{aligned}$$

where the inequality holds for each realization of $\mathbf{g}_{\text{PU}}$ and whatever is the probability distribution of $\tilde{v}_{\text{PU}}(\ell)$.[12] Consequently, we have $\mathsf{I}(\tilde{x}_{\text{PU,G}}(\ell); \tilde{y}_{\text{PU,NG}}(\ell)|\mathbf{g}_{\text{PU}}) \geq \mathsf{I}(\tilde{x}_{\text{PU,G}}(\ell); \tilde{y}_{\text{PU,G}}(\ell)|\mathbf{g}_{\text{PU}})$ and, by doing calculations similar to those of Appendix A [see, in particular, eqs. (33) and (35)], we get the lower bound (in nats/symbol) on the ergodic channel capacity of the PU [see eq. (12)]

$$\begin{aligned}
\mathsf{C}_{\text{PU}} \geq \mathsf{C}_{\text{PU,lower}} &\triangleq \lim_{M\to+\infty} \frac{1}{M}\sum_{\ell=0}^{M-1} \mathsf{I}\left(\tilde{x}_{\text{PU,G}}(\ell); \tilde{y}_{\text{PU,G}}(\ell)|\mathbf{g}_{\text{PU}}\right) \\
&= \lim_{M\to+\infty} \frac{1}{M}\sum_{\ell=0}^{M-1} \mathbb{E}_{\mathbf{g}_{\text{PU}}}\left\{\ln\left[1 + \frac{\mathsf{P}_{\text{PU}}|G_{\text{PU}}(\ell)|^2}{\sigma_{n_3}^2 + \sigma_{n_2}^2\sigma_{23}^2\sigma_{\text{SU,dirty}}^2}\right]\right\} \\
&= \mathbb{E}\left\{f\left[\Gamma_{3,\text{lower}}\left(|g_{23}|^2|x_{\text{SU}}|^2\right)\right]\right\}
\end{aligned}$$

where $\Gamma_{3,\text{lower}}(|g_{23}|^2|x_{\text{SU}}|^2)$ is a transformation of the RV $g_{23}x_{\text{SU}}$ defined in (16).

APPENDIX C
UPPER BOUND ON THE SU ERGODIC CHANNEL CAPACITY

An upper bound on $\mathsf{C}_{\text{SU,dirty}}$ can be obtained by assuming that the SRx has perfect knowledge of both the realization of the channel vector $\mathbf{g}_{\text{SU}}$ and the PU symbol block $\check{\mathbf{x}}_{\text{PU}}$. In this case, the channel output consists of the triplet $(\mathbf{y}_{\text{SU}}, \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}})$ and, thus, the mutual information between channel input and output (in nats/PU frame) is represented by $\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}, \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}})$. Owing to the statistical independence among x_{SU}, $\mathbf{g}_{\text{SU}}$, and $\check{\mathbf{x}}_{\text{PU}}$, which implies that $\mathsf{I}(x_{\text{SU}}; \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) = 0$, application of the chain rule [31] for mutual information allows one to write that $\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}, \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) = \mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}})$, i.e., it is equal to the conditional mutual information between x_{SU} and $\mathbf{y}_{\text{SU}}$, given $\mathbf{g}_{\text{SU}}$ and $\check{\mathbf{x}}_{\text{PU}}$. Similarly to (32), it results that

$$\begin{aligned}
\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) &= \mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}) \\
&\quad + \underbrace{\mathsf{I}(x_{\text{SU}}; \check{\mathbf{x}}_{\text{PU}}|\mathbf{y}_{\text{SU}}, \mathbf{g}_{\text{SU}})}_{\geq 0} \geq \mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}). \qquad (37)
\end{aligned}$$

Consequently, an upper bound $\mathsf{C}_{\text{SU,dirty,upper}}$ (in nats/symbol) on the ergodic channel capacity of the SU link is obtained by taking the supremum of $\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}})/M$ over all possible distributions of the symbol x_{SU} that satisfy the power constraint (3). Since conditional mutual information can be equivalently expressed as difference between conditional differential entropies [31], one has

$$\begin{aligned}
\mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) &= \mathsf{h}(\mathbf{y}_{\text{SU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) \\
&\quad - \mathsf{h}(\mathbf{y}_{\text{SU}}|x_{\text{SU}}, \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) = \mathsf{h}(\mathbf{y}_{\text{SU}} - g_{14}\mathbf{x}_{\text{PU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) \\
&\quad - \mathsf{h}(\mathbf{y}_{\text{SU}} - g_{14}\mathbf{x}_{\text{PU}}\,|x_{\text{SU}}, \mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) \\
&= \mathsf{I}(x_{\text{SU}}; \mathbf{y}_{\text{SU}} - g_{14}\mathbf{x}_{\text{PU}}|\mathbf{g}_{\text{SU}}, \check{\mathbf{x}}_{\text{PU}}) \qquad (38)
\end{aligned}$$

[12]The inequality is a consequence of two facts: (i) given $G_{\text{PU}}(\ell)$, the RVs $\tilde{x}_{\text{PU,G}}(\ell)$, and $G_{\text{PU}}(\ell)\tilde{x}_{\text{PU,G}}(\ell) + \tilde{v}_{\text{PU}}(\ell)$ are jointly circularly symmetric complex Gaussian since $\tilde{x}_{\text{PU,G}}(\ell)$ and $\tilde{v}_{\text{PU}}(\ell)$ are independent of each other and each one is circularly symmetric complex Gaussian; (ii) jointly Gaussian RVs maximize conditional differential entropy [31], [49].

$$\mathsf{MSE}(\mathbf{g}_{\rm SU}) \geq J^{-1}(\mathbf{g}_{\rm SU}) \triangleq \left\{ \mathbb{E}_{\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU}} \left[\left| \frac{\partial}{\partial x_{\rm SU}^*} \ln p(\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU}) \right|^2 \right] \right\}^{-1} \tag{41}$$

$$J(\mathbf{g}_{\rm SU}) = \mathbb{E}_{x_{\rm SU}|\mathbf{g}_{\rm SU}} \left\{ \mathbb{E}_{\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU}} \left[\left| \frac{\partial}{\partial x_{\rm SU}^*} \ln p(\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU}) \right|^2 \middle| x_{\rm SU} \right] \right\} + \mathbb{E}_{x_{\rm SU}} \left[\left| \frac{\partial}{\partial x_{\rm SU}^*} \ln p(x_{\rm SU}) \right|^2 \right] \tag{44}$$

where we have used the fact that subtracting a constant does not change differential entropy [31]. Strictly speaking, since SRx knows g_{14} and $\check{\mathbf{x}}_{\rm PU}$, it can decode $x_{\rm SU}$ by subtracting $g_{14}\mathbf{x}_{\rm PU}$ from (24), which amounts to a repetition coding [37] transmission scheme over a fast flat-fading channel with i.i.d. $\mathcal{CN}(0, \sigma_{n_4}^2)$ noise samples. According to (24), conditioned on $\mathbf{g}_{\rm SU}$ and $\check{\mathbf{x}}_{\rm PU}$, a sufficient statistic for detecting $x_{\rm SU}$ from $\mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}$ is given by the scalar

$$\begin{aligned} \check{y}_{\rm SU} &\triangleq (g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\rm PU})^{\rm H}(\mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}) \\ &= \|g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\rm PU}\|^2 x_{\rm SU} + (g_{24}g_{12}\mathbf{J}\check{\mathbf{x}}_{\rm PU})^{\rm H}\mathbf{v}_4 \end{aligned} \tag{39}$$

which is interpreted as an AWGN channel with SNR equal to $(\sigma_{\rm SU,dirty}^2 |g_{24}|^2 |g_{12}|^2 \sum_{m=0}^{M-1} |x_{\rm PU}(m-\tau_{\rm SU})|^2)/\sigma_{n_4}^2$. Since sufficient statistics preserve mutual information [31], one has $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU} - g_{14}\mathbf{x}_{\rm PU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU}) = \mathsf{I}(x_{\rm SU}; \check{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})$. The supremum of $\mathsf{I}(x_{\rm SU}; \check{y}_{\rm SU}|\mathbf{g}_{\rm SU}, \check{\mathbf{x}}_{\rm PU})$ over all distributions on $x_{\rm SU}$ satisfying (3) is attained [31] when $x_{\rm SU} \equiv x_{\rm SU,G} \sim \mathcal{CN}\left(0, \sigma_{\rm SU,dirty}^2\right)$, thus leading to the upper bound in (25).

Appendix D
Lower Bound on the SU Ergodic Channel Capacity

A general lower bound on $\mathsf{C}_{\rm SU,dirty}$ can be derived by linking the mutual information $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU})$ to the conditional symbol estimation error of the SU. Specifically, the SRx uses the observation vector $\mathbf{y}_{\rm SU}$ in (24) to produce a reliable estimate $\hat{x}_{\rm SU}$ of the symbol $x_{\rm SU}$. By virtue of the data processing theorem [31], any function of the channel output $\mathbf{y}_{\rm SU}$ cannot increase the information about $x_{\rm SU}$, i.e., $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}) \geq \mathsf{I}(x_{\rm SU}; \hat{x}_{\rm SU}|\mathbf{g}_{\rm SU})$, and using a technique similar to that used in [50], the following general lower bound on $\mathsf{I}(x_{\rm SU}; \hat{x}_{\rm SU}|\mathbf{g}_{\rm SU})$ can be obtained

$$\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}) \geq \mathsf{h}(x_{\rm SU}) - \mathbb{E}_{\mathbf{g}_{\rm SU}}\left\{\ln\left[\pi e \mathsf{MSE}(\mathbf{g}_{\rm SU})\right]\right\} \tag{40}$$

where, according to the maximum-entropy theorem for complex RVs [51], the inequality holds with equality if and only if $x_{\rm SU} - \hat{x}_{\rm SU}|\mathbf{g}_{\rm SU} \sim \mathcal{CN}[0, \mathsf{MSE}(\mathbf{g}_{\rm SU})]$, with $\mathsf{MSE}(\mathbf{g}_{\rm SU}) \triangleq \mathbb{E}_{\mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}}[|x_{\rm SU} - \hat{x}_{\rm SU}|^2]$ denoting the mean-square error (MSE) of the symbol estimate, given $\mathbf{g}_{\rm SU}$. Interestingly, eq. (40) provides a lower bound on $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU})$ that depends on the conditional mean-square value of the symbol estimation error and applies for *any* symbol estimation strategy.

By modeling $x_{\rm SU}$ as a random variable with a given *a priori* pdf $p(x_{\rm SU})$, whose particular realization has to be estimated, a lower bound on $\mathsf{MSE}(\mathbf{g}_{\rm SU})$ is given by the complex counterpart of the Bayesian Cramér-Rao inequality [46], [52] that is reported in (41), shown at the top of the page, where $p(\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU})$ is the conditional joint pdf of $\mathbf{y}_{\rm SU}$ and $x_{\rm SU}$, given $\mathbf{g}_{\rm SU}$, whereas $J(\mathbf{g}_{\rm SU})$ is referred to as the Bayesian Fisher information.[13] The lower bound (41) is valid for any $p(\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU})$ satisfying some regularity conditions in addition to a "weak unbiasedness" condition [52], which are typically fulfilled by Gaussian distributions [46]. The lower bound (40) is valid for any $\mathsf{MSE}(\mathbf{g}_{\rm SU})$ and, therefore, it also holds when $\mathsf{MSE}(\mathbf{g}_{\rm SU})$ is replaced with its minimum value $J^{-1}(\mathbf{g}_{\rm SU})$ given by (41), thus having

$$\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}) \geq \mathsf{h}(x_{\rm SU}) - \ln(\pi e) + \mathbb{E}_{\mathbf{g}_{\rm SU}}\left\{\ln\left[J(\mathbf{g}_{\rm SU})\right]\right\} \tag{42}$$

which is actually a lower bound if an efficient symbol estimator exists, i.e., it attains the Bayesian CRB.

To facilitate the derivation of the Fisher information $J(\mathbf{g}_{\rm SU})$ in a closed form, the PU symbols are modelled as $\check{\mathbf{x}}_{\rm PU} \equiv \check{\mathbf{x}}_{\rm PU,G} \sim \mathcal{CN}(\mathbf{0}_M, \mathsf{P}_{\rm PU}\mathbf{I}_{M+\tau_{\rm SU}})$. In such a case, it is verified that, given $x_{\rm SU}$ and $\mathbf{g}_{\rm SU}$, one has $\mathbf{y}_{\rm SU} \sim \mathcal{CN}[\mathbf{0}_M, \mathbf{R}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU})]$ [see eq. (24)], where

$$\begin{aligned} \mathbf{R}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU}) &\triangleq \mathbb{E}\left[\mathbf{y}_{\rm SU}\mathbf{y}_{\rm SU}^{\rm H}|x_{\rm SU}, \mathbf{g}_{\rm SU}\right] = \Omega(x_{\rm SU}, \mathbf{g}_{\rm SU})\mathbf{I}_M \\ &\quad + \Upsilon^*(x_{\rm SU}, \mathbf{g}_{\rm SU})\mathbf{B}^{\tau_{\rm SU}} + \Upsilon(x_{\rm SU}, \mathbf{g}_{\rm SU})\mathbf{F}^{\tau_{\rm SU}} \end{aligned} \tag{43}$$

is a tridiagonal Hermitian Toeplitz matrix, with $\Omega(x_{\rm SU}, \mathbf{g}_{\rm SU}) \triangleq (|g_{24}|^2|g_{12}|^2|x_{\rm SU}|^2 + |g_{14}|^2)\mathsf{P}_{\rm PU} + \sigma_{n_4}^2$, $\Upsilon(x_{\rm SU}, \mathbf{g}_{\rm SU}) \triangleq g_{24}\, g_{12}\, x_{\rm SU} g_{14}^* \mathsf{P}_{\rm PU}$, $\mathbf{B} \in \mathbb{R}^{M\times M}$ and $\mathbf{F} \in \mathbb{R}^{M\times M}$ being backward-shift and forward-shift matrices [53], respectively.[14] By applying the conditional expectation rule, using the fact that $p(\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU}) = p(\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU})p(x_{\rm SU}|\mathbf{g}_{\rm SU})$, remembering that $x_{\rm SU}$ and $\mathbf{g}_{\rm SU}$ are statistically independent, and accounting for the regularity conditions [46], [52], one obtains (44), shown at the top of the page, where, since $\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU} \sim \mathcal{CN}[\mathbf{0}_M, \mathbf{R}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU})]$, it follows that (see, e.g., [45])

$$\begin{aligned} &\mathbb{E}_{x_{\rm SU}|\mathbf{g}_{\rm SU}} \left\{ \mathbb{E}_{\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU}} \left[\left| \frac{\partial}{\partial x_{\rm SU}^*} \ln p(\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU}) \right|^2 |x_{\rm SU} \right] \right\} \\ &\quad = \mathbb{E}_{x_{\rm SU}|\mathbf{g}_{\rm SU}} \left\{ \mathrm{tr} \left[\mathbf{R}_{\rm SU}^{-1}(x_{\rm SU}, \mathbf{g}_{\rm SU}) \boldsymbol{\Sigma}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU}) \right.\right. \\ &\qquad \left.\left. \cdot\, \mathbf{R}_{\rm SU}^{-1}(x_{\rm SU}, \mathbf{g}_{\rm SU}) \boldsymbol{\Sigma}_{\rm SU}^{\rm H}(x_{\rm SU}, \mathbf{g}_{\rm SU}) \right] \right\} \end{aligned} \tag{45}$$

[13] Inequality (41) also holds for biased estimators, in contrast to the standard Cramér-Rao bound (CRB) where the Fisher information is obtained from (41) by replacing $p(\mathbf{y}_{\rm SU}, x_{\rm SU}|\mathbf{g}_{\rm SU})$ with $p(\mathbf{y}_{\rm SU}|x_{\rm SU}, \mathbf{g}_{\rm SU})$ and carrying out the ensemble average with respect to $\mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU}$.

[14] It results that $\mathbf{B}^{\rm T} = \mathbf{F}$ by construction [53].

with [see eq. (43)]

$$\begin{aligned}\mathbf{\Sigma}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU}) &\triangleq \partial/\partial x^*_{\rm SU} \mathbf{R}_{\rm SU}(x_{\rm SU}, \mathbf{g}_{\rm SU}) \\ &= \left(|g_{24}|^2|g_{12}|^2 x_{\rm SU}\mathbf{I}_M + g^*_{24}g^*_{12}g_{14}\mathbf{B}^{\tau_{\rm SU}}\right)\mathsf{P}_{\rm PU}\end{aligned} \tag{46}$$

and we have recalled that, for arbitrary matrices $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$, it results [45] $\mathbb{E}[\mathbf{x}^{\rm H}\mathbf{A}\,\mathbf{x}] = {\rm tr}(\mathbf{A}\,\Sigma)$ and, for $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_n, \Sigma)$, $\mathbb{E}[\mathbf{x}^{\rm H}\mathbf{A}\,\mathbf{x}\,\mathbf{x}^{\rm H}\mathbf{B}\mathbf{x}] = {\rm tr}(\mathbf{A}\Sigma){\rm tr}(\mathbf{B}\Sigma) + {\rm tr}(\mathbf{A}\Sigma\mathbf{B}\Sigma)$. Accounting for (42), (44), and (45), the lower bound in (25) comes from the two facts: (i) if $x_{\rm SU} \equiv x_{\rm SU,G} \sim \mathcal{CN}(0, \sigma^2_{\rm SU,dirty})$, then $\mathsf{h}(x_{\rm SU}) = \ln(\pi e) + \ln(\sigma^2_{\rm SU,dirty})$ [51] and $\mathbb{E}_{x_{\rm SU}}[|\partial/\partial x^*_{\rm SU} \ln p(x_{\rm SU})|^2] = 1/\sigma^2_{\rm SU,dirty}$; (ii) $\mathsf{C}_{\rm SU,dirty} \geq \mathsf{I}(x_{\rm SU,G}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU})/M$ since the Gaussian distribution may not maximize $\mathsf{I}(x_{\rm SU}; \mathbf{y}_{\rm SU}|\mathbf{g}_{\rm SU})$.

References

[1] M. McHenry, E. Livsics, T. Nguyen, and N. Majumdar, "XG dynamic spectrum access field test results [topics in radio communications]," *IEEE Commun. Mag.*, vol. 45, no. 6, pp. 51–57, Jun. 2007.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[3] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. New York, NY, USA: Cambridge Univ. Press, 2013.

[4] Y. Han, A. Pandharipande, and S. H. Ting, "Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4945–4950, Oct. 2009.

[5] R. Manna, R. H. Y. Louie, Y. Li, and B. Vucetic, "Cooperative amplify-and-forward relaying in cognitive radio networks," in *Proc. IEEE CROWNCOM*, Cannes, France, Jun. 2010, pp. 1–5.

[6] D. Hu and S. Mao, "Cooperative relay in cognitive radio networks: Decode-and-forward or amplify-and-forward," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.

[7] Q. Wang, D.-W. Yue, and F. C. M. Lau, "Outage performance and cooperative diversity under amplify and forward relaying in cognitive radio networks," *Wireless Pers. Commun.*, vol. 69, no. 2, pp. 891–914, Apr. 2012.

[8] Y. Han, S. H. Ting, and A. Pandharipande, "Cooperative spectrum sharing protocol with selective relaying system," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 62–67, Jan. 2012.

[9] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I & II," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.

[10] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[11] F. Verde and A. Scaglione, "Randomized space-time block coding for distributed amplify-and-forward cooperative relays," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 3030–3033.

[12] F. Verde, "Design of randomized space-time block codes for amplify-and-forward cooperative relaying," in *Proc. 5th ISCCSP*, Rome, Italy, May 2012, pp. 1–5.

[13] F. Verde, "Performance analysis of randomized space-time block codes for amplify-and-forward cooperative relaying," *IET Commun.*, vol. 67, no. 17, pp. 1883–1898, Nov. 2013.

[14] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Optimization of coherent amplify-and-forward cooperative transmissions in spatially-multiplexed MIMO-OFDM systems," in *Proc. EUCNC*, Bologna, Italy, Jun. 2014, pp. 1–5.

[15] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Error probability analysis of a MIMO system with noncoherent relaying and MMSE reception," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1198–1201, Jul. 2014.

[16] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Performance analysis of amplify-and-forward multiple-relay MIMO systems with ZF reception," *IEEE Trans. Veh. Technol.*, to be published. DOI 10.1109/TVT.2014.2349740.

[17] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Cognitive multiple access via cooperation: Protocol design and performance analysis," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3677–3696, Oct. 2007.

[18] T. Fujii and Y. Suzuki, "Ad-hoc cognitive radio-development to frequency sharing system by using multi-hop network," in *Proc. IEEE Int. Symp. DySPAN*, Baltimore, MD, USA, Nov. 2005, pp. 589–592.

[19] J. Jia, J. Zhang, and Q. Zhang, "Cooperative relay for cognitive radio networks," in *Proc. IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2304–2312.

[20] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. IEEE SIGCOMM*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[21] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[22] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[23] I. Marić, A. Goldsmith, and M. Médard, "Analog network coding in the high-SNR regime," in *Proc. IEEE WiNC*, Boston, MA, USA, Jun. 2010, pp. 1–6.

[24] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, Jun. 2002.

[25] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 563–593, Mar. 2003.

[26] W. Yu, A. Sutivong, D. Julian, T. M. Cover, and M. Chiang, "Writing on colored paper," in *Proc. IEEE ISIT*, Washington, DC, USA, Jun. 2001, p. 302.

[27] A. Khisti, U. Erez, and G. Wornell, "Writing on many pieces of dirty paper at once: The binary case," in *Proc. IEEE ISIT*, Chicago, IL, USA, Jun. 2004, p. 533.

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

[29] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[30] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[31] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[32] F. Verde, A. Scaglione, D. Darsena, and G. Gelli, "An amplify-and-forward scheme for cognitive radios," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 2743–2747.

[33] E. Biglieri, J. G. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2916–2692, Oct. 1998.

[34] L. Ozarow, S. Shamai, and A. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[35] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[36] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2001.

[37] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[38] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3413–3431, Oct. 2007.

[39] R. Ramaswamy, N. Weng, and T. Wolf, "Characterizing network processing delay," in *Proc. IEEE GLOBECOM*, Dallas, TX, USA, Dec. 2004, pp. 1629–1634.

[40] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[41] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receiver in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2999, Sep. 2001.

[42] D. S. Papailiopoulos, G. A. Elkheir, and G. N. Karystinos, "Maximum-likelihood noncoherent PAM detection," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1152–1159, Mar. 2013.

[43] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*. New York, NY, USA: Wiley, 2000.

[44] A. Papoulis, *Probability, Random Variables, and Stochastics Processes*, 3rd ed. Singapore: McGraw-Hill, 1991.

[45] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[46] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.
[47] R. M. Gray, "Toeplitz and circulant matrices: A review," *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.
[48] D. Bertsekas and R. G. Gallager, *Data Networks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1992.
[49] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Inf. Control*, vol. 37, no. 1, pp. 34–39, 1978.
[50] A. Vosoughi and A. Scaglione, "On the effect of receiver estimation error upon channel mutual information," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 459–472, Feb. 2006.
[51] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[52] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part I*. New York, NY, USA: Wiley, 1968.
[53] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

Francesco Verde (M'10–SM'14) was born in Santa Maria Capua Vetere, Italy, on June 12, 1974. He received the Dr. Eng. degree (*summa cum laude*) in electronic engineering from the Second University of Napoli, Italy, in 1998, and the Ph.D. degree in information engineering from the University of Napoli Federico II, in 2002. Since December 2002, he has been with the University of Napoli Federico II. He first served as an Assistant Professor of signal theory and mobile communications and, since December 2011, he has served as an Associate Professor of telecommunications with the Department of Electrical Engineering and Information Technology. His research activities include cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems, multicarrier modulation systems, space-time processing for cooperative and cognitive communications systems, and software-defined networks. He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS since November 2014. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from September 2010 to October 2014 and a Guest Editor of the *EURASIP Journal on Advances in Signal Processing* in 2010.

Anna Scaglione (M'99–F'11) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rome, "La Sapienza", Rome, Italy, in 1995 and 1999, respectively. She is a Professor with the Department of Electrical and Computer Engineering, Arizona State University, Tempe, AZ, USA. She was previously with the University of California at Davis, Davis, CA, USA, during 2008–2014, Cornell University, Ithaca, NY, USA, between 2001 and 2007, and the University of New Mexico, Albuquerque, NM, USA, between 2000 and 2001. Her expertise is in the broad area of signal processing for communication systems and networks. Her current research focuses on studying and enabling decentralized learning and signal processing in networks of sensors social networks, and on demand side management and reliable energy delivery. She was Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS, and served in many capacities for communication and signal processing journals and conferences. She was on the Board of Governors of the Signal Processing Society from 2012 to 2014. She is a corecipient of the 2013 IEEE Donald G. Fink Prize Paper Award, and the 2013 SPS society Young Author Best Paper award (with her student) first author of the paper that received the 2000 IEEE Signal Processing Transactions Best Paper Award. She has also received the NSF Career Award in 2002 and is corecipient of the SmartGridcomm 2014 Best Student Paper Award and the Ellersick Best Paper Award (MILCOM 2005).

Donatella Darsena (M'06) received the Dr. Eng. degree (*summa cum laude*) in telecommunications engineering in 2001, and the Ph.D. degree in electronic and telecommunications engineering in 2005, both from the University of Napoli Federico II, Italy.

From 2001 to 2002, she was an Engineer in the Telecommunications, Peripherals and Automotive Group, STMicroelectronics, Milano, Italy. Since 2005, she has been an Assistant Professor with the Department of Engineering, University of Napoli Parthenope, Italy. She has also held teaching positions at the University of Napoli Federico II, Italy. Her research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, her current interests are focused on equalization, channel identification, narrowband-interference suppression for multicarrier systems, space-time processing for cooperative communications systems and cognitive communications systems, and software-defined networks.

Giacinto Gelli was born in Napoli, Italy, on July 29, 1964. He received the Dr. Eng. degree (*summa cum laude*) in electronic engineering in 1990, and the Ph.D. degree in computer science and electronic engineering in 1994, both from the University of Napoli Federico II.

From 1994 to 1998, he was an Assistant Professor with the Department of Information Engineering, Second University of Napoli. Since 1998, he has been with the Department of Electrical Engineering and Information Technology, University of Napoli Federico II, first as an Associate Professor, and since November 2006 as a Full Professor of Telecommunications. He also held teaching positions at the University Parthenope of Napoli, Italy.

His research interests are in the broad area of signal and array processing for communications, with current emphasis on multicarrier modulation systems and space-time techniques for cooperative and cognitive communications systems.