Maximum Average Service Rate and Optimal Queue Scheduling of Delay-Constrained Hybrid Cognitive Radio in Nakagami Fading Channels

Jie Hu, *Student Member, IEEE*, Lie-Liang Yang, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—As a promising technique to improve achievable bandwidth efficiency, cognitive radio (CR) has attracted substantial research attention from both the academic and industrial communities. To improve the performance attained by the secondary user (SU), a novel hybrid CR system is proposed, which combines the conventional interweave and underlay paradigms to enhance the chance of the SU to access the spectrum. Queuing theory is invoked in this paper to analyze the impact of the primary user's maximum tolerable delay on the performance of the SU. Multiple queues are assumed for the SU, which is engaged in video communication. Apart from the Poisson traffic generation, we also model the classic Nakagami-m fading channel as a Poisson service process by utilizing the outage probability in the presence of cochannel interference. We optimize both the hybrid interweave/underlay procedure to maximize the average service rate $\mu_{S,\max}$ of the SU, as well as the queue's scheduling scheme, for the sake of minimizing the overall average delay (OAD). As a result, the OAD of the SU is reduced by up to 27% and 20%, compared with the proportion and round-robin schemes, respectively.**

***Index Terms*—Hybrid cognitive radio (CR) system, maximum average service rate (MASR) of secondary user (SU), multilayer video streams (VSs), Nakagami-m fading channel, optimal queue scheduling scheme.**

I. Introduction

TWO kinds of customers are supported in cognitive radio (CR), namely, the primary user (PU) and the secondary user (SU). Furthermore, three main paradigms are considered for CR systems, namely, overlay, underlay, and interweave paradigms [1]. Here, we focus our attention on the interweave and underlay paradigms. According to the *interweave paradigm*, the PUs are authorized to access the channel, whereas the SUs are only able to access it when the PUs release it. In the scenario that the PUs request resources, an SU's session must be temporarily paused until the PUs complete their transmission. In the *interweave paradigm*, SUs sense the activities of PUs and clinch every possible opportunity to carry on with their own transmissions. By contrast, according to the *underlay paradigm*, PUs and SUs are allowed to transmit their data at the same time. An interference threshold is set up at the PUs to guarantee that the PU's link is not unduly interfered by SUs. Their transmission would only be terminated if the interference imposed by SUs exceeds the threshold.

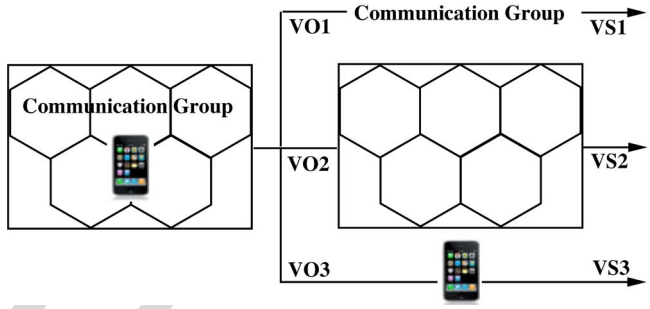

Fig. 1. Example of object-based MPEG-4 video transmission.

Although the transmission of PUs may be fully guaranteed with the aid of the *interweave paradigm*, only limited opportunities may be offered to SUs to convey their own data. Albeit SUs have improved chances of transmitting their own information to the destination using the *underlay paradigm*, the quality of service (QoS) of the PUs might not be as high as for the *interweave paradigm*. Can we strike an attractive compromise between these two paradigms so that the SUs have more opportunities to access the channels while the PUs may benefit from improved QoS? In this paper, a hybrid CR system is invoked to achieve this goal. In our novel system, SUs first detect the activity of PUs. If the channels are idle, the SUs may access them to transmit their own data. By contrast, if the channels are occupied by PUs, the SUs have a certain probability of accessing the channels while ensuring that the interference thresholds are not violated.

The Nakagami-m channel model has the advantage of characterizing a diverse range of multipath channels by a single fading parameter m [2]. Recently, supporting real-time video services subject to specific QoS constraints has become one of the essential requirements in wireless communications networks, where the video source is usually encoded into a number of streams in the application layer, representing the base and enhancement layers. An example of object-based MPEG-4 video transmission [3] is shown in Fig. 1, where different video objects are encoded into different video streams (VSs). The scheduling of the multilayer video queues at the SU's end

Manuscript received June 21, 2012; revised November 15, 2012 and December 20, 2012; accepted December 27, 2012. This work was supported in part by the RC-UK's India–UK Advanced Technology Centre, by the European Union's Concerto Project, by the China–UK Science Bridge, and by the China Scholarship Council. The review of this paper was coordinated by Prof. C. Lin.

The authors are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: jh10g11@ecs.soton.ac.uk; lly@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

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/TVT.2012.2237422

deserves further attention for the sake of achieving the best possible performance.

In [4], an adaptive resource-allocation scheme for multilayer VSs transmitted in wireless unicast/multicast scenarios is proposed, which inspires our work. The maximum throughput of an SU has been investigated in [5] under the interweave paradigm by finding the optimal transmit power of the SU in the scenario of a single queue conceived for the SU. In [6], this research is extended to the multiple PU and single SU CR system to explore the optimal transmit power and the relaying probability at the SU. The stability region has been already analyzed in the CR network supporting multiple PUs, multiple relays, and multiple SUs [7]. In [5]–[7], the interweave paradigm is employed, where the SUs have fewer opportunities to access the resources. All these studies assumed, however, that a single queue was set up for each of the users, which cannot faithfully represent the characteristics of multimedia communication. Furthermore, all results have been generated assuming queuing stability, where the specific QoS constraints such as the delay tolerance of lip-synchronized video communications were completely ignored.

Against this background, our novel contributions are as follows.

1) We amalgamate the interweave and underlay paradigms into a novel hybrid CR scheme characterized by the parameter $0 \leq \varepsilon \leq 1$, which is zero when the pure interweave paradigm is used and one for the pure underlay paradigm.
2) The Nakagami-m fading channel is used while modeling the service as a Poisson process by deriving the closed-form tail probability in the presence of cochannel interference. This allows us to use the classic $M/M/1$ queuing theory to investigate the delay performance.
3) The effects of both the PU's QoS requirement and the erroneous spectrum-sensing decisions on the performance of the SU are investigated.
4) The optimal parameter ε maximizing the average service rate of the SU is determined.
5) By exploiting the Lagrange optimization method, we derive the optimal closed-form solution for queue scheduling. Given this scheme, we are capable of minimizing the overall average delay (OAD) of multilayer VSs.

This paper is organized as follows. Our system model, including both the medium access control (MAC) and physical (PHY) layers, are described in Section II, whereas in Section III, the related queuing analysis is carried out for both the PU and the SU. In Section IV, the problem of finding the optimal solution is formulated and solved, followed by our numerical results in Section V. Finally, our conclusions are offered in Section VI.

II. System Model

Fig. 2 shows a CR system having a pair of PU source and PU destination (PD), as well as a pair of SU source and SU destination (SD). A novel hybrid CR policy is employed. Given a spectrum band, the SU first senses the activity of the PU. If this sensing process is reliable, the SU may access the spectrum

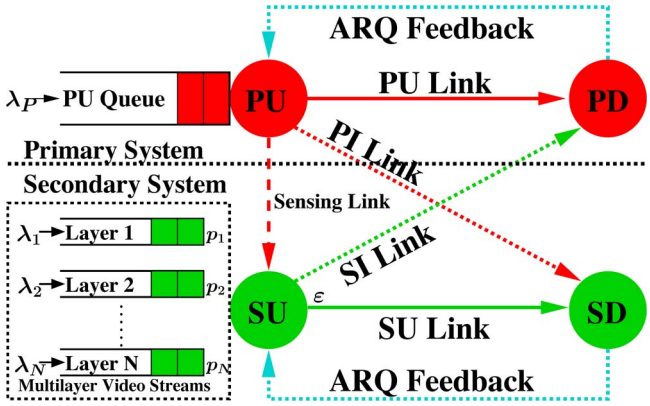

Fig. 2. System model with one PU and one SU sharing the same spectrum.

with a unity probability when the PU is idle. Furthermore, the SU may access the spectrum with a probability of $0 \leq \varepsilon \leq 1$, even if the PU is busy, provided that the PU's QoS is still guaranteed. If, however, this sensing process failed, the situation is reversed, and hence, the SU accesses the spectrum with a probability of ε when the PU is idle and with a unity probability when the PU is busy. Hence, erroneous spectrum sensing may lead to catastrophic performance degradation for both the PU and the SU. We will also consider the practical scenario of an erroneous spectrum-sensing decision and its effect on the performance of the SU. Let us now consider the details of the MAC and PHY layers.

A. MAC-Layer Model

Observe in Fig. 2 that a buffer is provided for the PU to store the packets, which cannot be immediately transmitted. Again, to support video communications for the SU, multiple buffers are provided to accommodate the different-priority queues generated by multilayer video encoding.

The basic unit of time in our system is a time slot (TS). The packets' arrival at the PU's buffer obeys the classic Poisson process with a mean of λ_P packets/TS. The QoS constraint of the PU is the total delay imposed by both the transmission and buffering delay, which should not exceed T_P, namely, the maximum tolerable delay of the PU. The packets' arrival at the buffer input of the SU also follows the Poisson process with a mean of $\{\lambda_n,\ n = 1, 2, \ldots, N\}$ packets/TS. Regardless of whether the PU is idle or busy, if the SU has already successfully accessed the spectrum, we should determine the probability of each queue at the SU completing its transmission of the stored packets. These probabilities may be denoted by $\{p_n,\ n = 1, 2, \ldots, N\}$. No specific QoS constraints are imposed at the SU. However, our prime goal is to find the optimal scheduling scheme and the associated hybrid parameter ε, which minimizes the OAD of all the queues at the SU.

We assume that every packet transmitted by the PU or the SU may be only successfully received by the corresponding destination in a TS when the signal-to-interference-plus-noise ratio (SINR) is above a specific threshold. In our system, automatic repeat request (ARQ) with an unlimited number of retransmissions is adopted to ensure that no packets are lost.

Furthermore, we assume that the acknowledgements are always successfully received by the transmitters.

Theorem 1: Given the successful reception probability μ of a single packet during a TS in the ARQ-aided system, the continuous transmit time obeys the exponential distribution with the mean of $1/\mu$.

Proof: The proof is provided in [8, App. A]. ■

As a benefit of the ARQ mechanism, a packet's departure from the buffer is synonymous to being served. The wireless channel may be considered as a "server" in the queuing analysis. As a result, the packets' departure process is also referred to as the service process.

B. PHY-Layer Model

Observe from Fig. 2 that there are five links in this hybrid CR system, namely, the PU link, the SU link, the sensing link, the PI link imposing interference from the PU on the SD, and the SI link inflicting interference from the SU upon the PD. In line with [11], a multipacket reception (MPR) model is introduced for the queuing analysis of the collision channel. We define the corresponding conditional probabilities for the MPR model of these five links.

- $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$ (or $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$): Only the PU's (or SU's) packet is successfully received at the PD (or SD) when only the PU (or SU) transmits.
- $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$ (or $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$): Only the PU's (or SU's) packet is successfully received at the PD (or SD) when both the PU and the SU transmit.
- $q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$: Only the PU's packet is successfully received at the SU when the PU transmits.

Next, we will briefly highlight the derivation of these conditional probabilities. The radio propagation between any pair of nodes is assumed to be affected by the independent stationary Nakagami-m flat-fading channels $h_i(t)$ having $E[|h_i(t)|^2]=1$ (t is the TS index). Hence, $Z=|h_i(t)|^2$ is a Gamma-distributed random variable (r.v.) having a mean of 1, which may be denoted by $Z\sim\mathrm{Gamma}(m,1/m)$ with the shape and scale parameters of m and $1/m$, respectively. Given the probability density function (pdf) $f_Z(z)$ of [9, eq. (3.39)], the tail distribution function (TDF) of Z is formulated as $P(Z>z)=\Gamma(m,mz)/\Gamma(m)$, where $\Gamma(\cdot,\cdot)$ represents the upper incomplete Gamma function, and m denotes the Nakagami-m fading parameter [15].

After being divided by the noise power, the normalized transmit power of the PU is defined as P_P, whereas the normalized transmit power of the SU is defined as P_S. The propagation path loss (PL) is denoted by Ω_i, and the SINR threshold that has to be exceeded for successful packet reception is denoted by β_i, while again, the Nakagami-m fading parameter is denoted by m_i, where i represents "P" for the PU link, "S" for the SU link, "PI" for the PI link, "SI" for the SI link, and "PS" for the sensing link spanning from the PU to the SU.

According to the TDF of the r.v. Z, we express the conditional probabilities of $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$, $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$, and $q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$ in the MPR model in the absence of interference thresholds as follows:

$$\begin{aligned}
q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} &= P\left[\frac{|h_P(t)|^2 P_P}{\Omega_P}>\beta_P\right]\\
&=\Gamma\left(m_P,\frac{m_P\beta_P\Omega_P}{P_P}\right)\Big/\Gamma(m_P)\\
q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}} &= P\left[\frac{|h_S(t)|^2 P_S}{\Omega_S}>\beta_S\right]\\
&=\Gamma\left(m_S,\frac{m_S\beta_S\Omega_S}{P_S}\right)\Big/\Gamma(m_S)\\
q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} &= P\left[\frac{|h_{\mathrm{PS}}(t)|^2 P_P}{\Omega_{\mathrm{PS}}}>\beta_{\mathrm{PS}}\right]\\
&=\frac{\Gamma\left(m_{\mathrm{PS}},m_{\mathrm{PS}}\beta_{\mathrm{PS}}\Omega_{\mathrm{PS}}/P_P\right)}{\Gamma(m_{\mathrm{PS}})}.
\end{aligned}$$

We introduce another two Gamma-distributed r.v.'s, namely, $X\sim\mathrm{Gamma}(m_X,1/m_X)$ with a pdf of $f_X(x)$ and $Y\sim\mathrm{Gamma}(m_Y,1/m_Y)$ with a pdf of $f_Y(y)$. Then, the problem of solving $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$ and $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$ becomes equivalent to solving the probability problem of $P[X>A+BY]$. Unfortunately, a closed-form solution cannot be derived for the general situation. Nonetheless, we may obtain the following results for two special cases:

$$P[X>A+BY]=\begin{cases}\Phi(A,B), & \text{if } m_X \text{ is a positive integer}\\ & \quad m_Y \text{ can be any real number}\\ \Phi'(A,B), & \text{if } m_Y \text{ is a positive integer}\\ & \quad m_X \text{ can be any real number.}\end{cases}$$

$$\begin{cases}\Phi(A,B)=\left(\frac{m_Y}{m_XB+m_Y}\right)^{m_Y}e^{-m_XA}\\ \qquad\cdot\sum_{n=0}^{m_X-1}\sum_{k=0}^{n}\frac{m_X^nB^kA^{n-k}}{k!(n-k)!(m_XB+m_Y)^k}\frac{\Gamma(m_Y+k)}{\Gamma(m_Y)}\\ \Phi'(A,B)=\frac{\Gamma(m_X,m_XA)}{\Gamma(m_X)}-\sum_{n=0}^{m_Y-1}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{m_Y}{B}\right)^n\\ \qquad\cdot\frac{e^{m_Y\frac{A}{B}}(-A)^{n-k}m_X^{m_X}}{n!\Gamma(m_X)}\cdot\frac{\Gamma[k+m_X,A(m_X+m_Y/B)]}{(m_X+m_Y/B)^{k+m_X}}.\end{cases}\tag{1}$$

For a detailed derivation of $\Phi(A,B)$ and $\Phi'(A,B)$, see [8, App. B]. Given this result, we may write down the other two conditional probabilities when subjected to the cochannel interference thresholds at the PD and the SD as follows:

$$\begin{aligned}
&q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\\
&=P\left[\frac{|h_P(t)|^2P_P/\Omega_P}{1+|h_{\mathrm{SI}}(t)|^2P_S/\Omega_{\mathrm{SI}}}>\beta_P\right]\\
&=P\left[|h_P(t)|^2>\frac{\beta_P\Omega_P}{P_P}+\frac{\beta_PP_S\Omega_P}{P_P\Omega_{\mathrm{SI}}}|h_{\mathrm{SI}}(t)|^2\right]\\
&=\begin{cases}\Phi\left(\frac{\beta_P\Omega_P}{P_P},\frac{\beta_PP_S\Omega_P}{P_P\Omega_{\mathrm{SI}}}\right), & \text{if } m_P \text{ is a positive integer}\\ \Phi'\left(\frac{\beta_P\Omega_P}{P_P},\frac{\beta_PP_S\Omega_P}{P_P\Omega_{SI}}\right), & \text{if } m_{SI} \text{ is a positive integer}\end{cases}
\end{aligned}\tag{2}$$

$$
\begin{aligned}
&q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} \\
&= P\left[\frac{|h_S(t)|^2 P_S/\Omega_S}{1+|h_{\mathrm{PI}}(t)|^2 P_P/\Omega_{\mathrm{PI}}} > \beta_S\right] \\
&= P\left[|h_S(t)|^2 > \frac{\beta_S\Omega_S}{P_S} + \frac{\beta_S P_P \Omega_S}{P_S\Omega_{\mathrm{PI}}}\,|h_{\mathrm{PI}}(t)|^2\right] \\
&= \begin{cases} \Phi\left(\frac{\beta_S\Omega_S}{P_S}, \frac{\beta_S P_P\Omega_S}{P_S\Omega_{\mathrm{PI}}}\right), & \text{if } m_S \text{ is a positive integer} \\ \Phi'\left(\frac{\beta_S\Omega_S}{P_S}, \frac{\beta_S P_P\Omega_S}{P_S\Omega_{\mathrm{PI}}}\right), & \text{if } m_{\mathrm{PI}} \text{ is a positive integer.} \end{cases}
\end{aligned} \tag{3}
$$

III. Queuing Analysis

The following tasks will be accomplished here: 1) The classic M/M/1 queuing is revisited. 2) The average service rate of the PU and the SU is attained for the original benchmark system. 3) A new system is defined for characterizing the interdependence of the PU's and SU's queues, and the corresponding average service rate of the PU and the SU is rederived.

A. $M/M/1$ Queuing System

In the M/M/1 queuing system, only a single server is invoked by the system. The packets' interarrival time follows the exponential distribution having a parameter λ, whereas the packets' interdeparture time obeys the same distribution having a parameter μ. An infinite buffer is assumed for guaranteeing that every single packet may be successfully transmitted, i.e., without any packet loss. According to [10], we can express the probability of the queue being empty and the total average delay, including the transmission and buffering time, as follows:

$$
P[Q=0] = 1-\lambda/\mu, \quad T = 1/(\mu-\lambda), \quad \text{where} \quad \lambda < \mu \tag{4}
$$

and we denote the random queue length and the total average delay by Q and T, respectively.

B. PU Link

According to the MAC-layer protocol of our hybrid CR system, the PU has the highest priority to access the channel. If the PU's queue is not empty, the PU would occupy the channel for its transmission, and the packet would be removed from the head of the queue when it is successfully received by the PD. However, the PU would suffer from the interference imposed by the activity of the SU. The service process of a PU's single packet follows the Bernoulli distribution. In analogy to the successful packet departure probability in the Bernoulli trial, the average service rate of the Poisson service process may be formulated for the PU link as

$$
\begin{aligned}
\mu_P = {} & P\{Q_S=0\}\cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} + P\{Q_S\neq 0\} \\
& \cdot\left(q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot\left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot(1-\varepsilon) + q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\cdot\varepsilon\right)\right. \\
& \left. + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\right)
\end{aligned} \tag{5}
$$

where Q_S represents the random queue length of the SU. Given the average service rate μ_P, the average arrival rate λ_P, and the QoS constraint stipulating that the total delay should not exceed the maximum delay tolerance T_P, we infer from (4) that the condition $0 \le \lambda_P \le \mu_P - 1/T_P$ must be obeyed.

C. SU Link

According to the MAC-layer protocol of the SU, before accessing the channel, the SU carries out spectrum sensing first. When the channel is released by the PU, after reliable sensing, the SU would transmit its own packets with a probability of unity (otherwise, with a probability of ε), which implies inefficient resource usage. When the channel is still occupied by the PU, following reliable sensing, the SU would transmit its own packets with a probability of ε (otherwise, with a probability of unity), which imposes strong interference on the PD. It is vital to guarantee that the transmission of the SU does not force the PU to violate its delay tolerance. We derive the probability of a packet being served, which is also identical to the average service rate μ_S of the SU's Poisson service process expressed as

$$
\begin{aligned}
\mu_S = {} & q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}\cdot P[Q_P=0] \\
& \cdot\left[q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\cdot\varepsilon\right] \\
& + q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}\cdot P[Q_P\neq 0] \\
& \cdot\left[q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot\varepsilon + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\right]
\end{aligned} \tag{6}
$$

where Q_P represents the random queue length of the PU.

D. Stochastic Dominance Principle

It may be shown from (5) and (6) that the average service rates of the SU and the PU depend on each other's queue sizes. Since these queues interact with each other, the average rate of the individual service processes cannot be directly determined. For the sake of circumventing this problem, the stochastic dominance principle of [12] is invoked to assist our analysis.

The concept of a so-called "dominant system" was defined in [12] by allowing a set of terminals having no packets in their transmit-buffer to continue transmitting hypothetical dummy packets. In this manner, the queues in the dominant system stochastically dominate the queues in the original system. This dominant system is defined here.

1) If we have $Q_S = 0$ and $Q_P = 0$, the SU transmits dummy packets with a unity probability.
2) If $Q_S = 0$ and $Q_P \neq 0$, the SU transmits dummy packets with a probability ε.

Note that the hypothetical dummy packets are introduced only for the sake of facilitating our performance analysis and for finding closed-form solutions to the optimum scheduling factor ε, but in reality, no dummy packets are assumed to be transmitted in the practical systems considered. As shown in [13], given the same initial conditions, the queuing performance of the dominant system transmitting hypothetical

dummy packets is capable of providing a tight approximation of the queuing performance of the original system. Since in practical systems no dummy packets are actually transmitted, no extra interference will be imposed on the system, and no extra power is required for their transmissions.

In the context of the extended dominant system, $\{Q_S \neq 0\}$ has a unity probability. Hence, (5) can be rewritten as

$$\mu_P = q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \cdot \left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} \cdot (1-\varepsilon) + q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} \cdot \varepsilon \right) + \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} \quad (7)$$

and the condition $0 \leq \lambda_P \leq \mu_P - 1/T_P$ is also rewritten as

$$0 \leq \lambda_P \leq q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} + \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} - 1/T_P - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} - q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SUh}\}}\right) \varepsilon. \quad (8)$$

According to (4), we may represent the probabilities of the PU's queue being not empty and empty by λ_P and μ_P, respectively, as shown in (7). Hence, we may rewrite (6) as

$$\mu_S = q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \varepsilon - \frac{\lambda_P}{\mu_P} \left[q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} \times \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \varepsilon - q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \varepsilon - q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \right]. \quad (9)$$

To simplify the derivations in the next section, some simple notations are introduced as follows: $C = q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$, $D = q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$, $E = q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$, $F = q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$, and $S = q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$.

IV. Problem Formulation and Solutions

Here, the following two problems are solved: 1) finding the optimal ε^* for maximizing the average service rate $\mu_{S,\max}$ of the SU and 2) finding the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$ to minimize the OAD associated with $\mu_{S,\max}$.

A. MASR of the SU

The optimal parameter ε^* will be found for the sake of maximizing the average service rate of the SU without violating the delay tolerance of the PU. The solution can be found by constructing the problem **P1** as

$$\arg\max_{\varepsilon}\ \mu_S(\varepsilon)$$
$$0 \leq \varepsilon \leq \xi = \frac{SC + (1-S)D - \lambda_P - 1/T_P}{S(C-D)} \leq 1 \quad (10)$$

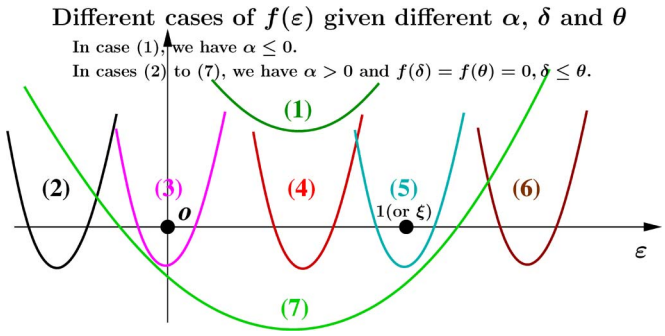

Fig. 3. Different cases of $f(\varepsilon)$ given different α, δ, and θ.

where the function $\mu_S(\varepsilon)$ is defined by (9), whereas the inequality condition in (10) is imposed for the sake of satisfying the specific delay tolerance of the PU, which is derived from (8).

To solve **P1**, we differentiate $\mu_S(\varepsilon)$ with respect to ε, yielding

$$\frac{d\mu_S}{d\varepsilon} = \frac{f(\varepsilon)}{[SC + (1-S)D - S(C-D)\varepsilon]^2} \quad (11)$$

where we define $f(\varepsilon) = E(1-S)[SC + (1-S)D - S(C-D)\varepsilon]^2 - \lambda_P \alpha$ and $\alpha = S(C-D)(E-F) + DE(1-S) - FSD$. Since the denominator of (11) is always positive, we can simply consider $f(\varepsilon)$ instead, which is a quadratic function having the shape shown in Fig. 3.

Let us consider the constraint shown in (8) more closely. Given $0 \leq \varepsilon \leq 1$, the right side of (8) reaches its minimum value of $(D - 1/T_P)$ when we have $\varepsilon = 1$. Therefore, if we have $\lambda_P \leq D - 1/T_P$, ε may assume any arbitrary value ranging from 0 to 1. By contrast, if $\lambda_P > D - 1/T_P$, then ε must be selected from the range determined by (10) in the context of the specific λ_P value.

Different scenarios have to be considered for finding the optimal values of ε^*, which lead to the maximum average service rate (MASR) $\mu_S(\varepsilon)$ of the SU formulated in (10).

Case 1 ($\lambda_P \leq D - 1/T_P$): In this case, as aforementioned, ε might assume any arbitrary value ranging from 0 to 1. Under this assumption, the problem may be further divided into the following two subcases, which will assist us in finding $\mu_{S,\max}$.

Subcase 1-1 ($\alpha \leq 0$): When $\alpha \leq 0$, the $f(\varepsilon)$ curve obeys the relative position shown by case (1) at the top of Fig. 3, in which case there are no real-valued solutions to $f(\varepsilon) = 0$. Hence, provided that $E(1-S) > 0$ holds, $f(\varepsilon)$ is always higher than 0, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function. Therefore, we have $\varepsilon^* = 1$ and $\mu_{S,\max} = \mu_S(1)$.

Subcase 1-2 ($\alpha > 0$): Second, when $\alpha > 0$, there are two real-valued solutions to $f(\varepsilon) = 0$, which are denoted by δ and θ, respectively, i.e.,

$$\delta = \frac{1}{S(C-D)} \left[SC + (1-D) - \sqrt{\frac{\lambda_P \left[S(C-D)(E-F) + DE(1-S) - FSD\right]}{E(1-S)}} \right] \quad (12)$$

$$\theta = \frac{1}{S(C-D)} \left[SC + (1-D) + \sqrt{\frac{\lambda_P \left[S(C-D)(E-F) + DE(1-S) - FSD \right]}{E(1-S)}} \right] \quad (13)$$

where we have $\delta \leq \theta$. In this case, depending on the values of δ and θ, the relative position of the $f(\varepsilon)$ curve corresponds to one of the scenarios (2)–(7) shown in Fig. 3, which are specifically considered in the following.

a) If $\delta \leq \theta < 0$, the $f(\varepsilon)$ curve obeys the positioning of case (2) in Fig. 3, where $f(\varepsilon) > 0$ holds within the range of $0 \leq \varepsilon \leq 1$, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Hence, $\varepsilon^* = 1$ maximizes the average service rate $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(1)$.

b) If $\delta < 0$ and $0 \leq \theta < 1$, the $f(\varepsilon)$ curve obeys the positioning of case (3) in Fig. 3. Hence, $f(\varepsilon) \leq 0$ holds within the range of $0 \leq \varepsilon \leq \theta$, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. However, within the range of $\theta < \varepsilon \leq 1$, $f(\varepsilon) > 0$ holds, and hence, $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. In this situation, the optimal value of ε^* is either 0 or 1. Hence, we have $\varepsilon^* = \arg\max_{\varepsilon \in \{0,1\}} \{\mu_S(\varepsilon)\}$, and the MASR becomes $\mu_{S,\max} = \max\{\mu_S(0), \mu_S(1)\}$.

c) If $0 \leq \delta < \theta < 1$, the $f(\varepsilon)$ curve corresponds to case (4) in Fig. 3. Therefore, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq \delta$, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Within the range of $\delta < \varepsilon \leq \theta$, however, $f(\varepsilon) \leq 0$ holds, which suggests that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Finally, in the range of $\theta < \varepsilon \leq 1$, $f(\varepsilon) > 0$ is satisfied, which indicates that $\mu_S(\varepsilon)$ becomes a monotonically increasing function of ε again. According to the preceding analysis, the pair of potential values that may maximize $\mu_S(\varepsilon)$ are $\varepsilon = \delta$ and $\varepsilon = 1$. Hence, the optimal value of ε^* is $\varepsilon^* = \arg\max_{\varepsilon \in \{\delta,1\}} \{\mu_S(\varepsilon)\}$, and the MASR is, hence, $\mu_{S,\max} = \max\{\mu_S(\delta), \mu_S(1)\}$.

d) If $0 \leq \delta < 1$ and $\theta \geq 1$, which results in an $f(\varepsilon)$ curve corresponding to case (5) in Fig. 3, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq \delta$, which implies that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Within the range of $\delta < \varepsilon \leq 1$, however, $f(\varepsilon) < 0$ holds, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Jointly considering these two situations, we conclude that $\varepsilon^* = \delta$ maximizes the average service rate $\mu(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(\delta)$.

e) If $\delta \geq 1$, we have encountered case (6) of Fig. 3. Correspondingly, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq 1$, which implies that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Hence, $\varepsilon^* = 1$ maximizes the average service rate of $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(1)$.

f) Finally, if $\delta < 0 < 1 \leq \theta$, $f(\varepsilon)$ is reminiscent of case (7) of Fig. 3. Consequently, $f(\varepsilon) \leq 0$ holds within the range of $0 \leq \varepsilon \leq 1$, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Hence, $\varepsilon^* = 0$ maximizes the average service rate $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(0)$.

Case 2 ($\lambda_P > D - 1/T_P$): If $\lambda_P > D - 1/T_P$, then ε may assume any arbitrary value ranging from 0 to ξ, where the value of ξ is restricted by the value of λ_P according to (10). Following a similar train of thought, by simply substituting "1" in Case 1 by "ξ," we have the following conclusions, which correspond to the cases (1)–(7) in Fig. 3, as we discussed earlier.

1) If $\alpha \leq 0$ (see case (1) in Fig. 3), or if $\alpha > 0$ and $\theta < 0$ (see case (2) in Fig. 3), or if $\alpha > 0$ and $\xi \leq \delta$ (see case (6) in Fig. 3), then we have $\varepsilon^* = \xi$, and the MASR becomes $\mu_{S,\max} = \mu_S(\xi)$.
2) If $\alpha > 0$, $\delta < 0$, and $0 \leq \theta < \xi$ (see case (3) in Fig. 3), then we have $\varepsilon^* = \arg\max_{\varepsilon \in \{0,\xi\}} \{\mu_S(\varepsilon)\}$, and the MASR becomes $\mu_{S,\max} = \max\{\mu_S(0), \mu_S(\xi)\}$.
3) If $\alpha > 0$ and $0 \leq \delta < \theta < \xi$ (see case (4) in Fig. 3), then we have $\varepsilon^* = \arg\max_{\varepsilon \in \{\delta,\xi\}} \{\mu_S(\varepsilon)\}$, and the MASR is $\mu_{S,\max} = \max\{\mu_S(\delta), \mu_S(\xi)\}$.
4) If $\alpha > 0$ and $0 \leq \delta < \xi \leq \theta$ (see case (5) in Fig. 3), then $\varepsilon^* = \delta$, and the MASR is $\mu_{S,\max} = \mu_S(\delta)$.
5) If $\alpha > 0$, $\delta < 0$, and $0 < \xi \leq \theta$ (see case (7) in Fig. 3), then $\varepsilon^* = 0$, and the MASR is formulated as $\mu_{S,\max} = \mu_S(0)$.

B. Optimal Queue Scheduling of the SU

Let us now further detail the derivation of the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$[1] conceived for minimizing the OAD of the SU's queues. Given (4), the average delay of each queue may be formulated as

$$\text{Delay}_n = 1/(p_n \cdot \mu_S(\varepsilon) - \lambda_n), \qquad n = 1, 2, \ldots, N \quad (14)$$

where p_n is the probability of the SU's nth queue transmitting its stored packets, and λ_n is the packets' average arrival rate.

Given the associated MASR $\mu_{S,\max}$ of the SU, we construct the second problem **P2**, i.e.,

$$\arg \min_{\{p_1, p_2, \ldots, p_N\}} \left\{ \frac{1}{N} \sum_{n=1}^{N} \frac{1}{p_n \cdot \mu_{S,\max} - \lambda_n} \right\} \quad (15)$$

$$\text{subject to} \quad \sum_{n=1}^{N} p_n = 1, \quad \lambda_n - p_n \cdot \mu_{S,\max} < 0. \quad (16)$$

Then, we may readily show the following.

1) The objective function (15) is convex over $\{p_n,\ n = 1, 2, \ldots, N\}$.
2) The first equality constraint in (16) is affine over $\{p_n,\ n = 1, 2, \ldots, N\}$, whereas the second inequality constraint in (16) is convex over $\{p_n,\ n = 1, 2, \ldots, N\}$.

Hence, **P2** is a convex problem [14], and the optimal solution may be obtained with the aid of the Lagrangian optimization method while additionally exploiting the Karush–Kuhn–Tucker (KKT) conditions [14], which were proposed for solving convex optimization problems similar to ours under an inequality constraint.

[1]The optimal shceduling scheme is fully specified when the probability p_n of transmitting the stream n is determined.

TABLE I
PHY-LAYER PARAMETER SETTING

	PL	Threshold	Fading	Norm Power
PU Link	Ω_P = 5 dB	β_P = 13.9 or 8.5 dB	m_P = 1 or 2	P_P = 26 dB
SU Link	Ω_S = 5 dB	β_S = 13.9 or 8.5 dB	m_S = 1 or 2	P_S = 26 dB
PI/SI Link	$\Omega_{PI/SI}$ = 15 dB	No threshold	$m_{SI/PI}$ = 1	$P_{P/S}$ = 26 dB
Sensing Link	Ω_{PS} = 2 dB	β_{PS} = 13.9 or 8.5 dB	m_{PS} = 1 or 2	P_P = 26 dB

The Lagrangian function is constructed as

$$J = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{p_n\mu_{S,\max}-\lambda_n} + \sum_{n=1}^{N}\varphi_n(\lambda_n - p_n\mu_{S,\max}) + \eta\left(\sum_{n=1}^{N}p_n - 1\right) \quad (17)$$

where $\eta \geq 0$ and $\varphi_n \geq 0,\ n = 1, 2, \ldots, N$ are the Lagrangian multipliers associated with the two conditions of (16), respectively. Then, the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$ and the Lagrangian multipliers of the optimization problem **P2** satisfy the following KKT conditions [14]:

$$\begin{cases} \left.\frac{\partial J}{\partial p_n}\right|_{p_n=p_n^*} = \frac{-\mu_{S,\max}}{N(p_n^*\mu_{S,\max}-\lambda_n)^2} - \varphi_n^*\mu_{S,\max} + \eta^* = 0 \\ \left.\frac{dJ}{d\eta}\right|_{p_n=p_n^*} = \sum_{n=1}^{N} p_n^* - 1 = 0 \\ \varphi_n^*\left(\lambda_n - p_n^*\mu_{S,\max}\right) = 0, \quad \varphi_n^* \geq 0 \text{ and } \eta^* \geq 0 \end{cases} \quad (18)$$

where $n = 1, 2, \ldots, N$. Naturally, according to the second constraint of (16), the only solution satisfying the third line of (18) is $\varphi_n^* = 0$. Substituting $\varphi_n^* = 0$ into the first line of (18), we derive p_n^* as

$$p_n^* = \frac{1}{\sqrt{N\eta^*\mu_{S,\max}}} + \frac{\lambda_n}{\mu_{S,\max}}. \quad (19)$$

Then, substituting (19) into the second line of (18), we arrive at

$$\eta^* = N\cdot\mu_{S,\max}\Big/\left(\mu_{S,\max} - \sum_{n=1}^{N}\lambda_n\right)^2. \quad (20)$$

Finally, substituting η^* of (20) into (19), we arrive at the optimal queue scheduling scheme characterized by

$$p_n^* = \frac{\mu_{S,\max} - \sum_{n=1}^{N}\lambda_n + N\lambda_n}{N\cdot\mu_{S,\max}}. \quad (21)$$

V. PERFORMANCE ANALYSIS

In our numerical results, we link the PHY-layer parameters to a realistic orthogonal frequency-division multiplexing scheme communicating over Nakagami-m fading channels without channel coding. We define the SINR threshold of the receiver for the sake of guaranteeing that the bit error ratio of the system is not higher than 10^{-2}. Hence, the threshold is determined by the channel model having different fading parameters. For example, according to [16, Fig. 5-5], if $m = 1$, the threshold at the receiver is 13.86 dB, whereas the threshold is 8.50 dB for $m = 2$.

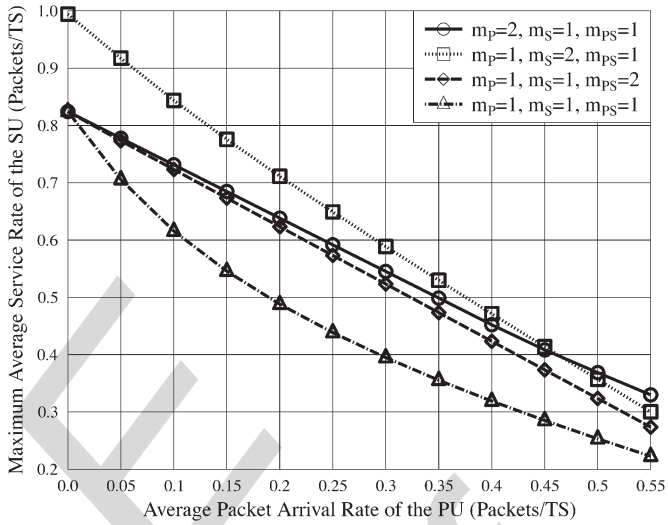

Fig. 4. Service rate of the SU versus the average packets arrival rate λ_P of the PU parameterized by the fading parameters.

We assume that the SU has five different queuing buffers in the MAC layer, where we set the average packet arrival rate of each queue, for example, to 0.01, 0.02, 0.03, 0.04, and 0.05 packet/TS, respectively.

Two other queue scheduling schemes are used as benchmarkers, namely, the round-robin scheme, where every queue of the SU has the same probability of transmitting packets stored in the buffer, and the proportional scheme, where the transmission probabilities are arranged to be proportional to the average arrival rates at the buffers.

A. Impact of the PU'S Average Packet Arrival Rate

In the PHY layer, the PL, the SINR threshold of the receiver associated with the Nakagami-m fading parameters, and the transmit power are set according to Table I. More particularly, there are four cases for the Nakagami-m parameters of the PU, the SU, and the sensing links, namely, 1) $m_P = 1$, $m_S = 1$, and $m_{PS} = 1$; 2) $m_P = 2$, $m_S = 1$, and $m_{PS} = 1$; 3) $m_P = 1$, $m_S = 2$, and $m_{PS} = 1$; and 4) $m_P = 1$, $m_S = 1$, and $m_{PS} = 2$.

In the MAC layer, the delay tolerance of the PU is $T_P = 5$ TS. We vary the average packet arrival rate λ_P from 0.1 to 0.5 packet/TS for characterizing its impacts on the MASR and the OAD in conjunction with three different queue scheduling schemes for the SU.

It may be readily shown from Fig. 4 that $\mu_{S,\max}$ linearly decays upon increasing λ_P. This is plausible since a higher λ_P results in an increased traffic load for the PU link, which, in turn, results in the PU's prolonged occupation of the

spectral band. This reduces the SU's chance of transmitting its own packets, which is reflected by the reduction of $\mu_{S,\max}$. Upon increasing the Nakagami-m parameters of the PU, SU, and sensing links, respectively, the MASR is improved. If we consider the higher m_P of the PU link, it is shown in Fig. 4 that the data stream of the PU may be transmitted more promptly and the duration of the PU's spectrum occupation would be shortened. Consequently, the SU's increased chances of transmitting its own data results in an increased MASR. Hence, the MASR is substantially improved. If we consider the higher m_S of the SU link characterized in Fig. 4, the MASR is also directly improved. For higher m_{PS} values, the MASR of Fig. 4 would be also improved due to the fact that the spectrum-sensing decisions become more reliable.

We may infer the following observations from Fig. 4 depending on the traffic load of the PU. We have the best performance gain upon improving the SU link if the PU's traffic load is lower than 0.45 packet/TS. By contrast, if the traffic load is higher than 0.45 packet/TS, improving the PU link may provide the best performance. Upon comparing the improvements of the sensing and PU links shown in Fig. 4, when the PU's traffic load is lower than 0.1 packet/TS, we infer that improving the sensing link may provide the same performance gain as improving the PU link. However, when the traffic load is higher than 0.1 packet/TS, improving the PU link is more influential. These observations may motivate us to design a more effective strategy for the sake of improving the performance of the SU, given different traffic loads of the PU. For example, given a light PU traffic load, improving the SU link is our best option to enhance overall performance. Furthermore, a reliable sensing strategy conceived for the SU may achieve the same overall performance as an improved PU link. By contrast, for a high traffic load, researchers should focus on directly improving the PU link.

B. Impact of the PU'S Delay Tolerance on the SU

Here, we consider the Nakagami-m parameters of $m_P = m_S = m_{PS} = 2$, and all the associated SINR thresholds are 8.5 dB, where the transmit power of the PU and the SU are $P_P = P_S = 18$ dB, whereas the others are the same as in Table I. We vary the delay tolerance T_P from 3 to 8 TS for the sake of investigating its impacts on the OAD in conjunction with three different queue scheduling schemes. Two scenarios, namely, $\lambda_P = 0.4$ and 0.5 packet/TS, are considered.

Observe furthermore in Fig. 5 that the SU benefits from increasing T_P and that a reduced λ_P decreases the OAD. The asymptotic values of the OAD for these three schemes may be determined when T_P tends to infinity. In Fig. 5, we focus our attention on the comparison of three different scheduling schemes. It can be seen that our proposed optimal scheme (OPT) performs better than the conventional schemes. The higher the traffic load of the PU, i.e., the higher λ_P and lower T_P, the more substantial the advantage of our optimal scheme becomes. For example, for $\lambda_P = 0.5$ packet/slot and $T_P = 3$ TS, our scheme has a 27% or 20% lower OAD than the proportional (PRO) or the round-robin (R-R) schemes, respectively, when considering the SU's queues.

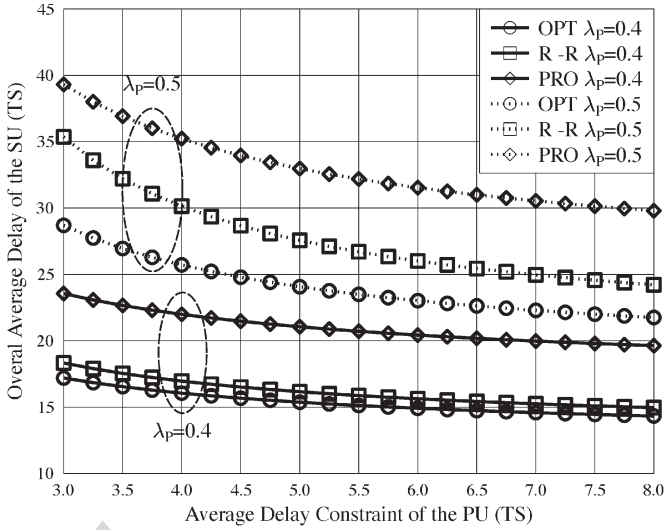

Fig. 5. Delay of the SU versus the delay tolerant T_P of the PU parameterized by the average packets arrival rate λ_P.

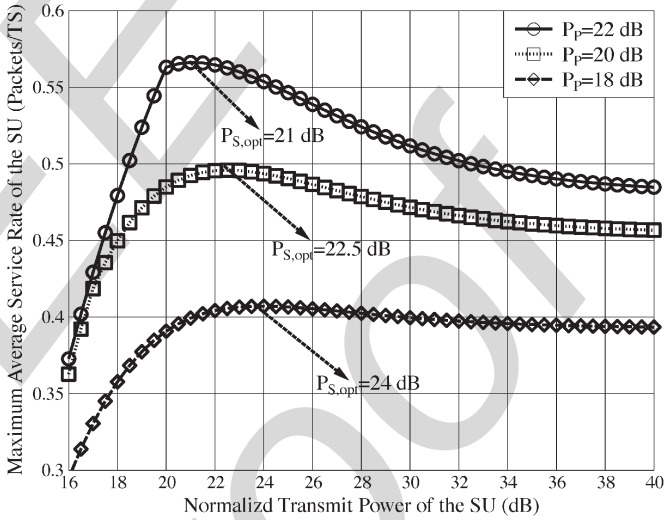

Fig. 6. Service rate versus the normalized transmit power of the SU parameterized by the normalized transmit power of the PU.

C. Impact of the SU's Normalized Transmit Power

In the PHY layer, the fading parameters are $m_P = m_S = m_{PS} = 2$, and all the associated thresholds are 8.5 dB. The normalized transmit power of the SU varies from 16 to 40 dB in the three different cases associated with the normalized transmit power of the PU, which are $P_P = 22, 20, 18$ dB. The other parameters are the same as in Table I. At the MAC layer, the average packet arrival rate is $\lambda_P = 0.5$ packet/TS, whereas the delay tolerance is $T_P = 5$ TS.

It may be observed in Fig. 6 that the MASR of the SU first increases upon increasing the normalized transmit power P_S, since increasing P_S improves the SINR performance of the SU link, which results in an improved MASR. However, beyond a specific point, the MASR is gradually reduced upon further increasing P_S. That is because a higher P_S imposes an increased interference upon the PU link, which reduces the PU's chance of emptying its own queues. As long as the PU occupies the resources, this reduces the chances of the SU to

access the system, and as a result, the MASR of the SU is reduced. Given a higher transmit power P_P of the PU, the MASR of the SU is also increased, and the optimal power P_S maximizing the MASR of the SU becomes explicit in Fig. 6.

VI. Conclusion

Our novel hybrid CR system amalgamated the interweave and underlay paradigms for the sake of enhancing the chances of the SU to access the system. The hybrid parameter ε was introduced for optimally blending the interweave and underlay paradigms, which was defined as the SU's probability of accessing the system when the PU is still transmitting. Several problems have been solved without violating the delay constraints of the PU: 1) The scenario relying on realistic imperfect sensing was considered. Under the assumption of Nakagami-m fading, we have modeled the service links by a Poisson service process. 2) The optimal hybrid parameter of ε^* was found for the MASR of the SU. 3) The most suitable queue scheduling scheme $\{p_i^*,\ i = 1, 2, \ldots, N\}$ was found for the sake of minimizing the OAD of the SU's multiple queues.

Our numerical results characterized the influence of the PHY- and MAC-layer parameters on both the delay imposed and the achievable average service rate of the SU. Several interesting observations have been made: 1) Given different traffic loads of the PU, different design strategies should be adopted for the sake of enhancing the MASR of the SU. 2) If the traffic load is high, our best strategy is to focus on improving the PU link, which is capable of attaining a more substantial overall performance gain, rather than improving the channel conditions of the SU link. If the traffic load is light, improving the SU link may achieve the best performance. Furthermore, improving the reliability of the SU's sensing scheme is capable of achieving the same overall performance gain as improving the PU link. 3) When the delay tolerance tends to infinity, the performance of the system becomes limited by the queue stability. 4) The SU's OAD relying on our optimal scheme gets up to 27% and 20% lower than that of the proportional and round-robin schemes. 5) The optimal transmit power of the SU may be found for the sake of maximizing the average service rate of the SU.

In our future work, the shadowing effect imposed by the wireless channel will also be taken into account. Furthermore, the SU's performance will be studied to select the most suitable transmit power.

References

[1] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[2] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*. Oxford, U.K.: Pergamon, 1960, pp. 3–36.

[3] L. Hanzo, P. J. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. Hoboken, NJ: Wiley, 2007.

[4] Q. Du and X. Zhang, "Statistical QoS provisionings for wireless unicast/multicast of multi-layer video streams," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.

[5] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Stable throughput of cognitive radios with and without relaying capability," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2351–2360, Dec. 2007.

[6] X. Bao, P. Martins, T. Song, and L. Shen, "Stable throughput analysis of multi-user cognitive cooperative systems," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–5.

[7] A. A. El-Sherif, A. K. Sadek, and K. J. R. Liu, "Opportunistic multiple access for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 704–714, Apr. 2011.

[8] J. Hu, L.-L. Yang, and L. Hanzo, "Appendix of the Submitted Paper "Maximum Average Service Rate and Optimal Queue Scheduling of Delay-Constrained Hybrid Cognitive Radio in Nakagami Fading Channels," Tech. Rep. [Online]. Available: http://eprints.soton.ac.uk/340414/

[9] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] P. V. Mieghem, *Performance Analysis of Communications Networks and Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[11] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted ALOHA with multipacket reception," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2636–2656, Jul. 2005.

[12] R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple-access system," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 918–930, Sep. 1988.

[13] W. Luo and A. Ephremides, "Stability of N interacting queues in random access systems," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1579–1587, Jul. 1999.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York: Academic, 2007.

[16] L.-L. Yang, *Multicarrier Communications*. Hoboken, NJ: Wiley, Jan. 2009.

Jie Hu (S'11) received the B.Eng. degree in communication engineering and the M.Eng. degree in communication and information system from Beijing University of Posts and Telecommunications, Beijing, China, in 2008 and 2011, respectively. He is currently working toward the Ph.D. degree with the Communication, Signal Processing and Control Group, University of Southampton, Southampton, U.K.

His research interests in wireless communications include cognitive radio and cognitive networks, queuing analysis, resource allocation and scheduling, ad hoc wireless networks, and mobile social networks.

Lie-Liang Yang (M'98–SM'02) received the B.Eng. degree in communications engineering from Shanghai Tiedao University, Shanghai, China, in 1988 and the M.Eng. and Ph.D. degrees in communications and electronics from Northern (Beijing) Jiaotong University, Beijing, China, in 1991 and 1997, respectively.

From June 1997 to December 1997, he was a Visiting Scientist with the Institute of Radio Engineering and Electronics, Academy of Sciences of the Czech Republic, Prague, Czech Republic. Since December 1997, he has been with the University of Southampton, Southampton, U.K., where he is currently a Professor with the School of Electronics and Computer Science. He has published more than 240 research papers in journals and conference proceedings, authored or coauthored three books, and published several book chapters. Details about his publications can be found at http://www-mobile.ecs.soton.ac.uk/lly/. His research has covered a wide range of topics in wireless communications, networking, and signal processing.

Dr. Yang is currently an Associate Editor of the IEEE Transactions on Vehicular Technology, the *Journal of Communications and Networks*, and the *Security and Communication Networks* journal.

Lajos Hanzo (M'91–SM'92–F'04) received the M.S. degree (with first-class honors) in electronics and the Ph.D. degree from the Technical University of Budapest, Budapest, Hungary, in 1976 and 1983, respectively, the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2004, and the "Doctor Honoris Causa" degree from the Technical University of Budapest in 2009.

During his 35-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. Since 2009, he has been a Chaired Professor with Tsinghua University, Beijing, China. He is currently directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry; the Engineering and Physical Sciences Research Council, U.K.; the European Information Society Technologies Program; and the Mobile Virtual Centre of Excellence, U.K. He is an enthusiastic supporter of industrial and academic liaison and offers a range of industrial courses. He has successfully supervised 80 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications totaling in excess of 10 000 pages, published more than 1250 research entries on IEEE Xplore, and presented keynote lectures. For further information on research in progress and associated publications, see http://www-mobile.ecs.soton.ac.uk/.

Dr. Hanzo is a Fellow of the Royal Academy of Engineering, U.K., a Fellow of the Institution of Electrical Engineers, and a Governor of the IEEE Vehicular Technology Society. He has been a Technical Program Committee Chair and a General Chair for IEEE conferences. During 2008–2012, he was the Editor-in-Chief of the IEEE Press. He has received a number of distinctions.

AUTHOR QUERY

NO QUERY.

Maximum Average Service Rate and Optimal Queue Scheduling of Delay-Constrained Hybrid Cognitive Radio in Nakagami Fading Channels

Jie Hu, *Student Member, IEEE*, Lie-Liang Yang, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—As a promising technique to improve achievable bandwidth efficiency, cognitive radio (CR) has attracted substantial research attention from both the academic and industrial communities. To improve the performance attained by the secondary user (SU), a novel hybrid CR system is proposed, which combines the conventional interweave and underlay paradigms to enhance the chance of the SU to access the spectrum. Queuing theory is invoked in this paper to analyze the impact of the primary user's maximum tolerable delay on the performance of the SU. Multiple queues are assumed for the SU, which is engaged in video communication. Apart from the Poisson traffic generation, we also model the classic Nakagami-m fading channel as a Poisson service process by utilizing the outage probability in the presence of cochannel interference. We optimize both the hybrid interweave/underlay procedure to maximize the average service rate $\mu_{S,\max}$ of the SU, as well as the queue's scheduling scheme, for the sake of minimizing the overall average delay (OAD). As a result, the OAD of the SU is reduced by up to 27% and 20%, compared with the proportion and round-robin schemes, respectively.**

***Index Terms*—Hybrid cognitive radio (CR) system, maximum average service rate (MASR) of secondary user (SU), multilayer video streams (VSs), Nakagami-m fading channel, optimal queue scheduling scheme.**

I. Introduction

TWO kinds of customers are supported in cognitive radio (CR), namely, the primary user (PU) and the secondary user (SU). Furthermore, three main paradigms are considered for CR systems, namely, overlay, underlay, and interweave paradigms [1]. Here, we focus our attention on the interweave and underlay paradigms. According to the *interweave paradigm*, the PUs are authorized to access the channel, whereas the SUs are only able to access it when the PUs release it. In the scenario that the PUs request resources, an SU's session must be temporarily paused until the PUs complete their transmission. In the *interweave paradigm*, SUs sense the activities of PUs and clinch every possible opportunity to carry on with their own transmissions. By contrast, according to the *underlay paradigm*, PUs and SUs are allowed to transmit their data at the same time. An interference threshold is set up at the PUs to guarantee that the PU's link is not unduly interfered by SUs. Their transmission would only be terminated if the interference imposed by SUs exceeds the threshold.

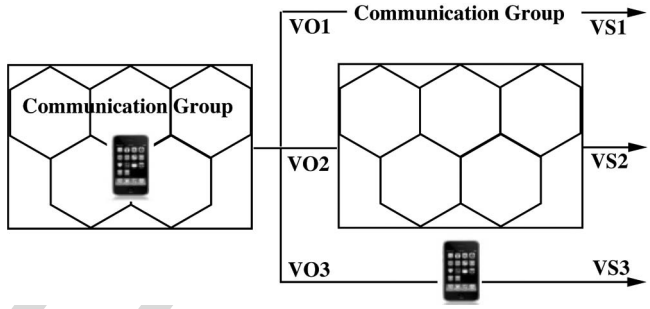

Fig. 1. Example of object-based MPEG-4 video transmission.

Although the transmission of PUs may be fully guaranteed with the aid of the *interweave paradigm*, only limited opportunities may be offered to SUs to convey their own data. Albeit SUs have improved chances of transmitting their own information to the destination using the *underlay paradigm*, the quality of service (QoS) of the PUs might not be as high as for the *interweave paradigm*. Can we strike an attractive compromise between these two paradigms so that the SUs have more opportunities to access the channels while the PUs may benefit from improved QoS? In this paper, a hybrid CR system is invoked to achieve this goal. In our novel system, SUs first detect the activity of PUs. If the channels are idle, the SUs may access them to transmit their own data. By contrast, if the channels are occupied by PUs, the SUs have a certain probability of accessing the channels while ensuring that the interference thresholds are not violated.

The Nakagami-m channel model has the advantage of characterizing a diverse range of multipath channels by a single fading parameter m [2]. Recently, supporting real-time video services subject to specific QoS constraints has become one of the essential requirements in wireless communications networks, where the video source is usually encoded into a number of streams in the application layer, representing the base and enhancement layers. An example of object-based MPEG-4 video transmission [3] is shown in Fig. 1, where different video objects are encoded into different video streams (VSs). The scheduling of the multilayer video queues at the SU's end

Manuscript received June 21, 2012; revised November 15, 2012 and December 20, 2012; accepted December 27, 2012. This work was supported in part by the RC-UK's India–UK Advanced Technology Centre, by the European Union's Concerto Project, by the China–UK Science Bridge, and by the China Scholarship Council. The review of this paper was coordinated by Prof. C. Lin.

The authors are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: jh10g11@ecs.soton.ac.uk; lly@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

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/TVT.2012.2237422

deserves further attention for the sake of achieving the best possible performance.

In [4], an adaptive resource-allocation scheme for multilayer VSs transmitted in wireless unicast/multicast scenarios is proposed, which inspires our work. The maximum throughput of an SU has been investigated in [5] under the interweave paradigm by finding the optimal transmit power of the SU in the scenario of a single queue conceived for the SU. In [6], this research is extended to the multiple PU and single SU CR system to explore the optimal transmit power and the relaying probability at the SU. The stability region has been already analyzed in the CR network supporting multiple PUs, multiple relays, and multiple SUs [7]. In [5]–[7], the interweave paradigm is employed, where the SUs have fewer opportunities to access the resources. All these studies assumed, however, that a single queue was set up for each of the users, which cannot faithfully represent the characteristics of multimedia communication. Furthermore, all results have been generated assuming queuing stability, where the specific QoS constraints such as the delay tolerance of lip-synchronized video communications were completely ignored.

Against this background, our novel contributions are as follows.

1) We amalgamate the interweave and underlay paradigms into a novel hybrid CR scheme characterized by the parameter $0 \leq \varepsilon \leq 1$, which is zero when the pure interweave paradigm is used and one for the pure underlay paradigm.
2) The Nakagami-m fading channel is used while modeling the service as a Poisson process by deriving the closed-form tail probability in the presence of cochannel interference. This allows us to use the classic $M/M/1$ queuing theory to investigate the delay performance.
3) The effects of both the PU's QoS requirement and the erroneous spectrum-sensing decisions on the performance of the SU are investigated.
4) The optimal parameter ε maximizing the average service rate of the SU is determined.
5) By exploiting the Lagrange optimization method, we derive the optimal closed-form solution for queue scheduling. Given this scheme, we are capable of minimizing the overall average delay (OAD) of multilayer VSs.

This paper is organized as follows. Our system model, including both the medium access control (MAC) and physical (PHY) layers, are described in Section II, whereas in Section III, the related queuing analysis is carried out for both the PU and the SU. In Section IV, the problem of finding the optimal solution is formulated and solved, followed by our numerical results in Section V. Finally, our conclusions are offered in Section VI.

II. System Model

Fig. 2 shows a CR system having a pair of PU source and PU destination (PD), as well as a pair of SU source and SU destination (SD). A novel hybrid CR policy is employed. Given a spectrum band, the SU first senses the activity of the PU. If this sensing process is reliable, the SU may access the spectrum

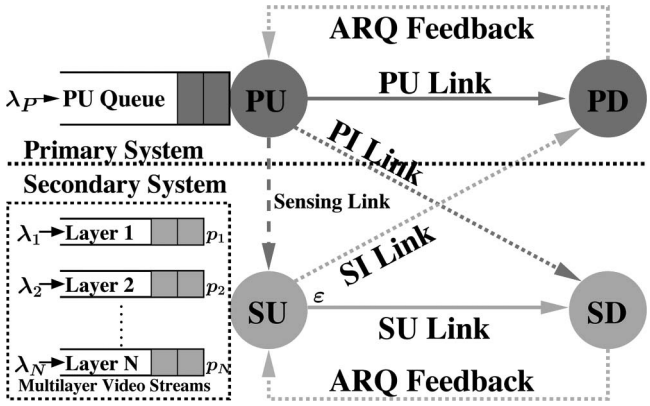

Fig. 2. System model with one PU and one SU sharing the same spectrum.

with a unity probability when the PU is idle. Furthermore, the SU may access the spectrum with a probability of $0 \leq \varepsilon \leq 1$, even if the PU is busy, provided that the PU's QoS is still guaranteed. If, however, this sensing process failed, the situation is reversed, and hence, the SU accesses the spectrum with a probability of ε when the PU is idle and with a unity probability when the PU is busy. Hence, erroneous spectrum sensing may lead to catastrophic performance degradation for both the PU and the SU. We will also consider the practical scenario of an erroneous spectrum-sensing decision and its effect on the performance of the SU. Let us now consider the details of the MAC and PHY layers.

A. MAC-Layer Model

Observe in Fig. 2 that a buffer is provided for the PU to store the packets, which cannot be immediately transmitted. Again, to support video communications for the SU, multiple buffers are provided to accommodate the different-priority queues generated by multilayer video encoding.

The basic unit of time in our system is a time slot (TS). The packets' arrival at the PU's buffer obeys the classic Poisson process with a mean of λ_P packets/TS. The QoS constraint of the PU is the total delay imposed by both the transmission and buffering delay, which should not exceed T_P, namely, the maximum tolerable delay of the PU. The packets' arrival at the buffer input of the SU also follows the Poisson process with a mean of $\{\lambda_n,\ n = 1, 2, \ldots, N\}$ packets/TS. Regardless of whether the PU is idle or busy, if the SU has already successfully accessed the spectrum, we should determine the probability of each queue at the SU completing its transmission of the stored packets. These probabilities may be denoted by $\{p_n,\ n = 1, 2, \ldots, N\}$. No specific QoS constraints are imposed at the SU. However, our prime goal is to find the optimal scheduling scheme and the associated hybrid parameter ε, which minimizes the OAD of all the queues at the SU.

We assume that every packet transmitted by the PU or the SU may be only successfully received by the corresponding destination in a TS when the signal-to-interference-plus-noise ratio (SINR) is above a specific threshold. In our system, automatic repeat request (ARQ) with an unlimited number of retransmissions is adopted to ensure that no packets are lost.

Furthermore, we assume that the acknowledgements are always successfully received by the transmitters.

Theorem 1: Given the successful reception probability μ of a single packet during a TS in the ARQ-aided system, the continuous transmit time obeys the exponential distribution with the mean of $1/\mu$.

Proof: The proof is provided in [8, App. A]. ■

As a benefit of the ARQ mechanism, a packet's departure from the buffer is synonymous to being served. The wireless channel may be considered as a "server" in the queuing analysis. As a result, the packets' departure process is also referred to as the service process.

B. PHY-Layer Model

Observe from Fig. 2 that there are five links in this hybrid CR system, namely, the PU link, the SU link, the sensing link, the PI link imposing interference from the PU on the SD, and the SI link inflicting interference from the SU upon the PD. In line with [11], a multipacket reception (MPR) model is introduced for the queuing analysis of the collision channel. We define the corresponding conditional probabilities for the MPR model of these five links.

- $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$ (or $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$): Only the PU's (or SU's) packet is successfully received at the PD (or SD) when only the PU (or SU) transmits.
- $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$ (or $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$): Only the PU's (or SU's) packet is successfully received at the PD (or SD) when both the PU and the SU transmit.
- $q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$: Only the PU's packet is successfully received at the SU when the PU transmits.

Next, we will briefly highlight the derivation of these conditional probabilities. The radio propagation between any pair of nodes is assumed to be affected by the independent stationary Nakagami-m flat-fading channels $h_i(t)$ having $E[|h_i(t)|^2]=1$ (t is the TS index). Hence, $Z=|h_i(t)|^2$ is a Gamma-distributed random variable (r.v.) having a mean of 1, which may be denoted by $Z\sim\mathrm{Gamma}(m,1/m)$ with the shape and scale parameters of m and $1/m$, respectively. Given the probability density function (pdf) $f_Z(z)$ of [9, eq. (3.39)], the tail distribution function (TDF) of Z is formulated as $P(Z>z)=\Gamma(m,mz)/\Gamma(m)$, where $\Gamma(\cdot,\cdot)$ represents the upper incomplete Gamma function, and m denotes the Nakagami-m fading parameter [15].

After being divided by the noise power, the normalized transmit power of the PU is defined as P_P, whereas the normalized transmit power of the SU is defined as P_S. The propagation path loss (PL) is denoted by Ω_i, and the SINR threshold that has to be exceeded for successful packet reception is denoted by β_i, while again, the Nakagami-m fading parameter is denoted by m_i, where i represents "P" for the PU link, "S" for the SU link, "PI" for the PI link, "SI" for the SI link, and "PS" for the sensing link spanning from the PU to the SU.

According to the TDF of the r.v. Z, we express the conditional probabilities of $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$, $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$, and $q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$ in the MPR model in the absence of interference thresholds as follows:

$$\begin{aligned}
q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} &= P\left[\frac{|h_P(t)|^2 P_P}{\Omega_P}>\beta_P\right]\\
&=\Gamma\left(m_P,\frac{m_P\beta_P\Omega_P}{P_P}\right)\Big/\Gamma(m_P)\\
q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}} &= P\left[\frac{|h_S(t)|^2 P_S}{\Omega_S}>\beta_S\right]\\
&=\Gamma\left(m_S,\frac{m_S\beta_S\Omega_S}{P_S}\right)\Big/\Gamma(m_S)\\
q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} &= P\left[\frac{|h_{\mathrm{PS}}(t)|^2 P_P}{\Omega_{\mathrm{PS}}}>\beta_{\mathrm{PS}}\right]\\
&=\frac{\Gamma\left(m_{\mathrm{PS}},m_{\mathrm{PS}}\beta_{\mathrm{PS}}\Omega_{\mathrm{PS}}/P_P\right)}{\Gamma(m_{\mathrm{PS}})}.
\end{aligned}$$

We introduce another two Gamma-distributed r.v.'s, namely, $X\sim\mathrm{Gamma}(m_X,1/m_X)$ with a pdf of $f_X(x)$ and $Y\sim\mathrm{Gamma}(m_Y,1/m_Y)$ with a pdf of $f_Y(y)$. Then, the problem of solving $q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$ and $q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$ becomes equivalent to solving the probability problem of $P[X>A+BY]$. Unfortunately, a closed-form solution cannot be derived for the general situation. Nonetheless, we may obtain the following results for two special cases:

$$P[X>A+BY]=\begin{cases}\Phi(A,B), & \text{if } m_X \text{ is a positive integer}\\ & \quad m_Y \text{ can be any real number}\\ \Phi'(A,B), & \text{if } m_Y \text{ is a positive integer}\\ & \quad m_X \text{ can be any real number.}\end{cases}$$

$$\begin{cases}\Phi(A,B)=\left(\frac{m_Y}{m_XB+m_Y}\right)^{m_Y}e^{-m_XA}\\ \qquad\cdot\sum_{n=0}^{m_X-1}\sum_{k=0}^{n}\frac{m_X^nB^kA^{n-k}}{k!(n-k)!(m_XB+m_Y)^k}\frac{\Gamma(m_Y+k)}{\Gamma(m_Y)}\\ \Phi'(A,B)=\frac{\Gamma(m_X,m_XA)}{\Gamma(m_X)}-\sum_{n=0}^{m_Y-1}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{m_Y}{B}\right)^n\\ \qquad\cdot\frac{e^{m_Y\frac{A}{B}}(-A)^{n-k}m_X^{m_X}}{n!\Gamma(m_X)}\cdot\frac{\Gamma[k+m_X,A(m_X+m_Y/B)]}{(m_X+m_Y/B)^{k+m_X}}.\end{cases}\tag{1}$$

For a detailed derivation of $\Phi(A,B)$ and $\Phi'(A,B)$, see [8, App. B]. Given this result, we may write down the other two conditional probabilities when subjected to the cochannel interference thresholds at the PD and the SD as follows:

$$\begin{aligned}
&q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\\
&=P\left[\frac{|h_P(t)|^2P_P/\Omega_P}{1+|h_{\mathrm{SI}}(t)|^2P_S/\Omega_{\mathrm{SI}}}>\beta_P\right]\\
&=P\left[|h_P(t)|^2>\frac{\beta_P\Omega_P}{P_P}+\frac{\beta_PP_S\Omega_P}{P_P\Omega_{\mathrm{SI}}}|h_{\mathrm{SI}}(t)|^2\right]\\
&=\begin{cases}\Phi\left(\frac{\beta_P\Omega_P}{P_P},\frac{\beta_PP_S\Omega_P}{P_P\Omega_{\mathrm{SI}}}\right), & \text{if } m_P \text{ is a positive integer}\\ \Phi'\left(\frac{\beta_P\Omega_P}{P_P},\frac{\beta_PP_S\Omega_P}{P_P\Omega_{SI}}\right), & \text{if } m_{SI} \text{ is a positive integer}\end{cases}
\end{aligned}\tag{2}$$

$$
\begin{aligned}
&q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} \\
&= P\left[\frac{|h_S(t)|^2 P_S/\Omega_S}{1+|h_{\mathrm{PI}}(t)|^2 P_P/\Omega_{\mathrm{PI}}} > \beta_S\right] \\
&= P\left[|h_S(t)|^2 > \frac{\beta_S\Omega_S}{P_S} + \frac{\beta_S P_P \Omega_S}{P_S\Omega_{\mathrm{PI}}}\,|h_{\mathrm{PI}}(t)|^2\right] \\
&= \begin{cases} \Phi\left(\frac{\beta_S\Omega_S}{P_S}, \frac{\beta_S P_P\Omega_S}{P_S\Omega_{\mathrm{PI}}}\right), & \text{if } m_S \text{ is a positive integer} \\ \Phi'\left(\frac{\beta_S\Omega_S}{P_S}, \frac{\beta_S P_P\Omega_S}{P_S\Omega_{\mathrm{PI}}}\right), & \text{if } m_{\mathrm{PI}} \text{ is a positive integer.} \end{cases}
\end{aligned} \tag{3}
$$

III. Queuing Analysis

The following tasks will be accomplished here: 1) The classic M/M/1 queuing is revisited. 2) The average service rate of the PU and the SU is attained for the original benchmark system. 3) A new system is defined for characterizing the interdependence of the PU's and SU's queues, and the corresponding average service rate of the PU and the SU is rederived.

A. $M/M/1$ Queuing System

In the M/M/1 queuing system, only a single server is invoked by the system. The packets' interarrival time follows the exponential distribution having a parameter λ, whereas the packets' interdeparture time obeys the same distribution having a parameter μ. An infinite buffer is assumed for guaranteeing that every single packet may be successfully transmitted, i.e., without any packet loss. According to [10], we can express the probability of the queue being empty and the total average delay, including the transmission and buffering time, as follows:

$$
P[Q=0] = 1-\lambda/\mu, \quad T = 1/(\mu-\lambda), \quad \text{where} \quad \lambda < \mu \tag{4}
$$

and we denote the random queue length and the total average delay by Q and T, respectively.

B. PU Link

According to the MAC-layer protocol of our hybrid CR system, the PU has the highest priority to access the channel. If the PU's queue is not empty, the PU would occupy the channel for its transmission, and the packet would be removed from the head of the queue when it is successfully received by the PD. However, the PU would suffer from the interference imposed by the activity of the SU. The service process of a PU's single packet follows the Bernoulli distribution. In analogy to the successful packet departure probability in the Bernoulli trial, the average service rate of the Poisson service process may be formulated for the PU link as

$$
\begin{aligned}
\mu_P = {} & P\{Q_S=0\}\cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} + P\{Q_S\neq 0\} \\
& \cdot\left(q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot\left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot(1-\varepsilon) + q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\cdot\varepsilon\right)\right. \\
& \left. + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}\right)
\end{aligned} \tag{5}
$$

where Q_S represents the random queue length of the SU. Given the average service rate μ_P, the average arrival rate λ_P, and the QoS constraint stipulating that the total delay should not exceed the maximum delay tolerance T_P, we infer from (4) that the condition $0 \le \lambda_P \le \mu_P - 1/T_P$ must be obeyed.

C. SU Link

According to the MAC-layer protocol of the SU, before accessing the channel, the SU carries out spectrum sensing first. When the channel is released by the PU, after reliable sensing, the SU would transmit its own packets with a probability of unity (otherwise, with a probability of ε), which implies inefficient resource usage. When the channel is still occupied by the PU, following reliable sensing, the SU would transmit its own packets with a probability of ε (otherwise, with a probability of unity), which imposes strong interference on the PD. It is vital to guarantee that the transmission of the SU does not force the PU to violate its delay tolerance. We derive the probability of a packet being served, which is also identical to the average service rate μ_S of the SU's Poisson service process expressed as

$$
\begin{aligned}
\mu_S = {} & q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}\cdot P[Q_P=0] \\
& \cdot\left[q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\cdot\varepsilon\right] \\
& + q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}\cdot P[Q_P\neq 0] \\
& \cdot\left[q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\cdot\varepsilon + \left(1-q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right)\right]
\end{aligned} \tag{6}
$$

where Q_P represents the random queue length of the PU.

D. Stochastic Dominance Principle

It may be shown from (5) and (6) that the average service rates of the SU and the PU depend on each other's queue sizes. Since these queues interact with each other, the average rate of the individual service processes cannot be directly determined. For the sake of circumventing this problem, the stochastic dominance principle of [12] is invoked to assist our analysis.

The concept of a so-called "dominant system" was defined in [12] by allowing a set of terminals having no packets in their transmit-buffer to continue transmitting hypothetical dummy packets. In this manner, the queues in the dominant system stochastically dominate the queues in the original system. This dominant system is defined here.

1) If we have $Q_S = 0$ and $Q_P = 0$, the SU transmits dummy packets with a unity probability.
2) If $Q_S = 0$ and $Q_P \neq 0$, the SU transmits dummy packets with a probability ε.

Note that the hypothetical dummy packets are introduced only for the sake of facilitating our performance analysis and for finding closed-form solutions to the optimum scheduling factor ε, but in reality, no dummy packets are assumed to be transmitted in the practical systems considered. As shown in [13], given the same initial conditions, the queuing performance of the dominant system transmitting hypothetical

dummy packets is capable of providing a tight approximation of the queuing performance of the original system. Since in practical systems no dummy packets are actually transmitted, no extra interference will be imposed on the system, and no extra power is required for their transmissions.

In the context of the extended dominant system, $\{Q_S \neq 0\}$ has a unity probability. Hence, (5) can be rewritten as

$$\mu_P = q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \cdot \left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} \cdot (1-\varepsilon) + q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} \cdot \varepsilon \right) + \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \cdot q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} \quad (7)$$

and the condition $0 \leq \lambda_P \leq \mu_P - 1/T_P$ is also rewritten as

$$0 \leq \lambda_P \leq q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} + \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}} - 1/T_P - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \left(q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}} - q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SUh}\}}\right) \varepsilon. \quad (8)$$

According to (4), we may represent the probabilities of the PU's queue being not empty and empty by λ_P and μ_P, respectively, as shown in (7). Hence, we may rewrite (6) as

$$\mu_S = q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \varepsilon - \frac{\lambda_P}{\mu_P} \left[q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} + q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU}\}} \times \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \varepsilon - q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}} \varepsilon - q^{\mathrm{SD}}_{\mathrm{SU}|\{\mathrm{SU,PU}\}} \left(1 - q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}\right) \right]. \quad (9)$$

To simplify the derivations in the next section, some simple notations are introduced as follows: $C = q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU}\}}$, $D = q^{(\mathrm{PD})}_{\mathrm{PU}|\{\mathrm{PU,SU}\}}$, $E = q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU}\}}$, $F = q^{(\mathrm{SD})}_{\mathrm{SU}|\{\mathrm{SU,PU}\}}$, and $S = q^{(\mathrm{SU})}_{\mathrm{PU}|\{\mathrm{PU}\}}$.

IV. Problem Formulation and Solutions

Here, the following two problems are solved: 1) finding the optimal ε^* for maximizing the average service rate $\mu_{S,\max}$ of the SU and 2) finding the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$ to minimize the OAD associated with $\mu_{S,\max}$.

A. MASR of the SU

The optimal parameter ε^* will be found for the sake of maximizing the average service rate of the SU without violating the delay tolerance of the PU. The solution can be found by constructing the problem **P1** as

$$\arg\max_{\varepsilon}\ \mu_S(\varepsilon)$$
$$0 \leq \varepsilon \leq \xi = \frac{SC + (1-S)D - \lambda_P - 1/T_P}{S(C-D)} \leq 1 \quad (10)$$

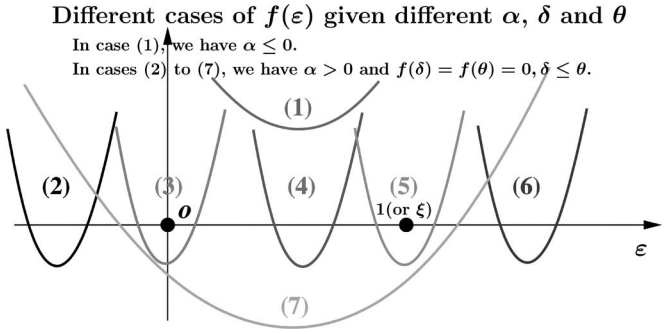

Fig. 3. Different cases of $f(\varepsilon)$ given different α, δ, and θ.

where the function $\mu_S(\varepsilon)$ is defined by (9), whereas the inequality condition in (10) is imposed for the sake of satisfying the specific delay tolerance of the PU, which is derived from (8).

To solve **P1**, we differentiate $\mu_S(\varepsilon)$ with respect to ε, yielding

$$\frac{d\mu_S}{d\varepsilon} = \frac{f(\varepsilon)}{[SC + (1-S)D - S(C-D)\varepsilon]^2} \quad (11)$$

where we define $f(\varepsilon) = E(1-S)[SC + (1-S)D - S(C-D)\varepsilon]^2 - \lambda_P \alpha$ and $\alpha = S(C-D)(E-F) + DE(1-S) - FSD$. Since the denominator of (11) is always positive, we can simply consider $f(\varepsilon)$ instead, which is a quadratic function having the shape shown in Fig. 3.

Let us consider the constraint shown in (8) more closely. Given $0 \leq \varepsilon \leq 1$, the right side of (8) reaches its minimum value of $(D - 1/T_P)$ when we have $\varepsilon = 1$. Therefore, if we have $\lambda_P \leq D - 1/T_P$, ε may assume any arbitrary value ranging from 0 to 1. By contrast, if $\lambda_P > D - 1/T_P$, then ε must be selected from the range determined by (10) in the context of the specific λ_P value.

Different scenarios have to be considered for finding the optimal values of ε^*, which lead to the maximum average service rate (MASR) $\mu_S(\varepsilon)$ of the SU formulated in (10).

Case 1 ($\lambda_P \leq D - 1/T_P$): In this case, as aforementioned, ε might assume any arbitrary value ranging from 0 to 1. Under this assumption, the problem may be further divided into the following two subcases, which will assist us in finding $\mu_{S,\max}$.

Subcase 1-1 ($\alpha \leq 0$): When $\alpha \leq 0$, the $f(\varepsilon)$ curve obeys the relative position shown by case (1) at the top of Fig. 3, in which case there are no real-valued solutions to $f(\varepsilon) = 0$. Hence, provided that $E(1-S) > 0$ holds, $f(\varepsilon)$ is always higher than 0, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function. Therefore, we have $\varepsilon^* = 1$ and $\mu_{S,\max} = \mu_S(1)$.

Subcase 1-2 ($\alpha > 0$): Second, when $\alpha > 0$, there are two real-valued solutions to $f(\varepsilon) = 0$, which are denoted by δ and θ, respectively, i.e.,

$$\delta = \frac{1}{S(C-D)} \left[SC + (1-D) - \sqrt{\frac{\lambda_P \left[S(C-D)(E-F) + DE(1-S) - FSD\right]}{E(1-S)}} \right] \quad (12)$$

$$\theta = \frac{1}{S(C-D)} \left[SC + (1-D) + \sqrt{\frac{\lambda_P \left[S(C-D)(E-F) + DE(1-S) - FSD \right]}{E(1-S)}} \right] \quad (13)$$

where we have $\delta \leq \theta$. In this case, depending on the values of δ and θ, the relative position of the $f(\varepsilon)$ curve corresponds to one of the scenarios (2)–(7) shown in Fig. 3, which are specifically considered in the following.

a) If $\delta \leq \theta < 0$, the $f(\varepsilon)$ curve obeys the positioning of case (2) in Fig. 3, where $f(\varepsilon) > 0$ holds within the range of $0 \leq \varepsilon \leq 1$, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Hence, $\varepsilon^* = 1$ maximizes the average service rate $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(1)$.

b) If $\delta < 0$ and $0 \leq \theta < 1$, the $f(\varepsilon)$ curve obeys the positioning of case (3) in Fig. 3. Hence, $f(\varepsilon) \leq 0$ holds within the range of $0 \leq \varepsilon \leq \theta$, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. However, within the range of $\theta < \varepsilon \leq 1$, $f(\varepsilon) > 0$ holds, and hence, $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. In this situation, the optimal value of ε^* is either 0 or 1. Hence, we have $\varepsilon^* = \arg\max_{\varepsilon \in \{0,1\}} \{\mu_S(\varepsilon)\}$, and the MASR becomes $\mu_{S,\max} = \max\{\mu_S(0), \mu_S(1)\}$.

c) If $0 \leq \delta < \theta < 1$, the $f(\varepsilon)$ curve corresponds to case (4) in Fig. 3. Therefore, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq \delta$, implying that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Within the range of $\delta < \varepsilon \leq \theta$, however, $f(\varepsilon) \leq 0$ holds, which suggests that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Finally, in the range of $\theta < \varepsilon \leq 1$, $f(\varepsilon) > 0$ is satisfied, which indicates that $\mu_S(\varepsilon)$ becomes a monotonically increasing function of ε again. According to the preceding analysis, the pair of potential values that may maximize $\mu_S(\varepsilon)$ are $\varepsilon = \delta$ and $\varepsilon = 1$. Hence, the optimal value of ε^* is $\varepsilon^* = \arg\max_{\varepsilon \in \{\delta,1\}} \{\mu_S(\varepsilon)\}$, and the MASR is, hence, $\mu_{S,\max} = \max\{\mu_S(\delta), \mu_S(1)\}$.

d) If $0 \leq \delta < 1$ and $\theta \geq 1$, which results in an $f(\varepsilon)$ curve corresponding to case (5) in Fig. 3, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq \delta$, which implies that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Within the range of $\delta < \varepsilon \leq 1$, however, $f(\varepsilon) < 0$ holds, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Jointly considering these two situations, we conclude that $\varepsilon^* = \delta$ maximizes the average service rate $\mu(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(\delta)$.

e) If $\delta \geq 1$, we have encountered case (6) of Fig. 3. Correspondingly, $f(\varepsilon) \geq 0$ holds within the range of $0 \leq \varepsilon \leq 1$, which implies that $\mu_S(\varepsilon)$ is a monotonically increasing function of ε. Hence, $\varepsilon^* = 1$ maximizes the average service rate of $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(1)$.

f) Finally, if $\delta < 0 < 1 \leq \theta$, $f(\varepsilon)$ is reminiscent of case (7) of Fig. 3. Consequently, $f(\varepsilon) \leq 0$ holds within the range of $0 \leq \varepsilon \leq 1$, which implies that $\mu_S(\varepsilon)$ is a monotonically decreasing function of ε. Hence, $\varepsilon^* = 0$ maximizes the average service rate $\mu_S(\varepsilon)$, which is $\mu_{S,\max} = \mu_S(0)$.

Case 2 ($\lambda_P > D - 1/T_P$): If $\lambda_P > D - 1/T_P$, then ε may assume any arbitrary value ranging from 0 to ξ, where the value of ξ is restricted by the value of λ_P according to (10). Following a similar train of thought, by simply substituting "1" in Case 1 by "ξ," we have the following conclusions, which correspond to the cases (1)–(7) in Fig. 3, as we discussed earlier.

1) If $\alpha \leq 0$ (see case (1) in Fig. 3), or if $\alpha > 0$ and $\theta < 0$ (see case (2) in Fig. 3), or if $\alpha > 0$ and $\xi \leq \delta$ (see case (6) in Fig. 3), then we have $\varepsilon^* = \xi$, and the MASR becomes $\mu_{S,\max} = \mu_S(\xi)$.
2) If $\alpha > 0$, $\delta < 0$, and $0 \leq \theta < \xi$ (see case (3) in Fig. 3), then we have $\varepsilon^* = \arg\max_{\varepsilon \in \{0,\xi\}} \{\mu_S(\varepsilon)\}$, and the MASR becomes $\mu_{S,\max} = \max\{\mu_S(0), \mu_S(\xi)\}$.
3) If $\alpha > 0$ and $0 \leq \delta < \theta < \xi$ (see case (4) in Fig. 3), then we have $\varepsilon^* = \arg\max_{\varepsilon \in \{\delta,\xi\}} \{\mu_S(\varepsilon)\}$, and the MASR is $\mu_{S,\max} = \max\{\mu_S(\delta), \mu_S(\xi)\}$.
4) If $\alpha > 0$ and $0 \leq \delta < \xi \leq \theta$ (see case (5) in Fig. 3), then $\varepsilon^* = \delta$, and the MASR is $\mu_{S,\max} = \mu_S(\delta)$.
5) If $\alpha > 0$, $\delta < 0$, and $0 < \xi \leq \theta$ (see case (7) in Fig. 3), then $\varepsilon^* = 0$, and the MASR is formulated as $\mu_{S,\max} = \mu_S(0)$.

B. Optimal Queue Scheduling of the SU

Let us now further detail the derivation of the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$[1] conceived for minimizing the OAD of the SU's queues. Given (4), the average delay of each queue may be formulated as

$$\text{Delay}_n = 1/(p_n \cdot \mu_S(\varepsilon) - \lambda_n), \qquad n = 1, 2, \ldots, N \quad (14)$$

where p_n is the probability of the SU's nth queue transmitting its stored packets, and λ_n is the packets' average arrival rate.

Given the associated MASR $\mu_{S,\max}$ of the SU, we construct the second problem **P2**, i.e.,

$$\arg \min_{\{p_1, p_2, \ldots, p_N\}} \left\{ \frac{1}{N} \sum_{n=1}^{N} \frac{1}{p_n \cdot \mu_{S,\max} - \lambda_n} \right\} \quad (15)$$

$$\text{subject to} \quad \sum_{n=1}^{N} p_n = 1, \quad \lambda_n - p_n \cdot \mu_{S,\max} < 0. \quad (16)$$

Then, we may readily show the following.

1) The objective function (15) is convex over $\{p_n,\ n = 1, 2, \ldots, N\}$.
2) The first equality constraint in (16) is affine over $\{p_n,\ n = 1, 2, \ldots, N\}$, whereas the second inequality constraint in (16) is convex over $\{p_n,\ n = 1, 2, \ldots, N\}$.

Hence, **P2** is a convex problem [14], and the optimal solution may be obtained with the aid of the Lagrangian optimization method while additionally exploiting the Karush–Kuhn–Tucker (KKT) conditions [14], which were proposed for solving convex optimization problems similar to ours under an inequality constraint.

[1]The optimal shceduling scheme is fully specified when the probability p_n of transmitting the stream n is determined.

TABLE I
PHY-LAYER PARAMETER SETTING

	PL	Threshold	Fading	Norm Power
PU Link	Ω_P = 5 dB	β_P = 13.9 or 8.5 dB	m_P = 1 or 2	P_P = 26 dB
SU Link	Ω_S = 5 dB	β_S = 13.9 or 8.5 dB	m_S = 1 or 2	P_S = 26 dB
PI/SI Link	$\Omega_{PI/SI}$ = 15 dB	No threshold	$m_{SI/PI}$ = 1	$P_{P/S}$ = 26 dB
Sensing Link	Ω_{PS} = 2 dB	β_{PS} = 13.9 or 8.5 dB	m_{PS} = 1 or 2	P_P = 26 dB

The Lagrangian function is constructed as

$$J = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{p_n\mu_{S,\max}-\lambda_n} + \sum_{n=1}^{N}\varphi_n(\lambda_n - p_n\mu_{S,\max}) + \eta\left(\sum_{n=1}^{N}p_n - 1\right) \quad (17)$$

where $\eta \geq 0$ and $\varphi_n \geq 0,\ n = 1, 2, \ldots, N$ are the Lagrangian multipliers associated with the two conditions of (16), respectively. Then, the optimal queue scheduling scheme $\{p_n^*,\ n = 1, 2, \ldots, N\}$ and the Lagrangian multipliers of the optimization problem **P2** satisfy the following KKT conditions [14]:

$$\begin{cases} \left.\frac{\partial J}{\partial p_n}\right|_{p_n=p_n^*} = \frac{-\mu_{S,\max}}{N(p_n^*\mu_{S,\max}-\lambda_n)^2} - \varphi_n^*\mu_{S,\max} + \eta^* = 0 \\ \left.\frac{dJ}{d\eta}\right|_{p_n=p_n^*} = \sum_{n=1}^{N} p_n^* - 1 = 0 \\ \varphi_n^*\left(\lambda_n - p_n^*\mu_{S,\max}\right) = 0, \quad \varphi_n^* \geq 0 \text{ and } \eta^* \geq 0 \end{cases} \quad (18)$$

where $n = 1, 2, \ldots, N$. Naturally, according to the second constraint of (16), the only solution satisfying the third line of (18) is $\varphi_n^* = 0$. Substituting $\varphi_n^* = 0$ into the first line of (18), we derive p_n^* as

$$p_n^* = \frac{1}{\sqrt{N\eta^*\mu_{S,\max}}} + \frac{\lambda_n}{\mu_{S,\max}}. \quad (19)$$

Then, substituting (19) into the second line of (18), we arrive at

$$\eta^* = N\cdot\mu_{S,\max}\Big/\left(\mu_{S,\max} - \sum_{n=1}^{N}\lambda_n\right)^2. \quad (20)$$

Finally, substituting η^* of (20) into (19), we arrive at the optimal queue scheduling scheme characterized by

$$p_n^* = \frac{\mu_{S,\max} - \sum_{n=1}^{N}\lambda_n + N\lambda_n}{N\cdot\mu_{S,\max}}. \quad (21)$$

V. PERFORMANCE ANALYSIS

In our numerical results, we link the PHY-layer parameters to a realistic orthogonal frequency-division multiplexing scheme communicating over Nakagami-m fading channels without channel coding. We define the SINR threshold of the receiver for the sake of guaranteeing that the bit error ratio of the system is not higher than 10^{-2}. Hence, the threshold is determined by the channel model having different fading parameters. For example, according to [16, Fig. 5-5], if $m = 1$, the threshold at the receiver is 13.86 dB, whereas the threshold is 8.50 dB for $m = 2$.

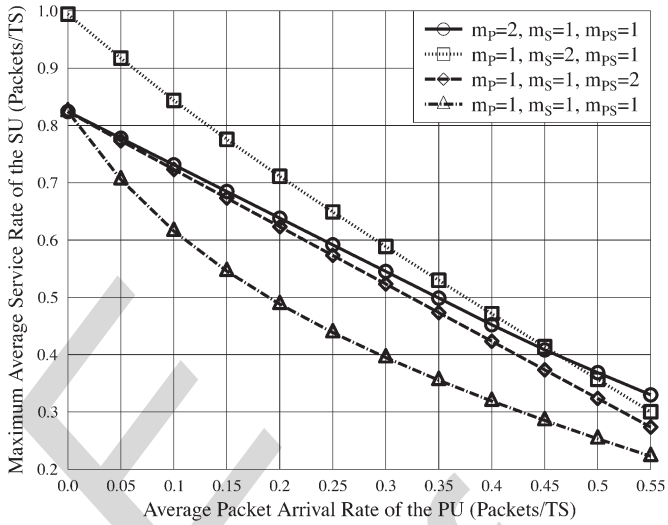

Fig. 4. Service rate of the SU versus the average packets arrival rate λ_P of the PU parameterized by the fading parameters.

We assume that the SU has five different queuing buffers in the MAC layer, where we set the average packet arrival rate of each queue, for example, to 0.01, 0.02, 0.03, 0.04, and 0.05 packet/TS, respectively.

Two other queue scheduling schemes are used as benchmarkers, namely, the round-robin scheme, where every queue of the SU has the same probability of transmitting packets stored in the buffer, and the proportional scheme, where the transmission probabilities are arranged to be proportional to the average arrival rates at the buffers.

A. Impact of the PU'S Average Packet Arrival Rate

In the PHY layer, the PL, the SINR threshold of the receiver associated with the Nakagami-m fading parameters, and the transmit power are set according to Table I. More particularly, there are four cases for the Nakagami-m parameters of the PU, the SU, and the sensing links, namely, 1) $m_P = 1$, $m_S = 1$, and $m_{PS} = 1$; 2) $m_P = 2$, $m_S = 1$, and $m_{PS} = 1$; 3) $m_P = 1$, $m_S = 2$, and $m_{PS} = 1$; and 4) $m_P = 1$, $m_S = 1$, and $m_{PS} = 2$.

In the MAC layer, the delay tolerance of the PU is $T_P = 5$ TS. We vary the average packet arrival rate λ_P from 0.1 to 0.5 packet/TS for characterizing its impacts on the MASR and the OAD in conjunction with three different queue scheduling schemes for the SU.

It may be readily shown from Fig. 4 that $\mu_{S,\max}$ linearly decays upon increasing λ_P. This is plausible since a higher λ_P results in an increased traffic load for the PU link, which, in turn, results in the PU's prolonged occupation of the

spectral band. This reduces the SU's chance of transmitting its own packets, which is reflected by the reduction of $\mu_{S,\max}$. Upon increasing the Nakagami-m parameters of the PU, SU, and sensing links, respectively, the MASR is improved. If we consider the higher m_P of the PU link, it is shown in Fig. 4 that the data stream of the PU may be transmitted more promptly and the duration of the PU's spectrum occupation would be shortened. Consequently, the SU's increased chances of transmitting its own data results in an increased MASR. Hence, the MASR is substantially improved. If we consider the higher m_S of the SU link characterized in Fig. 4, the MASR is also directly improved. For higher m_{PS} values, the MASR of Fig. 4 would be also improved due to the fact that the spectrum-sensing decisions become more reliable.

We may infer the following observations from Fig. 4 depending on the traffic load of the PU. We have the best performance gain upon improving the SU link if the PU's traffic load is lower than 0.45 packet/TS. By contrast, if the traffic load is higher than 0.45 packet/TS, improving the PU link may provide the best performance. Upon comparing the improvements of the sensing and PU links shown in Fig. 4, when the PU's traffic load is lower than 0.1 packet/TS, we infer that improving the sensing link may provide the same performance gain as improving the PU link. However, when the traffic load is higher than 0.1 packet/TS, improving the PU link is more influential. These observations may motivate us to design a more effective strategy for the sake of improving the performance of the SU, given different traffic loads of the PU. For example, given a light PU traffic load, improving the SU link is our best option to enhance overall performance. Furthermore, a reliable sensing strategy conceived for the SU may achieve the same overall performance as an improved PU link. By contrast, for a high traffic load, researchers should focus on directly improving the PU link.

B. Impact of the PU'S Delay Tolerance on the SU

Here, we consider the Nakagami-m parameters of $m_P = m_S = m_{PS} = 2$, and all the associated SINR thresholds are 8.5 dB, where the transmit power of the PU and the SU are $P_P = P_S = 18$ dB, whereas the others are the same as in Table I. We vary the delay tolerance T_P from 3 to 8 TS for the sake of investigating its impacts on the OAD in conjunction with three different queue scheduling schemes. Two scenarios, namely, $\lambda_P = 0.4$ and 0.5 packet/TS, are considered.

Observe furthermore in Fig. 5 that the SU benefits from increasing T_P and that a reduced λ_P decreases the OAD. The asymptotic values of the OAD for these three schemes may be determined when T_P tends to infinity. In Fig. 5, we focus our attention on the comparison of three different scheduling schemes. It can be seen that our proposed optimal scheme (OPT) performs better than the conventional schemes. The higher the traffic load of the PU, i.e., the higher λ_P and lower T_P, the more substantial the advantage of our optimal scheme becomes. For example, for $\lambda_P = 0.5$ packet/slot and $T_P = 3$ TS, our scheme has a 27% or 20% lower OAD than the proportional (PRO) or the round-robin (R-R) schemes, respectively, when considering the SU's queues.

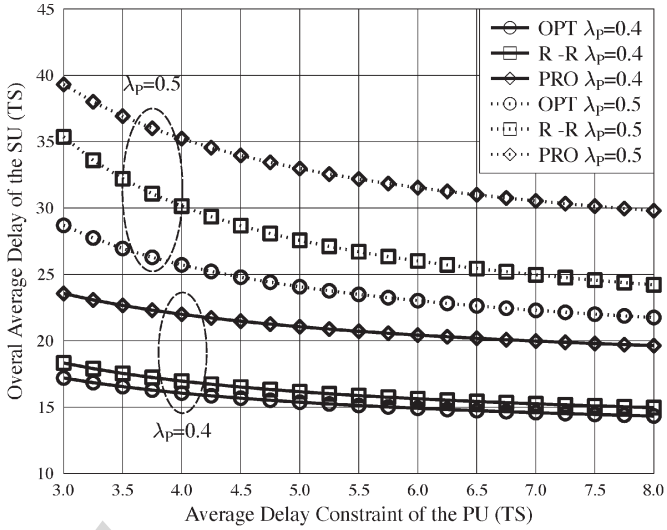

Fig. 5. Delay of the SU versus the delay tolerant T_P of the PU parameterized by the average packets arrival rate λ_P.

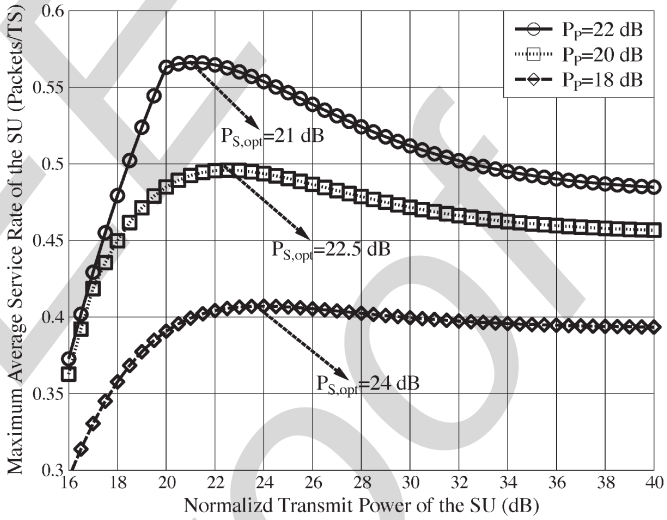

Fig. 6. Service rate versus the normalized transmit power of the SU parameterized by the normalized transmit power of the PU.

C. Impact of the SU's Normalized Transmit Power

In the PHY layer, the fading parameters are $m_P = m_S = m_{PS} = 2$, and all the associated thresholds are 8.5 dB. The normalized transmit power of the SU varies from 16 to 40 dB in the three different cases associated with the normalized transmit power of the PU, which are $P_P = 22, 20, 18$ dB. The other parameters are the same as in Table I. At the MAC layer, the average packet arrival rate is $\lambda_P = 0.5$ packet/TS, whereas the delay tolerance is $T_P = 5$ TS.

It may be observed in Fig. 6 that the MASR of the SU first increases upon increasing the normalized transmit power P_S, since increasing P_S improves the SINR performance of the SU link, which results in an improved MASR. However, beyond a specific point, the MASR is gradually reduced upon further increasing P_S. That is because a higher P_S imposes an increased interference upon the PU link, which reduces the PU's chance of emptying its own queues. As long as the PU occupies the resources, this reduces the chances of the SU to

access the system, and as a result, the MASR of the SU is reduced. Given a higher transmit power P_P of the PU, the MASR of the SU is also increased, and the optimal power P_S maximizing the MASR of the SU becomes explicit in Fig. 6.

VI. Conclusion

Our novel hybrid CR system amalgamated the interweave and underlay paradigms for the sake of enhancing the chances of the SU to access the system. The hybrid parameter ε was introduced for optimally blending the interweave and underlay paradigms, which was defined as the SU's probability of accessing the system when the PU is still transmitting. Several problems have been solved without violating the delay constraints of the PU: 1) The scenario relying on realistic imperfect sensing was considered. Under the assumption of Nakagami-m fading, we have modeled the service links by a Poisson service process. 2) The optimal hybrid parameter of ε^* was found for the MASR of the SU. 3) The most suitable queue scheduling scheme $\{p_i^*,\ i = 1, 2, \ldots, N\}$ was found for the sake of minimizing the OAD of the SU's multiple queues.

Our numerical results characterized the influence of the PHY- and MAC-layer parameters on both the delay imposed and the achievable average service rate of the SU. Several interesting observations have been made: 1) Given different traffic loads of the PU, different design strategies should be adopted for the sake of enhancing the MASR of the SU. 2) If the traffic load is high, our best strategy is to focus on improving the PU link, which is capable of attaining a more substantial overall performance gain, rather than improving the channel conditions of the SU link. If the traffic load is light, improving the SU link may achieve the best performance. Furthermore, improving the reliability of the SU's sensing scheme is capable of achieving the same overall performance gain as improving the PU link. 3) When the delay tolerance tends to infinity, the performance of the system becomes limited by the queue stability. 4) The SU's OAD relying on our optimal scheme gets up to 27% and 20% lower than that of the proportional and round-robin schemes. 5) The optimal transmit power of the SU may be found for the sake of maximizing the average service rate of the SU.

In our future work, the shadowing effect imposed by the wireless channel will also be taken into account. Furthermore, the SU's performance will be studied to select the most suitable transmit power.

References

[1] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[2] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*. Oxford, U.K.: Pergamon, 1960, pp. 3–36.

[3] L. Hanzo, P. J. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. Hoboken, NJ: Wiley, 2007.

[4] Q. Du and X. Zhang, "Statistical QoS provisionings for wireless unicast/multicast of multi-layer video streams," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.

[5] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Stable throughput of cognitive radios with and without relaying capability," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2351–2360, Dec. 2007.

[6] X. Bao, P. Martins, T. Song, and L. Shen, "Stable throughput analysis of multi-user cognitive cooperative systems," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–5.

[7] A. A. El-Sherif, A. K. Sadek, and K. J. R. Liu, "Opportunistic multiple access for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 704–714, Apr. 2011.

[8] J. Hu, L.-L. Yang, and L. Hanzo, "Appendix of the Submitted Paper "Maximum Average Service Rate and Optimal Queue Scheduling of Delay-Constrained Hybrid Cognitive Radio in Nakagami Fading Channels," Tech. Rep. [Online]. Available: http://eprints.soton.ac.uk/340414/

[9] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] P. V. Mieghem, *Performance Analysis of Communications Networks and Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[11] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted ALOHA with multipacket reception," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2636–2656, Jul. 2005.

[12] R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple-access system," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 918–930, Sep. 1988.

[13] W. Luo and A. Ephremides, "Stability of N interacting queues in random access systems," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1579–1587, Jul. 1999.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York: Academic, 2007.

[16] L.-L. Yang, *Multicarrier Communications*. Hoboken, NJ: Wiley, Jan. 2009.

Jie Hu (S'11) received the B.Eng. degree in communication engineering and the M.Eng. degree in communication and information system from Beijing University of Posts and Telecommunications, Beijing, China, in 2008 and 2011, respectively. He is currently working toward the Ph.D. degree with the Communication, Signal Processing and Control Group, University of Southampton, Southampton, U.K.

His research interests in wireless communications include cognitive radio and cognitive networks, queuing analysis, resource allocation and scheduling, ad hoc wireless networks, and mobile social networks.

Lie-Liang Yang (M'98–SM'02) received the B.Eng. degree in communications engineering from Shanghai Tiedao University, Shanghai, China, in 1988 and the M.Eng. and Ph.D. degrees in communications and electronics from Northern (Beijing) Jiaotong University, Beijing, China, in 1991 and 1997, respectively.

From June 1997 to December 1997, he was a Visiting Scientist with the Institute of Radio Engineering and Electronics, Academy of Sciences of the Czech Republic, Prague, Czech Republic. Since December 1997, he has been with the University of Southampton, Southampton, U.K., where he is currently a Professor with the School of Electronics and Computer Science. He has published more than 240 research papers in journals and conference proceedings, authored or coauthored three books, and published several book chapters. Details about his publications can be found at http://www-mobile.ecs.soton.ac.uk/lly/. His research has covered a wide range of topics in wireless communications, networking, and signal processing.

Dr. Yang is currently an Associate Editor of the IEEE Transactions on Vehicular Technology, the *Journal of Communications and Networks*, and the *Security and Communication Networks* journal.

Lajos Hanzo (M'91–SM'92–F'04) received the M.S. degree (with first-class honors) in electronics and the Ph.D. degree from the Technical University of Budapest, Budapest, Hungary, in 1976 and 1983, respectively, the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2004, and the "Doctor Honoris Causa" degree from the Technical University of Budapest in 2009.

During his 35-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. Since 2009, he has been a Chaired Professor with Tsinghua University, Beijing, China. He is currently directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry; the Engineering and Physical Sciences Research Council, U.K.; the European Information Society Technologies Program; and the Mobile Virtual Centre of Excellence, U.K. He is an enthusiastic supporter of industrial and academic liaison and offers a range of industrial courses. He has successfully supervised 80 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications totaling in excess of 10 000 pages, published more than 1250 research entries on IEEE Xplore, and presented keynote lectures. For further information on research in progress and associated publications, see http://www-mobile.ecs.soton.ac.uk/.

Dr. Hanzo is a Fellow of the Royal Academy of Engineering, U.K., a Fellow of the Institution of Electrical Engineers, and a Governor of the IEEE Vehicular Technology Society. He has been a Technical Program Committee Chair and a General Chair for IEEE conferences. During 2008–2012, he was the Editor-in-Chief of the IEEE Press. He has received a number of distinctions.

AUTHOR QUERY

NO QUERY.