On Medium Grain Scalable Video Streaming over Femtocell Cognitive Radio Networks

Donglin Hu and Shiwen Mao, *Senior Member, IEEE*

***Abstract*—Femtocells are shown highly effective on improving network coverage and capacity by bringing base stations closer to mobile users. In this paper, we investigate the problem of streaming scalable videos in femtocell cognitive radio (CR) networks. This is a challenging problem due to the stringent QoS requirements of real-time videos and the new dimensions of network dynamics and uncertainties in CR networks. We develop a framework that captures the key design issues and trade-offs with a stochastic programming problem formulation. In the case of a single FBS, we develop an optimum-achieving distributed algorithm, which is shown also optimal for the case of multiple non-interfering FBS's. In the case of interfering FBS's, we develop a greedy algorithm that can compute near-optimal solutions, and prove a closed-form lower bound on its performance. The proposed algorithms are evaluated with simulations, and are shown to outperform three alternative schemes with considerable margins.**

***Index Terms*—Cognitive radio, cross-layer optimization, femtocell, Medium Grain Scalable video, stochastic programming.**

I. Introduction

ACCORDING to a Cisco study, global mobile data traffic will increase 26-fold between 2010 and 2015 [2]. Furthermore, almost 66% of the mobile data will be video related by 2015, as driven by the proliferation of intelligent mobile devices and the compelling need for ubiquitous access to wireless multimedia content. Such drastic increase in wireless video traffic will significantly stress the capacity of existing and future wireless networks and will have far-reaching impacts on how future wireless networks are designed and operated.

Due to the use of open space as transmission medium, capacity of wireless networks are usually limited by interference. When a mobile user moves away from the base station, a considerably larger transmit power is needed to overcome attenuation, while causing interference to other users and deteriorating network capacity. To this end, femtocells provide an effective solution that brings network infrastructure closer to mobile users. A femtocell is a small (e.g., residential) cellular network, with a *femto base station* (FBS) connected to the owner's broadband wireline network [3]–[5]. The FBS serves approved users when they are within the coverage. Among the many benefits, femtocells are shown effective on improving network coverage and capacity [3]. Due to reduced distance, transmit power can be greatly reduced, leading to prolonged battery life, improved signal-to-interference-plus-noise ratio (SINR), and better spatial reuse of spectrum.

Femtocells have received significant interest from the wireless industry. Although highly promising, many important problems should be addressed to fully harvest their potential, such as interference mitigation, resource allocation, synchronization, and QoS provisioning [3], [4]. It is also critical for the success of this technology to support important applications such as real-time video streaming in femtocell networks.

In this paper, we investigate the problem of video streaming in femtocell cognitive radio (CR) networks. We consider a femtocell network consisting of a *macro base station* (MBS) and multiple FBS's. The femtocell network is co-located with a primary network with multiple licensed channels. The idea is to exploit CR and dynamic spectrum access to utilize spectrum opportunities in the licensed channels for streaming videos [6]. Femtocell subscribers (or, CR users) and FBS's sense licensed channels, and determine which channel(s) and which base station (i.e., the MBS or an FBS) to use for delivering video packets based on sensing results. The objective is to maximize the capacity of the femtocell CR network on carrying real-time video data, while bounding the interference to primary users.

This is a challenging problem due to the stringent QoS requirements of real-time videos and, on the other hand, the new dimensions of network dynamics (i.e., channel availability) and uncertainties (i.e., spectrum sensing and errors) found in CR networks. It also involves a long list of design factors that necessitates cross-layer optimization, such as spectrum sensing and errors, dynamic spectrum access, interference modeling and primary user protection, channel allocation, and video performance, among others.

We adopt Scalable Video Coding (SVC) in our system. SVC encodes a video into multiple substreams, subsets of which can be decoded to provide different quality levels for the reconstructed video [7]. Such scalability is very useful for video streaming systems, especially in CR networks, to accommodate heterogeneous channel availabilities and dynamic network conditions. We consider H.264/SVC medium grain scalable (MGS) videos, since MGS can achieve better rate-distortion performance over Fine-Granularity-Scalability (FGS), although it only has Network Abstraction Layer (NAL) unit-based granularity [7].

The unique femtocell network architecture and the scalable video allow us to develop a framework that captures the key design issues and trade-offs, and to formulate a

Manuscript received 10 March 2011; revised 1 September 2011. This work was supported in part by the US National Science Foundation under Grants CNS-0953513, ECCS-0802113, CNS-1145446 and DUE-1044021, and through the Wireless Internet Center for Advanced Technology at Auburn University. This work was presented in part at The 31st IEEE Int'l Conference on Distributed Computing Systems (ICDCS 2011) [1].

D. Hu and S. Mao are with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL.

Digital Object Identifier 10.1109/JSAC.2012.120413.

stochastic programming problem. It has been shown that the deployment of femtocells has a significant impact on the network performance [3]. In this paper, we examine three deployment scenarios. In the case of a single FBS, we apply *dual decomposition* to develop a distributed algorithm that can compute the optimal solution. In the case of multiple non-interfering FBS's, we show that the same distributed algorithm can be used to compute optimal solutions. In the case of multiple interfering FBS's, we develop a greedy algorithm that can compute near-optimal solutions, and prove a closed-form lower bound for its performance based on an *interference graph* model. The proposed algorithms are evaluated with simulations, and are shown to outperform three alternative schemes with considerable gains.

The remainder of this paper is organized as follows. The related work is discussed in Section II. The system model and preliminaries are given in Section III. We present the problem formulation and develop solution algorithms in Section IV. Simulation results are presented in Section V. Section VI concludes this paper.

II. Related Work

Femtocells have received considerable interest from both industry and academia. Comprehensive overviews of technical challenges, requirements, and some preliminary solutions to femtocell networks can be found in [3]–[5]. Since femtocells can use the same channels as conventional cellular networks, considerable research efforts were focused on interference analysis and mitigation [8]–[19]. In [8], a distributed utility-based SINR adaptation scheme at femtocells was proposed to alleviate cross-tier interference at the macrocell from co-channel femtocells. In [9], the authors proposed a fractional frequency reuse scheme to mitigate inter-femtocell interference. Co-channel interference analysis and cancellation were considered in [10]. In [11], partial GSM spectrum was reused by femtocells to mitigate the negative interference impact upon the macrocell. A solution was presented for frequency allocation and power configuration within the femtocell framework. In a recent work [12], a decentralized Q-learning approach was presented for interference management. It was shown that the Q-learning scheme yields significant gains in terms of coverage speed and precision.

Deploying femtocells by underlaying the macrocell has been proved to significantly improve indoor coverage and system capacity. However, interference mitigation in a two-tier heterogeneous network is a challenging problem. In [13], the interference from macrocell and femtocells was mitigated by a spatial channel separation scheme with codeword-to-channel mapping. In [14], the rate distribution in the macrocell was improved by subband partitioning and modest gains were achieved by interference cancellation. In [15], the interference was controlled by denying the access of femtocell base stations to protect the transmission of nearby macro base station. A novel algorithmic framework was presented in [16] for dynamic interference management to deliver QoS, fairness and high system efficiency in LTE-A femtocell networks. Requiring no modification of existing macrocells, CR was shown to achieve considerable performance improvement when applied to interference mitigation [17]. In [18], the orthogonal time-frequency blocks and transmission opportunities were allocated based on a safe/victim classification.

The high potential of CRs has attracted significant interest from the wireless community [6], [20]. However, the problem of video over CR networks has been addressed only in a few papers [21]–[25]. In [21], the authors presented a dynamic channel selection scheme for CR users to transmit videos over multiple channels. In [22], Ali and Yu considered transmitting a video stream over a CR link and provided a partially observable Markov decision process (POMDP) formulation and solution procedure. In [23], a distributed joint routing and spectrum sharing algorithm for video streaming applications over CR ad hoc networks was proposed and evaluated with simulations. In our prior work [24], [25], we considered scalable video streaming in an infrastructure-based CR network and a multi-hop CR network, respectively. We developed cross-layer formulation and effective algorithms for scheduling video data with proven optimality bounds.

III. System Model and Preliminaries

A. Spectrum and Network Model

We consider a spectrum consisting of $(M+1)$ channels, including one common, unlicensed channel (indexed as channel 0) and M licensed channels (indexed as channels 1 to M). The M licensed channels are allocated to a primary network, and the common channel is exclusively used by all CR users. We assume all the channels follow a synchronized time slot structure [6]. The capacity of each licensed channel is B_1 Mbps, while the capacity of the common channel is B_0 Mbps.

The channel states evolve independently, while the occupancy of each licensed channel follows a two-state discrete-time Markov process, as validated by several measurement studies [26], [27] and used in prior work [6], [28]. The network status in time slot t is denoted as $\vec{S}(t) = [S_1(t), S_2(t), \cdots, S_M(t)]$, where $S_m(t)$ is the status of channel m with either idle (when $S_m(t) = 0$) or busy (when $S_m(t) = 1$) states. Let P_m^{01} and P_m^{10} be the transition probability from state 0 to 1 and that from state 1 to 0 for channel m, respectively. The utilization of channel m with respect to primary user transmissions can be written as:

$$\eta_m = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} S_m(t) = \frac{P_m^{01}}{P_m^{01} + P_m^{10}}. \quad (1)$$

The femtocell CR network is illustrated in Fig. 1. There is an MBS and N FBS's deployed in the area to serve CR users. The N FBS's are connected to the MBS (and the Internet) via broadband wireline connections. Due to advances in antenna technology, it is possible to equip multiple antennas at the base stations. The MBS has one antenna that is always tuned to the common channel. Each FBS is equipped with multiple antennas (e.g., M) and is able to sense multiple licensed channels at the beginning of each time slot. There are K_i CR users in femtocell i, $i = 1, 2, \cdots, N$, and $\sum_{i=1}^{N} K_i = K$. Each CR user has a software radio transceiver, which can be tuned to any of the M+1 channels. A CR user will either connect to a nearby FBS using one or more of the licensed channels or to the MBS via the common channel.

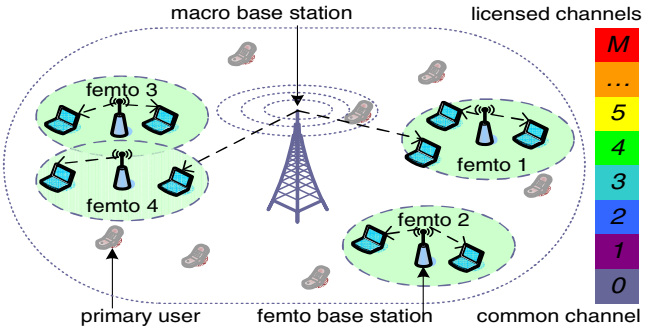

Fig. 1. A femtocell CR network with one MBS and four FBS's.

Although the CR users are mobile, we assume constant topology during a time slot. If the topology is changed during a time slot, the video transmission will only be interrupted for the time slot, since the proposed algorithms are executed in every time slot for new channel assignment and schedule.

B. Spectrum Sensing

The femtocell CR network is within the coverage of the infrastructure-based primary network. Both FBS's and CR users sense the channels to identify spectrum opportunities in each time slot. Each time slot consists of (i) a *sensing phase*, when CR users and FBS's sense licensed channels, (ii) a *transmission phase*, when CR users and FBS's attempt to access licensed channels, and (iii) an *acknowledgment phase*, when acknowledgments (ACK) are returned to the source.

We adopt a *hypothesis test* to detect channel availability. The null and alternative hypothesis are: $H_0^m := \{\text{channel } m \text{ is idle}\}$ and $H_1^m := \{\text{channel } m \text{ is busy}\}$. Two kinds of detection errors may occur: (i) *false alarm*, an idle channel is considered busy and a spectrum opportunity will be wasted; (ii) *miss detection*, a busy channel is considered idle, which may lead to collision with primary users. Let ϵ_i^m and δ_i^m be the false alarm and miss detection probabilities of the i-th sensing result Θ_i^m, respectively. We then have

$$\Pr\{\Theta_i^m = 1 | H_0^m\} = \epsilon_i^m \quad \text{and} \quad \Pr\{\Theta_i^m = 0 | H_1^m\} = \delta_i^m.$$

We assume ϵ_i^m and δ_i^m are known for all m, since these are the basic parameters when designing a spectrum sensor.

We assume that each CR user chooses one channel to sense in each time slot, since it only has one transceiver. The sensing results will be shared among CR users and FBS's via the common channel in the sensing phase. Given L sensing results on channel m, the conditional probability $P_m^A(\Theta_1^m, \cdots, \Theta_L^m)$ with which channel m is available is [1]

$$\begin{aligned} P_m^A(\vec{\Theta}_m) &= P_m^A(\Theta_1^m, \cdots, \Theta_L^m) = \Pr\{H_0^m | \Theta_1^m, \cdots, \Theta_L^m\} \\ &= \left[1 + \frac{\eta_m}{1-\eta_m} \prod_{i=1}^{L} \frac{(\delta_i^m)^{1-\Theta_i^m}(1-\delta_i^m)^{\Theta_i^m}}{(\epsilon_i^m)^{\Theta_i^m}(1-\epsilon_i^m)^{1-\Theta_i^m}}\right]^{-1}. \end{aligned} \quad (2)$$

The availability of channel m, i.e., $P_m^A(\Theta_1^m, \cdots, \Theta_L^m)$, can be computed iteratively by decomposing (2) as follows.

$$P_m^A(\Theta_1^m) = \left[1 + \frac{\eta_m}{1-\eta_m} \times \frac{(\delta_1^m)^{1-\Theta_1^m}(1-\delta_1^m)^{\Theta_1^m}}{(\epsilon_1^m)^{\Theta_1^m}(1-\epsilon_1^m)^{1-\Theta_1^m}}\right]^{-1} \quad (3)$$

$$\begin{aligned} &P_m^A(\Theta_1^m, \Theta_2^m, \cdots, \Theta_l^m) \\ &= \left\{1 + \left[\frac{1}{P_m^A(\Theta_1^m, \Theta_2^m, \cdots, \Theta_{l-1}^m)} - 1\right] \times \right. \\ &\left. \frac{(\delta_l^m)^{1-\Theta_l^m}(1-\delta_l^m)^{\Theta_l^m}}{(\epsilon_l^m)^{\Theta_l^m}(1-\epsilon_l^m)^{1-\Theta_l^m}}\right\}^{-1}, l = 2, \cdots, L. \end{aligned} \quad (4)$$

C. Opportunistic Channel Access

Let $D_m(t)$ be a decision variable for accessing channel m in time slot t. It is defined as

$$D_m(t) = \begin{cases} 0, & \text{if channel } m \text{ is considered to be idle} \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

We adopt a probabilistic approach: based on sensing results $\vec{\Theta}_m$, we have $D_m(t) = 0$ with probability $P_m^D(\vec{\Theta}_m)$ and $D_m(t) = 1$ with probability $1 - P_m^D(\vec{\Theta}_m)$. The probability $P_m^D(\vec{\Theta}_m)$ is determined as follows.

For primary user protection, the collision probability with primary users caused by CR users should be bounded. Let γ_m be the maximum allowable collision probability with primary users on channel m. We have the following condition:

$$\left[1 - P_m^A(\vec{\Theta}_m)\right] P_m^D(\vec{\Theta}_m) \leq \gamma_m. \quad (6)$$

For CR network throughput performance, $P_m^D(\vec{\Theta}_m)$ should be set as large as possible. Since it is between 0 and 1, we have

$$P_m^D(\vec{\Theta}_m) = \min\left\{\gamma_m / \left[1 - P_m^A(\vec{\Theta}_m)\right], 1\right\}. \quad (7)$$

Let $\mathcal{A}(t) := \{m | D_m(t) = 0\}$ be the set of available channels in time slot t. Then $G^t = \sum_{m \in \mathcal{A}(t)} P_m^A(\Theta_1^m)$ is the expected number of available channels. These channels will be accessed in the transmission phase of time slot t.

D. Channel Model

Without loss of generality, we consider independent block fading channels that is widely used in prior work [29]. The channel fading-gain process is piecewise constant on blocks of one time slot, and fading in different time slots are independent. Let $f_X^{i,j}(x)$ denote the *probability density function* of the received SINR X from a base station i at CR user j. We assume the packet can be successfully decoded if the received SINR exceeds a threshold H. The packet loss probability from base station i to CR user j is

$$P_{i,j} = \Pr\{X \leq H\} = \int_0^H f_X^{i,j}(x)dx = F_X^{i,j}(H), \quad (8)$$

where $F_X^{i,j}(H)$ is the cumulative density function of X.

In the case of correlated fading channels, which can be modeled as finite state Markov Chains [30], the packet loss probability in the next time slot can be estimated from the known state of the previous time slot and the transition probabilities. If the packet is successfully decoded, the CR user returns an ACK to the base station in the ACK phase. We assume ACKs are always successfully delivered.

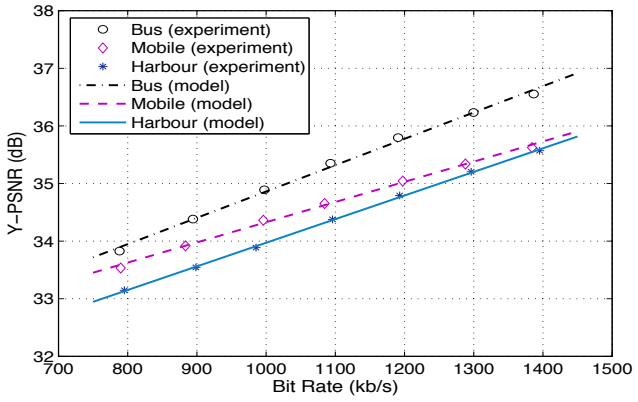

Fig. 2. Rate-distortion curves of three H.264/SVC MGS videos.

E. Video Performance Measure

We assume each active CR user receives a real-time video stream from either the MSB or an FSB. Without loss of generality, we adopt the MGS option of H.264/SVC, for scalability to accommodate the high variability of network bandwidth in CR networks.

Due to real-time constraint, each Group of Pictures (GOP) of a video stream must be delivered in the next T time slots. With MGS, enhancement layer NAL units can be discarded from a quality scalable bit stream, and thus packet-based quality scalable coding is provided. Our approach is to encode the video according to the maximum rate the channels can support. During transmission, only part of the MGS video gets transmitted as allowed by the current available channel bandwidth. The video packets are transmitted in decreasing order of their significance in decoding. When a truncated MGS video is received and decoded, the PSNR is computed by substituting the effective rate of the received MGS video into (9) given below, thus the original video is not required.

Without loss of generality, we assume that the last wireless hop is the bottleneck; video data is available at the MBS and FBS's when they are scheduled to be transmitted. The quality of reconstructed MGS video can be modeled as [7]:

$$W(R) = \alpha + \beta \times R, \tag{9}$$

where $W(R)$ is the average peak signal-to-noise ratio (PSNR) of the reconstructed video, R is the received data rate, α and β are constants depending on the video sequence and codec.

We verified (9) using an H.264/SVC codec and the *Bus*, *Mobile*, and *Harbour* test sequences. In Fig. 2, the markers are obtained by truncating the encoded video's enhancement layer at different positions to obtain different effective rates, while the curves are computed using (9). The curves fit well with measurements for the three sequences. It is worth noting that PSNR may not be a good measure of video quality as compared with alternative metrics such as MS-SSIM [31]. The main reason for choosing PSNR is that there is a closed-form model relating it to network level metrics–video rate. With the closed-form model, we can have a mathematical formulation of the scheduling/resource allocation problem, and derive effective algorithms. Should such closed-form models be available for MS-SSIM, it is possible to incorporate it into the optimization framework as well.

IV. MGS Video over Femtocell CR Networks

In this section, we address the problem of resource allocation for MGS videos over femtocell CR networks. We first examine the case of a single FBS, and then the more general case of multiple non-interfering or interfering FBS's. The algorithms for the single and non-interfering FBS cases are distributed ones and optimal. The algorithm for the interfering FBS case is a centralized one that can be executed at the MBS. To simplify notation, we omit the time slot index t for most of the variables in this Section. For example, x represents a variable for time slot t, x^- represents the variable in time slot $(t-1)$, and x^+ represents the variable in time slot $(t+1)$.

A. Case of Single FBS

1) Formulation: We first consider the case of a single FBS in the CR network, where the FBS can use all the G available channels to stream videos to K active CR users. Let w_j be the PSNR of CR user j at the beginning of time slot t and W_j the PSNR of CR user j at the end of time slot t. In time slot t, w_j is already known; W_j is a random variable that depends on channel condition and primary user activity; and w_j^+ is a *realization* of W_j. Let $\xi_{0,j}$ and $\xi_{1,j}$ indicate the random packet losses from the MBS and FBS, respectively, to CR user j in time slot t. That is, $\xi_{i,j}$ is 1 with probability $\bar{P}_{i,j} = 1 - P_{i,j}$ and 0 with probability $P_{i,j}$. Due to block fading channels, $P_{i,j}$'s do not change within the time slot.

Let $\rho_{0,j}$ and $\rho_{1,j}$ be the portions of time slot t when CR user j receives video data from the MBS and FBS, respectively. The average PSNR is computed every T time slots. We first have $W_j(0) = \alpha_j$, when $t = 0$. In each time slot t, the CR user receives $\xi_{0,j}\rho_{0,j}B_0$ bits through the MBS, and $\xi_{1,j}\rho_{1,j}GB_1$ bits through the FBS (assuming that OFDM is used), which contribute an increase of $\beta(\xi_{0,j}\rho_{0,j}B_0 + \xi_{1,j}\rho_{1,j}GB_1)/T$ to the total PSNR in this T time slot interval, according to (9). Therefore we have the following recursive relationship: $W_j = W_j^- + \beta(\xi_{0,j}\rho_{0,j}B_0 + \xi_{1,j}\rho_{1,j}GB_1)/T = W_j^- + \xi_{0,j}\rho_{0,j}R_{0,j} + \xi_{1,j}\rho_{1,j}GR_{1,j}$, where $R_{0,j} = \beta B_0/T$ and $R_{1,j} = \beta B_1/T$.

For proportional fairness, we aim to maximize the sum of the logarithms of the PSNRs of all CR users [32]. We formulate a *multistage stochastic programming problem* by maximizing the *expectation* of the logarithm-sum at time T.

$$\begin{aligned} \text{maximize: } & \textstyle\sum_{j=1}^{K} \mathbb{E}[\log(W_j(T))] \\ \text{subject to: } & W_j = W_j^- + \xi_{0,j}\rho_{0,j}R_{0,j} + \xi_{1,j}\rho_{1,j}GR_{1,j}, \\ & \qquad j = 1, \cdots, K,\ t = 1, \cdots, T \\ & \textstyle\sum_{j=1}^{K} \rho_{i,j} \le 1,\ \ i = 0, 1,\ t = 1, \cdots, T \\ & \rho_{i,j} \ge 0,\ i = 0, 1,\ j = 1, \cdots, K,\ t = 1, \cdots, T. \end{aligned} \tag{10}$$

$R_{0,j} = \beta_j B_0/T$ and $R_{1,j} = \beta_j B_1/T$ are constants for the j-th MGS video.

At the beginning of the last time slot T, a realization $\boldsymbol{\xi}_{[T-1]} = [\vec{\xi}_1, \vec{\xi}_2, \cdots, \vec{\xi}_{T-1}]$ is known, where $\vec{\xi}_t = [\xi_{0,1}^t, \xi_{0,2}^t, \cdots, \xi_{0,K}^t, \xi_{1,1}^t, \cdots, \xi_{1,K}^t]$, $t = 1, 2, \cdots, T-1$. It can be shown that the multistage stochastic programming problem (10) can be decomposed into T serial sub-problems, each to

be solved in a time slot t [24].

$$\begin{aligned} \text{maximize: } & \textstyle\sum_{j=1}^{K} \mathbb{E}\{\log(W_j)|\boldsymbol{\xi}_{[t-1]}\} \qquad (11)\\ \text{subject to: } & W_j = W_j^- + \xi_{0,j}\rho_{0,j}R_{0,j} + \xi_{1,j}\rho_{1,j}GR_{1,j}, \\ & \qquad\qquad\qquad\qquad j = 1, \cdots, K \\ & \textstyle\sum_{j=1}^{K} \rho_{i,j} \le 1, \;\; i = 0, 1 \\ & \rho_{i,j} \ge 0, \; i = 0, 1, \; j = 1, \cdots, K, \end{aligned}$$

where $\mathbb{E}\{\log(W_j)|\boldsymbol{\xi}_{[t-1]}\}$ denotes the *conditional expectation* of $\log(W_j)$ given realization $\boldsymbol{\xi}_{[t-1]}$. W_j^- is known given the realization. When $t = 1$, the conditional expectation becomes an unconditional expectation.

Since a CR user has only one transceiver, it can operate on either one or more licensed channels (i.e., connecting to the FBS) or the common channel (i.e., connecting to the MBS), but not both simultaneously. Assume CR user j operates on the common channel with probability p_j and one or more licensed channels with probability q_j. We then rewrite problem (11) as

$$\begin{aligned} \text{maximize: } & \textstyle\sum_{j=1}^{K} \big[p_j \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}R_{0,j}) + \qquad (12)\\ & \qquad\qquad q_j \bar{P}_{1,j} \log(W_j^- + \rho_{1,j}GR_{1,j})\big] \\ \text{subject to: } & \textstyle\sum_{j=1}^{K} \rho_{i,j} \le 1, \;\; i = 0, 1 \\ & p_j + q_j = 1, \;\; j = 1, \cdots, K \\ & \rho_{i,j}, \; p_j, \; q_j \ge 0, \;\; i = 0, 1, \; j = 1, \cdots, K. \end{aligned}$$

2) Properties: In this section, we analyze the formulated problem (12) and derive its properties. We have Lemmas 1, 2, and 3 and Theorem 1 and provide the proofs in the following.

Lemma 1: Problem (12) is a convex optimization problem.

Proof: First, it can be shown that the single term $p_j \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}R_{0,j}) + q_j \bar{P}_{1,j} \log(W_j^- + \rho_{1,j}GR_{1,j})$ is a concave function, because its *Hessian matrix* is negative semi-definite. Then, the objective function is concave since the sum of concave functions is also concave. Finally, all the constraints are linear. We conclude that problem (12) is convex with a unique optimal solution. ■

Lemma 2: If $[\rho, p, q]$ is a feasible solution to problem (12), then $[\rho, q, p]$ is also feasible.

Proof: Since $[\rho, p, q]$ is feasible, we have $p + q = 1$. Switching the two probabilities, we still have $q + p = 1$. Therefore, the derived new solution is also feasible. ■

Lemma 3: Let the optimal solution be $[\rho^*, p^*, q^*]$. If $p_j^* \ge q_j^*$, then $\bar{P}_{0,j} \log(W_j^- + \rho_{0,j}^* R_{0,j})$ is greater than or equal to $\bar{P}_{1,j} \log(W_j^- + \rho_{1,j}^* GR_{1,j})$. And vice versa.

Proof: Assume $\bar{P}_{0,j} \log(W_j^- + \rho_{0,j}^* R_{0,j})$ is less than $\bar{P}_{1,j} \log(W_j^- + \rho_{1,j}^* GR_{1,j})$. Since $p_j^* \ge q_j^*$, the sum $p_j^* \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}^* R_{0,j}) + q_j^* \bar{P}_{1,j} \log(W_j^- + \rho_{1,j}^* GR_{1,j})$ is smaller than the sum $q_j^* \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}^* R_{0,j}) + p_j^* \bar{P}_{1,j} \log(W_j^- + \rho_{1,j}^* GR_{1,j})$. Thus we can obtain an objective value larger than the optimum by switching the values of p_j^* and q_j^*, which is still feasible according to Lemma 2. This conflicts with the assumption that $[\rho^*, p^*, q^*]$ is optimal. The reverse statement can be proved similarly. ■

Theorem 1: Let the optimal solution be $[\rho^*, p^*, q^*]$. If $p_j^* > q_j^*$, then we have $p_j^* = 1$ and $q_j^* = 0$. Otherwise, we have $p_j^* = 0$ and $q_j^* = 1$.

Proof: If $p_j^* > q_j^*$, we have $\bar{P}_{0,j} \log(W_j^- + \rho_{0,j}^* R_{0,j}) \ge \bar{P}_{1,j} \log(W_j^- + \rho_{1,j}^* GR_{1,j})$ according to Lemma 3. Since the objective function is linear with respect to p_j and q_j, the optimal value can be achieved by setting p_j to its maximum value 1 and q_j to its minimum value 0. The reverse statement can be proved similarly. ■

According to Theorem 1, a CR user is connected to either the MBS or the FBS for the *entire* duration of a time slot in the optimal solution. That is, it does not switch between base stations during a time slot under optimal scheduling.

3) Distributed Solution Algorithm: To solve problem (12), we define non-negative *dual variables* $\lambda = [\lambda_0, \lambda_1]$ for the two inequality constraints. The *Lagrangian function* is

$$\begin{aligned} \mathcal{L}(p, \rho, \lambda) &= \textstyle\sum_{j=1}^{K} \big[p_j \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}R_{0,j}) + \\ &\quad (1 - p_j)\bar{P}_{1,j} \log(W_j^- + \rho_{1,j}GR_{1,j})\big] + \\ &\quad \lambda_0(1 - \textstyle\sum_{j=1}^{K} \rho_{0,j}) + \lambda_1(1 - \textstyle\sum_{j=1}^{K} \rho_{1,j}) \\ &= \textstyle\sum_{j=1}^{K} \mathcal{L}_j(p_j, \rho_{0,j}, \rho_{1,j}, \lambda_0, \lambda_1) + \lambda_0 + \lambda_1, \quad (13) \end{aligned}$$

where

$$\begin{aligned} \mathcal{L}_j(p_j, \rho_{0,j}, \rho_{1,j}, \lambda_0, \lambda_1) &= p_j \bar{P}_{0,j} \log(W_j^- + \rho_{0,j}R_{0,j}) + \\ & (1 - p_j)\bar{P}_{1,j} \log(W_j^- + \rho_{1,j}GR_{1,j}) - \lambda_0\rho_{0,j} - \lambda_1\rho_{1,j}. \end{aligned}$$

The corresponding problem can be decomposed into K sub-problems and solved iteratively. In Step $\tau \ge 1$, for given $\lambda_0(\tau)$ and $\lambda_1(\tau)$ values, each CR user j solves the following sub-problem using local information.

$$\begin{aligned} & [p_j^*(\tau), \rho_{0,j}^*(\tau), \rho_{1,j}^*(\tau)] \\ &= \underset{p_j, \rho_{0,j}, \rho_{1,j} \ge 0}{\arg\max} \; \mathcal{L}_j(p_j, \rho_{0,j}, \rho_{1,j}, \lambda_0(\tau), \lambda_1(\tau)). \qquad (14) \end{aligned}$$

There is a unique optimal solution since the objective function in (14) is concave. The CR users then exchange their solutions. The *master dual problem*, for given $p(\tau)$ and $\rho(\tau)$, is:

$$\begin{aligned} & \min_{\lambda \ge 0} \mathcal{L}(p(\tau), \rho(\tau), \lambda) \\ &= \textstyle\sum_{j=1}^{K} \mathcal{L}_j(p_j(\tau), \rho_{0,j}(\tau), \rho_{1,j}(\tau), \lambda_0, \lambda_1) + \lambda_0 + \lambda_1. \quad (15) \end{aligned}$$

Since the Lagrangian function is differentiable, the *gradient iteration* approach can be used.

$$\lambda_i(\tau + 1) = \Big[\lambda_i(\tau) - s \times \Big(1 - \textstyle\sum_{j=1}^{K} \rho_{i,j}^*(\tau)\Big)\Big]^+, \; i = 0, 1, \quad (16)$$

where s is a sufficiently small positive *step size* and $[\cdot]^+$ denotes the projection onto the nonnegative axis. The updated $\lambda_i(\tau+1)$ will again be used to solve the sub-problems, and so forth. Since the problem is convex, we have *strong duality*; the *duality gap* between the primal and dual problems is zero. The dual variables $\lambda(\tau)$ will converge to the optimal values as τ goes to infinity. Since the optimal solution to (14) is unique, the primal variables $p(\tau)$ and $\rho_{i,j}(\tau)$ will also converge to their optimal values when τ is sufficiently large.

The distributed solution procedure is presented in Table I. In the table, Steps 3–8 solve the sub-problem in (14); Step 9 updates the dual variables. The threshold ϕ is a prescribed small value with $0 \le \phi \ll 1$. The algorithm terminates when the dual variables are sufficiently close to the optimal values.

TABLE I
ALGORITHM FOR THE CASE OF SINGLE FBS

1: Set $\tau = 0$, $\lambda_0(0)$ and $\lambda_1(0)$ to some nonnegative value;
2: DO % (each CR user j executes Steps 3–8)
3: $\rho_{0,j}(\tau)=\left[\frac{\bar{P}_{0,j}}{\lambda_0(\tau)}-\frac{W_j^-}{R_{0,j}}\right]^+$, $\rho_{1,j}(\tau)=\left[\frac{\bar{P}_{1,j}}{\lambda_1(\tau)}-\frac{W_j^-}{R_{1,j}G}\right]^+$;
4: IF $\left[\bar{P}_{0,j}\log(W_j^- + \rho_{0,j}(\tau)R_{0,j}) - \lambda_0(\tau)\rho_{0,j}(\tau)\right] > \left[\bar{P}_{1,j}\log(W_j^- + \rho_{1,j}(\tau)GR_{1,j}) - \lambda_1(\tau)\rho_{1,j}(\tau)\right]$
5: Set $p_j(\tau) = 1$ and $\rho_{1,j}(\tau) = 0$;
6: ELSE
7: Set $p_j(\tau) = 0$ and $\rho_{0,j}(\tau) = 0$;
8: END IF
9: MBS updates $\lambda_i(\tau + 1)$ as in (16);
10: $\tau = \tau + 1$;
11: WHILE $\left(\sum_{i=0}^{1}(\lambda_i(\tau + 1) - \lambda_i(\tau))^2 > \phi\right)$

TABLE II
ALGORITHM FOR THE CASE OF MULTIPLE NON-INTERFERING FBS'S

1: Set $\tau = 0$, and $\lambda_0(0)$ and $\lambda_i(0)$ to some nonnegative values, for all i;
2: DO % (each CR user j executes Steps 3–8)
3: $\rho_{0,j}(\tau)=\left[\frac{\bar{P}_{0,j}}{\lambda_0(\tau)}-\frac{W_j^-}{R_{0,j}}\right]^+$, $\rho_{i,j}(\tau)=\left[\frac{\bar{P}_{i,j}}{\lambda_i(\tau)}-\frac{W_j^-}{R_{i,j}G}\right]^+$;
4: IF $\left[\bar{P}_{0,j}\log(W_j^- + \rho_{0,j}(\tau)R_{0,j}) - \lambda_0(\tau)\rho_{0,j}(\tau)\right] > \left[\bar{P}_{i,j}\log(W_j^- + \rho_{i,j}(\tau)GR_{i,j}) - \lambda_i(\tau)\rho_{i,j}(\tau)\right]$
5: Set $p_j(\tau) = 1$ and $\rho_{i,j}(\tau) = 0$;
6: ELSE
7: Set $p_j(\tau) = 0$ and $\rho_{0,j}(\tau) = 0$;
8: END IF
9: MBS updates $\lambda_i(\tau + 1)$ as in (18) and (19);
10: $\tau = \tau + 1$;
11: WHILE $\left(\sum_{i=0}^{N}(\lambda_i(\tau + 1) - \lambda_i(\tau))^2 > \phi\right)$

B. Case of Multiple Non-interfering FBS's

We next consider the case of $N > 1$ non-interfering FBS's. The coverages of the FBS's do not overlap with each other, as FBS 1 and 2 in Fig. 1. Consequently, each FBS can use all the available licensed channels without interfering other FBS's. Assume each CR user knows the nearest FBS and is associate with it. Let $\mathcal{U}_i$ denote the set of CR users associated with FBS i. The resource allocation problem becomes:

$$\begin{aligned}\text{maximize: } & \sum_{j=1}^{K} p_j\bar{P}_{0,j}\log(W_j^- + \rho_{0,j}R_{0,j}) + \\ & \sum_{i=1}^{N}\sum_{j\in\mathcal{U}_i} q_j\bar{P}_{i,j}\log(W_j^- + \rho_{i,j}GR_{i,j}) \qquad (17)\\ \text{subject to: } & \textstyle\sum_{j=1}^{K}\rho_{0,j} \le 1 \\ & \textstyle\sum_{j\in\mathcal{U}_i}\rho_{i,j} \le 1,\ \ i = 1,\cdots,N \\ & p_j + q_j = 1,\ \ j = 1,\cdots,K \\ & \rho_{i,j},\ p_j,\ q_j \ge 0,\ \ i = 1,\cdots,N,\ j = 1,\cdots,K.\end{aligned}$$

Since all the available channels can be allocated to each FBS with spatial reuse, problem (17) can be solved using the algorithm in Table I with some modified notation: $\rho_{1,j}(\tau)$ now becomes $\rho_{i,j}(\tau)$ and $\lambda_1(\tau)$ becomes $\lambda_i(\tau)$, $i = 1,\cdots,N$. The dual variables are iteratively updated as

$$\lambda_0(\tau+1) = \left[\lambda_0(\tau) - s\times\left(1 - \textstyle\sum_{j=1}^{K}\rho_{0,j}^*(\tau)\right)\right]^+ \qquad (18)$$

$$\lambda_i(\tau+1) = \left[\lambda_i(\tau) - s\times\left(1 - \textstyle\sum_{j\in\mathcal{U}_i}\rho_{i,j}^*(\tau)\right)\right]^+, \quad i = 1,\cdots,N. \qquad (19)$$

The modified solution algorithm is presented in Table II. As in the case of single FBS, the algorithm is jointly executed by the CR users and MBS, by iteratively updating the dual variables $\lambda_0(\tau)$ and $\lambda_i(\tau)$'s, and the resource allocations $\rho_{0,j}^*(\tau)$ and $\rho_{i,j}^*(\tau)$'s. It can be shown that the distributed algorithm can produce the optimal solution for problem (17).

C. Case of Multiple Interfering FBS's

1) Formulation: Finally, we consider the case of multiple interfering FBS's. Assume that the coverages of some FBS's overlap with each other, as FBS 3 and 4 in Fig. 1. They cannot use the same channel simultaneously, but have to compete for the available channels in the transmission phase. Define *channel allocation variables* $c_{i,m}$ for time slot t as:

$$c_{i,m} = \begin{cases} 1, & \text{if channel } m \text{ is allocated to FBS } i \\ 0, & \text{otherwise.}\end{cases} \qquad (20)$$

Given an allocation, the expected number of available channels for FBS i is $G_i = \sum_{m\in\mathcal{A}(t)} c_{i,m}P_m^A$.

We use *interference graph* to model the case of overlapping coverages, which is defined below.

Definition 1: An *interference graph* $G_I = (V_I, E_I)$ is an undirected graph where each vertex represents an FBS and each edge indicates interference between the two end FBS's.

For the example network given in Fig. 1, we can derive an interference graph as shown in Fig. 3. FBS 3 and 4 cannot use the same channel simultaneously, as summarized in the following lemma.

Lemma 4: If channel m is allocated to FBS i, the neighboring vertices of FBS i in the interference graph G_I, denoted as $\mathcal{R}(i)$, cannot use the same channel m simultaneously.

Further define index variables d_i^k as

$$d_i^k = \begin{cases} 1, & \text{if FBS } i \text{ is an endpoint of link } k \in G_I \\ 0, & \text{otherwise.}\end{cases} \qquad (21)$$

The interference constraint can be described as $\sum_{i=1}^{N} d_i^k c_{i,m} \le 1$, for $m = 0,\cdots,M$, and for all link $k \in G_I$. We then have the following problem formulation.

$$\begin{aligned}\text{maximize: } & \sum_{j=1}^{K} p_j\bar{P}_{0,j}\log(W_j^- + \rho_{0,j}R_{0,j}) + \qquad (22)\\ & \sum_{i=1}^{N}\sum_{j\in\mathcal{U}_i} q_j\bar{P}_{i,j}\log(W_j^- + \rho_{i,j}G_iR_{i,j}) \\ \text{subject to: } & \textstyle\sum_{j=1}^{K}\rho_{0,j} \le 1 \\ & \textstyle\sum_{j\in\mathcal{U}_i}\rho_{i,j} \le 1,\ \ i = 1,\cdots,N \\ & p_j + q_j = 1,\ \ j = 1,\cdots,K \\ & G_i = \textstyle\sum_{m\in\mathcal{A}(t)} c_{i,m}P_m^A,\ \ i = 1,\cdots,N\end{aligned}$$

Fig. 3. Interference graph for the femtocell CR network shown in Fig. 1.

TABLE III
CHANNEL ALLOCATION ALGORITHM FOR CASE OF INTERFERING FBS'S

1:	Initialize $\boldsymbol{c}$ to a zero matrix, FBS set $\mathcal{N} = \{1, \cdots, N\}$, and FBS-channel set $\mathcal{C} = \mathcal{N} \times \mathcal{A}(t)$;
2:	WHILE ($\mathcal{C}$ is not empty)
3:	Find FBS-channel pair $\{i', m'\}$, such that $\{i', m'\} = \arg\max_{\{i,m\}\in\mathcal{C}} \{Q(\boldsymbol{c} + \boldsymbol{e}_{i,m}) - Q(\boldsymbol{c})\}$;
4:	Set $\boldsymbol{c} = \boldsymbol{c} + \boldsymbol{e}_{i',m'}$;
5:	Remove $\{i', m'\}$ from $\mathcal{C}$;
6:	Remove $\mathcal{R}(i') \times m'$ from $\mathcal{C}$;
7:	END WHILE

$$\sum_{i=1}^{N} d_i^k c_{i,m} \leq 1, m = 0, \cdots, M, \text{for link } k \in G_I,$$
$$\rho_{i,j}, p_j, q_j, c_{i,m} \geq 0,$$
$$i = 1, \cdots, N,\ j = 1, \cdots, K,\ m = 0, \cdots, M.$$

2) Solution Algorithm : The optimal solution to problem (22) depends on the channel allocation variables $c_{i,m}$. Problem (22) can be solved with the algorithm in Table II if the $c_{i,m}$'s are known. Let $Q(\boldsymbol{c})$ be the suboptimal objective value for a given channel allocation $\boldsymbol{c}$, where $\boldsymbol{c} = [\vec{c}_1, \vec{c}_2, \cdots, \vec{c}_N]$ and $\vec{c}_i$ is a vector of elements $c_{i,m}$, for FBS i and channels $m \in \mathcal{A}(t)$. If all the FBS's are disjointedly distributed with no overlap, each FBS can use all the available channels. We have $c_{i,m} = 1$ for all i and $m \in \mathcal{A}(t)$, i.e., it is reduced to the case in Section IV-B.

To solve problem (22), we first apply a *greedy algorithm* to allocate the available channels in $\mathcal{A}(t)$ to the FBS's (i.e., to determine $\boldsymbol{c}$). We then apply the algorithm in Table II with the computed $\boldsymbol{c}$ to obtain a near-optimal solution. Let $\boldsymbol{e}_{i,m}$ be a matrix with 1 at position $\{i, m\}$ and 0 at all other positions, representing the allocation of channel $m \in \mathcal{A}(t)$ to FBS i. The greedy channel allocation algorithm is given in Table III, where the FBS-channel pair that can achieve the largest increase in $Q(\cdot)$ is chosen in each iteration. The worst case complexity of the greedy algorithm is $O(N^2M^2)$.

3) Performance Lower Bound: We next present a lower bound for the greedy algorithm. Let $e(l)$ be the l-th FBS-channel pair chosen in the greedy algorithm, and π_l denote the sequence $\{e(1), e(2), \cdots, e(l)\}$. The increase in object value (22) due to the l-th allocated FBS-channel pair is denoted as

$$\Delta_l := \Delta(\pi_l, \pi_{l-1}) = Q(\pi_l) - Q(\pi_{l-1}). \tag{23}$$

Since $Q(\pi_0) = Q(\emptyset) = 0$, we have

$$\begin{aligned} \sum_{l=1}^{L} \Delta_l &= Q(\pi_L) - Q(\pi_{L-1}) + \cdots + Q(\pi_1) - Q(\pi_0) \\ &= Q(\pi_L) - Q(\pi_0) = Q(\pi_L). \end{aligned}$$

For two FBS-channel pairs $e(l)$ and $e(l')$, we say $e(l)$ *conflicts with* $e(l')$ when there is an edge connecting the FBS in $e(l)$ and the FBS in $e(l')$ in the interference graph G_I, and the two FBS's choose the same channel. Let Ω be the global optimal solution. We define ω_l as the subset of Ω that conflicts with allocation $e(l)$ but not with the previous allocations $\{e(1), e(2), \cdots, e(l-1)\}$.

Lemma 5: Assume the greedy algorithm in Table III stops in L steps. The global optimal solution Ω can be partitioned into L non-overlapping subsets ω_l, $l = 1, 2, \cdots, L$.

Proof: According to the definition of ω_l, the L subsets of the optimal solution Ω do not intersect with each other. Assume the statement is false, then the union of these L subsets is not equal to the optimal set Ω. Let the *set difference* be $\omega_{L+1} = \Omega \backslash (\cup_{l=1}^{L} \omega_l)$. By definition, ω_{L+1} does not conflict with the existing L allocations $\{e(1), \cdots, e(L)\}$, meaning that the greedy algorithm can continue to at least the $(L+1)$-th step. This conflicts with the assumption that the greedy algorithm stops in L steps. It follows that $\Omega = \cup_{l=1}^{L} \omega_l$. ■

Let $\Delta(\pi_2, \pi_1) = Q(\pi_2) - Q(\pi_1)$ denote the difference between two feasible allocations π_1 and π_2. We next derive a lower bound on the performance of the greedy algorithm. We assume two properties for function $\Delta(\pi_2, \pi_1)$ in the following.

Property 1: Consider FBS-channel pair sets π_1, π_2, and σ, satisfying $\pi_1 \subseteq \pi_2$ and $\sigma \cap \pi_2 = \emptyset$. We have $\Delta(\pi_2 \cup \sigma, \pi_1 \cup \sigma) \leq \Delta(\pi_2, \pi_1)$.

Property 2: Consider FBS-channel pair sets π, σ_1, and σ_2 satisfying $\sigma_1 \cap \pi = \emptyset$, $\sigma_2 \cap \pi = \emptyset$, and $\sigma_1 \cap \sigma_2 = \emptyset$. We have $\Delta(\sigma_1 \cup \sigma_2 \cup \pi, \pi) \leq \Delta(\sigma_1 \cup \pi, \pi) + \Delta(\sigma_2 \cup \pi, \pi)$.

In Property 1, we have $\sigma \cap \pi_1 = \emptyset$ since $\pi_1 \subseteq \pi_2$ and $\sigma \cap \pi_2 = \emptyset$. This property states that the incremental objective value does not get larger as more channels are allocated and as the objective value gets larger. Property 2 states that the incremental objective value achieved by allocating multiple FBS-channel pair sets does not exceed the sum of the incremental objective values achieved by allocating each individual FBS-channel pair set. These are generally true for many resource allocation problems [32].

Since we choose the maximum incremental allocation in each step of the greedy algorithm, we have Lemma 6 that directly follows Step 3 in Table III.

Lemma 6: For any FBS-channel pair $\sigma \in \omega_l$, we have $Q(\pi_{l-1} \cup \sigma) - Q(\pi_{l-1}) = \Delta(\pi_{l-1} \cup \sigma, \pi_{l-1}) \leq \Delta_l$.

Lemma 7: Assume the greedy algorithm stops in L steps, we have

$$Q(\Omega) \leq Q(\pi_L) + \sum_{l=1}^{L} \sum_{\sigma\in\omega_l} \Delta(\sigma \cup \pi_{l-1}, \pi_{l-1}).$$

Proof: The following inequalities hold true according to the properties of the $\Delta(\cdot, \cdot)$ function:

$$\begin{aligned} &Q((\cup_{i=l+1}^{L} \omega_i) \cup \pi_l) = Q((\cup_{i=l+2}^{L} \omega_i) \cup \pi_l) + \\ &\qquad \Delta((\cup_{i=l+1}^{L} \omega_i) \cup \pi_l, (\cup_{i=l+2}^{L} \omega_i) \cup \pi_l) \\ &\leq Q((\cup_{i=l+2}^{L} \omega_i) \cup \pi_l) + \Delta(\omega_{l+1} \cup \pi_l, \pi_l) \\ &\leq Q((\cup_{i=l+2}^{L} \omega_i) \cup \pi_{l+1}) + \Delta(\omega_{l+1} \cup \pi_l, \pi_l) \\ &\leq Q((\cup_{i=l+2}^{L} \omega_i) \cup \pi_{l+1}) + \sum_{\sigma\in\omega_{l+1}} \Delta(\sigma \cup \pi_l, \pi_l). \end{aligned}$$

We have $\pi_0 = \emptyset$ and $\omega_{L+1} = \emptyset$ (see Lemma 5). With induction from $l = 0$ to $l = L - 1$, we have $Q((\cup_{i=1}^{L} \omega_i) \cup \emptyset) = Q(\Omega)$ and $Q(\Omega) \leq Q(\pi_L) + \sum_{l=1}^{L} \sum_{\sigma\in\omega_l} \Delta(\sigma \cup \pi_{l-1}, \pi_{l-1})$. ■

Lemma 8: The maximum size of ω_l is equal to the degree, in the interference graph G_I, of the FBS selected in the l-th step of the greedy algorithm, which is denoted as $D(l)$.

Proof: Once FBS i is allocated with channel m, the neighboring FBS's in G_I, $\mathcal{R}(i)$, cannot use the same channel m anymore due to the interference constraint. The maximum number of FBS-channel pairs that conflict with the selected FBS-channel pair $\{i, m\}$, i.e., the maximum size of ω_l, is equal to the degree of FBS i in G_I. ■

Then we have Theorem 2 that provides a lower bound on the objective value achieved by the greedy algorithm given in Table III.

Theorem 2: The greedy algorithm can achieve an objective value that is at least $\frac{1}{1+D_{max}}$ of the global optimum, where D_{max} is the maximum node degree in the interference graph G_I of the femtocell CR network.

Proof: According to Lemmas 7 and 8, we have:

$$Q(\Omega) \leq Q(\pi_L) + \sum_{l=1}^{L} D(l)\Delta_l = Q(\pi_L) + \bar{D}\sum_{l=1}^{L}\Delta_l = (1+\bar{D})Q(\pi_L), \quad (24)$$

where $\bar{D} = \sum_{l=1}^{L} D(l)\Delta_l / \sum_{l=1}^{L}\Delta_l$. The second equality is due to the facts that $\sum_{l=1}^{L}\Delta_l = Q(\pi_L)$.

To further simplify the bound, we replace $D(l)$ with the maximum node degree D_{max}. We then have $\bar{D} \leq \sum_{l=1}^{L} D_{max}\Delta_l / \sum_{l=1}^{L}\Delta_l = D_{max}$ and

$$\frac{1}{1+D_{max}} Q(\Omega) \leq Q(\pi_L) \leq Q(\Omega), \quad (25)$$

which provides a lower bound on the performance of the greedy algorithm. ■

When there is a single FBS in the CR network, we have $D_{max} = 0$ and $Q(\pi_L) = Q(\Omega)$ according to Theorem 2. The proposed algorithm produces the optimal solution. In the case of multiple non-interfering FBS's, we still have $D_{max} = 0$ and can obtain the optimal solution using the proposed algorithm. For the femtocell CR network given in Fig. 1 (with interference graph shown in Fig. 3), we have $D_{max} = 1$ and the low bound is a half of the global optimal. Note that (24) provides a tighter bound for the optimum than (25), but with higher complexity. These are interesting performance bounds since they bound the achievable video quality, an application layer performance measure, rather than lower layer metrics (e.g., bandwidth or time share).

V. Simulation Results

We evaluate the performance of the proposed algorithms using MATLAB and JSVM 9.13 Video codec. Two scenarios are used in the simulations: a single FBS CR network and a CR network with interfering FBS's. In every simulation, we compare the proposed algorithms with the following three more straightforward heuristic schemes:

- Heuristic 1 based on *equal allocation*: each CR user chooses the better channel (i.e., the common channel or a licensed channel) based on the channel conditions; time slots are equally allocated among active CR users;
- Heuristic 2 exploiting *multiuser diversity*: the MBS and each FBS chooses one active CR user with the best channel condition; the entire time slot is allocated to the selected CR user.
- *SCA-MAC* proposed in [33]: with this scheme, the successful transmission rate is evaluated based on channel packet loss rate and collision probability with primary users; the channel-user pair with the highest transmission probability is selected.

We choose SCA-MAC because it adopts similar models and assumptions as in this paper. Once the channels are selected, the same distributed algorithm is used for scheduling video data for all the three schemes.

We adopt the Raleigh block fading model and the packet loss probability is between [0.004, 0.028]. The frame rate is set to 30 fps and the GoP size is 16. The base layer mode is set to be AVC compatible. The motion search mode is set to Fast Search with search range 32. Each point in the figures presented in this section is the average of 10 simulation runs with different random seeds. We plot 95% confidence intervals in the figures, which are generally negligible.

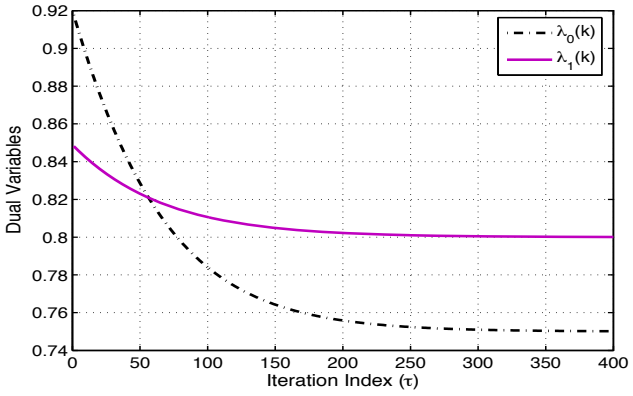

Fig. 4. Convergence of the two dual variables in the single FBS case.

A. Case of Single FBS

In the first scenario, there are $M = 8$ channels and the channel parameters P_{01}^m and P_{10}^m are set to 0.4 and 0.3, respectively, for all m. The maximum allowable collision probability γ_m is set to 0.2 for all m. There is one FBS and three active CR users. Three Common Intermediate Format (CIF, 352×288) video sequences are streamed to the CR users, i.e., *Bus* to CR user 1, *Mobile* to CR user 2, and *Harbor* to CR user 3. We have $T = 10$ as the delivery deadline. Both probabilities of false alarm ϵ and miss detection δ are set to 0.3 for all the FBS's and CR users, unless otherwise specified.

First we investigate the convergence of the distributed algorithm. The traces of the two dual variables are plotted in Fig. 4. To improve the convergence speed, the correlation in adjacent time slots can be exploited. In particular, we set the optimal values for the optimization variables in the previous time slot as the initialization values for the variables in the current time slot. By doing so, the convergence speed can be improved. It can be seen that both dual variables converge to their optimal values after 300 iterations. After convergence, the optimal solution for the primary problem can be obtained.

Our proposed scheme achieves the best performance among the three algorithms, with up to 4.3 dB improvement over the two heuristic schemes and up to 2.5 dB over SCA-MAC. Such gains are significant with regard to video quality, since a 0.5 dB difference is distinguishable by human eyes. Compared to the two heuristic schemes and SCA-MAC, the video quality of our proposed scheme is well balanced among the three users, indicating better fairness performance.

In Fig. 5, we examine the impact of the number of channels M on received video quality. First, we validate the video quality measure used in our formulation by comparing the PSNR value computed using (9) with that computed from real decoded video frames. The average PSNR for three received videos are plotted in the figure. It can be seen that the

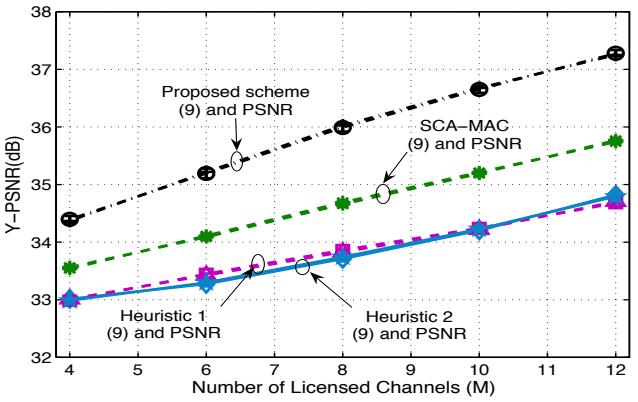

Fig. 5. Single FBS: received video quality vs. number of channels (computed with (9) and measured by PSNR).

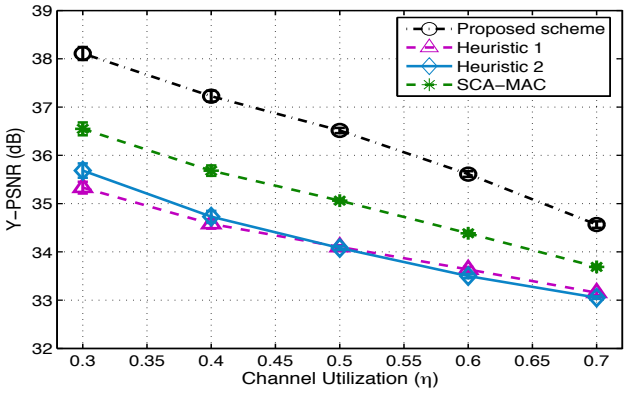

Fig. 7. Single FBS: received video quality vs. channel utilization.

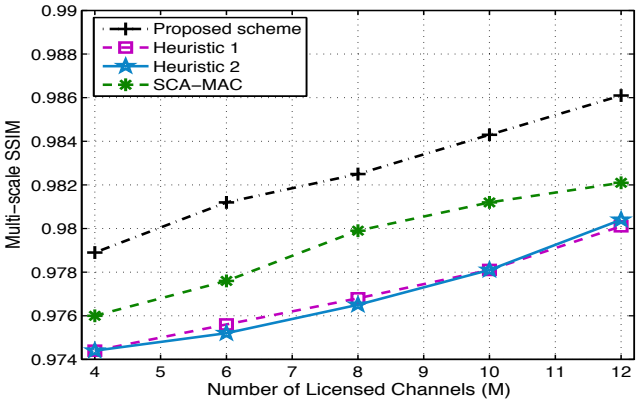

Fig. 6. Single FBS: received video quality vs. number of channels (measured by MS-SSIM).

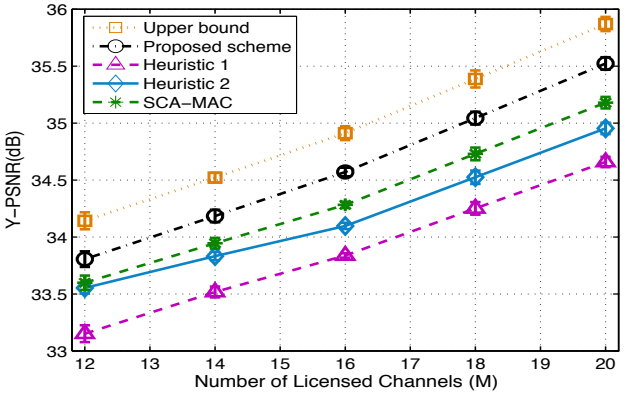

Fig. 8. Interfering FBS's: received video quality vs. number of channels.

real PSNRs are very close to those predicted by (9), with overlapping confidence intervals. This is also consistent with the results shown in Fig. 2. Second, as expected, the more licensed channels, the more spectrum opportunities for CR users and the higher PSNR for received videos. SCA-MAC performs better than two heuristics, but is inferior to the proposed scheme.

We also plot the MS-SSIM of the received videos at the three CR users in Fig. 6 [31]. Similar observations can be made from the MS-SSIM plot. All MS-SSIMs for the four curves are more than 0.97 and very close to 1. The proposed scheme still outperforms the other three schemes. In the remaining figures, we will use model predicted PSNR values, since the model (9) is sufficient to predict the real video quality.

In Fig. 7, we demonstrate the impact of channel utilization η on received video quality. The average PSNRs achieved by the four schemes are plotted when η is increased from 0.3 to 0.7. Intuitively, a smaller η allows more spectrum opportunities for video transmission. This is illustrated in the figure where all the three curves decrease as η gets larger. The performance of both heuristics are close and the proposed scheme achieves a gain about 3 dB over the heuristics and 2 dB over SCA-MAC.

We also compare the MGS and FGS videos while keeping other parameters identical. We find that MGS video achieves over 0.5 dB gain in video quality over FGS video. The results are omitted for brevity.

B. Case of Interfering FBS's

We next investigate the second scenario with three FBS's, and each FBS has three active CR users. Each FBS streams three different videos to the corresponding CR users. The coverages of FBS 1 and 2 overlap with each other, and the coverages of FBS 2 and 3 overlap with each other.

In Fig. 8, we examine the impact of the number of channels M on the received video quality. The average PSNRs of all the active CR users are plotted in the figure when we increase M from 12 to 20 with step size 2. As mentioned before, more channels imply more transmission opportunities for video transmission. In this scenario, heuristic 2 (with a multiuser diversity approach) outperforms heuristic 1 (with an equal allocation approach). But its PSNRs are still about $0.3 \sim 0.5$ dB lower that those of the proposed algorithm. The proposed scheme has up to 0.4 dB improvement over SCA-MAC. In Fig. 8, we also plot an upper bound on the optimal objective value, which is obtained as in (24). It can be seen that the performance of our proposed scheme is close to optimal solution since the gap between the upper bound and our scheme is generally small (about 0.5 dB).

Next, we examine the impact of sensing errors on the received video quality. In Fig. 9, we test five pairs of $\{\epsilon, \delta\}$

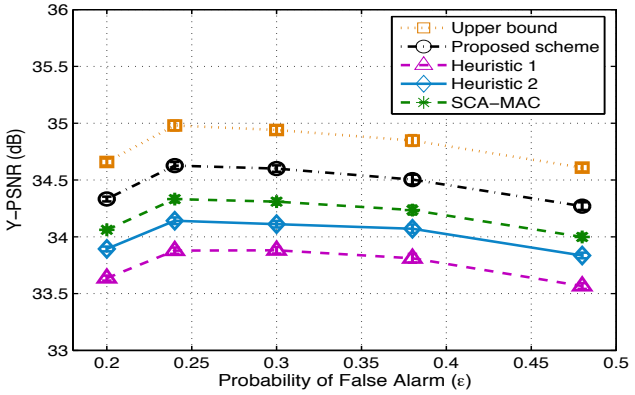

Fig. 9. Interfering FBS's: received video quality vs. sensing error probability.

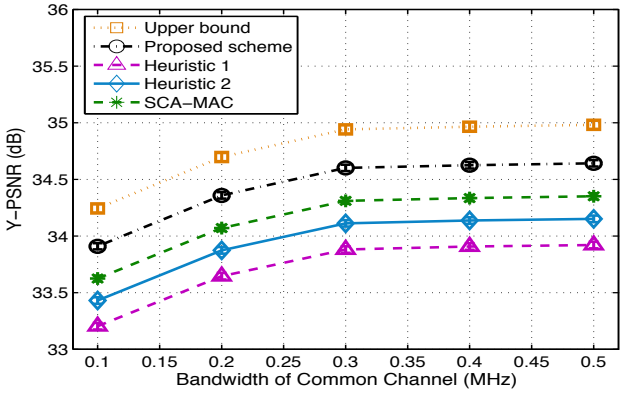

Fig. 10. Interfering FBS's: received video quality vs. bandwidth of the common channel.

values: {0.2,0.48}, {0.24,0.38}, {0.3,0.3}, {0.38,0.24}, and {0.48,0.2}. It is interesting to see that the performance of all the four schemes get worse when the probability of one of the two sensing errors gets large. We can trade-off between false alarm and miss detection probabilities to find the optimal operating point for the spectrum sensors. Moreover, the dynamic range of video quality is not big for the range of sensing errors simulated, compared to that in Fig. 8. This is because both sensing errors are modeled and treated in the algorithms. Again, our proposed scheme outperforms the two heuristic schemes and SCA-MAC with considerable margins for the entire range.

We also investigate the impact of the bandwidth of the common channel B_0. In this simulation, we fix B_1 at 0.3 Mbps and increase B_0 from 0.1 Mbps to 0.5 Mbps with step size 0.1 Mbps. The results are presented in Fig. 10. We notice that the average video quality increases rapidly as the common channel bandwidth is increased from 0.1 Mbps to 0.3 Mbps. Beyond 0.3 Mbps, the increases of the PSNR curves slow down and the curves get flat. This implies that a very large bandwidth for the common channel is not necessary, since the gain for additional bandwidth diminishes as B_0 gets large. Again, the proposed scheme outperforms the other three schemes and the gap between our scheme and the upper bound is small.

Next, we stop the distributed algorithm after a fixed amount of time, and evaluate the suboptimal solutions. In particular, we vary the duration of time slots, and let the distributed

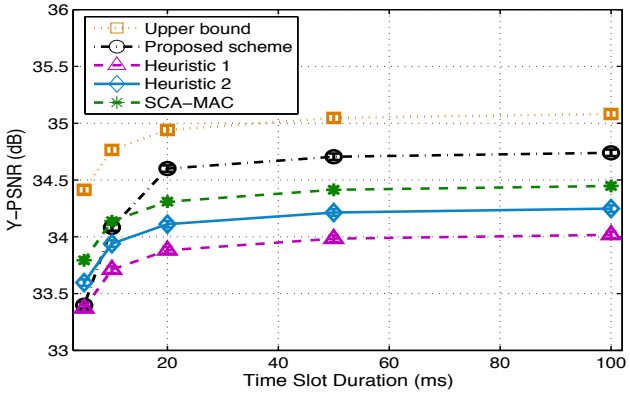

Fig. 11. Video quality achieved by the algorithms when they are only executed for 5% of the time slot duration.

algorithm run for 5% of the time slot duration at the beginning of the time slot. Then the solution obtained this way will be used for the video data transmissions. The results are presented in Fig. 11. It can be seen that when the time slot is 5 ms, the algorithm does not converge after $5\% \times 5 = 0.25$ ms and the PSNR produced by the distributed algorithm is close to that of Heuristic 1, and lower than those of Heuristic 2 and SCA-MAC. When the time slot is sufficiently large, the algorithm can get closer to the optimal and the proposed algorithm produces better video quality as compared to the two heuristic algorithms and SCA-MAC. Beyond 20 ms, the increase in PSNR is small since all the curves gets flat. Therefore the proposed algorithm could be useful even when there is no time for it to fully converge to the optimal.

During the simulations, we find the collision rate with primary users are strictly kept below the prescribed collision tolerance γ. These results are omitted for brevity.

VI. Conclusions

In this paper, we investigated the problem of streaming multiple MGS videos in a femtocell CR network. We formulated a multistage stochastic programming problem considering various design factors across multiple layers. We developed a distributed algorithm that can produce optimal solutions in the case of non-interfering FBS's, and a greedy algorithm for near-optimal solutions in the case of interfering FBS's with a proved lower bound. The proposed algorithms are evaluated with simulations and are shown to outperform three alternative schemes with considerable gains.

References

[1] D. Hu and S. Mao, "Resource allocation for medium grain scalable videos over femtocell cognitive radio networks," in *Proc. IEEE ICDCS'11*, Minneapolis, MN, June 2011, pp. 258–267.

[2] Cisco, "Visual Networking Index (VNI)," July 2010. [Online]. Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

[3] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

[4] R. Kim, J. S. Kwak, and K. Etemad, "WiMAX femtocell: requirements, challenges, and solutions," *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 84–91, Sept. 2009.

[5] I. Gven, S. Saunders, O. Oyman, H. Claussen, and A. Gatherer, "Femtocell networks," *EURASIP J. Wireless Comm. and Networking*, 2010, article ID 367878, 2 pages, doi:10.1155/2010/367878.

[6] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.
[7] M. Wien, H. Schwarz, and T. Oelbaum, "Performance analysis of SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1194–1203, Sept. 2007.
[8] V. Chandrasekhar, J. G. Andrews, T. Muharemovic, Z. Shen, and A. Gatherer, "Power control in two-tier femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4316–4328, Aug. 2009.
[9] H.-C. Lee, D.-C. Oh, and Y.-H. Lee, "Mitigation of inter-femtocell interference with adaptive fractional frequency reuse," in *Proc. IEEE ICC'10*, Cape Town, South Africa, May 2010, pp. 1–5.
[10] T. Alade, H. Zhu, and J. Wang, "Uplink co-channel interference analysis and cancellation in femtocell based distributed antenna system," in *Proc. IEEE ICC'10*, Cape Town, South Africa, May 2010, pp. 1–5.
[11] J. O'Carroll and L. D. Holger Claussen, "Partial GSM spectrum reuse for femtocells," in *Proc. IEEE PIMRC'09*, Tokyo, Japan, Sept. 2009, pp. 2111–2116.
[12] A. Galindo-Serrano, L. Giupponi, and M. Dohler, "Cognition and docition in OFDMA-based femtocell networks," in *Proc. IEEE GLOBECOM'10*, Miami, FL, Dec. 2010.
[13] F.-S. Chu and K.-C. Chen, "Mitigation of macro-femto co-channel interference by spatial channel separation," in *Proc. IEEE VTC-Spring'11*, Budapest, Hungary, May 2011, pp. 1–5.
[14] S. Rangan, "Femto-macro cellular interference control with subband scheduling and interference cancelation," in *Proc. IEEE GLOBECOM'10 Workshops*, Miami, FL, Dec. 2010, pp. 695–700.
[15] Z. Bharucha, H. Haas, G. Auer, and I. Cosovic, "Femto-cell resource partitioning," in *Proc. IEEE GLOBECOM Workshops*, Honolulu, HI, Dec. 2009, pp. 1–6.
[16] R. Madan, A. Sampath, A. Khandekar, J. Borran, and N. Bhushan, "Distributed interference management and scheduling in LTE-A femto networks," in *Proc. IEEE GLOBECOM'10*, Miami, FL, Dec. 2010, pp. 1–5.
[17] S.-M. Cheng, S.-Y. Lien, F.-S. Chu, and K.-C. Chen, "On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 40–47, 2011.
[18] S. Kaimaletu, R. Krishnan, S. Kalyani, N. Akhtar, and B. Ramamurthi, "Cognitive interference management in heterogeneous femto-macro cell networks," in *Proc. IEEE ICC'11*, Kyoto, Japan, June 2011, pp. 1–6.
[19] D. Hu and S. Mao, "Multicast in femtocell networks: a successive interference cancellation approach," in *Proc. IEEE GLOBECOM'11*, Huston, TX, Dec. 2011, pp. 1–6.
[20] Y. Zhao, S. Mao, J. Neel, and J. H. Reed, "Performance evaluation of cognitive radios: metrics, utility functions, and methodologies," *Proc. IEEE*, vol. 97, no. 4, pp. 642–659, Apr. 2009.
[21] H.-P. Shiang and M. van der Schaar, "Dynamic channel selection for multi-user video streaming over cognitive radio networks," in *Proc. IEEE ICIP'08*, San Diego, CA, Oct. 2008, pp. 2316–2319.
[22] S. Ali and F. Yu, "Cross-layer QoS provisioning for multimedia transmissions in cognitive radio networks," in *Proc. IEEE WCNC'09*, Budapest, Hungary, Apr. 2009, pp. 1–5.
[23] L. Ding, S. Pudlewski, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Distributed spectrum sharing for video streaming in cognitive radio ad hoc networks," in *Intl. Workshop on Cross-layer Design in Wireless Mobile Ad Hoc Networks*, Niagara Falls, Canada, Sept. 2009, pp. 1–13.
[24] D. Hu and S. Mao, "Streaming scalable videos over multi-hop cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3501–3511, Nov. 2010.
[25] D. Hu, S. Mao, Y. Hou, and J. Reed, "Fine grained scalability video multicast in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 334–344, Apr. 2010.
[26] A. Motamedi and A. Bahai, "MAC protocol design for spectrum-agile wireless networks: stochastic control approach," in *Proc. IEEE DySPAN'07*, Dublin, Ireland, Apr. 2007, pp. 448–451.
[27] S. Geirhofer, L. Tong, and B. Sadler, "Cognitive medium access: constraining interference based on experimental models," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 95–105, Jan. 2008.
[28] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 118–129, Jan. 2008.
[29] T. Rappaport, *Wireless Communications: Principles & Practice*, 2nd ed. Indianapolis, IN: Prentice Hall PTR, 2001.
[30] Q. Zhang and S. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1688–1692, Nov. 1999.
[31] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment using structural distortion measurement," *Signal Processing: Image Commun.*, no. 2, pp. 121–132, Feb. 2004.
[32] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, Mar. 1998.
[33] A. Hsu, D. Wei, and C. Kuo, "A cognitive MAC protocol using statistical channel allocation for wireless ad-hoc networks," in *Proc. IEEE WCNC'07*, Hong Kong, P.R. China, Mar. 2007, pp. 105–110.

Donglin Hu received the M.S. degree from Tsinghua University, Beijing, China, in 2007 and the B.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China in 2004, respectively, all in electrical engineering. He also received the M.S. degree in Probability and Statistics from Auburn University, Auburn, AL, in 2011. Since 2007, he has been pursuing a Ph.D. degree in the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL. His research interests include cognitive radio networks, femtocell networks, network modeling, cross-layer design, performance analysis, and algorithm optimization for wireless networks and multimedia communications.

Shiwen Mao (S'99-M'04-SM'09) received Ph.D. in electrical and computer engineering from Polytechnic University, Brooklyn, NY (now Polytechnic Institute of New York University) in 2004. He was a research staff member with IBM China Research Lab from 1997 to 1998. He was a Research Scientist in the Bradley Department of Electrical and Computer Engineering at Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA from 2003 to 2006. Currently, he is an Associate Professor in the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL.

His research interests include cross-layer optimization of wireless networks and multimedia communications, with current focus on cognitive radio, femtocell, free space optical and 60GHz networks. He is on the Editorial Board of IEEE Transactions on Wireless Communications, IEEE Communications Surveys and Tutorials, Elsevier Ad Hoc Networks Journal, Wiley International Journal of Communication Systems, and ICST Transactions on Mobile Communications and Applications. He chairs the Interest Group on Cross-layer Design for Multimedia Communications of IEEE Communications Society's Multimedia Communications Technical Committee.

Dr. Mao is a coauthor of *TCP/IP Essentials: A Lab-Based Approach* (Cambridge University Press, 2004). He received the US National Science Foundation *Faculty Early Career Development Award (CAREER)* in 2010, *The 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems*, and *The Best Paper Runner-up Award* at The Fifth International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine) 2008. He also received *Auburn Alumni Council Research Awards for Excellence–Junior Award* in 2011 and two *Auburn Author Awards* in 2011. Dr. Mao holds one US patent.