Rayleigh fading and path loss. We demonstrated that the lower bound provides an accurate approximation of the exact BEP at low SNRs, whereas the upper bound is accurate at medium to high SNRs. Our analytical expressions are further evaluated asymptotically to characterize the cooperative diversity order in the high-SNR regime.

References

[1] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[2] L. Daewon, S. Hanbyul, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: Deployment scenarios and operational challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[3] P. Marsch and G. Fettweis, "Uplink CoMP under a constrained backhaul and imperfect channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1730–1742, Jun. 2011.

[4] A. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[5] S. Hanly and P. Whiting, "Information-theoretic capacity of multi-receiver networks," in *Telecommunications Systems*. Amsterdam, The Netherlands: Baltzer, Jan. 1993, pp. 1–42.

[6] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1894, Nov. 1997.

[7] S. Shamai and A. D. Wyner, "Uplink macro diversity of limited backhaul cellular network," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3457–3478, Aug. 2009.

[8] E. Aktas, J. Evans, and S. Hanly, "Distributed decoding in a cellular multiple-access channel," *IEEE Commun. Mag.*, vol. 7, no. 1, pp. 241–250, Jan. 2008.

[9] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "The effect of macrodiversity on the performance of MLD in flat Rayleigh/Rician fading," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1764–1767, Nov. 2012.

[10] S. Verdu, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[11] I. Grashteyn and I. Ryzhik, *Table of Integrals, Series, and Products,* 7th ed. Burlington, MA, USA: Academic, 2007.

[12] W. Zhengdao and G. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[13] C. Kyongkuk and Y. Dongweon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, Jul. 2002.

On Optimality of Myopic Policy for Opportunistic Access With Nonidentical Channels and Imperfect Sensing

Kehao Wang, Lin Chen, and Quan Liu

***Abstract*—We consider the access problem in a multichannel opportunistic communication system with imperfect sensing, where the state of each channel evolves as a nonidentical and independently distributed Markov process. This problem can be cast into a restless multiarmed bandit (RMAB) problem, which is intractable for its exponential computation complexity. A promising approach that has attracted much research attention is the consideration of an easily myopic policy that maximizes the immediate reward by ignoring the impact of the current policy on future reward. Specially, we formalize a family of generic functions, which is referred to as g-regular functions, characterized by three axioms, and then establish a set of closed-form conditions for the optimality of the myopic policy and illustrate the engineering implications behind the obtained results.**

***Index Terms*—Imperfect detection, myopic policy, opportunistic spectrum access, restless multiarmed bandit (RMAB).**

I. Introduction

We consider an opportunistic multichannel communication system with nonidentical but independent channels in which a user is limited to sense and transmit only on a subset of them each time. Given that the detection of a channel state is not perfect, the fundamental optimization problem we address is how the user can exploit past imperfect detection information and the stochastic properties of channels to maximize its utility (e.g., expected throughput) by switching among channels opportunistically.

A. General Context and Related Work

The decision problem can be cast into a restless multiarmed bandit (RMAB) problem, which is proved to be pSPACE-hard [1], and very few results are reported on the structure of the optimal policy due to its high complexity. Thus, the myopic strategy with a simple and tractable structure has recently attracted extensive research attention, which consists of sensing the channels to maximize the expected immediate reward, also called the greedy policy, while ignoring the impact of the current decision on future reward.

Following the research line on the myopic policy, for the case of perfect sensing, Zhao *et al.* [2] established the structure of the myopic policy, analyzed the performance, and partly obtained the optimality for the case of independent and identically distributed (i.i.d.) channels. Ahmad *et al.* [3] derived the optimality of the myopic sensing policy for the positively correlated i.i.d. channels when the user is limited

Manuscript received March 17, 2013; revised July 18, 2013; accepted September 11, 2013. Date of publication October 17, 2013; date of current version June 12, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61303027, by the China Postdoctoral Science Foundation under Grant 2013M531753, by the Agence Nationale de la Recherche (ANR) under Grant Green-Dyspan (ANR-12-IS03), and by the Keygrant Project of the Chinese Ministry of Education under Grant 313042. The review of this paper was coordinated by Prof. D. B. da Costa.

K. Wang and Q. Liu are with the Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan University of Technology, 430070 Hubei, China (e-mail: Kehao.wang@whut.edu.cn; quanliu@whut.edu.cn).

L. Chen is with the Laboratoire de Recherche en Informatique, Department of Computer Science, University of Paris-Sud, Orsay 91405, France (e-mail: lin.chen@lri.fr).

Digital Object Identifier 10.1109/TVT.2013.2285713

TABLE I
SUMMARY OF RELATED WORK ON MYOPIC POLICY

	i.i.d. channels	non i.i.d. channels
Perfect sensing	[2], [3], [4]	[5]
Imperfect sensing	[7], [8]	this paper

to accessing one channel (i.e., $k = 1$) each time and further extended the optimality to the case of sensing multiple i.i.d. channels ($k > 1$) [4]. In our previous work [5], we extended i.i.d. channels [3], [6] to nonidentical and independently distributed channels and focused on a family of generic and important utility functions, namely, *regular* function, and derived a set of closed-form conditions to guarantee the optimality of the myopic policy.

For the more challenging case of imperfect sensing, there exist only limited results on the optimality of the myopic policy. Liu *et al.* [7] proved the optimality of the myopic policy for the case of two channels with a particular utility function and conjectured it for arbitrary number of channels N. In [8], we extended the optimality of the myopic policy for i.i.d. channels from perfect sensing to imperfect sensing and, as a consequence, derived closed-form conditions to guarantee the optimality of the myopic sensing policy for arbitrary N and for a regular function. Furthermore, the belief vector was divided into "value belief vector" and "policy belief vector," and then, a set of closed-form conditions was established for the myopic policy of opportunistic spectrum access [9].

B. Summary of Contributions

Our study presented in this paper builds upon and extends our previous work [5], [8]. Specially, focusing on the case of imperfect sensing, we perform an analytical study on the optimality of the myopic policy for the considered decision problem. The contributions of this paper, given the current state of the art and compared with our previous work [5], [8], are twofold.

1) We generalize the third axiom in [5] to cover a much larger class of reward functions, particularly the logarithmic and exponential functions, and further derive the optimality conditions in this more general context with [5] being a degenerated case.
2) We derive the optimality conditions of the myopic policy with *imperfect* channel observation and *nonidentical* and independently distributed channels. In this regard, the main technical obstacle overcome in this paper is the dependence of the belief value on both the evolution of the channel and the detection outcome, which leads to indeterministic channel state transition and nonlinear propagation of the belief vector.

It is worth noting that despite the vital importance, very few work has been done on the impact of imperfect observation on the performance of the myopic policy. To our knowledge, [7] and [8] are the only analysis pertinent to our study, but they assume i.i.d. channels. However, the analysis in this paper elevates this assumption by considering the heterogeneous channels, which requires an original analysis on the optimality, as detailed later in this paper. Table I compares the results presented in this paper and those obtained in the literature.

The rest of this paper is organized as follows. Our model is formulated in Section II with the g-regular function defined in Section III. Section IV studies the optimality of the myopic sensing policy. Finally, this paper is concluded in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a synchronous multichannel opportunistic communication system where the user is allowed to sense k ($1 \le k \le N$) of N channels (denoted by $\mathcal{N}$) at each slot t. Each channel is characterized by the discrete two-state (good/1, bad/0) Markov chain, and the transition probabilities of channel i are $\{p_{rs}^i\}_{r,s=0,1}$. We assume $p_{11}^i > p_{01}^i (1 \le i \le N)$ to characterize the gradual evolution of the states of the channels. Let $\mathcal{A}(t)$ denote the set of channels chosen by the user at slot t, where $\mathcal{A}(t) \subseteq \mathcal{N}$ and $|\mathcal{A}(t)| = k$. Let $\mathbf{S}(t) \triangleq [S_1(t), \ldots, S_N(t)]$ denote the channel state vector, where $S_i(t) \in \{0, 1\}$ is the state of channel i in slot t, and let $\mathbf{S}'(t) \triangleq \{S_i'(t), i \in \mathcal{A}(t)\}$ denote the sensing outcome vector, where $S_i'(t) = 0$ (1, respectively) means that channel i is sensed bad (good) in slot t. We are interested in the imperfect-sensing case, i.e., a good channel may be sensed bad, and vice versa, which is characterized by two system parameters: the false-alarm rate $\epsilon_i(t)$ and the miss-detection rate $\zeta_i(t)$, which are formally defined as follows:

$$\epsilon_i(t) \triangleq \Pr\{S_i'(t) = 0 | S_i(t) = 1\}$$

$$\zeta_i(t) \triangleq \Pr\{S_i'(t) = 1 | S_i(t) = 0\}.$$

In our analysis, we consider the case where the values of $\epsilon_i(t)$ and $\zeta_i(t)$ are independent w.r.t. t and i, that is, $\epsilon_i(t) = \epsilon$ and $\zeta_i(t) = \zeta$.

Let $\mathcal{B}(t)$ be the set of channels on which the receiver tunes its k radios. To ensure $\mathcal{A}(t) = \mathcal{B}(t)$ at any time, channel selections should be based on common observations. Without considering the communicating cost, a simple approach is given in [7] as follows. When the receiver successfully receives a packet from a channel, it sends an acknowledgement (ACK) to the transmitter over the same channel at the end of the slot. The absence of an ACK (NACK) signifies that the transmitter does not transmit over this channel or does transmit but the channel is busy in this slot. We assume that an ACK is received without error.[1]

Obviously, by sensing only k out of N channels, the user cannot observe the state information of the whole system and has to infer the channel states from its past decision and observation history so as to make its future decision. To this end, we define the *channel state belief vector* (hereinafter referred to as *belief vector* for briefness) $\Omega(t) \triangleq \{\omega_i(t), i \in \mathcal{N}\}$, where $0 \le \omega_i(t) \le 1$ is the conditional probability that channel i is in state good (i.e., $S_i(t) = 1$) at slot t given all past states, actions, and observations. Therefore, given the action $\mathcal{A}(t)$ and the observations $\{ACK_i(t) \in \{0, 1\} : i \in \mathcal{A}(t)\}$, the belief vector can be recursively updated using Bayes' rule, as shown in the following equation:

$$\omega_i(t+1) = \begin{cases} p_{11}^i, & i \in \mathcal{A}(t), ACK_i(t) = 1 \\ \tau_i(\varphi(\omega_i(t))), & i \in \mathcal{A}(t), ACK_i(t) = 0 \\ \tau_i(\omega_i(t)), & i \notin \mathcal{A}(t). \end{cases} \quad (1)$$

Note that the belief update under $ACK_i(t) = 0$ results from the fact that the receiver cannot distinguish a failed transmission, i.e., collides with the primary traffic with probability $\zeta(1 - \omega_i(t))$ from no transmission with probability $\epsilon\omega_i(t) + (1 - \zeta)(1 - \omega_i(t))$ [7]. For convenience, we introduce two operators ($\varphi(\omega_i) = \epsilon\omega_i(t)/(\epsilon\omega_i(t) + 1 - \omega_i(t))$ and $\tau_i(\omega_i(t)) \triangleq \omega_i(t) \cdot p_{11}^i + (1 - \omega_i(t)) \cdot p_{01}^i$.

Remark: We would like to emphasize that in contrast to the perfect-sensing case, where $\omega_i(t+1)$ is a linear function of $\omega_i(t)$, in the imperfect-sensing case, the mapping from $\omega_i(t)$ to $\omega_i(t+1)$ is no longer linear due to the sensing error [cf. the second line of (1)]. Therefore, an original study on the optimality of the myopic sensing policy is particularly required since the aforementioned difference

[1] Although wireless transmissions are prone to failure due to noise, perfect ACK/NACK signaling can be justified by the relative strong encoding of control messages with respect to data transmissions due to the different amount of the encoded information.

makes the analysis [2]–[5], [10], [11] infeasible in the imperfect-sensing case.

A sensing policy π specifies a sequence of functions $\pi = [\pi_1, \pi_2, \ldots, \pi_T]$, where π_t maps the belief vector $\Omega(t)$ to the action $\mathcal{A}(t)$ (i.e., the set of channels to sense) in slot $t : \pi_t : \Omega(t) \to \mathcal{A}(t), |\mathcal{A}(t)| = k$.

Given the imperfect-sensing context, we are interested in the user's optimization problem to find the optimal sensing policy π^* that maximizes the expected accumulated discounted reward over a finite horizon, i.e.,

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[\sum_{t=1}^{T} \beta^{t-1} R\left(\pi_t\left(\Omega(t)\right)\right) \middle| \Omega(1)\right] \tag{2}$$

where $R(\pi_t(\Omega(t)))$ is the reward collected in slot t under the sensing policy π_t with the initial belief vector $\Omega(1)$,[2] and $0 \le \beta \le 1$ is the discount factor characterizing the fact that future reward is less important than the immediate reward. By treating the belief value of each channel as the state of each arm of a bandit, the user's optimization problem can be cast into a two-state RMAB problem.

Theoretically, the optimal policy can be obtained by solving the given dynamic programming. Unfortunately, due to the impact of the current action on the future reward and the unaccountable space of the belief vector, obtaining the optimal solution directly from the given recursive equations is computationally prohibitive. Therefore, we focus on the implementable myopic sensing policy, maximizing the immediate reward, formally defined as follows.

Definition 1: Let $F(\Omega_A(t)) \triangleq \mathbb{E}[R_{\pi_t}(\Omega(t))]$ denote the expected immediate reward obtained in slot t under action $\mathcal{A}(t)$ with $\Omega_A(t) \triangleq \{\omega_i(t), i \in \mathcal{A}(t)\}$, the myopic sensing policy $\widetilde{\mathcal{A}}(t)$ consists of sensing a subset of k channels that maximizes $F(\Omega_A(t))$, i.e., $\widetilde{\mathcal{A}}(t) \triangleq \text{argmax}_{\mathcal{A}(t)\subseteq\mathcal{N}} F(\Omega_A(t))$.

For convenience of the next analysis, we state some properties of $\tau_i(\omega_i(t))$ and $\varphi(\omega_i(t))$.

Lemma 1: Given $p_{01}^i < p_{11}^i$ $(i \in \mathcal{N})$, we have:

1) $\tau_i(\omega_i(t))$ and $\varphi(\omega_i(t))$ monotonically increase with $\omega_i(t)$;
2) $p_{01}^i \le \tau_i(\omega_i(t)) \le p_{11}^i, \forall\, 0 \le \omega_i(t) \le 1$.

Proof: Noticing $\tau_i(\omega_i(t)) = (p_{11}^i - p_{01}^i)\omega_i(t) + p_{01}^i$ and $\varphi(\omega_i) = (\epsilon\omega_i(t)/\epsilon\omega_i(t) + 1 - \omega_i(t))$, Lemma 1 holds. ∎

III. g-Decomposability Function

Here, we first introduce g-decomposability and then defines a family of generic functions, namely, *g-regular* functions, which serve as a basis for further analysis on the structure and the optimality of the myopic sensing policy. By slightly abusing the notations without introducing ambiguity, we drop the time index of $\omega_i(t)$ $(\mathcal{A}(t))$ and use $\omega_i(t)$ $(\mathcal{A}(t))$ and ω_i $(\mathcal{A})$ interchangeably.

The following axiom significantly extends the axiom of decomposability in [5] to cover a much larger range of utility functions in combination with symmetry and monotonicity.

Axiom 1 [g-Decomposability]: A function $f(\Omega_A) : [0,1]^k \to \mathbb{R}$ is decomposable if there exists a continuous, differential, and increasing function $g : [0,1] \to [0,\infty)$ and a constant c such that for any $i \le k$

$$f(\omega_1, \ldots, \omega_i, \ldots, \omega_k) = c \cdot g(\omega_i) f(\omega_1, \ldots, 1, \ldots, \omega_k) + c \cdot (1 - g(\omega_i))\, f(\omega_1, \ldots, 0, \ldots, \omega_k).$$

[2]If no information on the initial system state is available, each entry of $\Omega(1)$ can be set to the stationary distribution $\omega_0^i = (p_{01}^i)/(1 + p_{01}^i - p_{11}^i)$, $1 \le i \le N$.

It is insightful to note that the axiom of g-decomposability significantly extends the axiom of decomposability in [5] by covering a much larger range of utility functions, particularly the logarithmic functions (e.g., $f(\Omega_A) = \sum_{i=1}^k \log_a(1+\omega_i)$ $(a > 1)$, where $c = (1/\log_2 a)$, $g(\omega_i) = \log_2(1+\omega_i))$ and the power functions (e.g., $f(\Omega_A) = \sum_{i=1}^k \omega_i^a$, $a > 0$, where $c = 1$, $g(\omega_i) = \omega_i^a$). By setting $g(\omega_i) = \omega_i$ and $c = 1$, Axiom 1 degenerates to the Axiom of decomposability in [5].

In the following, we characterize a family of generic functions, which are referred to as *g-regular* functions, defined as follows.

Definition 2 [g-Regular Function]: A function is called g-regular if it satisfies symmetry, monotonicity, and g-decomposability.

When g-regularity holds, the myopic sensing policy, which is defined in Definition 1, consists of sensing the k channels with the largest belief values, which is easy to obtain by monotonicity.

IV. Analysis on Optimality of Myopic Policy Under Imperfect Sensing

Here, we set up by defining the pseudo value function and then derive the closed-form conditions to guarantee the optimality of the myopic policy.

A. Pseudo Value Function

Armed with three axioms, we first define *pseudo value function* in the imperfect-sensing case, which is crucial in the study of the optimality of the myopic sensing policy.

Definition 3: The pseudo value function, which is denoted by $W_t^{\mathcal{A}}(\Omega(t))$ $(1 \le t \le T,\ t+1 \le r \le T)$, is recursively defined as follows:

$$\begin{cases} W_T\left(\Omega(T)\right) = F\left(\Omega_{\widetilde{A}}(T)\right) \\ W_r\left(\Omega(r)\right) = F\left(\Omega_{\widetilde{A}}(r)\right) + \beta \sum_{\mathcal{E}\subseteq\widetilde{\mathcal{A}}(r)} \\ \qquad\qquad \times Pr\left(\widetilde{\mathcal{A}}(r), \mathcal{E}\right) W_{r+1}\left(\Omega_{\mathcal{E}}(r+1)\right) \\ W_t^{\mathcal{A}}\left(\Omega(t)\right) = F\left(\Omega_A(t)\right) \\ \qquad + \underbrace{\beta \sum_{\mathcal{E}\subseteq\mathcal{A}(t)} Pr(\mathcal{A}(t), \mathcal{E}) W_{t+1}\left(\Omega_{\mathcal{E}}(t+1)\right)}_{\Gamma^{\mathcal{A}}(\Omega(t))} \end{cases} \tag{3}$$

where $\Omega_{\mathcal{E}}(t+1)$ and $\Omega_{\mathcal{E}}(r+1)$ are generated by $\langle\Omega(t), \mathcal{A}(t), \mathcal{E}\rangle$ and $\langle\Omega(r), \widetilde{\mathcal{A}}(r), \mathcal{E}\rangle$, respectively, according to (1), and $Pr(\mathcal{M}, \mathcal{E}) \triangleq \prod_{i\in\mathcal{E}}(1-\epsilon)\omega_i(t) \prod_{j\in\mathcal{M}\setminus\mathcal{E}}[1-(1-\epsilon)\omega_j(t)]$.

The pseudo value function gives the expected accumulated discounted reward of the following sensing policy: In slot t sense the channels in $\mathcal{A}(t)$ and then sense the channels in $\widetilde{\mathcal{A}}(r)(t+1 \le r \le T)$ (i.e., adopt the myopic policy from slots $t+1$ to T). If $\mathcal{A}(t) = \widetilde{\mathcal{A}}(t)$, $W_t^{\mathcal{A}}(\Omega(t))$ is the reward with the myopic policy adopted from slots t to T.

B. Myopic Sensing Policy: Condition of Optimality

For convenience of discussion, we introduce some notations.

1) $p_{11}^{\max} \triangleq \max_{i\in\mathcal{N}}\{p_{11}^i\}$, $p_{01}^{\min} \triangleq \min_{i\in\mathcal{N}}\{p_{01}^i\}$;
2) $\delta_p^{\max} \triangleq \max_{i\in\mathcal{N}}\{p_{11}^i - p_{01}^i\}$, $\delta_p^{\min} \triangleq \min_{i\in\mathcal{N}}\{p_{11}^i - p_{01}^i\}$;
3) $g'_{\max} \triangleq \max_{p_{01}^{\min} \le \omega \le p_{11}^{\max}}\{dg(\omega)/d\omega\}$;
4) $g'_{\min} \triangleq \min_{p_{01}^{\min} \le \omega \le p_{11}^{\max}}\left\{dg(\omega)/d\omega\right\}$;
5) $\Omega_{-i} \triangleq \{\omega_j : j \in \mathcal{N}, j \ne i\}$: the belief vector except ω_i;

6) $\omega_{-i} \triangleq \{\omega_j : j \in \mathcal{A}, j \neq i\}$, $\Delta_{\max}(\Delta_{\min}) \triangleq \max(\min)_{\omega_{-i}\in[0,1]^{N-1}} \{F(1, \omega_{-i}) - F(0, \omega_{-i})\}$.

We start by showing the following important lemma (Lemma 2) and then establish the sufficient condition under which the optimality of the myopic sensing policy is ensured. In Lemma 2, we consider $\Omega_l = [\omega_1, \ldots, \omega_l, \ldots, \omega_N]$ and $\Omega'_l = [\omega_1, \ldots, \omega'_l, \ldots, \omega_N]$, which differ only in one element $\omega'_l \geq \omega_l$. Let $\mathcal{A}'$ and $\mathcal{A}$ denote the k channels with the largest belief values in Ω'_l and Ω_l, respectively, where $\mathcal{A}'$ differs from $\mathcal{A}$ in channel 1 at most. Lemma 2 gives the upper and lower bounds of $W_t^{\mathcal{A}'}(\Omega'_l) - W_t^{\mathcal{A}}(\Omega_l)$.

Lemma 2: If $F(\Omega_A)$ is g-regular, $\omega_l \leq \omega_{l'}$ ($l \in \mathcal{N}$) and $1 \leq t \leq T$, the following conditions hold.

1) If $l \in \mathcal{A}'$ and $l \in \mathcal{A}$, then

$$c \cdot (\omega'_l - \omega_l)\, g'_{\min}\Delta_{\min} \leq W_t^{\mathcal{A}'}(\Omega'_l) - W_t^{\mathcal{A}}(\Omega_l)$$
$$\leq c \cdot (\omega'_l - \omega_l)\, g'_{\max}\Delta_{\max} \sum_{i=0}^{T-t} \beta^i \left(\delta_p^{\max}\right)^i.$$

2) If $l \notin \mathcal{A}'$ and $l \notin \mathcal{A}$, then $0 \leq W_t^{\mathcal{A}'}(\Omega'_l) - W_t^{\mathcal{A}}(\Omega_l) \leq c \cdot (\omega'_l - \omega_l) g'_{\max}\Delta_{\max} \sum_{i=1}^{T-t} \beta^i (\delta_p^{\max})^i$.
3) If $l \in \mathcal{A}'$ and $l \notin \mathcal{A}$, then $0 \leq W_t^{\mathcal{A}'}(\Omega'_l) - W_t^{\mathcal{A}}(\Omega_l) \leq c \cdot (\omega'_l - \omega_l) g'_{\max}\Delta_{\max} \sum_{i=0}^{T-t} \beta^i (\delta_p^{\max})^i$.

Proof: The proof is given in Appendix. ■

Remark: We would like to emphasize that imperfect sensing brings about the nonlinear propagation of the belief vector such that it renders the LP relaxation [10], [11] for perfect sensing infeasible to provide guaranteed approximation ratio algorithms for the imperfect-sensing case. Therefore, any original study on the optimality of this problem is particularly required. Under this context, we observe from the proof that the problem in this paper is isomorphic with that in [5] from the perspective of *decomposability*, that is to say, $W_t^{\mathcal{A}}(\Omega(t))$ is decomposable. Decomposability makes the proof seem similar to that in [5]; however, in fact, there exists some significant difference in the critical induction of the proof, i.e., inequalities (4)–(6) differing from the counterparts in [5] and improving the upper bound (when $l \notin \mathcal{A}'$, $l \notin \mathcal{A}$) and leading to different conditions in Lemma 3 compared with [5].

In the following lemma, we consider $\mathcal{A}_l$ and $\mathcal{A}_m$ differing in one element, that is, $\mathcal{A}_l \setminus \{l\} = \mathcal{A}_m \setminus \{m\}$, $l \in \mathcal{A}_l$ and $m \in \mathcal{A}_m$ and $\omega_l > \omega_m$, and establish a sufficient condition such that $W_t^{\mathcal{A}_l}(\Omega) > W_t^{\mathcal{A}_m}(\Omega)$.

Lemma 3: If $(g'_{\min}\Delta_{\min}/g'_{\max}\Delta_{\max}) \geq \sum_{i=1}^{T-1} \beta^i (\delta_p^{\max})^i$ and $F(\Omega_A)$ is g-regular, then $W_t^{\mathcal{A}_l}(\Omega) > W_t^{\mathcal{A}_m}(\Omega)$.

Proof: Let Ω' denote the set of channel belief values with $\omega'_l = \omega_m$ and $\omega'_i = omega_i$ for $\forall i \neq l$ and $i \in \mathcal{N}$, then $W_t^{\mathcal{A}_l}(\Omega') = W_t^{\mathcal{A}_m}(\Omega')$. By Lemma 2, we have

$$\begin{aligned}
&W_t^{\mathcal{A}_l}(\Omega) - W_t^{\mathcal{A}_m}(\Omega)\\
&\quad= \left[W_t^{\mathcal{A}_l}(\Omega) - W_t^{\mathcal{A}_l}(\Omega')\right] - \left[W_t^{\mathcal{A}_m}(\Omega) - W_t^{\mathcal{A}_m}(\Omega')\right]\\
&\quad> c(\omega_l - \omega_m) g'_{\min}\Delta_{\min}\\
&\quad\quad - c \cdot (\omega_l - \omega_m) g'_{\max}\Delta_{\max} \sum_{i=1}^{T-t} \beta^i \left(\delta_p^{\max}\right)^i\\
&\quad\geq c(\omega_l - \omega_m) g'_{\max}\Delta_{\max}\\
&\quad\quad \cdot \left[\frac{g'_{\min}}{g'_{\max}} \cdot \frac{\Delta_{\min}}{\Delta_{\max}} - \sum_{i=1}^{T-1} \beta^i \left(\delta_p^{\max}\right)^i\right] \geq 0.
\end{aligned}$$

■

The following theorem states the optimality of the myopic sensing policy under imperfect sensing.

Theorem 1: The myopic sensing policy is optimal if the following two conditions hold: 1) The expected slot reward function $F(\Omega_A)$ is g-regular, and 2) $(g'_{\min}\Delta_{\min}/g'_{\max}\Delta_{\max}) \geq \sum_{i=1}^{T-1} \beta^i (\delta_p^{\max})^i$.

Proof: We prove the theorem by backward induction. The theorem trivially holds for $t = T$. Assume that it holds for $T, T-1, \ldots, t+1$. We now show that it holds for t. To this end, assume, by contradiction, that given the belief vector $\Omega(t) \triangleq \{\omega_{i_1}, \ldots, \omega_{i_N}\}$ and $\widetilde{\mathcal{A}}(t) = \{1, \ldots, k\}$, the optimal sensing policy is to sense channels $\{i_1, \ldots, i_k\} \neq \widetilde{\mathcal{A}}(t)$ at slot t and to sense $\widetilde{\mathcal{A}}(r)(t+1 \leq r \leq T)$, there must exist i_m and i_l at slot t such that $m \leq k < l$ and $\omega_{i_m} < \omega_{i_k} \leq \omega_{i_l}$. It then follows from Lemma 3 that $W_t^{\{i_1,\ldots,i_k\}}(\Omega) < W_t^{\{i_1,\ldots,i_{m-1},i_l,i_{m+1},\ldots,i_k\}}(\Omega)$, that is, sensing channels $\{i_1, \ldots, i_k\}$ at slot t is not optimal, which contradicts the assumption. This contradiction completes our proof. ■

Theorem 1 generalizes the results with perfect sensing [5, Th. 1] by: 1) covering a much larger class of reward functions of practical importance, including the logarithmic and power functions, and 2) establishing the optimalty conditions in the generic case with a sensing error. The following theorem further establishes the optimality conditions in the asymptotic case $T \to \infty$, which follows straightforwardly from Theorem 1 by noticing that $\sum_{i=1}^{\infty} x^i = x/(1-x)$ for any $x \in (0, 1)$.

Theorem 1: In the infinite-horizon case $T \to \infty$, the myopic policy is optimal if the following conditions hold: 1) The expected slot reward function $F(\Omega_A)$ is g-regular; 2) $\beta \leq (g'_{\min}\Delta_{\min}/(g'_{\min}\Delta_{\min} + g'_{\max}\Delta_{\max})\delta_p^{\max})$.

C. Discussion

It is insightful to compare the results established in this paper with those presented in [8]. To this end, we consider the scenario where a user is limited to sense k of N i.i.d. channels and gets one unit of reward if the sensed channel is in the good state, i.e., the utility function can be formulated as $F(\Omega_A) = \sum_{i \in A}[(1-\epsilon)\omega_i]$. It can be checked that $\Delta_{\min} = \Delta_{\max} = 1 - \epsilon$. We can then verify that when $\epsilon < (p_{01}(1-p_{11})/P_{11}(1-p_{01}))$, it holds that $(\Delta_{\min}/\Delta_{\max}[(1-\epsilon)(1-p_{01}) + (\epsilon(p_{11}-p_{01})/1-(1-\epsilon)(p_{11}-p_{01}))]) > 1$. Therefore, when conditions 1 and 2 of [8, Th. 1] holds, the myopic sensing policy is always optimal for any β. Applying the notations of this paper, we have $c = 1$, $g(\omega) = \omega$ and $\Delta_{\min} = \Delta_{\max} = 1 - \epsilon$. Furthermore, it holds that the myopic policy is optimal for any β and ϵ if $\delta_p^{\max} \leq 0.5$ according to Theorem 2. Compare the optimal conditions established in this paper and those obtained in [8] for i.i.d. channels, we observe that the conditions derived in this paper are stricter in terms of transition probability ($\delta_p^{\max} \leq 0.5$ in our paper) but looser in terms of false-alarm rate ($\epsilon < (p_{01}(1-p_{11})/P_{11}(1-p_{01}))$ in [8]). The stricter constraint on the transition probability is due to the employed method in this paper, which sacrifices part of the optimality to cover the case of nonidentical channels, whereas the looser constraint on the sensing error comes from the fact that all the channels are only categorized into sensed channels and nonsensed channels, while the analysis in [8] relies on the sorting of channels in their belief values.

When the g-regular function is degenerated into a regular function, the optimal condition in Theorem 2 is also degenerated into that of [5, Th. 2]; however, it can cover the scenario with a sensing error. In other words, whether or not there exists a sensing error, the condition to guarantee the optimality of the myopic policy is the same. This can be explained as follows. Although the sensing error brings about the nonlinearity of the system dynamic update, this kind of nonlinearity does not change the *decomposability* [5], [8] of the value function, which is the cornerstone of deriving the closed-formed condition of optimality. Therefore, the RMAB considered in this paper is isomorphic with that in [5] from the perspective of *decomposability*.

V. Conclusion

We have investigated the optimality of the myopic policy in a multichannel access system with imperfect sensing and developed three axioms to characterize a family of important functions, i.e., g-regular functions. By performing a mathematical analysis based on the developed axioms, we have obtained closed-form conditions to guarantee the optimality of the myopic policy. As future work, a natural direction we are pursuing is the investigation of the RMAB problem with multiple players with potential conflicts among them and the study of the structure and the optimality of the myopic policy in that context.

Appendix
Proof of Lemma 2

We prove the lemma by backward induction. For slot T, noticing that $W_T^{\mathcal{A}}(\Omega) = F(\Omega_A)$ and that $g'_{\min} \leq (g(\omega) - g(\omega')/\omega - \omega') \leq g'_{\max}$ for any $p_{01}^{\min} \leq \omega' \leq \omega \leq p_{11}^{\max}$, the following statements hold.

1) For $l \in \mathcal{A}', l \in \mathcal{A}$, it holds that

$$\begin{aligned} c\cdot(\omega_l' - \omega_l)\, g'_{\min}\Delta_{\min} &\leq W_T^{\mathcal{A}'}(\Omega_l') - W_T^{\mathcal{A}}(\Omega_l) \\ &\leq c\cdot\left[g(\omega_l') - g(\omega_l)\right]\Delta_{\max} \\ &\leq c\cdot(\omega_l' - \omega_l)\, g'_{\max}\Delta_{\max}. \end{aligned}$$

2) For $l \notin \mathcal{A}'$, $l \notin \mathcal{A}$, it holds that $W_T^{\mathcal{A}'}(\Omega_l') - W_T^{\mathcal{A}}(\Omega_l) = 0$.

3) For $l \in \mathcal{A}'$, $l \notin \mathcal{A}$, by the definition of the myopic policy, there exists at least one channel m such that $\omega_l' \geq \omega_m \geq \omega_l$. It then holds that

$$\begin{aligned} 0 \leq c\cdot(\omega_l' - \omega_l)\, g'_{\min}\Delta_{\min} &\leq W_T^{\mathcal{A}'}(\Omega_l') - W_T^{\mathcal{A}}(\Omega_l) \\ &\leq c\cdot\left[g(\omega_l') - g(\omega_m)\right]\Delta_{\max} \\ &\leq c\cdot\left[g(\omega_l') - g(\omega_l)\right]\Delta_{\max} \\ &\leq c\cdot(\omega_l' - \omega_l)\, g'_{\max}\Delta_{\max}. \end{aligned}$$

Therefore, Lemma 2 holds for slot T. Assume that Lemma 2 holds for $T, \ldots, t+1$. We now prove the lemma for slot t.

We first prove the first case: $l \in \mathcal{A}'$ and $l \in \mathcal{A}$.

Considering the whole realization of the belief vector and developing on channel l, we have

$$\begin{aligned} &\Gamma^{\mathcal{A}'}(\Omega'(t)) - \Gamma^{\mathcal{A}}(\Omega(t)) \\ &= \sum_{\mathcal{E}\subseteq\mathcal{A}(t)\setminus\{l\}} \prod_{i\in\mathcal{E}}(1-\epsilon)\omega_i(t) \prod_{j\in\mathcal{A}(t)\setminus\mathcal{E}\setminus\{l\}}\left[1-(1-\epsilon)\omega_j(t)\right] \\ &\quad\times\Big\{(1-\epsilon)\left(\omega_l'(t) - \omega_l(t)\right) \\ &\quad\times\left[W_{t+1}\left(\Omega_{-l}; p_{11}^l\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right)\right] \\ &\quad+\left(1-(1-\epsilon)\omega_l(t)\right)\left[W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right)\right. \\ &\qquad\left.- W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi(\omega_l)\right)\right)\right]\Big\}. \end{aligned} \tag{4}$$

Next, we derive the bound of $W_{t+1}(\Omega_{-l}; p_{11}^l) - W_{t+1}(\Omega_{-l}; \tau_l(\varphi(\omega_l')))$ through three cases[3] as follows.

Case 1: If $l \in \mathcal{A}'(t+1)$ and $l \in \mathcal{A}(t+1)$, according to the induction hypothesis, we have

$$\begin{aligned} 0 &\leq c\cdot\left(p_{11}^l - \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) g'_{\min}\Delta_{\min} \\ &\leq W_{t+1}\left(\Omega_{-l}; p_{11}^l\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) \\ &\leq c\cdot\left(p_{11}^l - \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned}$$

Case 2: If $l \notin \mathcal{A}'(t+1)$ and $l \notin \mathcal{A}(t+1)$, according to the induction hypothesis, we have

$$\begin{aligned} 0 &\leq W_{t+1}\left(\Omega_{-l}; p_{11}^l\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) \\ &\leq c\cdot\left(p_{11}^l - \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) g'_{\max}\Delta_{\max}\sum_{i=1}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned}$$

Case 3: If $l \in \mathcal{A}'(t+1)$ and $l \notin \mathcal{A}(t+1)$, according to the induction hypothesis, we have

$$\begin{aligned} 0 &\leq W_{t+1}\left(\Omega_{-l}; p_{11}^l\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) \\ &\leq c\cdot\left(p_{11}^l - \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned}$$

Combining the three cases, we obtain

$$\begin{aligned} 0 &\leq W_{t+1}\left(\Omega_{-l}; p_{11}^l\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) \\ &\leq c\cdot\left(p_{11}^l - \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i \\ &= c\cdot\left[1 - \frac{\epsilon\omega_l'}{1-(1-\epsilon)\omega_l'}\right] \\ &\quad\times\left(p_{11}^l - p_{01}^l\right) g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned} \tag{5}$$

According to Lemma 1, we have $\tau_l(\varphi(\omega_l')) \geq \tau_l(\varphi(\omega_l))$ when $\omega_l' \geq \omega_l$. Thus, we have the bounds of $W_{t+1}(\Omega_{-l}; \tau_l(\varphi(\omega_l'))) - W_{t+1}(\Omega_{-l}; \tau_l(\varphi(\omega_l)))$ by similar induction as follows:

$$\begin{aligned} 0 &\leq W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi\left(\omega_l'\right)\right)\right) - W_{t+1}\left(\Omega_{-l}; \tau_l\left(\varphi(\omega_l)\right)\right) \\ &\leq c\cdot\left(\tau_l\left(\varphi\left(\omega_l'\right)\right) - \tau_l\left(\varphi(\omega_l)\right)\right) g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i \\ &= c\cdot\frac{\epsilon\left(\omega_l' - \omega_l\right)\left(p_{11}^l - p_{01}^l\right) g'_{\max}\Delta_{\max}}{\left[1-(1-\epsilon)\omega_l'\right]\left[1-(1-\epsilon)\omega_l\right]}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned} \tag{6}$$

Combining (4)–(6), and $p_{11}^l - p_{01}^l \leq \delta_p^{\max}$, we have

$$\begin{aligned} 0 &\leq \Gamma^{\mathcal{A}'}(\Omega'(t)) - \Gamma^{\mathcal{A}}(\Omega(t)) \\ &\leq c\cdot(\omega_l' - \omega_l)\,\delta_p^{\max} g'_{\max}\Delta_{\max}\sum_{i=0}^{T-t-1}\beta^i\left(\delta_p^{\max}\right)^i. \end{aligned}$$

[3]It can be noted that the case $l \notin \mathcal{A}'(t+1)$ and $l \in \mathcal{A}(t+1)$ is impossible according to the definition of myopic policy and (3).

Since

$$c \cdot (\omega_l' - \omega_l)\, g_{\min}' \Delta_{\min} \le F(\Omega_{A'}(t)) - F(\Omega_A(t)) \le c \cdot (\omega_l' - \omega_l)\, g_{\max}' \Delta_{\max}$$

we have

$$\begin{aligned} & c \cdot (\omega_l' - \omega_l)\, g_{\min}' \Delta_{\min} \\ & \le W_t^{\mathcal{A}'}(\Omega_l') - W_t^{\mathcal{A}}(\Omega_l) \\ &= F(\Omega_{A'}(t)) - F(\Omega_A(t)) + \beta \left(\Gamma^{\mathcal{A}'}(\Omega'(t)) - \Gamma^{\mathcal{A}}(\Omega(t)) \right) \\ & \le c \cdot (\omega_l' - \omega_l)\, \Delta_{\max} g_{\max}' \left[1 + \beta \cdot \delta_p^{\max} \sum_{i=0}^{T-t-1} \beta^i \left(\delta_p^{\max}\right)^i \right] \\ &= c \cdot (\omega_l' - \omega_l)\, g_{\max}' \Delta_{\max} \sum_{i=0}^{T-t} \beta^i \left(\delta_p^{\max}\right)^i . \end{aligned}$$

We thus complete the proof of the first part ($l \in \mathcal{A}'$ and $l \in \mathcal{A}$) of Lemma 2.

Regarding the second case $l \notin \mathcal{A}'$ and $l \notin \mathcal{A}$ and the third case $l \in \mathcal{A}'(t)$ and $l \notin \mathcal{A}(t)$, the proof is basically an adaptive extension of the proof of [5, Lemma 4] by incorporating the imperfect observation on the channel state; hence, it is omitted here.

References

[1] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queueing network control," *Math. Oper. Res.*, vol. 24, no. 2, pp. 293–305, May 1999.

[2] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: Structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 5413–5440, Dec. 2008.

[3] S. H. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, Sep. 2009.

[4] S. Ahmad and M. Liu, "Multi-channel opportunistic access: A case of restless bandits with multiple plays," in *Proc. Allerton Conf.*, Monticello, Il, USA, Sep.–Oct. 2009, pp. 1361–1368.

[5] K. Wang and L. Chen, "On optimality of myopic policy for restless multi-armed bandit problem: An axiomatic approach," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 300–309, Jan. 2012.

[6] K. Wang, Q. Liu, and L. Chen, "On optimality of greedy policy for a class of standard reward function of restless multi-armed bandit problem," *IET Signal Process.*, vol. 6, no. 6, pp. 584–593, Aug. 2012.

[7] K. Liu, Q. Zhao, and B. Krishnamachari, "Dynamic multichannel access with imperfect channel state detection," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2795–2807, May 2010.

[8] K. Wang, L. Chen, Q. Liu, and K. Al Agha, "On optimality of myopic sensing policy with imperfect sensing in multi-channel opportunistic access," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3854–3862, Sep. 2013.

[9] K. Wang, L. Chen, and Q. Liu, "Opportunistic spectrum access by exploiting primary user feedbacks in underlay cognitive radio systems: An optimality analysis," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 5, pp. 869–882, Oct. 2013.

[10] S. Guha and K. Munagala, "Approximation algorithms for partial-information based stochastic control with Markovian rewards," in *Proc. IEEE Symp. FOCS*, Providence, RI, USA, Oct. 2007, pp. 483–493.

[11] S. Guha and K. Munagala, "Approximation algorithms for restless bandit problems," in *Proc. ACM-SIAM SODA*, New York, NY, USA, Jan. 2009, pp. 28–37.

Optimization of an Amplify-and-Forward Relay Network Considering Time Delay and Estimation Error in Channel State Information

Ying Zhang, Huapeng Zhao, *Member, IEEE*, and Chuanyi Pan

***Abstract*—This paper presents the optimization of an amplify-and-forward (AF) relay network with time delay and estimation error in channel state information (CSI). The CSI time delay and estimation error are modeled by the channel time variation model and stochastic error model, respectively. The conditional probability density function of the ideal CSI upon the estimated CSI is computed based on these two models, and it is used to derive the conditional expectation of the mean square error (MSE) between estimated and desired signals upon estimated CSI, which is minimized to optimize the beamforming and equalization coefficients. Computer simulations show that the proposed method obtains lower bit error rate (BER) than the conventional minimum MSE and the maxmin SNR strategies when CSI contains time delay and estimation error.**

***Index Terms*—Amplify and forward (AF), conditional expectation, estimation error, minimum mean square error (MMSE), outdated channel state information (CSI), relay network.**

I. Introduction

Relay technique is capable of extending communication range and coverage, and it has received extensive study in recent years. To maximize the benefit of relay technique, various protocols were proposed, which include fixed relaying schemes, selection relaying schemes, and incremental relaying schemes [1]–[3]. Meanwhile, it is important to optimize the relay beamforming and equalization coefficients to obtain optimal quality of communication. Perfect channel state information (CSI) was assumed in many existing relay network optimization methods [4]–[8]. In practice, the CSI is usually obtained through estimation, and the estimated CSI is different from the ideal CSI due to potential noises such as estimation error and feedback error. Therefore, the design of relay network with imperfect CSI has attracted much interest recently [9]–[15].

In [13], two kinds of models were presented to describe CSI imperfectness, namely, the stochastic error model and the norm-bounded error model, which are applicable to CSI estimation and quantization errors, respectively. In addition to estimation and quantization errors, the channel time delay may also contribute to the CSI imperfectness, which is known as the outdated CSI phenomena. Here, the time delay includes delays due to CSI estimation and feedback. For point-to-point communication system, the modeling of CSI estimation error and time delay was introduced in [16]. Meanwhile, relay selection with outdated CSI estimation has been recently studied extensively [17]–[20]. It was shown that CSI time delay affected the performance

Manuscript received February 18, 2013; revised July 8, 2013 and November 12, 2013; accepted November 22, 2013. Date of publication November 26, 2013; date of current version June 12, 2014. The work of Y. Zhang and C. Pan was supported by the National Natural Science Foundation of China for Young Scholars under Grant 61101094 and Grant 61201275. The review of this paper was coordinated by Prof. W. A. Hamouda.

Y. Zhang and C. Pan are with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: superlapin007@hotmail.com).

H. Zhao is with the Department of Electronics and Photonics, Institute of High Performance Computing, Singapore 138632 (e-mail: huapengzhao@ieee.org).

Digital Object Identifier 10.1109/TVT.2013.2292939