"©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE."

Resource Allocation for OFDMA-Based Cognitive Radio Multicast Networks With Primary User Activity Consideration

Duy T. Ngo, *Student Member, IEEE*, Chintha Tellambura, *Senior Member, IEEE*, and Ha H. Nguyen, *Senior Member, IEEE*

***Abstract*—This paper considers the primary user activity or the subchannel availability in optimally distributing the available resources for an orthogonal frequency-division multiple-access (OFDMA) cognitive radio multicast network. For this purpose, a risk-return model is presented, and a general rate-loss function, which gives a reduction in the attainable throughput whenever primary users reoccupy the temporarily accessible subchannels, is introduced. Taking the maximization of the expected sum rate of secondary multicast groups as the design objective, an efficient joint subcarrier and power-allocation scheme is proposed. Specifically, the design problem is solved via a dual optimization method under constraints on the tolerable interference thresholds at individual primary user's frequency bands. It is shown that as the number of subcarriers gets large (which is often the case in practice), the dual-domain solution becomes globally optimum with regard to the primal problem. More attractively, the "practically optimal" performance of this approach is achieved with a substantially lower complexity, which is only linear in the total number of subcarriers as opposed to exponential complexity typically required by a direct search method. Our proposed design is valid for unicast and multicast transmissions and is applicable for a wide range of rate-loss functions, among which, the linear function is a special case. The superiority of the dual scheme is thoroughly verified by numerical examples.**

***Index Terms*—Cognitive radio, dual optimization, loss function, orthogonal frequency-division multiple access (OFDMA), resource allocation, wireless multicast.**

Manuscript received May 11, 2009; revised October 14, 2009. First published February 25, 2010; current version published May 14, 2010. The work of D. T. Ngo was supported in part by the Alberta Ingenuity Foundation Student Scholarship and in part by the Alberta Informatics Circle of Research Excellence Information and Communication Technology Graduate Student Award. The review of this paper was coordinated by Dr. H. Harada.

D. T. Ngo was with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada. He is now with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: trongduy@ece.ualberta.ca; duy.ngo@mail.mcgill.ca).

C. Tellambura is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: chintha@ece.ualberta.ca).

H. H. Nguyen is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: ha.nguyen@usask.ca).

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.2010.2042827

I. Introduction

In November 2002, the Federal Communications Commission published a report on the current management of the precious radio spectrum resource in the U.S. One of the main findings stated in the report is the following [1]:

> "In many bands, spectrum access is a more significant problem than physical scarcity of spectrum, in large part due to legacy command-and-control regulation that limits the ability of potential spectrum users to obtain such access."

Simply put, it has been confirmed that much of the licensed spectrum lies idle at any given time and location and that the spectrum shortage results from the spectrum-management policy rather than the physical scarcity of usable frequencies. Spectrum utilization can thus be significantly improved by allowing secondary users to access spectrum holes unoccupied by the primary users at given locations and times. Cognitive radio [2], [3] has been identified as an efficient technology to promote this idea by exploiting the existence of the spectrum portions unoccupied by the primary (or licensed) users. Potentially, while the primary users have priority access to the spectrum, the secondary (or unlicensed or cognitive) users have restricted access, subject to a constrained degradation on the primary users' performance [4]. In spectrum-sharing environments, the key design challenges of a cognitive radio network are, therefore, to guarantee protection of the primary users from excessive interference induced by the secondary users and to meet some quality-of-service requirements for the latter [5], [6].

On the other hand, spectrum pooling is an opportunistic spectrum access approach that enables public access to the already licensed frequency bands [7], [8]. The basic idea is to merge spectral ranges from different spectrum owners (for example, military trunked radios) into a common pool, from which the secondary users may temporarily rent spectral resources during idle periods of licensed users. In effect, the licensed system does *not* need to be changed while the secondary users access the unused resources. Among the many possible technologies for unlicensed users' transmission in spectrum-pooling radio systems, orthogonal frequency-division multiplexing (OFDM) has already been widely recognized as a highly promising candidate mainly due to its great flexibility in dynamically allocating the unused spectrum among secondary users, as well as its ability to monitor the spectral activities of licensed users

0018-9545/$26.00 © 2010 IEEE

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

at no extra cost [9]. However, it has been shown that employing OFDM also affects the performance of a cognitive radio network, for instance, causing mutual interference between the primary and secondary users due to the non-orthogonality of the respective transmitted signals [10], [11].

Resource allocation for OFDM-based cognitive radio networks has been examined in [12], where an optimal scheme derived via Lagrangian formulation is proposed to maximize the downlink capacity of a single cognitive user while guaranteeing the interference to the primary user being below a specified threshold. The work in [13] extends [12] to multiuser scenarios, in which a discrete sum rate of the secondary network is maximized constrained on the interference to the primary user bands, as well as on the total transmitted power. Subject to the per-subchannel power constraints (due to primary users interference limits), the study in [14] proposes a partitioned iterative water-filling algorithm that enhances the capacity of an OFDM cognitive radio system. Further, the issue of downlink channel assignment and power control for frequency-division multiple-access-based cognitive networks has been addressed in [15], wherein a set of base stations (BSs) makes opportunistic spectrum access in order to serve the fixed-location wireless users within their cells. To maximize the total number of active users that can be supported while guaranteeing the minimum signal-to-interference-plus-noise ratio (SINR) requirements of secondary users, as well as protecting the primary users, suboptimal schemes are suggested for the formulated mixed-integer program. Considering networks with the coexistence of multiple primary and secondary links through an orthogonal frequency-division multiple-access (OFDMA)-based air interface, [16] utilizes the dual framework from [17] to provide centralized and distributed algorithms that improve the total achievable sum rate of secondary networks subject to interference constraints specified at the primary users' receivers.

While previous related studies implicitly assume that the designated spectrum for secondary usage is fixed and always available, the work in [18] investigates another important aspect of subchannel availability or primary user activity in an OFDM cognitive radio system. Here, cognitive radio can be realized as a risky environment where the licensed users may, at any time, come back and take up the frequency bands currently available for secondary access. In such scenarios, the power already invested by unlicensed users in those bands becomes wasted. By referring to a risk-return model and upon defining a general rate-loss function that gives a decrease in total throughput whenever primary users reoccupy the temporarily accessible subchannels, a problem of optimally allocating power for one *single* cognitive user is formulated by incorporating the reliability or availability of OFDM subchannels. For the special case of linear rate loss, the problem belongs to the class of convex optimization, and a multilevel water-filling (MLW) solution has been provided in [18]. However, for other types of rate-loss functions, the design problem becomes highly nonconvex, and hence, this solution is no longer applicable.

Different from all the aforementioned works that only consider unicast transmission, this paper studies resource allocation in a secondary OFDMA-based multicast network, where the patterns of primary user activities on the available radio spectrum are dynamic. As an efficient means of transmitting the same content to multiple receivers while minimizing network resource usage, multicasting [19], [20] is clearly an attractive transmission technique for secondary networks who only have limited access to the available spectrum. However, in *multigroup* settings, the problem of joint subcarrier assignment and power distribution usually turns out to be of nonconvex structure. This makes the solution derived in [18] no longer valid, even for the linear rate-loss model. As well, performing a direct search in the primal domain to find the global optimal solutions is certainly impractical in these cases because the computational complexity of such an approach is prohibitively demanding. Motivated by the shortcomings of the existing designs, we propose in this paper a dual-optimization scheme to efficiently solve the challenging resource allocation in a cognitive OFDMA network consisting of multiple multicast groups.

Adopting a similar risk-return model from [18] to account for the primary user activities, our proposed subcarrier-assignment and power-allocation solution targets the maximization of the expected sum rate of all secondary users in an OFDMA-based cognitive radio multicast network while satisfying the tolerable interference level imposed on individual licensed users. Specifically, the original nonconvex optimization problem is effectively solved in the dual domain with the global optimum obtained in the limit as the number of subcarriers goes to infinity. More significantly, it is shown that the proposed approach has only linear complexity in the total number of subcarriers, resulting in a huge reduction in computational burden. These features are certainly attractive for practical OFDMA-based systems that deploy a large number of subcarriers. Further, the dual approach presented here is valid for both unicast and multicast scenarios and is applicable for a wide range of rate-loss functions, among which, linear is a special case. As well, the mutual interference between secondary and primary networks, which is an important factor, is explicitly quantified. The effects of adjacent subcarrier-nulling technique [10], which is used to decrease the mutual interference, on the proposed design are also analyzed.

The rest of this paper is organized as follows: Section II presents the system model under consideration. In addition, formulated in this section is the resource-allocation problem for OFDMA-based multicast secondary networks, where primary user activities are taken into account. Section III introduces the dual-optimization method, which is an effective approach when dealing with a large class of multicarrier resource-allocation problems. In Section IV, an iterative scheme derived from the dual-optimization framework to resolve the design problem is proposed. Section V provides numerical examples to verify the performance of the devised solution. Finally, Section VI concludes this paper.

II. System Model and Problem Formulation

A. System Model

We consider a primary BS that transmits (not necessarily OFDM) signals to its N primary users, each of which occupies a predetermined frequency band $B_P^{(n)}$ $(n = 1, \ldots, N)$ in the

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

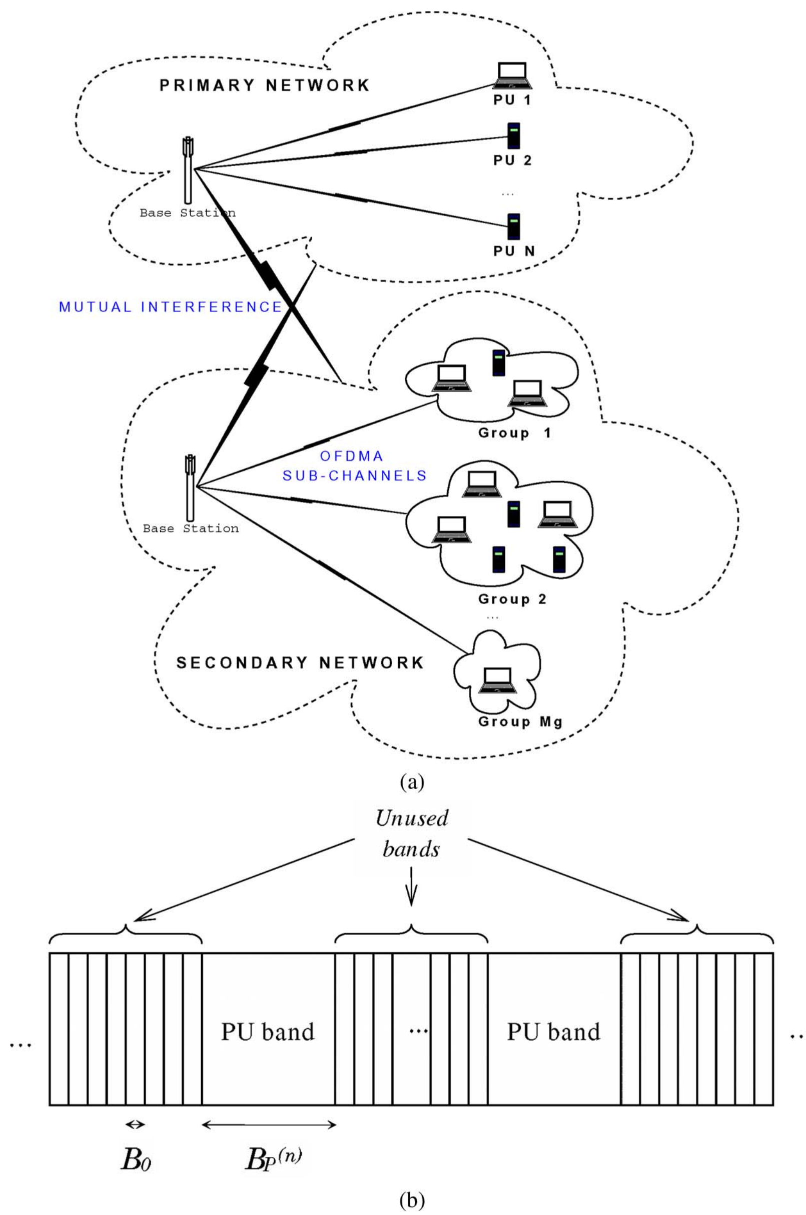

Fig. 1. System model. (a) Coexistence of primary and secondary networks in cognitive radio. (b) Distribution of available spectrum to primary and secondary users.

available spectrum. To implement efficient opportunistic spectrum access, a secondary BS is permitted to employ K OFDM subcarriers to transmit G downlink traffic flows, each of them to one distinct multicast group consisting of secondary users, over the temporarily unused frequency bands. Information regarding the availability of these bands is made known at the secondary BS either by means of signaling from the primary BS or as the result of spectrum sensing performed by the secondary BS itself. Notice that since licensed users have priority access to the radio spectrum, the unused frequency bands need to be handed back to the primary network upon request at any time. Therefore, depending on the activity of primary users, there is a chance that the temporarily unused spectrum will become reoccupied.

Assume that each secondary user receives one traffic flow at a time, and hence, it belongs to only one multicast group. Let M_g and $|M_g|$ $(g = 1, \ldots, G)$ denote the user set of group g and its cardinality, respectively. The gth group is unicast if $|M_g| = 1$, whereas it is multicast if $|M_g| > 1$. Thus, the system framework presented here is applicable to both unicast and multicast transmissions. Clearly, all the secondary users belong to the set $M = \bigcup_{g=1}^{G} M_g$, and $|M| = \sum_{g=1}^{G} |M_g|$ is the total number of users in the cognitive multicast network. Let B denote the total bandwidth available for secondary usage and also assume that each subchannel has an equal bandwidth of $B_s = B/K$. The system setup is depicted in Fig. 1(a), with the distribution of accessible spectrum shown in Fig. 1(b).

As a consequence of having two coexisting networks, the OFDM signals from the secondary BS, which are intended for its own serviced users, might interfere with the reception at the primary users' receivers. Upon defining $\check{P}_{m,k}$ as the power spent for transmitting to secondary user m in group g on subcarrier k and denoting T_s as the OFDM symbol duration, the power spectral density (PSD) of the subcarrier-k signal can

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

be modeled as $\Phi_k(f) = \check{P}_{m,k} T_s (\sin \pi f T_s / \pi f T_s)^2,\ m \in M_g$. Then, the interference caused by this signal onto primary user n is given as [10]

$$\begin{aligned} I_k^{(n)} &= \left| g_{\mathrm{SP}}^{(n)} \right|^2 \int\limits_{d_k^{(n)} - B_P^{(n)}/2}^{d_k^{(n)} + B_P^{(n)}/2} \Phi_k(f) df \\ &= \check{P}_{m,k} \left\{ \left| g_{\mathrm{SP}}^{(n)} \right|^2 T_s \int\limits_{d_k^{(n)} - B_P^{(n)}/2}^{d_k^{(n)} + B_P^{(n)}/2} \left(\frac{\sin \pi f T_s}{\pi f T_s} \right)^2 df \right\} \\ &= \check{P}_{m,k} \check{I}_k^{(n)} \end{aligned} \tag{1}$$

where $d_k^{(n)} = |f_k - f_n|$ represents the spectral distance between subcarrier k and center frequency f_n of primary user n, and $g_{\mathrm{SP}}^{(n)}$ denotes the channel from secondary BS to primary user n. Clearly, interference $I_k^{(n)}$ depends on the transmitted power $\check{P}_{m,k}$, the channel coefficient $g_{\mathrm{SP}}^{(n)}$, and the spectral distance $d_k^{(n)}$.

In addition, the coexistence of primary users and multicast groups of secondary users may cause interference induced by the signals from the primary BS, which are destined to primary users, onto the secondary users' frequency bands. Let $g_{m,k}^{PS}$ be the channel coefficient from primary BS to secondary user m in group g on subcarrier k, and let $\Phi_{\mathrm{PU}}^{(n)}(e^{jw})$ be the PSD of the signal transmitted from primary BS to primary user n. Then, the interference power caused by this signal onto secondary user $m \in M_g$ on subcarrier k can be computed as [10]

$$J_{m,k}^{(n)} = \left| g_{m,k}^{PS} \right|^2 \int\limits_{d_k^{(n)} - B_s/2}^{d_k^{(n)} + B_s/2} \mathcal{E}\left\{ I_K(w) \right\} dw \tag{2}$$

where $\mathcal{E}\{I_K(w)\} = (1/2\pi K) \int_{-\pi}^{\pi} \Phi_{\mathrm{PU}}^{(n)}(e^{jw})((\sin(w-\phi)K/2)/(\sin(w-\phi)/2))^2 d\phi$ is the PSD of the primary user n's signal after K-Fast-Fourier-transform (FFT) processing.

In this paper, the resource allocation of the secondary network is accomplished in a centralized manner with perfect channel state information of all primary and secondary users in the system being assumed (for example, via training and feedbacks from the users through dedicated channels). Similar to [18] and [21]–[23], we further assume a slow-fading channel model such that the channel conditions remain unchanged during the allocation period. This model is particularly valid for high-data-rate systems and/or environments with reduced degrees of mobility, where the channel gains do not vary too significantly over time. With the perfect link information available, it is therefore possible to determine the maximum rate at which an individual secondary user can reliably receive data, as well as the corresponding subcarrier over which the data will be transmitted. The channel SINR (CSINR) of the secondary user $m \in M_g$ on subcarrier k can be shown to be

$$\alpha_{m,k} = \frac{\left| h_{m,k}^{SS} \right|^2}{\Gamma \left(N_0 B_s + \sum_{n=1}^{N} J_{m,k}^{(n)} \right)} \tag{3}$$

where $h_{m,k}^{SS}$ is the corresponding channel coefficient, and N_0 is the one-sided PSD of the additive white Gaussian noise. The parameter Γ represents the signal-to-noise ratio gap to the capacity limit, which is a function of the desired bit error rate, coding gain, and noise margin [24]. The maximum attainable rate of secondary user $m \in M_g$ on subcarrier k is then

$$\check{r}_{m,k} = \frac{B_s}{B} \log_2(1 + \alpha_{m,k} \check{P}_{m,k}). \tag{4}$$

An attractive feature of wireless multicast is that multicast data can be transmitted from the BS to multiple mobile users only through a single transmission. However, while all users within a multicast group receive the same rate from the BS, the main issue arises from the mismatch data rates attainable by individual users of that group whose link conditions are typically asymmetric. If the BS transmits rate higher than the maximum rate that a user can handle, then that user is not able to decode *any* of the transmitted data at all. In this paper, the conventional multicast transmission approach is followed by enforcing the secondary BS to transmit at the lowest rate of all the users within a group, which is determined by the user with the smallest CSINR [25]. This assures that the multicast services can be provided to *all* the subscribed users. Although it is also possible to adopt other techniques to overcome the issue of data rate mismatch (for instance, exploiting the hierarchy in multicast data together with the assumption of multidescription coding (MDC) [26]), such solutions are limited to multimedia applications. Therefore, MDC approaches are not pursued in this paper, where, instead, the conventional approach will be followed to deal with a more general class of applications. With the conventional multicast transmission, let

$$\gamma_{g,k} = \min_{m \in M_g} \alpha_{m,k} \tag{5}$$

be the equivalent CSINR of group g on subcarrier k. Then, the maximum rate at which all secondary users of group g are able to decode the data transmitted on that same subcarrier is

$$r_{g,k}^{\max} = \frac{B_s}{B} \log_2(1 + \gamma_{g,k} P_{g,k}) \tag{6}$$

where $P_{g,k}$ denotes the power allocated to group g on subcarrier k. Since all the secondary users in a group receive the same rate, the aggregate rate transmitted to group g on subcarrier k is scaled by the group size as

$$R_{g,k} = \sum_{m \in M_g} r_{g,k}^{\max} = |M_g|\, r_{g,k}^{\max}. \tag{7}$$

B. Problem Formulation With Primary User Activity Consideration

In a cognitive radio environment, there is likely a delay from the moment that a channel is made available for secondary usage to the time that the secondary network is fully aware of that accessibility. The time delay could be due to, for example, the efficiency of spectrum-sensing algorithms performed by the cognitive network. This effect is of particular concern if

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

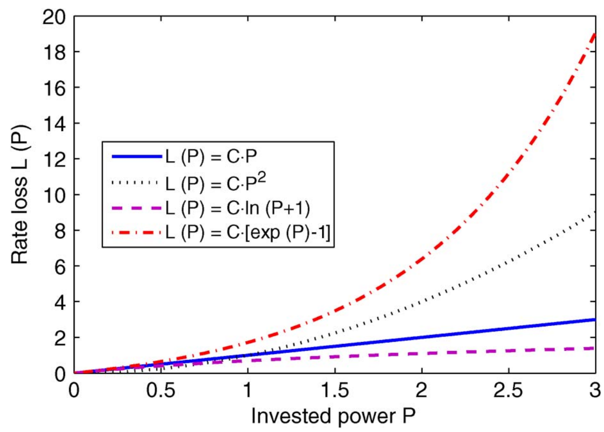

Fig. 2. Some common rate-loss functions.

the patterns of spectrum usage by primary users are greatly dynamic, for instance, frequent occurrences of releasing and reoccupying certain bands. Consequently, it is possible that the present resource-allocation process, carried out by secondary BS at the current time frame t, is indeed based on an already-obsolete information, which is only valid at time $t - \Delta t$ in the past (regarding, for example, locations of spectrum holes, link conditions, interference, etc.). This is because during that time delay interval $\Delta t > 0$, primary users may have come back and taken up the subchannels once available for secondary access, making the current allocation no longer optimal. In many cases, this effect, which is essentially caused by the dynamics of primary users, can be highly severe.

To account for the primary user activities (or equivalently, the availability of OFDM subchannels), we refer to the risk-return model in which the power allocated to a frequency band is considered an investment in that band [18]. In this model, the cognitive radio environment can be thought of as a risky environment, where the primary users may return to take up the available band at any time. In such cases, the secondary users' power investment in that band is wasted. This represents a loss in the data rate achieved by the secondary users, probably due to, for instance, better allocation schemes that could have been utilized or an increase in the amount of interference caused to primary users when the unused bands are reoccupied. To model this loss, we define a rate loss $L(P)$ that is a function of the power invested by the cognitive network. Strictly, $L(P)$ is required to satisfy the following two conditions.

1) $L(P) > 0$ for $P > 0$.
2) $L(P) = 0$ for $P = 0$.

The study in [18] assumes a linear rate-loss function $L(P) = C \cdot P$, where C is the normalized average cost per unit power for the secondary users to utilize the resource. While this function models quite good approximations of many practical cases, it also simplifies the analysis and provides better insights into the design problems. Certainly, many other types of rate-loss functions, shown in Fig. 2, are also possible [27], e.g., quadratic $L(P) = C \cdot P^2$, exponential $L(P) = C \cdot [\exp(P) - 1]$, and logarithmic $L(P) = C \cdot \ln(P + 1)$.

Given the probability ϕ_k that the subchannel k is taken up by the primary users in the current time frame, the expected rate loss can be written as

$$\mathcal{E}\{\Delta R_{g,k}\} = \phi_k L(P_{g,k}). \tag{8}$$

Then, the expected rate transmitted from secondary BS to group g on subcarrier k becomes

$$\begin{aligned}\mathcal{E}\{R_{g,k}\} &= R_{g,k} - \mathcal{E}\{\Delta R_{g,k}\}\\ &= \frac{|M_g|}{K}\log_2(1+\gamma_{g,k}P_{g,k}) - \phi_k L(P_{g,k}).\end{aligned} \tag{9}$$

The goal of this paper is to devise a subcarrier assignment and power allocation policy that maximizes the expected sum rate of all multicast groups of secondary users while satisfying constraints on the tolerable interference level of each individual primary user. Specifically, the design problem can be formulated as follows:

$$\max_{\{P_{g,k}\}} \sum_{g=1}^{G}\sum_{k=1}^{K} \frac{w_g|M_g|}{K}\log_2(1+\gamma_{g,k}P_{g,k}) - \phi_k L(P_{g,k}) \tag{10}$$

$$\text{s.t.} \quad \sum_{g=1}^{G}\sum_{k=1}^{K} P_{g,k}\check{I}_k^{(n)} \leq I_{\text{th}}^{(n)}, \quad n = 1,\ldots,N \tag{11}$$

$$P_{g,k} \geq 0, \quad g = 1,\ldots,G, \quad k = 1,\ldots,K \tag{12}$$

$$P_{g,k}P_{g',k} = 0 \quad \forall g' \neq g. \tag{13}$$

In this formulation, weight $w_g \geq 0$ reflects the priority designated to group g and is obliged to satisfy $\sum_{g=1}^{G} w_g = 1$. Constraint (11) expresses the tolerable interference level at the receiver of primary user n, with $I_{\text{th}}^{(n)}$ representing the interference threshold. Constraints (12) and (13) enforce a disjoint subchannel assignment in OFDMA systems, that is, one subcarrier is permitted to be assigned to at most one group at a time [21].

It is noteworthy that the optimization problem in (10)–(13) is NP-hard, since it requires the allocation of an optimal set of subcarriers to each multicast group of secondary users. The water-filling procedure [18], which was devised for a much simpler scenario with one single cognitive user and a linear rate loss model, is no longer applicable here. Further, the complexity needed to directly resolve this combinatorial problem increases at least exponentially with the number of subcarriers K. Such prohibitively high computational effort is required even for a simplified case, as discussed in the Appendix. Moreover, the multiple constraints in (11) make it even more challenging to derive an analytical solution for the problem in (10)–(13).

In the following sections, we will first introduce the dual optimization method for nonconvex multicarrier resource allocation. Then, we will show how the difficult optimization problem in (10)–(13) can effectively be resolved in the dual domain with virtually zero duality gap. Global optimal solutions can be obtained in the limit as the number of subcarriers goes to infinity. In addition, we establish that the complexity of the proposed dual scheme is only linear in the total number of subcarriers. This represents a significant reduction in computational burden at the BS, where it is desirable to rapidly find the optimal solutions to mitigate the fluctuations of wireless channels.

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

III. Dual Optimization of Non-Convex Multicarrier Resource Allocation

Consider the problem of optimally allocating resources in a multicarrier system with M users and K subcarriers. The objective and constraints of the optimization consist of a number of individual functions, each corresponding to one of the K subcarriers, and can be expressed as

$$\begin{aligned} \max_{\{\mathbf{x}_k\}\in\mathcal{R}^M} \quad & \sum_{k=1}^{K} f_k(\mathbf{x}_k) \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathbf{h}_k(\mathbf{x}_k) \le \mathbf{P} \end{aligned} \tag{14}$$

where $f_k(\cdot)$ are $\mathcal{R}^M \to \mathcal{R}$ (not necessarily concave) functions, $\mathbf{h}_k(\cdot)$ are $\mathcal{R}^M \to \mathcal{R}^N$ (not necessarily convex) functions, and the constant $\mathbf{P}$ denotes the N-vector of constraints.

The idea of dual optimization is to solve (14) by first forming its Lagrangian dual, which is defined as [28]

$$\mathcal{L}\left(\{\mathbf{x}_k\}, \boldsymbol{\lambda}\right) = \sum_{k=1}^{K} f_k(\mathbf{x}_k) - \boldsymbol{\lambda}^T \left(\sum_{k=1}^{K} \mathbf{h}_k(\mathbf{x}_k) - \mathbf{P} \right) \tag{15}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T \succeq \mathbf{0}$ is a vector of Lagrange dual variables.

Then, upon defining the dual objective as $D(\boldsymbol{\lambda}) = \max_{\{\mathbf{x}_k\}} \mathcal{L}(\{\mathbf{x}_k\}, \boldsymbol{\lambda})$, the dual optimization problem in (14) becomes

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & D(\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \boldsymbol{\lambda} \succeq \mathbf{0}. \end{aligned} \tag{16}$$

The Lagrange dual problem is a convex optimization problem that can very efficiently be solved in practice. This is the case whether the primal problem is convex or not. Nevertheless, solving a dual problem is not always equivalent to solving the primal problem. Let f^* and D^* denote the primal and dual optimal values, respectively. Then, the difference $d = D^* - f^*$ is defined as the optimal duality gap. It has been proven from duality theory that $d \ge 0$ always holds. In particular, when $f_k(\mathbf{x}_k)$'s are concave and $\mathbf{h}_k(\mathbf{x}_k)$'s are convex (that is, (14) is convex), a strong duality is guaranteed, which implies $d = 0$. In such cases, the primal and dual problems have the same optimal value, and thus, the globally optimal solution can be derived in the dual domain via Lagrangian decomposition. However, this gap in general is not always zero, and the optimal solution of the dual problem only gives the best upper bound on that of the primal.

Interestingly enough, it has been proven in [17] and [29] that even if the multicarrier optimization problem in (14) is nonconvex, the duality gap is zero if either of the following conditions is met.

Condition 1: $\mathbf{x}_k^*(\boldsymbol{\lambda}) = \arg\max_{\mathbf{x}_k} \mathcal{L}(\{\mathbf{x}_k\}, \boldsymbol{\lambda})$ is continuous at optimal $\boldsymbol{\lambda}^*$.

Condition 2: The optimal value of $\sum_{k=1}^{K} f_k(\mathbf{x}_k)$ is concave in $\mathbf{P}$.

In particular, it has been shown that for nonconvex multicarrier optimization problems with the general form of (14), the concavity requirement in Condition 2, which is called *frequency-sharing* condition in [17], is always satisfied when the number of subcarriers goes to infinity. The proofs of this important result are given in [17], whereas a more general theoretical justification can be found in [28, Sec. 5.1.6]. Significantly, the result implies that the original challenging nonconvex problem can efficiently be solved in the dual domain with a virtually negligible duality gap for a realistically large number of subcarriers.

IV. Practically Optimal Subcarrier and Power Allocation via Dual Method

It can be observed that the particular structure of the problem in (10)–(13) satisfies the frequency-sharing condition, and hence, its global optimum can be obtained in the dual domain by an iterative method at a significantly reduced computational complexity [17]. In brief, for a fixed Lagrange dual variable set, it is possible to first decompose (10)–(13) by Lagrangian into several unconstrained per-tone power allocation subproblems, each of which can be solved by water filling or exhaustive search. Once the optimal distribution of powers is found for all subcarriers, the Lagrange dual variables are updated by a subgradient-based method. The procedure is repeated until convergence, and the optimal solution of subcarrier and power allocation obtained in the dual domain becomes that of the primal problem (10)–(13) as the number of subcarriers tends to be large.

A. Proposed Dual Design

The exclusive channel-assignment constraint [see (12) and (13)] can be expressed as $P_{g,k} \in \mathcal{S}$, where the domain $\mathcal{S}$ is defined as $\mathcal{S} = \{P_{g,k} \ge 0; g = 1, \ldots, G, k = 1, \ldots, K | P_{g,k} P_{g',k} = 0, \forall g' \ne g\}$. Over the domain $\mathcal{S}$, the Lagrangian of problems (10) and (11) is given as

$$\begin{aligned} \mathcal{L}\left(\{P_{g,k}\}, \boldsymbol{\lambda}\right) = & \sum_{g=1}^{G} \sum_{k=1}^{K} \frac{w_g |M_g|}{K} \log_2(1 + \gamma_{g,k} P_{g,k}) \\ & - \phi_k L(P_{g,k}) - \sum_{n=1}^{N} \lambda_n \left(\sum_{g=1}^{G} \sum_{k=1}^{K} P_{g,k} \check{I}_k^{(n)} - I_{\text{th}}^{(n)} \right) \end{aligned} \tag{17}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T \succeq \mathbf{0}$ is the vector of dual variables.

Now, thanks to the disjoint subchannel constraint in OFDMA-based systems, the Lagrange dual function in (17) can be decomposed into K independent optimization problems, i.e., one for each subcarrier k, as follows:

$$\begin{aligned} D(\boldsymbol{\lambda}) &= \max_{\{P_{g,k}\}\in\mathcal{S}} \mathcal{L}\left(\{P_{g,k}\}, \boldsymbol{\lambda}\right) \\ &= \sum_{k=1}^{K} D_k(\boldsymbol{\lambda}) + \sum_{n=1}^{N} \lambda_n I_{\text{th}}^{(n)} \end{aligned} \tag{18}$$

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

where the per-subcarrier problem is

$$D_k(\boldsymbol{\lambda}) = \max_{\{P_{g,k}\}\in\mathcal{S}} \sum_{g=1}^{G} \left\{ \frac{w_g|M_g|}{K} \log_2(1+\gamma_{g,k}P_{g,k}) - \left(\sum_{n=1}^{N} \lambda_n \check{I}_k^{(n)} P_{g,k} + \phi_k L(P_{g,k}) \right) \right\} \quad (19)$$

for $k = 1, \ldots, K$.

For each subcarrier k, there is at most one $P_{g,k} > 0$ for all $g = 1, \ldots, G$. Therefore, the optimal group assignment for subcarrier k can be found by first deriving G optimal power allocations, i.e., one for each of the total G groups, and then selecting the value that maximizes $D_k(\boldsymbol{\lambda})$. Assume that multicast group g is active on subcarrier k. For a fixed $\boldsymbol{\lambda}$, if the rate loss is linear with respect to the invested power, the objective of the maximization in (19) becomes a concave function of $P_{g,k}$. From the Karush–Kuhn–Tucker (KKT) conditions [28], the optimal power allocation can then be devised as

$$P_{g,k}^* = \left(\frac{1}{\gamma_{0,k}} - \frac{1}{\gamma_{g,k}} \right)^+ \quad (20)$$

where $(\cdot)^+ = \max(\cdot, 0)$. Apparently, this is a form of water filling, where the water level is

$$\gamma_{0,k} = \frac{K\left(C\phi_k + \sum_{n=1}^{N} \lambda_n \check{I}_k^{(n)}\right) \log 2}{w_g|M_g|}. \quad (21)$$

For a quadratic loss, the objective of the maximization in (19) is also a concave function. A closed-form solution can also be derived from the KKT conditions as

$$P_{g,k}^* = \left[\frac{1}{4\gamma_{g,k}C\phi_k} \left(\frac{8w_g|M_g|}{K\log 2}\gamma_{g,k}^2 C\phi_k + x^2\gamma_{g,k}^2 + 4C^2\phi_k^2 - 4x\gamma_{g,k}C\phi_k \right)^{\frac{1}{2}} - \left(\frac{1}{2\gamma_{g,k}} + \frac{x}{4C\phi_k} \right) \right]^+ \quad (22)$$

where $x = \sum_{n=1}^{N} \lambda_n \check{I}_k^{(n)}$.

If the rate loss is exponential, then the objective function is still concave with respect to $P_{g,k}$. However, an analytical solution is not available in this case. Instead, numerical methods can be utilized to obtain the optimal power allocation. For logarithmic rate-loss types, the objective turns out to be one of maximizing the difference of two concave functions (or equivalently, minimizing the difference of two convex functions). Many algorithms recently developed in the area of d.c. programming [30] may be applied to solve the resulting problem. In the worst case, if the per-subcarrier problem has no special structure (mainly due to the types of rate loss function), an exhaustive search is typically required to determine its solution. Nevertheless, since this problem is unconstrained, it is much easier to handle than the original problem.

Once the optimal power for each group has been found, by searching over all G possible group assignments for subcarrier k, the optimal value of (19) is actually

$$D_k^*(\boldsymbol{\lambda}) = \max_{g} \left\{ \frac{w_g|M_g|}{K} \log_2\left(1+\gamma_{g,k}P_{g,k}^*\right) - \left(\sum_{n=1}^{N} \lambda_n \check{I}_k^{(n)} P_{g,k}^* + \phi_k L\left(P_{g,k}^*\right) \right) \right\} \quad (23)$$

for $k = 1, \ldots, K$. This is achieved when the power allocation on subcarrier k is $P_{g^*,k} = P_{g^*,k}^*$ and $P_{g,k} = 0$ for all $g \neq g^*$, where g^* represents the group being allocated the subcarrier k. From (20)–(22), it is worth mentioning that the allocation depends not only on the CSINR and the number of group users but also on the availability of subchannel k, as represented by ϕ_k.

After (23) has been solved for all subcarriers $(k = 1, \ldots, K)$, the overall Lagrange dual function $D(\boldsymbol{\lambda})$ in (18) can be evaluated for the fixed $\boldsymbol{\lambda}$. Finally, it remains to find $\boldsymbol{\lambda}^* \succeq \mathbf{0}$ that minimizes $D(\boldsymbol{\lambda})$. This can efficiently be done by a subgradient-based method that iteratively updates $\boldsymbol{\lambda}$ until its convergence. By definition, a vector $\mathbf{d}$ is called a subgradient of $D(\boldsymbol{\lambda})$ at $\boldsymbol{\lambda}$ if for all $\boldsymbol{\mu} \succeq \mathbf{0}$ [28]

$$D(\boldsymbol{\mu}) = D(\boldsymbol{\lambda}) + (\boldsymbol{\mu} - \boldsymbol{\lambda})^T \mathbf{d}. \quad (24)$$

Proposition 1: Let $P_{g,k}^*$ be the optimizing variable in the maximization problem in the definition of $D(\boldsymbol{\lambda})$. For the optimization problem in (10)–(13) with dual objective defined in (18), the following choice of $\mathbf{d} = [d_1, d_2, \ldots, d_n]^T$ with

$$d_n = I_{\text{th}}^{(n)} - \sum_{g=1}^{G}\sum_{k=1}^{K} P_{g,k}^* \check{I}_k^{(n)} \quad (25)$$

is a subgradient of $D(\boldsymbol{\lambda})$.

Proof: For any $\boldsymbol{\mu} \succeq \mathbf{0}$, it can be shown that

$$\begin{aligned} D(\boldsymbol{\mu}) &= \max_{\{P_{g,k}\}\in\mathcal{S}} \mathcal{L}\left(\{P_{g,k}\}, \boldsymbol{\mu}\right) \\ &\geq \mathcal{L}\left(\left\{P_{g,k}^*\right\}, \boldsymbol{\mu}\right) \\ &= D(\boldsymbol{\lambda}) + \sum_{n=1}^{N} (\mu_n - \lambda_n) \\ &\quad \cdot \left(I_{\text{th}}^{(n)} - \sum_{g=1}^{G}\sum_{k=1}^{K} P_{g,k}^* \check{I}_k^{(n)} \right) \end{aligned} \quad (26)$$

which verifies the definition of the subgradient in (24). ■

The basic idea of the subgradient method is to design a step-size sequence to update $\boldsymbol{\lambda}$ in the subgradient direction. For our problem of interest, the update may be performed as follows:

$$\lambda_n^{(t+1)} = \left(\lambda_n^{(t)} - \delta^{(t)} \left(I_{\text{th}}^{(n)} - \sum_{g=1}^{G}\sum_{k=1}^{K} P_{g,k} \check{I}_k^{(n)} \right) \right)^+ \quad (27)$$

for $n = 1, \ldots, N$, where $\delta^{(t)} > 0$ is a sequence of scalar step sizes. This subgradient update is guaranteed to converge to the optimal $\boldsymbol{\lambda}^*$ as long as $\delta^{(t)}$ is chosen to be sufficiently small. Some popular choices include a constant step size $\delta^{(t)} = \varepsilon > 0$ or diminishing rules $\delta^{(t)} = \beta/t$ and $\delta^{(t)} = \beta/\sqrt{t}$ for

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

TABLE I
SUBCARRIER AND POWER ALLOCATION BY DUAL OPTIMIZATION METHOD

THE PROPOSED DUAL SCHEME	
1:	**Initialize** $(\lambda_1, \ldots, \lambda_N)$
2:	**Repeat**
3:	for $k = 1, \ldots, K$
4:	compute $P^*_{g,k}$ for all groups $g = 1, \ldots, G$
5:	pick group g^* that gives minimum of $D_k(\boldsymbol{\lambda})$ as in (23)
6:	assign subcarrier k to group g^*
7:	set $P_{g^*,k} := P^*_{g^*,k}$ and $P_{g,k} := 0,\ \forall g \neq g^*$
8:	end
9:	update $(\lambda_1, \ldots, \lambda_N)$ according to (27)
10:	**Until** convergence of $\boldsymbol{\lambda}$

some constant $\beta > 0$. At the point of convergence, the sum interference induced by transmission from secondary BS to all of its multicast groups to each primary user's frequency band also converges, and (positive) optimal powers have been distributed to eligible multicast groups.

The overall proposed dual scheme is summarized in Table I. It is important to point out that as the number of subcarriers goes to infinity, the gap between primal and dual optimal solutions vanishes quickly to zero. In practice, OFDM systems employ a very large number of subcarriers (for example, as many as 3780); thus, the optimal solution, which is obtained in the dual domain by the proposed scheme, virtually becomes a global optimum of the primal problem in (10)–(13) with a negligible duality gap. Evidently, this demonstrates the practical optimality achieved by our proposed design.

B. Complexity Analysis

For a fixed $\boldsymbol{\lambda}$, solving (18) requires $\mathcal{O}(KG)$ executions. With an appropriate choice of step size, the subgradient method used to update $\boldsymbol{\lambda}$ converges in Δ iterations, which is a typically small number. The total complexity of the proposed dual scheme is therefore $\mathcal{O}(KG\Delta)$, which is only *linear* in the number of subcarriers. Since the total number of subcarriers is often large in practical scenarios, a huge reduction in computational burden is expected from the proposed dual scheme as compared with, at least, $\mathcal{O}(KG^K)$ operations by the optimal primal domain solution (for a simple case with only $N = 1$ primary user and zero rate loss, see the Appendix). Certainly, this is a highly desirable feature of adaptive algorithms designed for wireless communication systems, where resolutions often need to be found within a very short time due to the dynamics of wireless channels.

C. Effects of Adjacent Subcarrier Nulling Technique

The study in [10] proposes the method of dynamically deactivating subcarriers as a countermeasure to reduce the amount of interference from secondary to primary bands. Essentially, the suggested approach provides flexible guard bands between primary and secondary users by nulling subcarriers adjacent to the primary users' bands. However, this benefit comes at the cost of sacrificing bandwidth and, consequently, throughput of secondary users. It is therefore critical to balance the contradicting requirements of reducing interference and achieving the highest possible throughput of secondary users.

In the context of this paper, nulling adjacent subcarriers reduces the available degree of freedom, which is the number of available subcarriers for possible transmission from the secondary BS to its own users, and in turn leads to a decrease in the throughput achieved by all cognitive multicast groups. On the other hand, since the number of subcarriers K decreases, the computational complexity required by the proposed dual scheme to find the optimal solutions is lower. As well, more power can be distributed into the far-away subcarriers for a given interference threshold $I^{(n)}_{\text{th}}$. Together with the effects of multicarrier and multiuser diversity, this power distribution may compensate the consequences of bandwidth reduction.

V. PERFORMANCE EVALUATION

A. Simulation Setup and Assumptions

We consider a wireless system in which the primary BS communicates with N primary users. The primary user frequency bands are predetermined in the available spectrum. All the primary user signals are assumed to be elliptically filtered white noise with equal amplitude $P_{\text{PU}} = 1$. To exploit temporary unused spectrum holes, a secondary BS is also allowed to simultaneously transmit to $G = 2$ multicast groups, each respectively consisting of $|M_1| = 5$, $|M_2| = 3$ secondary users. The number of OFDM subcarriers used by the secondary BS is K, and the unused frequency bands are located on the sides of the already occupied bands. Moreover, the probability that primary users reoccupy unused subchannel k is assumed to be equal $\phi_k = \phi$ for all k.

We perform our numerical experiments in MATLAB 2009a environment on a PC equipped with Intel Pentium 4 processor (3.40-GHz CPU speed and 2-GB RAM). In the computation of attainable sum rates, 100 sets of independent channel gains $\{g^{(n)}_{\text{SP}}\}$, $\{g^{PS}_{m,k}\}$, and $\{h^{SS}_{m,k}\}$ are randomly generated according to the Rayleigh distribution. The average channel gains, the noise power of each subcarrier, the OFDM symbol duration, and the individual subcarrier bandwidth are all normalized to 1. It is further assumed that perfect coding is employed, which means that $\Gamma = 1$. Since all the spectral distances $d^{(n)}_k$ can be determined, it is possible to compute the interferences $\check{I}^{(n)}_k$, $J^{(n)}_{m,k}$, and the CSINR of individual secondary user $\alpha_{m,k}$. Then, the equivalent CSINR of group M_g on subcarrier k is simply $\gamma_{g,k} = \min_{m\in M_g} \alpha_{m,k}$. Unless stated otherwise, both groups are assumed to have equal priority $w_1 = w_2 = 0.5$.

As seen before, the choice of a linear rate-loss function makes the problem more straightforward to analyze, and a water-filling solution of power can be obtained as in (20) and (21). Indeed, any other form of rate-loss function only leads to a difference in the resolution of the per-subcarrier problems, whereas all other steps in the proposed dual scheme still remain unchanged. Therefore, the numerical examples are only performed for the case of linear rate-loss function $L(P_{g,k}) = C \cdot P_{g,k}$.

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

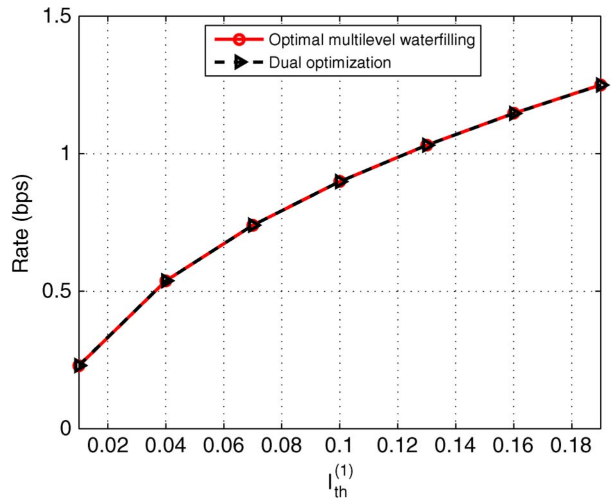

Fig. 3. Comparison of dual optimization and MLW methods ($N = 1$, $K = 8$, and $L(P_{g,k}) = 0$).

TABLE II
AVERAGE CONVERGENCE TIME FOR $N = 1$, $K = 8$, $L(P_{g,k}) = 0$, AND $I_{\text{th}} = 0.1$

Tolerance ϵ	10^{-3}	10^{-4}	10^{-5}	10^{-6}
CPU time (in seconds)	1.6405	9.4899	14.3542	23.3693

B. Numerical Results

To confirm the practical optimality achieved by the proposal, we first study a simple case with $N = 1$ primary user, $K = 8$ OFDM subcarriers, zero rate loss $L(P_{g,k}) = 0$, and $I_{\text{th}} = 0.1$. We compare the performance of the new design with that of the optimal primal-domain MLW solution derived in the Appendix, which is indeed based on the work of [31]. Fig. 3 displays the actual achieved throughput by both the dual optimization (with tolerance $\epsilon = 10^{-6}$ and step size $\delta^{(t)} = 10/\sqrt{t}$) and the MLW. Clearly, the two rate curves are almost indistinguishable. Notice that the duality gap is already insignificant even with only eight active subcarriers. In addition, we present in Table II the average computational time of the dual algorithm in our experiments. While it takes around 38.5 s of CPU time for the MLW algorithm to locate the globally optimal solution, that for the dual counterpart is fewer than 23.5 s, representing an almost 40% reduction. Furthermore, as the dual algorithm offers more flexibility to control the level of accuracy required, the convergence time can be as low as 1.6 s by setting $\epsilon = 10^{-3}$. It should also be pointed out that the required computational time for the proposed design depends on the choice of step size $\delta^{(t)}$. Fig. 4 shows that the dual method converges in only a few tens of iterations by setting $\delta^{(t)} = 10/\sqrt{t}$, whereas it may take up to 1000 iterations to converge with $\delta^{(t)} = 10/t$. This emphasizes the importance of selecting the appropriate step size to speed up the convergence process.

Next, we examine the case of multiple primary users $N = 2$ and a positive rate loss $L(P_{g,k}) > 0$. To better observe the effects of multiuser and multicarrier diversity, a larger number of OFDM subcarriers $K = 36$ will be employed. Since it is, in this case, too complex to carry out an exhaustive direct search to find the optimal solutions of (10)–(13), we instead verify

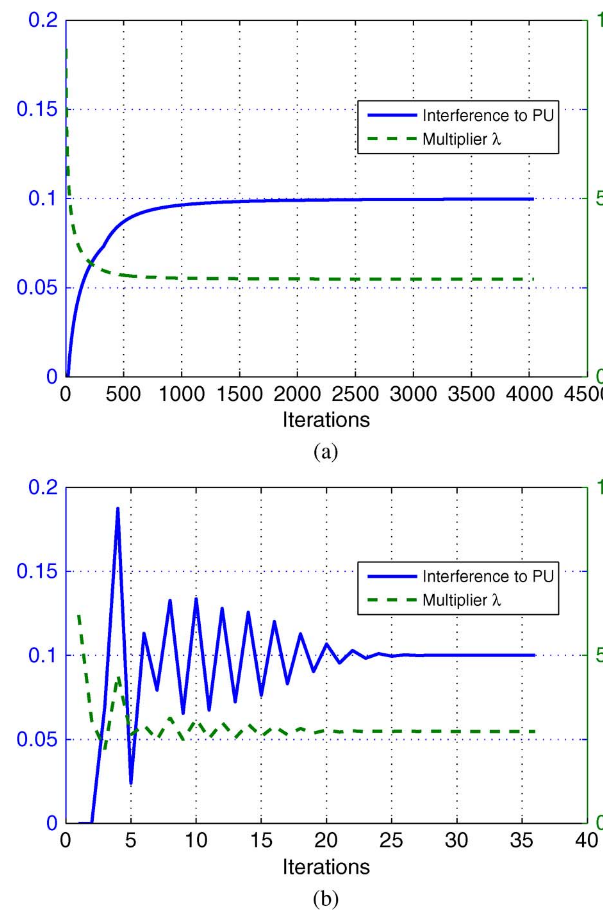

Fig. 4. Convergence process of the proposed dual scheme with different choices of step size. (a) Step size $\delta = 10/t$. (b) Step size $\delta = 10/\sqrt{t}$.

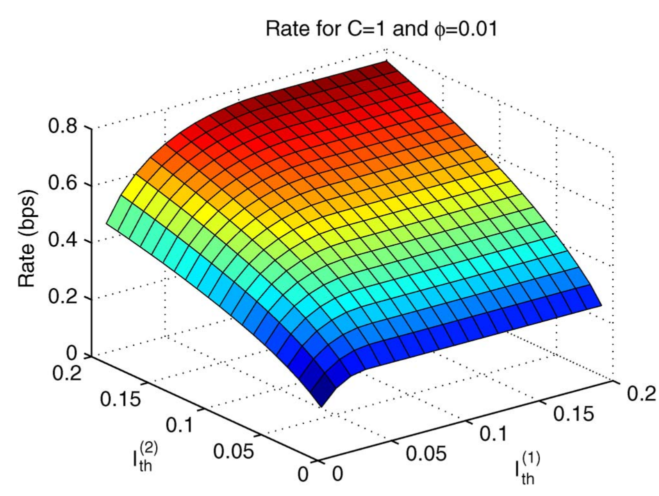

Fig. 5. Concavity of the optimal throughput for $N = 2$, $K = 36$, and $L(P_{g,k}) > 0$.

that Condition 2 stated in Section III is met. For $C = 1.0$ and $\phi = 0.01$, Fig. 5 demonstrates that the total expected rate sum at optimality is indeed a concave function of $\bar{\mathbf{I}}_{\text{th}} = [I_{\text{th}}^{(1)}, I_{\text{th}}^{(2)}]$. It is further expected that the concavity of optimal throughput with respect to $\bar{\mathbf{I}}_{\text{th}}$ will become even more visible as the number of subcarriers K is much larger than 36. From Condition 2, this observation implies a negligible duality gap and, again, indicates that the solution obtained by our proposal is virtually

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

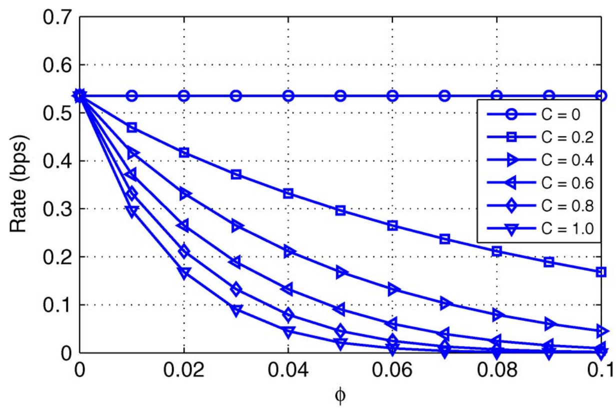

Fig. 6. Effect of varying rate loss parameters (with fixed $I_{\mathrm{th}} = 0.10$).

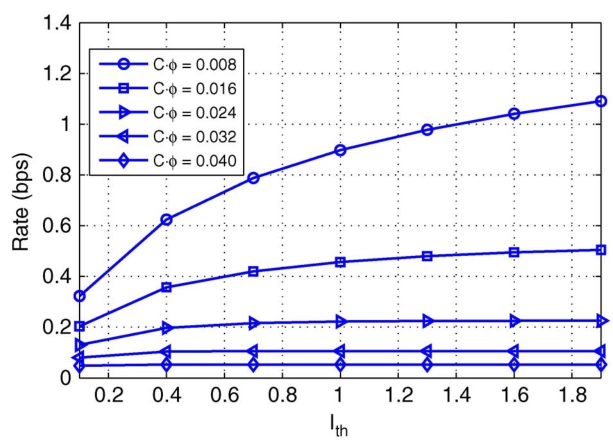

Fig. 7. Saturation effect of total throughput as rate loss increases.

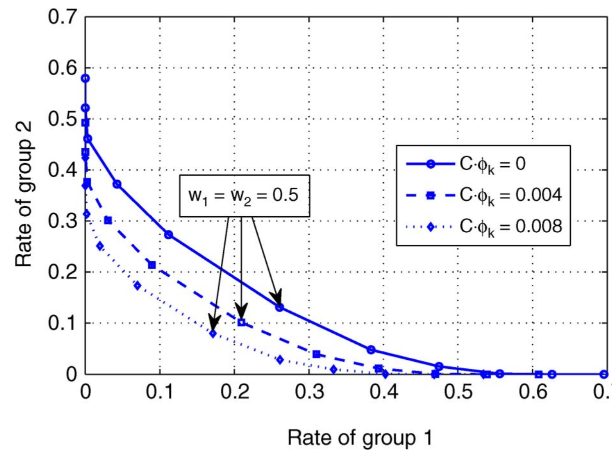

Fig. 8. Effect of varying weights w_g (with fixed $I_{\mathrm{th}} = 0.10$).

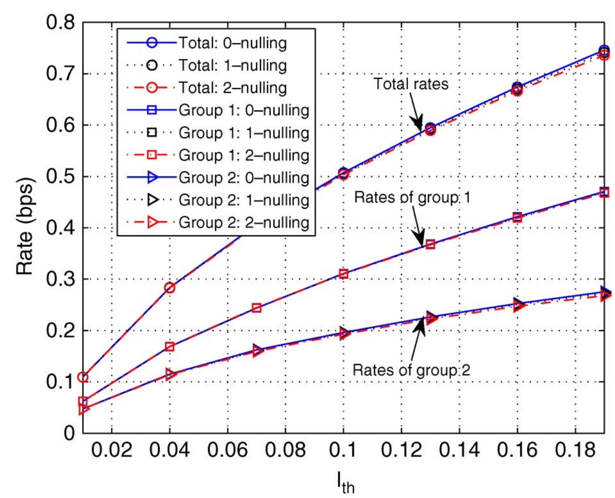

Fig. 9. Effect of adjacent subcarrier nulling (with zero rate loss assumed).

the primal global optimum. In what follows, all the simulation results are therefore presented for this more practical scenario with $N = 2$ primary users and $K = 36$ OFDM subcarriers.

With these parameters, we assess the consequences of varying the rate-loss parameters (i.e., C and ϕ) on the system throughput. From Fig. 6, it is apparent that an increase in rate loss will result in a decrease in the attained data rates. In particular, with $C = 1.0$, the throughput of the cognitive radio multicast network approaches zero at $\phi = 0.08$ and beyond. This implies that if the OFDM subchannels are too busy (i.e., large ϕ) or secondary access of the available resources is too costly (i.e., large C), then no secondary transmission is beneficial. Moreover, the saturation effect of total sum rates in the case of substantial rate loss is exhibited in Fig. 7. In this situation, no matter how much the primary users can tolerate the interference from the secondary network, it is impossible to further improve the system throughput of the latter. Essentially, this signifies the importance of considering the primary users' activities on the available frequency bands in any cognitive radio design.

As well, we examine the effect of the weight w_g on the achievable rates. Given a certain rate loss and for equal weights $w_1 = w_2 = 0.5$, it is apparent from Fig. 8 that group 1 may achieve a higher rate than that by group 2. Indeed, this is an expected result from (20) and (21). The group with greater member set $|\mathcal{M}_g|$ is likely to be allocated more power and thus have a better chance to occupy a particular subcarrier. In many cases, this leads to the unfair allocation of the available spectrum. One common way to overcome this issue is to assign different weights to individual groups, each of which having different levels of priority. By varying the values of w_g, the region of achieved throughput can be traced out for different rate loss values. Fig. 8 also shows that the attained data-rate region gets smaller as the value of the rate loss becomes larger. Moreover, by observing the top-left and bottom-right parts of these regions, it is clear that the group with lower priority (i.e., being assigned a smaller weight) may be allocated no resources at all.

Finally, to evaluate the effect of the adjacent subcarrier nulling technique, Fig. 9 plots the total system throughput (assuming no rate loss), as well as individual group rates obtained by the proposed scheme with and without nulling adjacent subcarriers on each side of the primary users' bands. Note that only deactivating one and two adjacent subcarriers are considered in the simulations. It can be observed from Fig. 9 that the attainable rates in the one- and two-nulling cases actually decrease since the adjacent subcarriers are assigned zero power, even when their respective channel conditions are very good. However, the degradation is minor in our numerical experiments. We speculate that this is due to two main reasons: 1) The number of deactivated subcarriers is small compared

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

with the total available spectrum, and 2) more power has been distributed into the far-away subcarriers, effectively compensating the effect of reducing the amount of accessible bandwidth.

VI. Conclusion

We have proposed a dual scheme for the allocation of subcarriers and power to maximize the expected throughput of a secondary network employing OFDMA subject to tolerable interference at the primary users in cognitive radio settings. The solution also takes into account the subchannel availability or the primary users' activities by incorporating a rate-loss function in the design. Global optimality can be achieved by the devised scheme for a realistically large number of OFDM subcarriers. Further, the proposed dual optimization method can handle both unicast and multicast transmissions, and its complexity is only linear in the number of subcarriers. The effects of nulling adjacent subcarriers on the proposed design have also been investigated. Numerical results have confirmed the potential benefits of our proposed approach.

Appendix

In this Appendix, we establish that although it is possible to directly derive an optimal solution for the design problem (10)–(13) in the primal domain, the complexity of such an approach is actually exponential in the number of OFDM subcarriers. For simplicity, let us consider the case of $N = 1$ primary user and zero rate loss $L(P_{g,k}) = 0$. The optimization problem can now be reduced to

$$\max_{\{P_{g,k}\}} \quad \sum_{g=1}^{G}\sum_{k=1}^{K} \frac{w_g|M_g|}{K}\log_2(1+\gamma_{g,k}P_{g,k}) \tag{28}$$

$$\text{s.t.} \quad \sum_{g=1}^{G}\sum_{k=1}^{K} P_{g,k}\check{I}_k \le I_{\text{th}}^{(1)} \tag{29}$$

$$P_{g,k} \ge 0, \quad g = 1,\ldots,G, \quad k = 1,\ldots,K \tag{30}$$

$$P_{g,k}P_{g',k} = 0 \quad \forall g' \ne g. \tag{31}$$

Let S_g denote the set of subcarriers allocated to group g. For any fixed channel assignment S_g, the problem in (28)–(31) is convex, and thus, its optimal solution can be determined from the KKT conditions as follows:

$$P_{g,k}^{*} = \left(\frac{w_g|M_g|}{K\check{I}_k\log 2}\cdot\mu - \frac{1}{\gamma_{g,k}}\right)^{+} \tag{32}$$

$$\mu = \frac{I_{\text{th}}^{(1)} + \sum_{g=1}^{G}\sum_{k\in\{S_g:P_{g,k}>0\}}\frac{\check{I}_k}{\gamma_{g,k}}}{\sum_{g=1}^{G}\frac{w_g|M_g||S_g|}{K\log 2}}. \tag{33}$$

Clearly, this is a form of MLW, wherein the number of used OFDM subchannels needs to be optimized until all powers are positive [31]. As finding optimal subchannel assignment among G groups of secondary users requires G^K searches, the overall optimization requires $\mathcal{O}(KG^K)$ operations, which is exponentially complex.

In addition, notice that the analytical solution in the foregoing derivation is made possible thanks to the many simplified assumptions. In the presence of a positive rate-loss function and multiple primary users, the optimal search in the primal domain would be far more complicated. This emphasizes the need to have more suitable approaches to efficiently solve the design problem (10)–(13).

Acknowledgment

The authors would like to thank the anonymous reviewers, whose invaluable comments helped improve the presentation of this paper.

References

[1] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Fed. Commun. Comm., Washington, DC, Tech. Rep. ET Docket No. 02-135, Nov. 2002.
[2] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[4] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.
[5] T. Clancy, "Formalizing the interference temperature model," *Wireless Commun. Mobile Comput.*, vol. 7, no. 9, pp. 1077–1086, Nov. 2007.
[6] Y. Xing, C. N. Mathur, M. Haleem, R. Chandramouli, and K. Subbalakshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comput.*, vol. 6, no. 4, pp. 423–433, Apr. 2007.
[7] T. Weiss and F. Jondral, "Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. S8–14, Mar. 2004.
[8] U. Berthold, F. Jondral, S. Brandes, and M. Schnell, "OFDM-based overlay systems: A promising approach for enhancing spectral efficiency [Topics in radio communications]," *IEEE Commun. Mag.*, vol. 45, no. 12, pp. 52–58, Dec. 2007.
[9] S. Chuprun, J. Kleider, and C. Bergstrom, "Emerging software defined radio architectures supporting wireless high data rate OFDM," in *Proc. IEEE RAWCON*, 1999, pp. 117–120.
[10] T. Weiss, J. Hillenbrand, A. Krohn, and F. Jondral, "Mutual interference in OFDM-based spectrum pooling systems," in *Proc. IEEE VTC*, May 2004, vol. 4, pp. 1873–1877.
[11] B. Farhang-Boroujeny and R. Kempter, "Multicarrier communication techniques for spectrum sensing and communication in cognitive radios," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 80–85, Apr. 2008.
[12] G. Bansal, M. Hossain, and V. Bhargava, "Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4710–4718, Nov. 2008.
[13] T. Qin and C. Leung, "Fair adaptive resource allocation for multiuser OFDM cognitive radio systems," in *Proc. 2nd Int. Conf. CHINACOM*, Aug. 2007, pp. 115–119.
[14] P. Wang, M. Zhao, L. Xiao, S. Zhou, and J. Wang, "Power allocation in OFDM-based cognitive radio systems," in *Proc. IEEE GLOBECOM*, Dec. 2007, pp. 4061–4065.
[15] A. T. Hoang and Y.-C. Liang, "Downlink channel assignment and power control for cognitive networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3106–3117, Aug. 2008.
[16] P. Cheng, Z. Zhang, H.-H. Chen, and P. Qiu, "Optimal distributed joint frequency, rate and power allocation in cognitive OFDMA systems," *IET Commun.*, vol. 2, no. 6, pp. 815–826, Jul. 2008.
[17] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
[18] Z. Hasan, E. Hossain, C. Despins, and V. K. Bhargava, "Power allocation for cognitive radios based on primary user activity in an OFDM system," in *Proc. IEEE GLOBECOM*, Dec. 2008, pp. 1–6.
[19] U. Varshney, "Multicast over wireless networks," *Commun. ACM*, vol. 45, no. 12, pp. 31–37, Dec. 2002.
[20] H. Won, H. Cai, D. Eun, K. Guo, R. Netravali, I. Rhee, and K. Sabnani, "Multicast scheduling in cellular data networks," in *Proc. IEEE INFOCOM*, 2007, pp. 1172–1180.
[21] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.

[22] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.
[23] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.
[24] A. J. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.
[25] J. Liu, W. Chen, Z. Cao, and K. B. Letaief, "Dynamic power and subcarrier allocation for OFDMA-based wireless multicast systems," in *Proc. IEEE ICC*, May 2008, pp. 2607–2611.
[26] C. Suh and J. Mo, "Resource allocation for multicast services in multicarrier wireless communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 27–31, Jan. 2008.
[27] V. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, Sep. 1999.
[28] D. P. Bertsekas, *Nonlinear Programming,* 2nd ed. Boston, MA: Athena Scientific, 1999.
[29] P. Hande, Z. Shengyu, and C. Mung, "Distributed rate allocation for inelastic flows," *IEEE/ACM Trans. Netw.*, vol. 15, no. 6, pp. 1240–1253, Dec. 2007.
[30] H. Tuy, "D.C. optimization: Theory, methods and algorithms," in *Handbook of Global Optimization*. Norwell, MA: Kluwer, 1995, pp. 149–216.
[31] L. Hoo, J. Tellado, and J. Cioffi, "FDMA-based multiuser transmit optimization for broadcast channels," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2000, vol. 2, pp. 597–602.

Duy T. Ngo (S'08) received the B.Eng. degree (with first-class honors) in telecommunications engineering from the University of New South Wales (UNSW), Sydney, NSW, Australia, in 2007 and the M.Sc. degree in communication engineering from the University of Alberta, Edmonton, AB, Canada, in 2009. He is currently working toward the Ph.D. degree in electrical engineering with the Broadband Communications Research Laboratory, McGill University, Montreal, QC, Canada.

His research interests include applications of optimization in the resource allocation of wireless communications systems.

Mr. Ngo received the Australian Development Scholarship for his undergraduate studies in Australia. In 2006, he was the recipient of the National Information and Communication Technology Australia Telecommunications Excellence Award. Upon graduation from UNSW, he was awarded the University Medal—the highest honor given to the best graduand. At the University of Alberta, he was the holder of the Alberta Ingenuity Foundation Student Scholarship and the Informatics Circles of Research Excellence Information and Communication Technology Graduate Student Award.

Chintha Tellambura (M'97–SM'02) received the B.Sc. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.Sc. degree in electronics from the University of London, London, U.K., in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993.

He was a Postdoctoral Research Fellow with the University of Victoria from 1993 to 1994 and the University of Bradford, Bradford, U.K., from 1995 to 1996. From 1997 to 2002, he was with Monash University, Melbourne, VIC, Australia. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His research interests include coding, communication theory, modulation, equalization, and wireless communications.

Dr. Tellambura is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the Area Editor on wireless communications theory and systems for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was the Chair of the Communication Theory Symposium at the 2005 Global Communications Conference in St. Louis, MO.

Ha H. Nguyen (M'01–SM'05) received the B.Eng. degree in electrical engineering from Hanoi University of Technology, Hanoi, Vietnam, in 1995, the M.Eng. degree in electrical engineering from the Asian Institute of Technology, Bangkok, Thailand, in 1997, and the Ph.D. degree in electrical engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2001.

Since 2001, he has been with the Department of Electrical Engineering, University of Saskatchewan, Saskatoon, SK, Canada, where he is currently a Full Professor. He is the holder of adjunct appointments with the Department of Electrical and Computer Engineering, University of Manitoba, and TRLabs, Saskatoon. From October 2007 to June 2008, he was a Senior Visiting Fellow with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, Australia. His research interests include digital communications, spread-spectrum systems, and error-control coding. He is a coauthor, with E. Shwedyk, of the textbook *A First Course in Digital Communications* (Cambridge University Press, 2009).

Dr. Nguyen is a Registered Member of the Association of Professional Engineers and Geoscientists of Saskatchewan. He is currently an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on May 18,2010 at 19:50:43 UTC from IEEE Xplore. Restrictions apply.