

# Power Allocation in Spectrum Sharing Cognitive Radio Networks with Quantized Channel Information

YuanYuan He, *Student Member, IEEE*, and Subhrakanti Dey, *Senior Member, IEEE*

**Abstract**—We consider a wideband spectrum sharing system where a secondary user can access a number of orthogonal frequency bands each licensed to a distinct primary user. We address the problem of optimum secondary transmit power allocation for its ergodic capacity maximization subject to an average sum (across the bands) transmit power constraint and individual average interference constraints on the primary users. The major contribution of our work lies in considering quantized channel state information (CSI) (for the vector channel space consisting of all secondary-to-secondary and secondary-to-primary channels) at the secondary transmitter as opposed to the prevalent assumption of full CSI in most existing work. It is assumed that a central entity called a cognitive radio network manager has access to the full CSI information from the secondary and primary receivers and designs (offline) an optimal power codebook based on the statistical information (channel distributions) of the channels and feeds back the index of the codebook to the secondary transmitter for every channel realization in real-time, via a delay-free noiseless limited feedback channel. A modified Generalized Lloyds-type algorithm (GLA) is designed for deriving the optimal power codebook, which is proved to be globally convergent and empirically consistent. An approximate quantized power allocation (AQPA) algorithm is presented, that performs very close to its GLA based counterpart for large number of feedback bits and is significantly faster. We also present an extension of the modified GLA based quantized power codebook design algorithm for the case when the feedback channel is noisy. Numerical studies illustrate that with only 3-4 bits of feedback per band, the modified GLA based algorithms provide secondary ergodic capacity very close to that achieved by full CSI and with only as little as 4 bits of feedback per band, AQPA provides a comparable performance, thus making it an attractive choice for practical implementation.

**Index Terms**—Cognitive radio, spectrum sharing, fading channels, ergodic capacity, finite-rate feedback.

## I. INTRODUCTION

**R**ADIO spectrum is a limited and precious natural resource, which, traditionally, is licensed to users by regulatory authorities in a very rigid manner where in order to avoid interference, the licensed owner has an exclusive right to access the allocated frequency band [1]. Consequently, as the number of wireless communication systems and services grows, the availability of vacant spectrum becomes severely

scarce. However, recent measurements by the Federal Communications Commission reveal that many portions of spectrum are mostly under utilized or even unoccupied. This led to the idea of cognitive radio (CR) technology, originally introduced by J. Mitola [2], which holds tremendous promise to dramatically improve the efficiency of spectral utilization. The key idea behind CR is that an unlicensed/secondary user (SU) is allowed to communicate over the frequency band originally licensed to a primary user (PU), as long as the transmission of SU does not generate unfavorable impact on the operation of PU. Effectively, three categories of CR network paradigms have been proposed: interweave, overlay, and underlay [3]. In the underlay systems, which is the focus of this paper, the SU can transmit even when the PU is present, but the transmitted power of SU should be controlled properly so as to ensure that the resulting interference does not degrade the received signal quality of PU to an undesirable level [4] by imposing the so called interference temperature [1] constraints at PU (average or peak interference power (AIP/PIP) constraint). This type of CR is also known as the 'spectrum sharing' [1] model.

The behavior of capacities of various types of additive white Gaussian noise (AWGN) channels under received-power constraints at the PU receiver (PU-RX) was first studied in [6], which showed that for point-to-point non-fading AWGN channels, the capacity performances with transmit and received power constraints are essentially similar. The ergodic capacity of narrow band spectrum sharing model with one SU and one or more PU under either AIP or PIP constraint at PU-RX in various fading environments was studied by [1], illustrating that in a fading environment, spectrum access opportunity for the SU significantly increases compared to the AWGN case. In [7], the authors studied optimum power allocation for three different capacity notions under both AIP and PIP constraints. The optimal power allocation strategies for maximizing secondary ergodic capacity and outage capacity under various combinations of secondary transmit power constraints and interference constraints were derived in [4].

Most of the above results assume perfect knowledge of full CSI including the SU-TX to PU-RX channels, which is hard to realize in practice. A few recent papers have emerged that address this concern by investigating capacity analysis with imperfect CSI. The effect of imperfect channel information for the secondary to primary channels has been investigated in [8] by considering the channel information as a noisy estimate of the true CSI, and [9] has proposed a practical design paradigm for cognitive beamforming based on finite-rate cooperative

Paper approved by S. Roy, the Editor for Cross-Layer Wireless/Cognitive Radio Networks of the IEEE Communications Society. Manuscript received June 5, 2010; revised November 7, 2010 and January 15, 2011.

Y. He and S. Dey are with the Department of Electrical and Electronic Engineering, University of Melbourne, Vic. 3010, Australia (e-mail: {yyhe, s.dey}@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/TCOMM.2011.042811.100337

feedback from the PU-RX to the SU-TX. Another recent work [10] also considers imperfect CSI for the SU-TX to PU-RX channel in the form of noisy channel estimate and quantized channel information and investigates the effect of such imperfect CSI on the capacity performance of the secondary user, while assuming that the SU-TX has full knowledge of the SU-TX to SU-RX channel. Finally, [11] studies the issue of channel quantization for resource allocation via the framework of utility maximization in OFDMA based cognitive radio networks, but does not investigate the joint channel partitioning and rate/power codebook design problem. Indeed, the lack of a rigorous and systematic design methodology for quantized resource allocation algorithms in the context of cognitive radio networks forms the key motivation for our work.

In this paper, we consider the uplink of an *infrastructure-based* wideband spectrum sharing system where one SU shares  $M$  different frequency bands with  $M$  PU's, each PU using a separate band. We address the problem of ergodic capacity maximization of the secondary user subject to an average sum (across the bands) transmit power constraint on the secondary user and individual average interference constraints on the primary users, using quantized channel information about the vector channel space consisting of all SU-TX to SU-RX (contained in the SU Base station (SU-BS)) channels and all SU-TX to PU-RX (contained in the primary base station (PU-BS)) channels. To this end, we assume the availability of an entity called CR network manager who has access to the full CSI including all secondary-to-secondary and secondary-to-primary channels via (possibly fibre-optic) links with the primary and secondary base stations, which in turn are assumed to have receiver side full CSI of the secondary-to-primary and secondary-to-secondary channels, respectively. The CR network manager designs (offline) an optimal power codebook based on the statistical information (channel distributions) of the channels and in real-time, feeds back the index of the codebook to the secondary transmitter for every channel realization, via a  $B$ -bits per band finite-rate feedback link. The secondary transmitter then uses the corresponding power code vector for its transmission.

We make the following key contributions: (1) First and foremost, we present a modified Generalized Lloyd's type algorithm (GLA) for designing the optimal power codebook using quantized channel information. For easier exposition, we focus on the narrowband case first and present the quantized power allocation algorithm, where we prove that the modified GLA based power codebook design algorithm is globally convergent and empirically consistent. We provide a number of useful and interesting properties of the quantized powers. Then we present a complete description of the optimal power codebook design algorithm for the wideband spectrum sharing case under the average transmit power and average interference constraints. We believe this paper is the first to provide a systematic quantized power allocation algorithm with limited feedback for the spectrum sharing scenario in cognitive radio. (2) Although an offline algorithm, GLA based quantizer designs usually require a large number of training samples and can be computationally expensive. We therefore design an approximate quantized power allocation algorithm

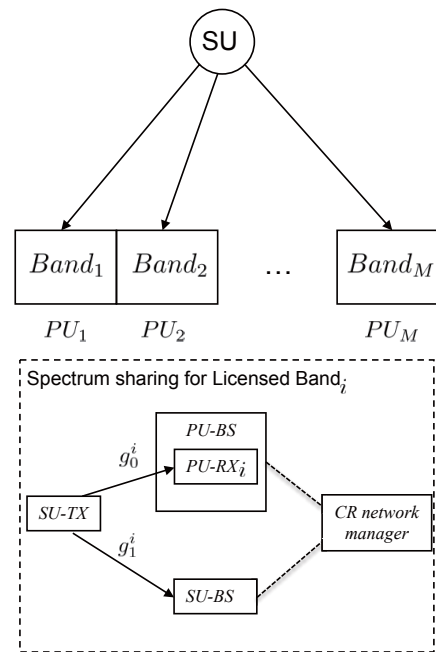


Fig. 1. System model for wideband spectrum-sharing scenario.

based on the derived properties of the power codebook, which is computationally significantly faster and is seen to have a far superior performance compared to other suboptimal algorithms. (3) We then generalize the modified GLA based algorithm for quantized power allocation algorithm to the case where the limited feedback channel is noisy but memoryless. (4) We present a comprehensive set of numerical results which illustrate that (i) the modified GLA-based power codebook can achieve a secondary ergodic capacity with only 3-4 bits of feedback at each band, that is very close to the capacity with full CSI, (ii) the performance of the approximate quantized power allocation algorithm is almost indistinguishable from that of the GLA-based algorithm with  $B \geq 4$  bits per band of feedback and (iii) how the performance of the quantized power allocation degrades when the noisy feedback channel error probability increases.

The rest of the paper is organized as follows. Section II presents the system model and assumptions about the spectrum sharing problem with limited feedback. In Section III, we present the modified GLA based quantized power codebook design algorithms along with associated convergence results and some useful properties of the quantized power code vectors. The approximate quantized power allocation algorithm (AQPA) and two other suboptimal algorithms are also presented. In Section IV, we extend the modified GLA based power codebook design algorithm to a noisy limited feedback channel model. Numerical results are presented in Section V and finally, concluding remarks and possible extensions are presented in Section VI. All proofs are relegated to the Appendix unless otherwise mentioned.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider the uplink of a wideband spectrum sharing scenario in an infrastructure-based cognitive radio network

with one SU and Multiple PUs, as shown in Fig. 1, where a SU is allowed to use  $M$  parallel orthogonal frequency bands ( $Band_1$  to  $Band_M$ ) which are individually licensed to  $PU_1, \dots, PU_M$  respectively. Regardless of the ON/OFF status of  $PU_i$ , SU uses the  $i$ -th channel as long as the impact of the secondary transmission does not substantially degrade the received signal quality of  $PU_i$ . It is assumed that the channels between the secondary transmitter (SU-TX) and secondary receiver (SU-RX) and those between the secondary transmitter and the each primary receiver are all block fading channels. Let  $g_0^i \in \mathbb{R}_+$  and  $g_1^i \in \mathbb{R}_+$  denote the real-valued instantaneous channel power gains for the link between the SU-TX and the receiver of  $PU_i$  (at the PU-BS) and  $i$ -th channel between the SU-TX and SU-RX (at the SU-BS) <sup>1</sup>, respectively, where  $\mathbb{R}_+$  denotes the set of nonnegative real numbers. These channels are assumed to be stationary ergodic with absolutely continuous probability density functions (pdf)  $f_0(g_0^i)$  and  $f_1(g_1^i)$ . For analytical simplicity, the interference from  $PU_i$ -TX to SU-RX is neglected (similarly as in [1], [4]). In the case where the interference caused by the primary transmitter at the secondary receiver is significant, the SU ergodic capacity results derived in this paper can be taken as upper bounds on the actual capacity under primary-induced interference. All  $g_0^i$  and  $g_1^i$  ( $i = 1, \dots, M$ ) are statistically mutually independent, and without loss of generality <sup>2</sup> (*w.l.o.g.*) are assumed to have unity mean. Similarly, additive noises for each channel are independent Gaussian random variables with zero mean and unit variance *w.l.o.g.* When  $M = 1$ , this system becomes a typical narrowband spectrum sharing model considered in [1][12][4].

Given a channel realization  $\mathbf{g}_0 \triangleq \{g_0^1, \dots, g_0^M\}$  and  $\mathbf{g}_1 \triangleq \{g_1^1, \dots, g_1^M\}$ , we assume that a channel state information (CSI)  $\eta(\mathbf{g}_0, \mathbf{g}_1)$  is available at the SU-TX. The power allocated at the SU-TX on the  $M$  parallel SU links is represented by the vector  $\mathbf{p}(\eta(\mathbf{g}_0, \mathbf{g}_1)) = \{p_1(\eta(\mathbf{g}_0, \mathbf{g}_1)), \dots, p_M(\eta(\mathbf{g}_0, \mathbf{g}_1))\}$ , the ergodic capacity of the SU for this wideband spectrum sharing system can be expressed as

$$C = \frac{1}{M} \sum_{i=1}^M E \left[ \log \left( 1 + g_1^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \right) \right] \quad (1)$$

where, for simplicity, we have ignored the factor  $\frac{1}{2}$  at the front of the capacity expression and  $\log$  represents the natural logarithm. Note also that the factor  $1/M$  in front of the capacity formula above is not strictly necessary for the problem formulation and since  $M$  is fixed for the problem, inclusion or exclusion of it does not change the solution techniques and the results (except for a scaling factor). However, channel capacity is usually expressed in terms of per degree of freedom and in order to have a fair comparison of capacities for various values

<sup>1</sup>Fig. 1 also shows that the PU-BS and SU-BS are connected (possibly via fibre-optic links) to a central controller called the CR network manager, the existence of which is crucial in designing the power allocation algorithms in the quantized feedback case (see Section III for further details).

<sup>2</sup>Note that here the secondary-to-primary and secondary-to-secondary fading channel power gains are modelled as  $g_p^i = m_0^i g_0^i$  and  $g_s^i = m_1^i g_1^i$ , respectively where  $m_0^i, m_1^i$  are the corresponding means, and  $g_0^i, g_1^i$  have unity mean. It can be easily seen from (2) that any non-unity mean  $m_0^i$  can be absorbed into the average interference threshold  $Q_{avg}^i$  and any non-unity mean  $m_1^i$  can be absorbed into  $p_i(\cdot)$ , the secondary transmission power on the  $i$ -th band.

of  $M$ , it is common practice to normalize it by the number of degrees of freedom, in this case,  $M$ . A similar formulation can be found in the paper [5].

A common way to protect PU's received signal quality is by imposing either an average or a peak interference power (AIP/PIP) constraint at PU-RX [1][12][4], although there are other forms of PU quality of service constraints such as PU's capacity loss and PU's outage probability [13]. It was shown in [12] that an AIP constraint is more favorable than a peak constraint especially in the context of transmission over fading channels, since the AIP constraint is more flexible and can achieve a larger SU capacity with less PU capacity loss than that achieved by PIP.

Motivated by this observation, we consider the following optimal power allocation problem that maximizes the ergodic capacity of SU in a wideband spectrum sharing scenario, under an AIP constraint at each  $PU_i$ -RX and an average sum transmit power constraint (ATP) for the SU, given by,

$$\begin{aligned} \max_{p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \geq 0, \forall i} \quad & \frac{1}{M} \sum_{i=1}^M E \left[ \log \left( 1 + g_1^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \right) \right] \\ \text{s.t.} \quad & E \left[ g_0^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \right] \leq Q_{avg}^i, \quad \forall i, \\ & \frac{1}{M} \sum_{i=1}^M E \left[ p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \right] \leq P_{avg}. \end{aligned} \quad (2)$$

When full CSI is available at the SU-Tx (i.e.  $\eta(\mathbf{g}_0, \mathbf{g}_1) = (\mathbf{g}_0, \mathbf{g}_1)$ ), the optimization problem is convex and the corresponding optimal secondary transmitter power allocation policy is given by the following Theorem.

**Theorem 1:** With perfect channel information  $\eta(\mathbf{g}_0, \mathbf{g}_1) = (\mathbf{g}_0, \mathbf{g}_1)$  at the SU-TX, the optimal power allocation for problem (2) is given by the equation at the top of the next page, where  $(x)^+ = \max(x, 0)$  and  $\lambda^f, \mu_i^f$  are the nonnegative Lagrange multipliers associated with the ATP constraint and the AIP constraint of  $PU_i$  respectively, and condition (a) corresponds to the case  $\lambda^f = 0, \mu_i^f > 0, \forall i$  and  $\mu_i^f$  is determined by solving  $E \left[ g_0^i \left( \frac{1}{\mu_i^f g_0^i} - \frac{1}{g_1^i} \right)^+ \right] = Q_{avg}^i$ , condition (b) corresponds to the case  $\lambda^f > 0, \mu_i^f = 0$  where  $\lambda^f$  is determined by solving  $\frac{1}{M} \sum_{i=1}^M E \left[ \left( \frac{1}{\lambda^f} - \frac{1}{g_1^i} \right)^+ \right] = P_{avg}$ , and condition (c) corresponds to the case  $\lambda^f > 0, \mu_i^f > 0$  such that  $\lambda^f$  and  $\mu_i^f$  are determined by solving  $\frac{1}{M} \sum_{i=1}^M E \left[ \left( \frac{1}{\lambda^f + \mu_i^f g_0^i} - \frac{1}{g_1^i} \right)^+ \right] = P_{avg}$  and  $E \left[ g_0^i \left( \frac{1}{\lambda^f + \mu_i^f g_0^i} - \frac{1}{g_1^i} \right)^+ \right] = Q_{avg}^i$ .

This result is a minor extension of the corresponding narrowband result in [4] and follows using the necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions for the convex optimization problem. For a proof of this result, and associated special cases and numerical studies, see an extended online version of this paper [14].

### III. OPTIMUM QUANTIZED POWER CONTROL WITH FINITE-RATE FEEDBACK

The assumption of full CSI at the SU-TX (especially that of  $\mathbf{g}_0$ ) is usually unrealistic in practical systems. In this section, we are therefore interested in designing power allocation

$$p_i^*(\mathbf{g}_0, \mathbf{g}_1) = \begin{cases} \left(\frac{1}{\mu_i^f g_0^i} - \frac{1}{g_1^i}\right)^+ & \text{iff } P_{avg} \geq \frac{1}{M} \sum_{i=1}^M E \left[ \left(\frac{1}{\mu_i^f g_0^i} - \frac{1}{g_1^i}\right)^+ \right] \quad (a) \\ \left(\frac{1}{\lambda^f} - \frac{1}{g_1^i}\right)^+ & \text{iff } E \left[ \left(\frac{1}{\lambda^f} - \frac{1}{g_1^i}\right)^+ \right] \leq Q_{avg}^i \quad (b) \\ \left(\frac{1}{\lambda^f + \mu_i^f g_0^i} - \frac{1}{g_1^i}\right)^+ & \text{otherwise} \quad (c) \end{cases}$$

schemes based on quantized  $(\mathbf{g}_0, \mathbf{g}_1)$  information acquired via a no-delay and error-free feedback link with limited rate. As shown in Fig. 1, here we assume that there is a central controller termed as the CR network manager who can obtain perfect information on  $\mathbf{g}_1$  from SU-RX at the SU base station and perfect information on  $\mathbf{g}_0$  from the PU base station, possibly over fibre-optic links and then forward some appropriately quantized CSI to SU-TX (and SU-RX for decoding purposes) through a finite-rate feedback link. Note that existence of such central controllers is also assumed quite commonly in literature on multi-cell MIMO or macro-diversity based systems with cooperative base stations in a primary network, where several base stations are assumed to be connected to a central controller via a backhaul link so that information about out-of-cell interference can be obtained resulting in higher capacity [15], [16]. If cognitive radio networks are to be successful, it is imperative that the primary users cooperate with the CR service provider at some level. In our model, we assume this cooperation in terms of primary channel information sharing with the CR network manager via the PU-BS. Any cost incurred by the CR service provider as a result of obtaining this information from the PU-BS can be recovered by charging the secondary users a nominal price. This assumption is further (and perhaps more crucially) justified by the fact that having full CSI about the joint channel space  $\mathbf{g}_0, \mathbf{g}_1$  allows the CR network manager to design a joint channel quantizer and power codebook that has a far superior performance than quantizing  $\mathbf{g}_0$  and  $\mathbf{g}_1$  separately in the absence of this central controller, which is clearly suboptimal (see Section V on *Numerical Results* for further details).

In order to formulate the optimal power allocation with quantized channel feedback, we first make the observation that due to convexity of the original problem (2), it can be solved by the Lagrange duality method, namely, by solving the dual problem of (2):  $\min_{\lambda^f \geq 0, \mu_i^f \geq 0, \forall i} g(\lambda^f, \{\mu_i^f\}) + \lambda^f P_{avg} + \sum_{i=1}^M \mu_i^f Q_{avg}^i$ , where the Lagrange dual function  $g(\lambda^f, \{\mu_i^f\})$  is given by

$$\max_{p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \geq 0, \forall i} \frac{1}{M} \sum_{i=1}^M E[\log(1 + g_1^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)))] - \lambda^f p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) - M \mu_i^f g_0^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)). \quad (3)$$

Here  $g(\lambda^f, \{\mu_i^f\})$  can be decomposed into  $M$  parallel sub-problems, which for band  $i, i = 1, \dots, M$  is given by (4) at the top of the next page, which implies given  $\lambda^f$  and  $\mu_i^f, i = 1, 2, \dots, M$ , we can solve the above problem individu-

ally for each band. This observation motivates us to formulate Problem (2) with quantization channel feedback individually for each band as described below.

Under the network modelling assumptions described above, given a  $B$ -bit per band limited feedback link between the CR network manager and the the SU-TX, the power codebook for  $i$ -th band ( $i = 1, \dots, M$ ) given by  $\mathcal{P}_i = \{p_{i1}, \dots, p_{iL}\}$  of cardinality  $L = 2^B$ , is designed by the CR network manager off-line purely on the basis of the statistics of  $g_0^i, g_1^i$ . These codebooks  $\mathcal{P}_1, \dots, \mathcal{P}_M$  are known *a priori* by both SU-TX and CR network manager. For the  $i$ -th band, the vector space of  $(g_0^i, g_1^i)$  is thus partitioned into  $L$  regions  $\mathcal{R}_1^i, \dots, \mathcal{R}_L^i$  using a quantizer  $\mathcal{Q}_i$  (such that the codebook element  $p_{ij}$  represents the power level used in  $\mathcal{R}_j^i$ ). For the  $i$ -th band, the CR network manager maps the current instantaneous  $(g_0^i, g_1^i)$  information into one of  $L$  integer indices and sends the corresponding index to the SU-TX via the feedback link of rate  $B = \log_2 L$  (e.g., if the current  $(g_0^i, g_1^i)$  falls in  $\mathcal{R}_j^i$ , then index  $j$  for  $i$ -th band will be conveyed back to SU-TX). The SU-TX will use the associated power codebook element  $p_{ij}$  as the transmission power to adapt its transmission strategy for the band  $i$ .

*Remark 1:* Note that it is possible to consider different feedback bit rates for different bands and our analysis can be adapted to this scenario, but for simplicity and also due to the fact that all bands are assumed to be statistically *i.i.d.*, we use identical  $B$  bits of feedback for each band.

Let  $Pr(\mathcal{R}_j^i), E[\bullet|\mathcal{R}_j^i]$  denote  $Pr((g_0^i, g_1^i) \in \mathcal{R}_j^i)$  (the probability that  $(g_0^i, g_1^i)$  falls in the region  $\mathcal{R}_j^i$ ) and  $E[\bullet|(g_0^i, g_1^i) \in \mathcal{R}_j^i]$ , respectively. Then the secondary ergodic capacity maximization problem (2) with limited feedback can be formulated as

$$\begin{aligned} \max_{p_{ij} \geq 0, \forall i, j} & \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E[\log(1 + g_1^i p_{ij})|\mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \right) \\ \text{s.t.} & \sum_{j=1}^L E[g_0^i p_{ij}|\mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \leq Q_{avg}^i, \forall i, \\ & \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E[p_{ij}|\mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \right) \leq P_{avg}. \end{aligned} \quad (5)$$

Our objective is thus the joint optimization of the channel partition regions and the power codebooks such that the ergodic capacity of SU is maximized under the above average transmit power and average interference constraints.

$$\max_{p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) \geq 0, \forall i} E[\log(1 + g_1^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1))) - \lambda^f p_i(\eta(\mathbf{g}_0, \mathbf{g}_1)) - M \mu_i^f g_0^i p_i(\eta(\mathbf{g}_0, \mathbf{g}_1))] \quad (4)$$

### A. Narrowband Spectrum-Sharing Case

For ease of exposition, we first look at the relatively simpler case of  $M = 1$  (where SU shares a narrowband spectrum with only one PU). For simplicity (with some abuse of notation), let  $p_j, g_1, g_0, Q_{avg}, \mathcal{R}_j$  represent  $p_{1j}, g_1^1, g_0^1, Q_{avg}^1, \mathcal{R}_j^1$  respectively. Thus problem (5) with  $M = 1$  becomes,

$$\begin{aligned} \max_{p_j \geq 0, \forall j} & \sum_{j=1}^L E[\log(1 + g_1 p_j) | \mathcal{R}_j] Pr(\mathcal{R}_j) \\ \text{s.t.} & \sum_{j=1}^L E[g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j) \leq Q_{avg}, \\ & \sum_{j=1}^L p_j Pr(\mathcal{R}_j) \leq P_{avg} \end{aligned} \quad (6)$$

We solve the problem (6) based on the Lagrange duality method. First we write the Lagrangian of above problem as  $L(P, \lambda, \mu) = \sum_{j=1}^L E[\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j) + \lambda P_{avg} + \mu Q_{avg}$  where  $\lambda$  and  $\mu$  are the nonnegative Lagrange multipliers associated with the ATP constraint and AIP constraint respectively. The Lagrange dual function  $g(\lambda, \mu)$  is defined as

$$\max_{p_j \geq 0} \sum_{j=1}^L E[\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j) \quad (7)$$

and the corresponding dual problem is  $\min_{\lambda \geq 0, \mu \geq 0} g(\lambda, \mu) + \lambda P_{avg} + \mu Q_{avg}$ .

We first consider solving the above primal optimization problem with fixed  $\lambda$  and  $\mu$ . To this end, we employ an algorithm similar to a Generalized Lloyd Algorithm (GLA) [17], [18] to design an optimal power codebook for (7), which is based on two optimality conditions: 1) optimum channel partitioning for a given codebook, also called the nearest neighbor condition (NNC) in the context of traditional vector quantization (VQ), and 2) optimum codebook design for a given partition, also known as the centroid condition (CC) (in the context of VQ) [18]. GLA is usually initialized with a random choice of codebook, and then the above two conditions are iterated until some pre-specified convergence criterion is met. The same procedure is used here for designing an optimal quantizer  $\mathcal{Q}$ , but the design criterion for our case is minimizing the difference between the capacity with perfect CSI and the capacity with quantized power allocation under the given constraints. This amounts to designing an optimal power codebook  $\mathcal{Q}$  that maximizes the Lagrangian function for quantized CSI,  $\sum_{j=1}^L E[\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j)$ . We call the corresponding quantized power allocation algorithm for a given  $\lambda, \mu$  as a *modified GLA*.

In practice, this modified GLA is implemented using a sufficiently large number of training samples (channel realizations for  $g_0, g_1$ ). Beginning with a random initial codebook, one can design the optimal partitions using the fact that the optimal

partitions satisfy  $\mathcal{R}_j = \{(g_0, g_1) : (\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j) \geq (\log(1 + g_1 p_n) - \lambda p_n - \mu g_0 p_n), \forall n \neq j\}$  where  $\mathcal{R}_j$  is the corresponding partition region for power level  $p_j$  in the codebook, and ties are broken arbitrarily. Once the optimal partitions are designed, the new optimal power codebook is found by solving for  $\arg\max_{p_j \geq 0} E[\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j), \forall j = 1, 2, \dots, L$ . Given a partition, this optimization problem is convex and by using the KKT conditions, one can obtain the optimal power as  $\max(p_j^*, 0)$ , where  $p_j^*$  is the solution to the equation  $E[\frac{g_1}{1 + g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j] = 0$ . These two steps are repeated until the resulting ergodic capacity converges within a pre-specified accuracy. One needs to note that GLA cannot in general guarantee global optimality, since the two optimality conditions (NNC and CC) mentioned above are just necessary conditions [18]. Thus it is very likely that the our resulting quantizer is only locally optimal. While convergence (to a local optimum) of our modified GLA follows immediately by noting that the Lagrangian  $\sum_{j=1}^L E[\log(1 + g_1 p_j) - \lambda p_j - \mu g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j)$  is non-decreasing at each iteration and is upper bounded (due to the finite average transmit power and average interference constraints), it is important and instructive to state a more formal result along the lines of [19]. Since GLA is initialized with a random codebook and the optimal partitions and codevectors are found using training samples drawn from empirical distributions, it is crucial that GLA is globally convergent with respect to the choice of initial codebooks and empirically consistent. For more formal definitions of these two properties, see [19]. Under the assumption of absolutely continuous fading distributions for  $g_0, g_1$  and mild regularity assumptions satisfied by these distributions, one can show that the modified GLA satisfies the conditions for global convergence and empirical consistency stated in [19] and thus we have the following result:

*Theorem 2:* The modified GLA that solves the optimization problem (7) satisfies the global convergence and empirical consistency properties of [19].

Next, we present some useful properties of the optimal power solutions obtained via the modified GLA. We use the partitions  $\mathcal{R}_1, \dots, \mathcal{R}_L$  and the corresponding power levels  $p_1, \dots, p_L$  to denote the optimal solutions after convergence. First, we state the following Lemma without proof (the proof is straightforward and can be found in [14]).

*Lemma 1:* Given partitions  $\mathcal{R}_1, \dots, \mathcal{R}_L$  and the corresponding power level  $p_1, \dots, p_L$ , (where  $\mathcal{R}_j$  and  $\mathcal{R}_{j+1}, \forall j \in \{1, \dots, L-1\}$  are adjacent regions and  $p_j \neq p_{j+1}$ ), the boundary between any two adjacent regions  $\mathcal{R}_j$  and  $\mathcal{R}_{j+1}$  is given by,  $g_1 = \frac{e^{(\lambda + \mu g_0)(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{(\lambda + \mu g_0)(p_j - p_{j+1})}}$ . When  $\mu \neq 0$ ,  $g_1$  is a monotonically increasing convex function of  $g_0$  and as  $g_1 \rightarrow \infty, g_0 \rightarrow \frac{1}{\mu} \left( \frac{\log(\frac{p_j}{p_{j+1}})}{p_j - p_{j+1}} - \lambda \right)$ .

*Remark 2:* In case  $\lambda > 0, \mu = 0$ , the AIP constraint is inactive and the ATP constraint is satisfied with equality. In

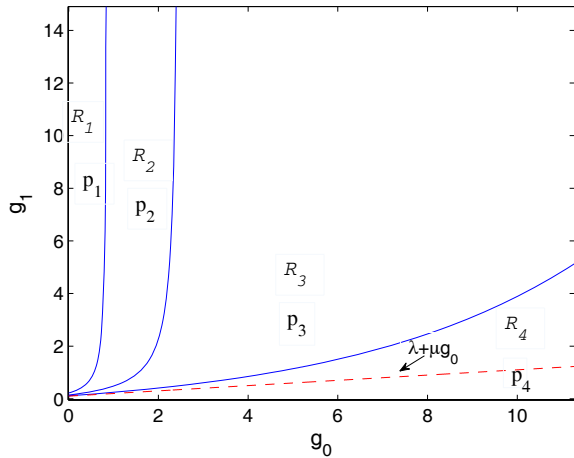


Fig. 2. The structure of optimum partition regions with  $B = 2$  bits of feedback given  $\lambda = \mu = 0.1$ ,  $P_{av} = 5.6336$  dB,  $Q_{av} = 3.4492$  dB.

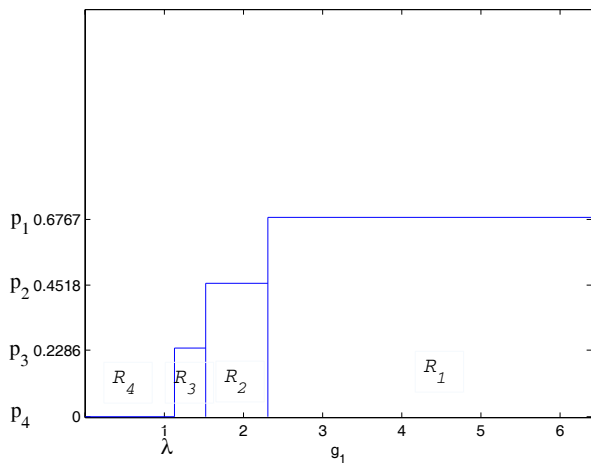


Fig. 3. The structure of optimum partition regions with  $B = 2$  bits of feedback given  $\lambda = 1, \mu = 0$ ,  $P_{av} = -8.3298$  dB.

this case, the boundary between any two adjacent regions  $\mathcal{R}_j$  and  $\mathcal{R}_{j+1}$  becomes  $g_1 = \frac{e^{\lambda(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{\lambda(p_j - p_{j+1})}}$ . Clearly, Problem (5) reduces to an ergodic capacity maximization problem with quantized channel information. For the narrowband case, it becomes a scalar quantization problem involving quantizing  $g_1$  only. Note that while for the narrowband case, this no longer pertains to a cognitive radio problem, these properties of the optimal quantized power allocation scheme are still important for the wideband case ( $M > 1$ ). This is due to the fact that in the wideband case, it is possible that for a specific (say the  $i$ -th) channel, the AIP constraint is inactive ( $\mu_i > 0$ ) while  $\lambda > 0$ . See Section III-B for further details.

We now give an example to illustrate what the optimum partition regions actually look like. For this example,  $g_0$  and  $g_1$  are both exponentially distributed (Rayleigh fading) with unit mean and  $L = 4$  (2 bits of feedback). The optimum partition regions are as shown in Fig. 2 for  $\lambda > 0, \mu > 0$ , and Fig. 3 for  $\lambda > 0, \mu = 0$ .

We obtain the following properties for the optimal quantized

power levels where (as illustrated in Fig. 2) the regions  $\mathcal{R}_1, \mathcal{R}_2, \dots$  etc. are sequentially numbered, with  $\mathcal{R}_1$  being the region closest to the  $g_1$  axis and  $\mathcal{R}_L$  being the region closest to the  $g_0$  axis. Note that these properties apply regardless of whether  $\mu > 0$  or  $\mu = 0$ .

*Theorem 3:* i).  $p_1 > \dots > p_L$

ii). All boundaries between any two adjacent partitions satisfy  $g_1 > \lambda + \mu g_0$ .

iii). Given  $B$  bits of feedback (or  $L = 2^B$  regions), for the first  $L-1$  regions, we always have strictly positive power, i.e.  $p_1 > \dots > p_{L-1} > 0$ , whereas  $p_L$  is simply nonnegative, i.e.  $p_L \geq 0$ .

iv). When  $\lambda + \mu \geq 1$  (note that if  $\lambda = 0, \mu \geq 1$  implies  $Q_{avg} < 1$ , and if  $\mu = 0, \lambda \geq 1$  corresponds to  $P_{avg} < 1$ ), we always have  $p_L = 0$ . In addition, when  $L$  (the number of quantized regions) is sufficiently large, no matter what  $\lambda, \mu$  are,  $p_L$  must be 0. Additionally, as  $L \rightarrow \infty$  the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches  $g_1 = \lambda + \mu g_0$  and  $\lim_{L \rightarrow \infty} p_{L-1} = 0$ .

*Proof:* See Proof in Appendices B-E.

*Remark 3:* The above properties of optimal quantized power values are interesting for two reasons. From property ii), it is clear that  $(g_0, g_1) \in \mathcal{R}_j$  for  $j = 1, 2, \dots, L-1$  satisfy the property  $g_1 > \lambda + \mu g_0$  whereas for the region  $\mathcal{R}_L$ , this property may or may not be satisfied. Since the quantized power values in the first  $L-1$  regions are strictly positive, it is easy to relate this property to the corresponding property of the full CSI based optimal power value which is strictly positive if and only if when  $g_1 > \lambda^f + \mu^f g_0$ . Also, as  $L \rightarrow \infty$ , the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches  $g_1 = \lambda + \mu g_0$ , thus making this relationship between the quantized power allocation scheme and the full CSI power allocation scheme stronger. Finally, property iv) allows one to obtain an approximate quantized power allocation scheme (AQPA) for large  $L$  by setting  $p_L = 0$  and taking the limit as  $p_{L-1} \rightarrow 0$ . This is particularly useful as the modified GLA becomes computationally intensive for large  $L$ , whereas AQPA provides a performance that is extremely close to that of the modified GLA, while requiring very little computation time. A detailed description of the AQPA is provided in Section III-C followed by illustrative numerical simulations in Section V.

Based on Lemma 1 and Theorem 3, one can also obtain the optimal quantized power values for Problem (7) by solving the following set of nonlinear equations for  $p_1, p_2, p_3, \dots, p_L$ :

$$E \left[ \frac{g_1}{1 + g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j \right] = 0, \quad j = 1, \dots, L, \quad (8)$$

where  $p_L = \max(0, p_L)$ , and if  $\mu \neq 0$ ,  $E \left[ \frac{g_1}{1 + g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j \right] = \int_{c_j}^{\infty} \int_{r_{j-1}}^{r_j} \left( \frac{g_1}{1 + g_1 p_j} - (\lambda + \mu g_0) \right) f(g_0) f(g_1) dg_0 dg_1$ , with  $c_j = \frac{e^{\lambda(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{\lambda(p_j - p_{j+1})}}$ ,  $j = 1, \dots, L-1, c_L = 0$  and  $r_j = \frac{1}{\mu} \left( \frac{\log \frac{p_j * g_1 + 1}{p_{j+1} * g_1 + 1}}{p_j - p_{j+1}} - \lambda \right)$ ,  $j = 1, \dots, L-1, r_0 = 0, r_L = \infty$ . When  $\mu = 0$ ,  $E \left[ \frac{g_1}{1 + g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j \right] = \int_{c_j}^{c_{j-1}} \left( \frac{g_1}{1 + g_1 p_j} - \lambda \right) f(g_1) dg_1$ , with  $c_0 = \infty$ . Note that (8) can be solved efficiently by a suitable nonlinear equation solver. This particular solution methodology will be very useful in the case of AQPA (see Section III.C for further details).

We now solve the dual problem for finding the optimal values  $\lambda$  and  $\mu$ . Since the dual function (7) is convex in  $\lambda, \mu$ , we can find their optimal values by using an iterative sub-gradient based method [20], where  $\lambda$  and  $\mu$  are updated until convergence using  $\lambda^{l+1} = [\lambda^l - \alpha^l (P_{avg} - \sum_{j=1}^L p_j Pr(\mathcal{R}_j))]^+$  and  $\mu^{l+1} = [\mu^l - \beta^l (Q_{avg} - \sum_{j=1}^L E[g_0 p_j | \mathcal{R}_j] Pr(\mathcal{R}_j))]^+$  respectively, where  $l$  is the iteration number,  $\alpha^l, \beta^l$  are positive scalar step sizes for the  $l$ -th iteration and satisfy  $\sum_{l=1}^{\infty} \alpha_l = \infty$ ,  $\sum_{l=1}^{\infty} \beta_l = \infty$ ,  $\sum_{l=1}^{\infty} \alpha_l^2 < \infty$  and  $\sum_{l=1}^{\infty} \beta_l^2 < \infty$ , and  $[x]^+ \triangleq \max(x, 0)$ . Note that this method is guaranteed to converge to the global optimum of the dual function even though the primal problem is non-convex [20]. One can then repeat the modified GLA based algorithm for finding a locally optimum quantized power values for fixed  $\lambda, \mu$  and the sub-gradient based method for updating  $\lambda, \mu$  as described above until an overall convergence criterion is met. Note that since the primal problem is non-convex, the resultant power allocation and codebook design can only be guaranteed to be locally optimal. An algorithmic format for this procedure is provided for the more general wideband ( $M > 1$ ) case in the next subsection.

*Remark 4:* Note that the idea of using a Lagrangian based cost function as a modified distortion measure for optimizing via GLA is not new. It has been used for combined adaptive power control and beamforming for MIMO link capacity optimization with limited feedback in [21] and more recently for optimal power and rate allocation and scheduling in TDMA based wireless sensor networks with limited feedback in [22]. However, these papers do not consider average interference constraints and therefore their results cannot be applied to the problem of power allocation in cognitive radio networks with limited feedback. Generic convergence results for the specific GLA used in [21], [22] are also presented in these papers. However, what is unique in our paper (apart from the novel application to quantized power allocation for secondary throughput maximization in cognitive radio) is that the detailed global convergence and empirical consistency result presented in Theorem 2 and more importantly the properties of the quantized power allocation scheme detailed in Theorem 3, which are specific to the case of the cognitive radio problem. Finally, these properties are used to derive a novel approximate power allocation algorithm (AQPA), which is significantly faster than the GLA based algorithm and as will be seen later, performs very close to it with more than 4 bits of feedback.

### B. Wideband Spectrum-Sharing Case

The above algorithm for the narrowband case can be easily extended to the wideband case corresponding to the problem (5). For this scenario, the Lagrangian function becomes,

$$\begin{aligned} L(P, \lambda, \mathbf{u}) = & \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E \left[ \log(1 + g_1^i p_{ij}) | \mathcal{R}_j^i \right] Pr(\mathcal{R}_j^i) \right) \\ & - \lambda \left( \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E[p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \right) - P_{avg} \right) \\ & - \sum_{i=1}^M \mu_i \left( \sum_{j=1}^L E[g_0^i p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) - Q_{avg}^i \right) \end{aligned} \quad (9)$$

where  $\lambda$  and  $\mu_i$  are the nonnegative Lagrange multipliers associated with the ATP constraint and  $i$ th AIP constraint respectively. The Lagrange dual function  $g(\lambda, \{\mu_i'\})$  is defined as

$$\begin{aligned} \max_{p_{ij} \geq 0 \forall i,j} \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^L E[\log(1 + g_1^i p_{ij}) - \lambda p_{ij} \\ - \mu_i' g_0^i p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \end{aligned} \quad (10)$$

where  $\mu_i' = M\mu_i$ ,  $\forall i$ , and the dual problem is  $\min_{\lambda \geq 0, \mu_i' \geq 0, \forall i} g(\lambda, \{\mu_i'\}) + \lambda P_{avg} + \sum_{i=1}^M \frac{\mu_i'}{M} Q_{avg}^i$ .

Similar to the narrowband case, we first consider the problem (10) to obtain  $g(\lambda, \{\mu_i'\})$  with given  $\lambda$  and  $\{\mu_i'\}$ . As discussed before, problem (10) can be decomposed into  $M$  parallel subproblems, where for each band  $i, i = 1, \dots, M$

$$\begin{aligned} \max_{p_{ij} \geq 0 \forall j} \sum_{j=1}^L E[\log(1 + g_1^i p_{ij}) - \lambda p_{ij} \\ - \mu_i' g_0^i p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \end{aligned} \quad (11)$$

is defined as the sub-dual function  $g_i(\lambda, \mu_i')$  and  $g(\lambda, \{\mu_i'\}) = \frac{1}{M} \sum_{i=1}^M g_i(\lambda, \mu_i')$ . This kind of duality method is also known as the 'dual decomposition algorithm' [23]. Since each sub-problem (11) is similar to the problem (7) for the narrowband case and can be similarly solved by using a modified GLA.  $\lambda$  and  $\{\mu_i'\}$  can be also obtained in a manner similar to the narrowband case. These two steps are then repeated until a satisfactory convergence criterion is met. Due to the increased complexity resulting from the presence of multiple bands, we provide below a description of the overall optimization algorithm (Algorithm 1) for solving (5).

#### Algorithm 1:

- 1) Let  $\lambda = 0$ , then all  $\mu_i', i = 1, \dots, M$  must satisfy  $\mu_i' > 0$ . For each *band* <sub>$i$</sub> , starting with a random initial value for  $\mu_i'$ , obtain the corresponding optimal power codebook  $\mathcal{P}_i = \{p_{i1}, \dots, p_{iL}\}$  using a modified GLA, then update  $\mu_i'$  by using an iterative subgradient method  $\mu_i'(l+1) = [\mu_i'(l) - \beta_i^l (Q_{avg}^i - \sum_{j=1}^L E[g_0^i p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i))]^+$ , where  $l$  denotes the iteration number,  $\beta_i^l > 0$  is scalar step size for  $l$ -th iteration satisfying  $\sum_{l=1}^{\infty} \beta_i^l = \infty$  and  $\sum_{l=1}^{\infty} (\beta_i^l)^2 < \infty \forall i = 1, 2, \dots, M$ . Repeat these two steps until convergence resulting in  $M$  power codebooks  $\{\mathcal{P}_1, \dots, \mathcal{P}_M\}$  (one for each band). With these codebooks, if  $\frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E[p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \right) \leq P_{avg}$ , it is an optimal power codebook and stop; otherwise go to step 2).
- 2) If 1) is not satisfied, we must have  $\lambda > 0$ . Starting with a random initial value for  $\lambda$ : for each  $i$ , use the modified GLA to find an optimal power codebook first with  $\mu_i' = 0$ . If  $\sum_{j=1}^L E[g_0^i p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \leq Q_{avg}^i$ , then the corresponding optimal codebook  $\mathcal{P}_i = \{p_{i1}, \dots, p_{iL}\}$  is a locally optimal solution for this  $i$ -th subproblem, otherwise, we must have  $\mu_i' > 0$ , the optimal value of which can be found by using an iterative subgradient method as described in step 1). The optimal value of  $\lambda$  can be obtained by a similar iterative subgradient based method given by the equation at the top of the next page,

$$\lambda^{l+1} = [\lambda^l - \alpha^l \left( P_{avg} - \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L E[p_{ij} | \mathcal{R}_j^i] Pr(\mathcal{R}_j^i) \right) \right)]^+$$

where again  $l$  is the iteration number,  $\alpha^l > 0$  is a scalar step size for the  $l$ -th iteration satisfying  $\sum_{l=1}^{\infty} \alpha^l = \infty$  and  $\sum_{l=1}^{\infty} (\alpha^l)^2 < \infty$ . Repeat the modified GLA based step for finding a local optimum for the quantized powers and the subgradient based updates for  $\lambda$  and  $\mu'_i$ ,  $i = 1, 2, \dots, M$  until convergence and the final codebook will be a locally optimal codebook for the wideband spectrum sharing Problem (5).

*Remark 5: Convergence of Algorithm 1:* Note that it is straightforward to extend the global convergence and empirical consistency results of Theorem 2 to the wideband case for fixed values of  $\lambda$  and  $\mu'_i$ ,  $i = 1, 2, \dots, M$ . As noted in Section III.A for the narrowband case, the iterative subgradient based methods for updating  $\lambda$  and  $\mu'_i$  converge to the globally optimal values corresponding to the dual function due to the convexity of the dual function with respect to the Lagrange multipliers [20]. Thus, Algorithm 1 converges to a local optimum (since convergence of the modified GLA can be guaranteed to a local optimum only) of the quantized power values  $\{p_{i1}, \dots, p_{iL}\}$ ,  $i = 1, 2, \dots, M$ .

*Remark 6:* Theorem 3 also holds for the wideband case in the sense that the properties i)-iv) hold for each  $\{p_{i1}, p_{i2}, \dots, p_{iL}\}$ ,  $\forall i = 1, 2, \dots, M$  with  $\mu$  replaced by  $\mu_i$ ,  $i = 1, 2, \dots, M$  and  $\lambda$  representing the Lagrange multiplier associated with the average sum power constraint in (5).

### C. Approximate Quantized Power Allocation Algorithm (AQPA)

Although an offline algorithm, the complexity of modified GLA for determining the optimal quantized power is very high for even a moderately large value of  $L$ . This is due to the fact that the optimal channel partitions and the corresponding optimal power codebook are obtained via empirically generating a large number of channel realizations as training samples. As  $L$  increases, the number of training samples required will also increase. Thus motivated, we use part iv) of Theorem 3 to derive a low-complexity suboptimal scheme for implementing the modified GLA for large  $L$  values. Below we describe this scheme for the narrowband case. A similar scheme for the wideband case can be designed accordingly.

Note that part iv) of Theorem 3 states that as  $L \rightarrow \infty$ ,  $p_L = 0$  and  $p_{L-1} \rightarrow 0$ . Applying these approximations to (8) allows us to obtain an approximate but computationally efficient algorithm (called approximate quantized power allocation algorithm (AQPA)) for large  $L$ . AQPA first solves  $E[\frac{g_1}{1+g_1 p_{L-1}} - (\lambda + \mu g_0) | \mathcal{R}_{L-1}] = 0$  for  $p_{L-2}$  by substituting  $p_L = 0$  and taking the limit  $p_{L-1} \rightarrow 0$ , which, if  $\mu > 0$ , is equivalent to solving  $\int_{\lambda}^{\infty} \int_{\frac{g_1 - \lambda}{\mu}}^{\frac{\mu}{\mu}} \left( \frac{1}{\mu} \left( \frac{\log(1+g_1 p_{L-2})}{p_{L-2}} - \lambda \right) (g_1 - (\lambda + \mu g_0)) f(g_0) f(g_1) dg_0 dg_1 = 0 \right.$  for  $p_{L-2}$ . When  $\mu = 0$ , it is equivalent to solving for  $p_{L-2}$  from  $\int_{\lambda}^{\frac{e^{\lambda p_{L-2} - 1}}{p_{L-2}}} (g_1 - \lambda) f(g_1) dg_1 = 0$ . Note that the above equations (for both

$\mu > 0$  and  $\mu = 0$ ) involve only one variable:  $p_{L-2}$  and are thus straightforward to solve. One can then recursively compute  $p_{L-3}, p_{L-4}, \dots$ , by using the optimality condition (8) for the regions  $\mathcal{R}_{L-2}, \mathcal{R}_{L-3}, \dots$ , respectively, in the reverse direction. These equations can be solved by appropriate nonlinear equation solvers and do not require the use of large number of training samples. Thus AQPA is significantly faster than GLA and is applicable to the case of large number of feedback bits. Note however, as this is an approximate algorithm only, the performance of this algorithm becomes comparable to modified GLA only for large values of  $L$ . Numerical results presented in Section V illustrate that AQPA performs extremely well for  $L \geq 16$ . Note also that AQPA will be a suitable algorithm to use if any of the system specifications (such as channel statistics or  $P_{avg}, Q_{avg}$  etc.) changed and the quantized power values needed to be recalculated.

*Other suboptimal algorithms:* For comparison purposes, we also propose two other suboptimal methods for finding quantized power allocation in the narrowband case (extension to the wideband case is obvious). (1) In the first method, we quantize  $g_0$  and  $g_1$  separately (i.e. separate scalar quantizations) by minimizing their corresponding distortion  $\sum_{n=1}^{L_1} E[(g_0 - g'_{0n})^2 | R_n] Pr(R_n)$  and  $\sum_{k=1}^{L_2} E[(g_1 - g'_{1k})^2 | R'_k] Pr(R'_k)$  respectively, with the Lloyd Algorithm, where  $g'_{0n}, g'_{1k}$  are the reconstruction points for  $g_0$  and  $g_1$  respectively, and  $L_1 \times L_2 = L$ . We then use the resulting (locally) optimal channel quantization regions to solve  $E[\frac{g_1}{1+g_1 p_{nk}} - (\lambda + \mu g_0) | R_n, R'_k] = 0$  for finding the optimal power allocation  $p_{nk}$  for the region where  $g_0 \in R_n, g_1 \in R'_k$ . We call this method as "separate channel quantization" (SCQ). (2) In the second method, we jointly quantize  $g_0$  and  $g_1$  by minimizing the distortion  $\sum_{m=1}^L E[(g_0 - g'_{0m})^2 + (g_1 - g'_{1m})^2 | R_m] Pr(R_m)$  with Lloyd Algorithm, and then use the resultant optimal channel quantization regions to solve  $E[\frac{g_1}{1+g_1 p_m} - (\lambda + \mu g_0) | R_m] = 0$  for finding the optimal power allocation  $p_m$ . We call it "joint channel quantization" (JCQ). Numerical results illustrate that AQPA significantly outperforms these two suboptimal methods.

## IV. OPTIMUM QUANTIZED POWER ALLOCATION WITH NOISY LIMITED FEEDBACK

In the previous section, we assumed ideal error-free feedback in the limited feedback model. However, feedback channel noise can result in unavoidable erroneous feedback, which can cause the SU-TX to select an incorrect quantized power vector, resulting in an inferior ergodic capacity performance compared to the case of noise-free feedback. In this section, we allow noise in the limited feedback channel model and study the ergodic capacity maximization problem (5) with noisy limited feedback. Note that in general modelling of feedback errors in a quantized CSI feedback system is a challenging problem. In the analysis to follow, we make some simplifying assumptions regarding the feedback errors



and operating conditions in order to formulate a tractable problem. The noisy feedback link, assumed to be memoryless, is characterized by the index transition probabilities  $\rho_{kj}$ , ( $k, j = 1, \dots, L$ ), which is defined as the probability of receiving index  $k$  at the SU-TX, given index  $j$  was sent from the CR network manager. For each band, given  $B = \log_2 L$  bits feedback, denote binary representation of index  $k$  and  $j$  as  $k_1 k_2 \dots k_B$  and  $j_1 j_2 \dots j_B$  respectively, where  $k_n, j_n \in \{0, 1\}$  for  $n = 1, \dots, B$ , and  $k_1, j_1$  represent the most significant bit. For each band with  $B$  bits feedback, we model the noisy feedback channel as  $B$  independent uses of a binary symmetric channel with crossover probability  $q_f$  for every feedback bit. Since bit errors are assumed to be independent,  $\rho_{kj} = \prod_{n=1}^B \rho_{k_n j_n} = q_f^{d_{k,j}} (1 - q_f)^{B - d_{k,j}}$ , where  $d_{k,j}$  is the Hamming distance between the binary representations of  $k$  and  $j$  [24].

Thus problem (5) with noisy limited feedback can be reformulated as (12) at the top of the next page.

Note that for each band the binary codewords representing the feedback indices for a power codebook of size  $L$  can be designed in  $L!$  different ways. In general, finding the optimal index assignment scheme is computationally prohibitive and sub-optimal or randomized schemes are preferred. However, it was shown in [25] in the context of capacity optimization for MIMO links with noisy limited feedback that when the channel quantizers and the precoder adaptation are jointly optimized for a given index assignment, all index assignment schemes are equally good (see Lemma 2 in [25]). The proof of Lemma 2 in [25] is directly applicable to our scenario due to the specific discrete memoryless nature of the feedback channel and hence all index assignment schemes are equally good for our noisy feedback model as well. Due to space restrictions, we do not go into further details regarding this. Instead, we simply concentrate on finding the optimum CSI partitions  $\mathcal{R}_j^i, \forall j$  and power codebook  $\mathcal{P}_i$  for the  $i$ -th band,  $i = 1, \dots, M$  that jointly optimize the ergodic capacity of SU under the long term average transmit power constraint and average interference constraint given by (12), for a fixed index assignment scheme (which can be arbitrarily chosen). Again, to keep things simple, we look at the narrowband spectrum-sharing case ( $M=1$ ). Using the simplified notations  $\mathcal{R}_j, p_j, j = 1, 2, \dots, L$ , and  $g_1, g_0, Q_{avg}$ , we write the Lagrangian for Problem (12) with  $M = 1$  as  $L(P, \lambda, \mu) = \sum_{j=1}^L \sum_{k=1}^L E[\log(1 + g_1 p_k) - \lambda p_k - \mu g_0 p_k | \mathcal{R}_j] \rho_{kj} Pr(\mathcal{R}_j) + \lambda P_{avg} + \mu Q_{avg}$ , where  $\lambda$  and  $\mu$  are the nonnegative Lagrange multipliers associated with the ATP constraint and AIP constraint respectively. The Lagrange dual function  $g(\lambda, \mu)$  is defined as  $\max_{p_k \geq 0, \forall k, \mathcal{R}_j, \forall j} \sum_{j=1}^L \sum_{k=1}^L E[\log(1 + g_1 p_k) - \lambda p_k - \mu g_0 p_k | \mathcal{R}_j] \rho_{kj} Pr(\mathcal{R}_j)$  and the corresponding dual problem is  $\min_{\lambda \geq 0, \mu \geq 0} g(\lambda, \mu) + \lambda P_{avg} + \mu Q_{avg}$ . It is obvious that this optimization problem with fixed  $\lambda$  and  $\mu$  can be easily solved using another modified GLA, (termed as modified GLA-2 to distinguish it from the noise free case) resulting in a locally optimum power codebook. For this power codebook, the optimal values  $\lambda$  and  $\mu$  can then be obtained via subgradient based methods similar to the ones for the noise free case. These two steps are then repeated until a satisfactory convergence criterion is met. Further details and the extension to the wideband case are omitted due to space restrictions.

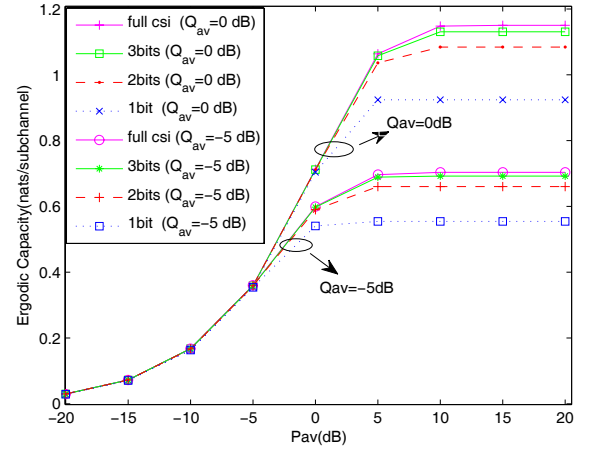


Fig. 4. SU Ergodic capacity with quantized power allocation (using GLA) with one PU for  $Q_{av} = -5$  dB and  $Q_{av} = 0$  dB.

## V. NUMERICAL RESULTS

In this section, we will evaluate the performance of the designed power allocation strategies via numerical simulations. We implement a wideband spectrum sharing system with one SU and  $M$  independent frequency bands (each band is originally licensed to a PU), where all the channels involved are assumed to undergo Rayleigh fading, namely all  $\mathbf{g}_0$  and  $\mathbf{g}_1$  are exponentially distributed with unit mean. For each simulation, 100,000 randomly generated channel realizations for each  $\mathbf{g}_0$  or  $\mathbf{g}_1$  are used. Any reference to the number of feedback bits should be interpreted in a *per band* sense.

Fig. 4 shows the capacity performance of SU sharing a narrowband spectrum with one PU with limited feedback and illustrates the effect of increasing the number of feedback bits on the capacity performance. For comparison, we also plot the corresponding capacity performance with full CSI. The striking observation from Fig. 4 is that introducing one extra bit of feedback substantially reduces the gap with capacity based on perfect CSI. This property is not very obvious when  $P_{av}$  is small, for example when  $P_{av} \leq -5$  dB ( $P_{av} \leq 0$  dB) for  $Q_{av} = -5$  dB ( $Q_{av} = 0$  dB). But with increasing  $P_{av}$ , it becomes more pronounced. To be specific, for  $Q_{av} = -5$  dB case, at  $P_{av} = 10$  dB, with 1 bit, 2 bits and 3 bits of feedback, the percentage capacity loss is approximately 21.23%, 6.21% and 1.62% respectively, and for both  $Q_{av} = -5$  dB and  $Q_{av} = 0$  dB cases, only 3 bits feedback can result in secondary ergodic capacity very close to that with full CSI. This is very encouraging since only a small number of bits of feedback are required to achieve close performance to the full CSI case. It can be also seen that the capacity performance with a higher AIP threshold ( $Q_{av} = 0$  dB) outperforms that with a lower AIP threshold ( $Q_{av} = -5$  dB), as expected. A similar behaviour can be also observed in Fig. 5 for a  $M = 4$  wideband spectrum sharing case with 1, 2, 3 bits of feedback and full CSI performance respectively, where  $(Q_{av1}, Q_{av2}, Q_{av3}, Q_{av4}) = (-10$  dB,  $-5$  dB,  $0$  dB,  $5$  dB).

In Fig. 6 we compare the performance of AQPA with modified GLA, where SU shares the spectrum with four PUs ( $M =$

$$\begin{aligned}
& \max_{p_{ik} \geq 0, \mathcal{R}_j^i} \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L \sum_{k=1}^L E[\log(1 + g_1^i p_{ik}) | \mathcal{R}_j^i] \rho_{kj} Pr(\mathcal{R}_j^i) \right) \\
& \text{s.t.} \quad \sum_{j=1}^L \sum_{k=1}^L E[g_0^i p_{ik} | \mathcal{R}_j^i] \rho_{kj} Pr(\mathcal{R}_j^i) \leq Q_{avg}^i, \forall i, \\
& \frac{1}{M} \sum_{i=1}^M \left( \sum_{j=1}^L \sum_{k=1}^L E[p_{ik} | \mathcal{R}_j^i] \rho_{kj} Pr(\mathcal{R}_j^i) \right) \leq P_{avg}
\end{aligned} \tag{12}$$

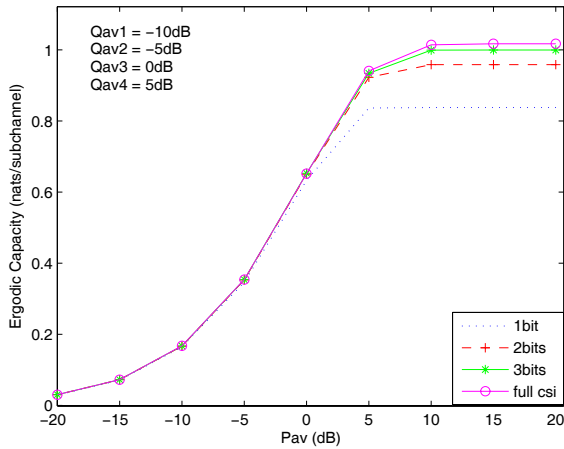


Fig. 5. SU Ergodic capacity performance with quantized power allocation (GLA) for four PUs ( $M = 4$ ) under various number of feedback bits per band.

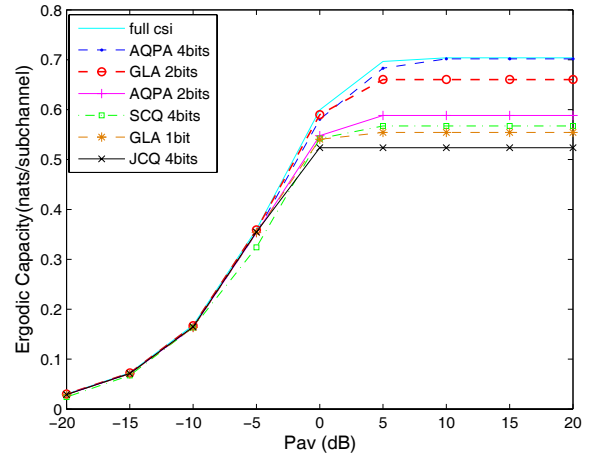


Fig. 7. Comparison of capacity performance of AQPA with two other suboptimal methods ( $M = 1$ ).

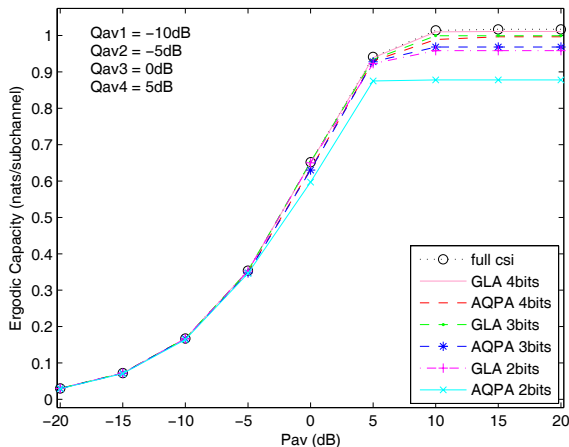


Fig. 6. Capacity performance of AQPA with four PUs ( $M = 4$ , feedback bits here refer to bits per band).

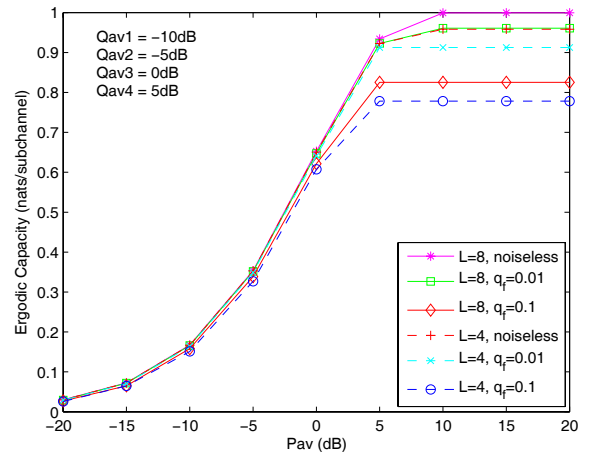


Fig. 8. Capacity performance of noisy limited feedback with four PUs ( $M=4$ ) and different BSC crossover probabilities (number of feedback regions  $L$  here refers to  $L$  regions per band).

4) and the AIP constraint thresholds ( $Q_{av1}, Q_{av2}, Q_{av3}, Q_{av4}$ ) = (-10 dB, -5 dB, 0 dB, 5 dB). It is illustrated that with the same number of bits of feedback per band, the gap between AQPA and modified GLA becomes smaller as  $L$  increases. For example, when  $P_{av} = 15$  dB, the capacity loss by using AQPA instead of GLA is about 8.38%, 3.12% and 1.42% for 2 bits, 3 bits and 4 bits of feedback respectively. It is clearly seen that

AQPA with 4 bits of feedback can almost approach the full CSI performance. In order to determine the speedup factor of AQPA compared to GLA for a fixed  $\lambda$  and  $\mu$  with  $M = 4$  and 4 bits of feedback, AQPA and GLA were implemented in MATLAB (version 7.10.0.499 (R2010a)) on an Intel Core 2 Duo processor (CPU T9600 with a clock speed of 2.80 GHz and a memory of 4 GB). It was seen that GLA (with

100,000 training samples) took approximately 7395 seconds or just over 2 hours whereas AQPA took only 5.44 seconds to achieve comparable levels of accuracy. Furthermore, as shown in Fig. 7, we also compare capacity performance of AQPA with the two other proposed possible suboptimal methods (SCQ and JCQ) for the narrowband case with  $Q_{av} = -5$  dB. For the SCQ case, various combinations of  $L_1, L_2$  such that  $L_1 \times L_2 = L$  are investigated and the one with the best performance is reported for every value of  $P_{av}$ . We can easily observe that even with 4 bits feedback, the performance of both JCQ and SCQ are worse than AQPA with only 2 bits of feedback, which further confirms the efficiency of AQPA.

Finally, we investigate the SU ergodic capacity performance with noisy limited feedback in Fig. 8, with  $M = 4$  and  $(Q_{av1}, Q_{av2}, Q_{av3}, Q_{av4}) = (-10$  dB,  $-5$  dB,  $0$  dB,  $5$  dB). It can be observed that as the feedback becomes less reliable (the crossover probability  $q_f$  increases), significant capacity performance degradation occurs, especially in high  $P_{avg}$ . For example, when  $P_{avg} = 10$  dB, for 3 (2) bits feedback, a noisy feedback channel with  $q_f = 0.01$  and  $q_f = 0.1$  can result in approximately 3.843% (4.769%) and 17.394% (18.783%) capacity loss respectively, compared to the noise-free case. This clearly illustrates that as the quality of feedback link degrades, the benefit of designing an optimal power codebook diminishes rapidly.

## VI. CONCLUSIONS AND EXTENSIONS

We have derived quantized power allocation algorithms for a wideband spectrum sharing system with one secondary user and multiple primary users, each licensed to use a separate frequency band, each band modelled as independent block fading channels. The objective has been to maximize the SU ergodic capacity under an average sum transmit power constraint and individual average interference constraints at the PU receivers. Modified Generalized Lloyd-type algorithms (GLA) have been derived and various properties of the quantized power allocation laws have been presented, along with a rigorous convergence and consistency proof of the modified GLA based algorithm. By appropriately exploiting the properties of the quantized power values for large number of bits of feedback, we have also derived approximate quantized power allocation algorithms that perform very close to the modified GLA based algorithms but are significantly faster. Finally, we have presented an extension of the modified GLA based quantized power allocation algorithm to the case of noisy feedback channels. Future work will include deriving expressions for asymptotic (as the number of feedback bits goes to infinity) capacity loss with quantized power allocation, consideration of primary interference at the secondary receiver and quantized power allocation with other types of interference constraints at the primary receiver.

### APPENDIX

#### A. Proof of Theorem 2

For the modified GLA, one can define a distortion measure  $d((g_0, g_1), p) = -(\log(1 + g_1 p) - \lambda p - \mu g_0 p)$ . For such non-difference distortion measures, following [26], one can ensure

nonnegativity of the distortion measure by introducing a modified distortion measure as  $\hat{d}((g_0, g_1), p) = d((g_0, g_1), p) - \min_p d((g_0, g_1), p)$ . Since  $d((g_0, g_1), p)$  is a convex function of  $p$  for fixed  $(g_0, g_1)$ , we get the unique minimum  $p^* = (\frac{1}{\lambda + \mu g_0} - \frac{1}{g_1})^+$ , thus  $\min_p d((g_0, g_1), p) = d((g_0, g_1), p^*)$ . Therefore we have  $\hat{d}((g_0, g_1), p) \geq 0$ . Since  $d((g_0, g_1), p^*)$  is constant for a given  $(g_0, g_1)$ , thus using the distortion measure  $\hat{d}((g_0, g_1), p)$  instead of  $d((g_0, g_1), p)$  does not affect the results of modified GLA. One can easily show that  $\hat{d}$  satisfies the following properties: (1)  $\hat{d}$  is continuous and  $\hat{d} \in [0, \infty)$ , (2)  $\hat{d}((g_0, g_1), p)$  is a convex function of  $p$  for each fixed  $(g_0, g_1)$ , (3) for each  $(g_0, g_1)$  and some  $(g_0, g_1)$  not identically equal to  $(g_0, g_1)$ ,  $\hat{d}((g_0, g_1), p) \rightarrow \infty$ , as  $(g_0, g_1) \rightarrow (g_0, g_1)$  and  $\|p\| \rightarrow \infty$ , and (4) the partition boundaries in the channel space  $(g_0, g_1)$  have zero probability.

Properties 1), 2) and 3) are easy to show and the proofs here are omitted. Property 4) holds due to the assumption of continuous fading channels in this work. Note that this is also a necessary condition for a codebook to be optimal for a given partition [18]. Note also that the popular fading distributions such as Rayleigh, Rician and Nakagami and Log-normal etc. all satisfy the absolutely continuity assumption. It is then easy to show that for these types of fading scenarios, the cumulative distribution function (cdf) of  $(g_0, g_1)$ , denoted by  $F$ , satisfies the following properties [19]: (5)  $F$  contains no singular-continuous part and (6)  $\int \hat{d}((g_0, g_1), p) dF(g_0, g_1) < \infty$  for each  $p$  (implying a finite average distortion).

Next, let  $\mathbf{g}$  denote  $(g_0, g_1)$ . Noting that  $\{\mathbf{g}(\omega)\}$  is a stationary ergodic sequence with a cdf  $F$ , and letting  $F_{n,\omega}$  be the empirical distribution function of the first  $n$  members of the sequence [19], one can show that for almost every  $\omega$ ,  $\{F_{n,\omega}\}$  and  $F$  satisfy (see Lemma 4 of [19]) (7)  $\{F_n\}$  converges weakly to the  $F$  and (8)  $\lim_n \int \hat{d}((g_0, g_1), p) dF_n(g_0, g_1) = \int \hat{d}((g_0, g_1), p) dF(g_0, g_1)$ , for every  $p$ .

Hence, from [19], we can conclude that the modified GLA satisfies properties 1) to 8). Therefore, Lemmas 1-3 of [19] are applicable to the modified GLA designed in this paper and so the modified GLA satisfies the global convergence and empirical consistency properties as defined in [19].

#### B. Proof of Theorem 3 i)

We need to prove that for any two adjacent regions  $\mathcal{R}_j$  and  $\mathcal{R}_{j+1}$ ,  $j = 1, \dots, L-1$ ,  $p_j > p_{j+1}$ . Given an arbitrary  $g_0$  satisfying  $0 \leq g_0 < \frac{1}{\mu} (\frac{\log(\frac{p_j}{p_{j+1}})}{p_j - p_{j+1}} - \lambda)$  (assuming  $\mu > 0$ ), suppose there is a point  $(g_0, g_1^a) \in \mathcal{R}_j$  and a point  $(g_0, g_1^c) \in \mathcal{R}_{j+1}$  (neither of these two points is on the boundary), and let  $(g_0, g_1^b)$  denote the point on the boundary corresponding to the same  $g_0$ , which from Lemma 1, is given by  $g_1^b = \frac{e^{(\lambda + \mu g_0)(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{(\lambda + \mu g_0)(p_j - p_{j+1})}}$ . Then, we have  $g_1^a > g_1^b > g_1^c$ . Now suppose  $p_j < p_{j+1}$ . Since  $(g_0, g_1^a) \in \mathcal{R}_j$ , we have  $\log(1 + g_1^a p_j) - \lambda p_j - \mu g_0 p_j \geq \log(1 + g_1^a p_{j+1}) - \lambda p_{j+1} - \mu g_0 p_{j+1}$ . As  $p_j < p_{j+1}$ , we have  $(\lambda + \mu g_0)(p_{j+1} - p_j) \geq \log(\frac{1 + g_1^a p_{j+1}}{1 + g_1^a p_j})$  which implies  $e^{(\lambda + \mu g_0)(p_{j+1} - p_j)} - 1 \geq g_1^a (p_{j+1} - p_j) e^{(\lambda + \mu g_0)(p_{j+1} - p_j)}$ . We also have  $g_1^b = \frac{e^{(\lambda + \mu g_0)(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{(\lambda + \mu g_0)(p_j - p_{j+1})}} = \frac{e^{(\lambda + \mu g_0)(p_{j+1} - p_j)} - 1}{p_{j+1} - p_j e^{(\lambda + \mu g_0)(p_{j+1} - p_j)}}$ . Note that  $p_{j+1} > p_j$  implies  $e^{(\lambda + \mu g_0)(p_{j+1} - p_j)} - 1 > 0$ . Since  $g_1^b > 0$ ,

we have  $p_{j+1} - p_j e^{(\lambda + \mu g_0)(p_{j+1} - p_j)} > 0$ . Combining the above two results, we obtain  $g_1^a \leq \frac{e^{(\lambda + \mu g_0)(p_{j+1} - p_j)} - 1}{p_{j+1} - p_j e^{(\lambda + \mu g_0)(p_{j+1} - p_j)}} = g_1^b$  which is a contradiction to  $g_1^a > g_1^b$ . Similarly, we can also prove that if  $p_j < p_{j+1}$ , we have  $g_1^c \geq g_1^b$  which is a contradiction to  $g_1^c < g_1^b$ . Thus we must have  $p_j > p_{j+1}$ .

### C. Proof for Theorem 3 ii)

From Lemma 1, the boundary between any two adjacent regions  $\mathcal{R}_j$  and  $\mathcal{R}_{j+1}$  is given by

$$\begin{aligned} g_1 &= \frac{e^{(\lambda + \mu g_0)(p_j - p_{j+1})} - 1}{p_j - p_{j+1} e^{(\lambda + \mu g_0)(p_j - p_{j+1})}} \\ &= \frac{e^{(\lambda + \mu g_0)p_j} - e^{(\lambda + \mu g_0)p_{j+1}}}{p_j e^{(\lambda + \mu g_0)p_{j+1}} - p_{j+1} e^{(\lambda + \mu g_0)p_j}} \\ &= (\lambda + \mu g_0) \frac{e^{(\lambda + \mu g_0)p_\epsilon} (p_j - p_{j+1})}{p_j e^{(\lambda + \mu g_0)p_{j+1}} - p_{j+1} e^{(\lambda + \mu g_0)p_j}} \\ &> \lambda + \mu g_0 \end{aligned} \quad (13)$$

where the last equality follows from the mean value theorem for some  $p_\epsilon \in (p_{j+1}, p_j)$ . The last inequality holds since we have  $p_j e^{(\lambda + \mu g_0)p_\epsilon} > p_j e^{(\lambda + \mu g_0)p_{j+1}}$  and  $-p_{j+1} e^{(\lambda + \mu g_0)p_\epsilon} > -p_{j+1} e^{(\lambda + \mu g_0)p_j}$ . By rearranging, we get  $\frac{e^{(\lambda + \mu g_0)p_\epsilon} (p_j - p_{j+1})}{p_j e^{(\lambda + \mu g_0)p_{j+1}} - p_{j+1} e^{(\lambda + \mu g_0)p_j}} > 1$ .

### D. Proof of Theorem 3 iii)

Given a fixed channel partitioning scheme, the optimal quantized power for  $\mathcal{R}_j$  is obtained as  $p_j = \max(p_j^*, 0), \forall j$ , where  $p_j^*$  is determined by solving the equation  $E[\frac{g_1}{1+g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j] = 0$ . We can see that if  $E[g_1 | \mathcal{R}_j] \leq E[(\lambda + \mu g_0) | \mathcal{R}_j]$ , then to satisfy the equation,  $p_j^* < 0$ , implying  $p_j = \max(p_j^*, 0) = 0$ . On the other hand, if  $E[g_1 | \mathcal{R}_j] > E[(\lambda + \mu g_0) | \mathcal{R}_j]$ ,  $p_j^*$  has to be strictly positive in order to satisfy the optimality equation, implying  $\max(p_j^*, 0) = p_j^*$ . We know from Theorem 3 ii) that all boundaries between any two adjacent regions have a lower bound given by  $g_1 > \lambda + \mu g_0$ , i.e. for any given  $(g_0, g_1)$  belonging to any of the first  $L - 1$  regions,  $g_1 > \lambda + \mu g_0$ . Thus for the first  $L - 1$  regions,  $E[g_1 | \mathcal{R}_j] Pr\{\mathcal{R}_j\} > E[(\lambda + \mu g_0) | \mathcal{R}_j] Pr\{\mathcal{R}_j\}$ . Therefore the optimal quantized power in the first  $L - 1$  regions is strictly positive. This cannot be said however for  $p_L$  as for  $\mathcal{R}_L$ , we cannot guarantee  $g_1 > \lambda + \mu g_0$  for any given  $(g_0, g_1)$  pair in that region. It is thus possible to have  $p_L$  to be 0. The next result shows under what circumstances one can have  $p_L$  to be exactly 0.

### E. Proof for Theorem 3 iv)

*Step 1:* We know from Theorem 3 iii) that we always have  $E[\frac{g_1}{1+g_1 p_j} - (\lambda + \mu g_0) | \mathcal{R}_j] = 0, j = 1, \dots, L - 1$ , and for the region  $\mathcal{R}_L$ , this equation may not be satisfied when  $p_L = 0$ . Let us assume that  $p_L > 0$ . Then we have  $\sum_{j=1}^L E[\lambda + \mu g_0 | \mathcal{R}_j] Pr\{\mathcal{R}_j\} = \sum_{j=1}^L E[\frac{g_1}{1+g_1 p_j} | \mathcal{R}_j] Pr\{\mathcal{R}_j\}$ , implying  $\lambda + \mu = \sum_{j=1}^L E[\frac{g_1}{1+g_1 p_j} | \mathcal{R}_j] Pr\{\mathcal{R}_j\} < \sum_{j=1}^L E[g_1 | \mathcal{R}_j] Pr\{\mathcal{R}_j\} = 1$  since  $\sum_{j=1}^L E[g_i | \mathcal{R}_j] Pr\{\mathcal{R}_j\} = E[g_i] = 1$ , for  $i = 0, 1$ . Hence if  $\lambda + \mu \geq 1$ , we must have  $p_L = 0$ .

From the optimality equation, one can write  $p_i = \frac{E[\frac{g_1 p_i}{1+g_1 p_i} | \mathcal{R}_i]}{\lambda + \mu E[g_0 | \mathcal{R}_i]}$  when  $p_i > 0$ , it is obvious that  $p_i < \frac{1}{\lambda + \mu E[g_0 | \mathcal{R}_i]}$ ,  $i = 1, 2, \dots, L - 1$ . Since  $p_L \geq 0$ , this is also true for region  $\mathcal{R}_L$ . Therefore when  $\mu \neq 0$ ,  $\mu Q_{avg} = \mu \sum_{i=1}^L p_i E[g_0 | \mathcal{R}_i] Pr(\mathcal{R}_i) < \sum_{i=1}^L \frac{\mu E[g_0 | \mathcal{R}_i]}{\lambda + \mu E[g_0 | \mathcal{R}_i]} Pr(\mathcal{R}_i) < \sum_{i=1}^L Pr(\mathcal{R}_i) = 1$ . Similarly, if  $\lambda \neq 0$ ,  $\lambda P_{avg} < 1$ . Thus  $\mu > 1$  implies  $Q_{av} < 1$  and  $\lambda > 1$  implies  $P_{av} < 1$ .

*Step 2:* Next, we will show that no matter what  $\lambda, \mu$  is,  $p_L$  must be zero for a sufficiently large  $L$  and  $\lim_{L \rightarrow \infty} p_{L-1} = 0$ . First, we will prove that as  $L \rightarrow \infty$ , the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches its limiting boundary  $g_1 = \frac{\lambda + \mu g_0}{1 - (\lambda + \mu g_0) \delta^*}$ , where  $\delta^* = \lim_{L \rightarrow \infty} p_L$ . Given  $p_1 > \dots > p_L \geq 0$ , it is clear that the sequence  $\{p_i\}$ ,  $i = 1, 2, \dots, L$  is a monotonically decreasing sequence bounded below, therefore it must converge to its greatest-lower bound  $\delta^*$  ( $\delta^* = \lim_{L \rightarrow \infty} p_L \geq 0$ ) as  $L \rightarrow \infty$ . Therefore, it can be easily shown that for an arbitrarily small  $\epsilon > 0$ , we can always find a sufficiently large  $L$  such that  $p_{L-1} - p_L < \epsilon$ . Thus, as  $L \rightarrow \infty$ ,  $(p_{L-1} - p_L) \rightarrow 0$ . Using this result, we can show that the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches the limiting boundary  $g_1 = \frac{\lambda + \mu g_0}{1 - (\lambda + \mu g_0) \delta^*}$  (or  $\lambda + \mu g_0 = \frac{g_1}{1 + g_1 \delta^*}$ ) as  $L \rightarrow \infty$ , (since this boundary can be written as  $\lambda + \mu g_0 = \frac{\log(\frac{1+g_1 p_{L-1}}{1+g_1 p_L})}{p_{L-1} - p_L}$ ,

and  $\lim_{L \rightarrow \infty} (\lim_{(p_{L-1} - p_L) \rightarrow 0} \frac{\log(\frac{1+g_1 p_{L-1}}{1+g_1 p_L})}{p_{L-1} - p_L}) = \lim_{L \rightarrow \infty} \frac{g_1}{1+g_1 p_L} = \frac{g_1}{1+g_1 \delta^*}$ ). Now, suppose there exists a pair  $(\lambda, \mu)$  such that  $p_L > 0$  for any arbitrarily large  $L$  (implying  $\delta^* > 0$ ). Thus for any  $L$ ,  $p_L$  satisfies  $E[\frac{g_1}{1+g_1 p_L} - (\lambda + \mu g_0) | \mathcal{R}_L] = 0$ . From (1), we have as  $L \rightarrow \infty$ , the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches its limit  $\lambda + \mu g_0 = \frac{g_1}{1+g_1 \delta^*}$ . Note that for a finite value of  $L$ , the region  $\mathcal{R}_L$  can be divided into two parts  $\mathcal{R}_{L1}$  and  $\mathcal{R}_{L2}$  where  $\mathcal{R}_{L1}$  corresponds to  $\frac{\log(\frac{1+g_1 p_{L-1}}{1+g_1 p_L})}{p_{L-1} - p_L} \leq \lambda + \mu g_0 < \frac{g_1}{1+g_1 \delta^*}$  and  $\mathcal{R}_{L2}$  corresponds to  $\frac{g_1}{1+g_1 \delta^*} \leq \lambda + \mu g_0 < \infty$ . As  $L$  becomes arbitrarily large, the region  $\mathcal{R}_{L1}$  becomes vanishingly small, and one obtains  $E(\lambda + \mu g_0 | \mathcal{R}_L) > E(\frac{g_1}{1+g_1 \delta^*} | \mathcal{R}_L) \geq E(\frac{g_1}{1+g_1 p_L} | \mathcal{R}_L)$  for a sufficiently large  $L$ , which is a contradiction to the KKT optimality condition for  $p_L > 0$ . Hence no matter what  $\lambda, \mu$  are,  $p_L$  must be zero for a sufficiently large  $L$ . And  $\delta^* = \lim_{L \rightarrow \infty} p_L = 0$ . Finally,  $\delta^* = 0$  implies the boundary between  $\mathcal{R}_{L-1}$  and  $\mathcal{R}_L$  approaches  $g_1 = \lambda + \mu g_0$  as  $L \rightarrow \infty$ , and since as  $L \rightarrow \infty$ ,  $(p_{L-1} - p_L) \rightarrow 0$  and  $p_L = 0$ , we have  $\lim_{L \rightarrow \infty} p_{L-1} = 0$ .

## REFERENCES

- [1] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 649–658, Feb. 2007.
- [2] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications," in *Proc. IEEE Int. Workshop Mobile Multimedia Commun.*, Nov. 1999, pp. 3–10.
- [3] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.
- [4] X. Kang, Y. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 940–950, Feb. 2009.

- [5] J. Luo, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation for parallel fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2594–2611, 2005.
- [6] M. Gastpar, "On capacity under received-signal constraints," in *Proc. 42nd Annual Allerton Conf. Commun., Control Computing*, Sep. 2004.
- [7] L. Musavian and S. Aissa, "Capacity and power allocation for spectrum-sharing communications in fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 148–156, Jan. 2009.
- [8] L. Musavian and S. Aissa, "Fundamental capacity limits of cognitive radio in fading environments with imperfect channel information," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3472–3480, Nov. 2009.
- [9] K. Huang and R. Zhang "Cooperative feedback for multi-antenna cognitive radio networks," submitted for publication.
- [10] H. A. Suraweera, P. J. Smith, and M. Shafi, "Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge," *IEEE Trans. Veh. Technol.*, accepted for publication.
- [11] A. G. Marques, X. Wang, and G. B. Giannakis, "Dynamic resource management for cognitive radios using limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3651–3666, Sep. 2009.
- [12] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.
- [13] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102–114, Mar. 2010.
- [14] Y. Y. He and S. Dey, "Spectrum sharing in cognitive radio with quantized channel information." Available: <http://arxiv.org/abs/1006.0619>
- [15] R. Chen, J. G. Andrews, R. W. Heath, and A. Ghosh, "Uplink power control in multi-cell spatial multiplexing wireless systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2700–2711, July 2007.
- [16] B. L. Ng, J. S. Evans, S. Hanly, and D. Aktas, "Distributed downlink beamforming with cooperative base stations," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5491–5499, Dec. 2008.
- [17] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
- [18] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.
- [19] M. Sabin, and R. Gray, "Global convergence and empirical consistency of the generalized Lloyd algorithm," *IEEE Trans. Inf. Theory*, vol. 32, no. 2, pp. 148–155, Mar. 1986.
- [20] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.
- [21] V. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62–70, Jan. 2004.
- [22] X. Wang, A. G. Marques, and G. B. Giannakis, "Power-efficient resource allocation and quantization for TDMA using adaptive transmission and limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4470–4485, Sep. 2008.
- [23] L. Zhang, Y. Xin, and Y. Liang, "Optimal power allocation for multiple access channels in cognitive radio networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2008, pp. 1550–2252.
- [24] S. Ekbatani, F. Etemadi, and H. Jafarkhani, "Throughput maximization over slowly fading channels using quantized and erroneous feedback," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2528–2533, Sep. 2009.
- [25] T. Wu and V. K. N. Lau, "Robust precoder adaptation for MIMO links with noisy limited feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1640–1649, Apr. 2009.
- [26] T. Linder and R. Zamir, "High-resolution source coding for non-difference distortion measures: the rate-distortion function," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 533–547. Mar. 1999.



**Yuanyuan He** (S'08) received her B.S. degree in communication engineering from Yanshan University, China, in 2005 and the M.Eng degree in electronic engineering from the Harbin Institute of Technology, China, in 2007. She is currently pursuing the Ph.D. degree in the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Australia. Her research interests lie in the area of wireless communications including resource allocation, cognitive radio, and limited feedback.



**Subhrakanti Dey** (M'96-SM'06) was born in Calcutta, India, in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996. He has been with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since February 2000, where he is currently a full Professor. From September 1995 to September 1997, and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a post-doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control. Dr. Dey currently serves on the Editorial Board of *Elsevier Systems and Control Letters*. He was also an Associate Editor for the *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* during 2005–2007 and an Associate Editor for the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* during 2007–2010. He is a Senior Member of IEEE.