Exploiting Orthogonally Dual-Polarized Antennas in Cooperative Cognitive Radio Networking

Bin Cao, Student Member, IEEE, Hao Liang, Student Member, IEEE, Jon W. Mark, Life Fellow, IEEE, and Qinyu Zhang, Senior Member, IEEE

Abstract—This work is concerned with enhancement of spectrum utilization by using polarization enabled two-phase cooperation between primary users (PUs) and secondary users (SUs) in cooperative cognitive radio networking (CCRN). The use of orthogonally dual-polarized antennas (ODPAs) enables concurrent transmissions of multiple independent signals of PUs and SUs, and interference suppression via polarization zero-forcing and polarization filtering to obtain significant performance improvement. To maximize a weighted sum throughput of PUs and SUs under energy/power constraints, the problem is formulated and solved based on a multi-timescale Markov decision process, and two modified backward iteration algorithms are devised to attain the optimal policies. Numerical results validate the effectiveness of the proposed CCRN framework, showing that the obtained policy outperforms both greedy and random ones.

Index Terms—Cooperative cognitive radio networking, ODPA, Markov decision process, backward iteration.

I. INTRODUCTION

SPECTRUM is statically and exclusively allocated to dedicated networks and services, i.e., only licensed users, also referred to as primary users (PUs), can access the assigned spectrum. However, the legacy fixed spectrum access leads to significant spectral underutilization owing to the sporadic use of spectrum. To address this issue, dynamic spectrum access has been proposed [1]–[4]. The methodology of dynamic spectrum access is to enable unlicensed users, referred to as secondary users (SUs), to opportunistically use licensed spectrum without causing harmful interference to the PUs. One key enabling technique for dynamic spectrum access is the cognitive radio equipped by SUs. A network consists of cognitive radios is referred to as a cognitive radio network (CRN).

Manuscript received April 15, 2012; revised August 31, 2012 and November 9, 2012.

Bin Cao and Q. Zhang (corresponding author) are with the Communication Engineering Research Center (CERC), School of Electronics and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, Guangdong, P. R. China, 518055 (e-mail: caobin@hitwh.edu.cn; qzy@hit.edu.cn).

H. Liang and J. W. Mark are with the Center for Wireless Communications, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1 (e-mail: {hliang, jwmark}@uwaterloo.ca).

This research has been supported in part by the Key Project of Chinese National Programs for Fundamental Research and Development under Grant No. 2009CB320402, the State Key Program of National Natural Science of China under Grant No. 61032003, the Chinese Scholarship Council, and a Discovery Grant under Grant No. RGPIN77779 from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Digital Object Identifier 10.1109/JSAC.2013.13111

CRN can be generally categorized into interweave CRN, underlay CRN, and overlay CRN, respectively [5].

Since there are no connections between PUs and SUs in both interweave and underlay modes, i.e., SUs are transparent to PUs, mutual benefit based cooperation between PUs and SUs is a promising way to implement dynamic spectrum access. In overlay CRN, PUs and SUs dynamically select appropriate partners for cooperative relaying to create a win-win situation. The rationale behind this framework is that PUs have the need to improve primary transmissions and SUs have the need to access spectrum for secondary transmissions. Furthermore, the cooperation mechanism can be an alternative way for SUs to access spectrum if no idle spectrum holes can be detected in the spectrum sensing based framework.

The focus of this work is on exploiting user cooperation between PUs and SUs to attain overlay CRN, also known as cooperative cognitive radio networking (CCRN) [6]–[15]. In CCRN, a PU is free to select one or more SUs as relays, and SUs obtain transmission opportunities from the PU as rewards. Considerable cooperation frameworks have been proposed for CCRN, such as three-phase time division multiple access based CCRN [6]–[8], two-phase frequency division multiple access based CCRN [9], two-phase space division multiple access based CCRN [12], and two-phase quadrature signaling based CCRN [13], [14].

All the above-mentioned schemes deal with cooperation within a single frame in which the channel is assumed to experience spatial flat fading. Both SUs and PUs aim to optimize the cooperation performance in a greedy manner. However, for an energy-constrained network, these schemes can no longer be optimal. Another challenging issue is that each cooperation should be finalized within one frame, while a new relay selection procedure is required at the beginning of the next frame. A trade-off between the cooperative diversity and relay selection overhead should be addressed.

To tackle the aforementioned challenging issues, we propose to utilize the polarization property of electromagnetic waves in CCRN. In the system networked by PUs and SUs, mobile primary transmitters (PTs) communicate with their intended primary receivers (PRs) (e.g., primary base stations), and mobile secondary transmitters (STs) need to communicate with secondary receivers (SRs) (e.g., secondary access points). Both PTs and STs are energy-constrained. SUs are able to avoid interference to PUs and other cooperating SUs by only leveraging polarization filtering and polarization zero-
forcing capabilities provided by orthogonally dual-polarized antennas (ODPAs), given that PUs still equip with legacy uni-polarized antennas. The other reason that we introduce polarization in CCRN is due to the limitation of MIMO based frameworks in some applications with physical size constraint. Although the MIMO technology can offer significant enhancement in throughput and link range without additional bandwidth and/or transmit power, the application of MIMO is limited by the hardware cost and physical size. The capability and degrees of freedom provided by MIMO systems are determined by sufficient space interval of antenna elements. This results in a large hardware size of SUs.

One obvious advantage of using ODPA is that there is no spacing requirement for the two antennas since they are co-located, which makes ODPA more suitable for devices with size and cost limitation [16]. This enables SUs to relay for PUs while concurrently accessing the same spectrum to transmit their own information in cost- and space-efficient manners. Another byproduct of using ODPA is the fact that polarization fading in terms of cross-polar discrimination (XPD) changes more slowly than its spatial counterpart, multi-path fading [17]. The use of ODPA enables SUs to keep using the same polarization states to attain interference suppression in a large time-scale. This indicates that polarization based scheme is more stable than MIMO based scheme with respect to (w.r.t.) time variations. The corresponding resource allocation problem is formulated as a multi-timescale dynamic programming in terms of maximizing a weighted sum throughput subject to power constraints. The optimal policies of the frame- and superframe-level resource allocation problems are proved and analyzed. Modified backward iteration algorithms and the associated numerical and simulations validate the effectiveness of the proposed scheme, showing that the proposed policy outperforms both greedy and random power policies. To the best of our knowledge, this is the first work that utilizes polarization into energy-constrained CCRN.

The remainder of the paper is organized as follows. Section II describes the system model and fundamentals on polarization. In Section III, the resource allocation is formulated as a weighted sum throughput maximization problem. The proposed allocation policies and detailed analysis are presented in Section IV. Modified backward iteration algorithms and the associated numerical results are given in V, followed by concluding remarks and future work in Section VI.

II. SYSTEM MODEL

A. Networking Architecture and System Description

Consider a system consists of both PUs and SUs. Mobile PTs communicate with PRs, and mobile STs need to communicate with SRs, while SUs are networked in a CCRN fashion. PTs and STs are battery-powered, while PRs and SRs are AC-powered. In addition, PUs and SUs transceivers work at half-duplex mode, and SUs use the decode-and-forward mode to relay PUs’ packets. The cooperators fully trust each other, i.e., there are no malicious users and misbehaviors after the cooperation is established. Furthermore, it is assumed that SUs utilize the capability provided by ODPAs in terms of vertical (V) polarization and horizontal (H) polarization to access PUs’ spectrum for secondary transmissions coexisting with associated primary transmissions. The PUs are equipped with the traditional vertically uni-polarized antenna. In this context, the links between SUs, from SUs to PUs, from PUs to SUs, and between PUs are VH-to-VH, VH-to-V, V-to-VH, and V-to-V transmissions in terms of polarization, respectively, where VH-to-VH means the transmitter and receiver both have ODPAs in terms of vertical and horizontal polarization, and similar for VH-to-V, V-to-VH, and V-to-V.

B. Representation of Polarization

In a right-handed x-y-z Cartesian coordinate system with the z-coordinate representing the propagating direction, and the x-coordinate and y-coordinate representing H- and V-polarized vectors, the two-dimensional signal \( x(t) \) radiated by ODPA can be expressed in the form of Jones vector as

\[
x(t) = \left[ \begin{array}{c} x_H(t) \\ x_V(t) \end{array} \right] = \left[ \begin{array}{c} \cos \varepsilon \\ \sin \varepsilon \exp(j\delta) \end{array} \right] x(t) = U x(t),
\]

(1)

where \( x(t) \) is the waveform of \( x(t) \), \( \varepsilon \) is the polarized angle which denotes amplitude relationship between V- and H-polarized components, i.e., \( \varepsilon = \arctan(|x_V(t)|/|x_H(t)|) \), and \( \delta \) describes phase difference between them, i.e., \( \delta = \arg\{x_V(t)\} - \arg\{x_H(t)\} \), and \( U = [\cos \varepsilon, \sin \varepsilon \exp(j\delta)]^T \) is the transmitted polarization state of \( x(t) \).

Due to imperfect antenna cross-polar isolation and cross-polar ratio caused by propagation mediums [16], the effect
of depolarization which is determined by these two factors, should be considered. The degree of depolarization is described by a cross-polar discrimination (XPD), and XPD varies slowly compared with multi-path fading when signals propagate in the wireless channel [16], [17]. For simplicity, we further assume that ODPA is comprised of linearly polarized antennas, so that cross-polar isolation and cross-polar ratio can be decoupled. The ODPA used in our work is with infinite cross-polar isolation, then cross-polar ratio is the only factor determines XPD. Therefore, the depolarization effect can be described by a 2x2 matrix \( D \) [16], given by

\[
D = \begin{bmatrix}
D_{HH} & D_{HV} \\
D_{VH} & D_{VV}
\end{bmatrix}.
\]  

(2)

Furthermore, we have \( D_{HH} = D_{HV} = 0 \) for VH-to-VH transmissions and \( D_{VH} = D_{HH} = 0 \) for V-to-V and V-to-VH transmissions. Throughout this work, the XPD is assumed to be known by SUs. Interested readers can refer to [17] for more details about XPD estimation.

C. Fundamentals of Exploiting Polarization for CCRN

As one feasible application of polarization processing to suppress co-channel interference, polarization filtering has been widely studied and applied in radar and communication systems [18]. The main principle of polarization filtering is based on the assumption that the polarization state of interference is different from that of the desired signal. In this work, oblique projection polarization filtering is used by SUs to separate secondary and primary signals [19], [20], and polarization zero-forcing is used to suppress SUs signals.

As shown in Fig. 1, PT sends \( x_p(t) \) to PR, and ST sends \( x_s(t) \) to SR. The polarization states used by PT and ST are denoted by \( U_P = [0, \pm 1]^T \) and \( U_S = [\cos \varepsilon P, \pm \sin \varepsilon P]^T \), respectively. According to (1), the received signal \( y(t) \) at SR can be expressed as

\[
y(t) = h_{SS} U_S x_S(t) + h_{PS} U_P x_P(t) + w,
\]  

(3)

where \( h_{SS} = h_{SS} D_{SS} \) (resp. \( h_{PS} = h_{PS} D_{PS} \)) is the composite channel fading matrix containing spatial fading coefficient \( h_{SS} \) (resp. \( h_{PS} \)) and polarization fading in terms of depolarization matrix \( D_{SS} \) (resp. \( D_{PS} \)) of ST to SR (resp. PT to SR). Matrix \( w \) represents the two-dimensional additive noise with zero mean, and covariance matrix \( \sigma^2 I \).

Define \( D_{SS} U_S = d_S \) and \( D_{PS} U_P = d_P \), the oblique projection operator that projects vectors onto subspace \( \langle d_S \rangle \) along \( \langle d_P \rangle \) can be written as [23]

\[
E_{d_S,d_P} = d_S (d_S^H \mathbf{R}_{d_P}^- d_S)^{-1} d_S^H \mathbf{R}_{d_P}^- d_P,
\]  

(4)

where \( \mathbf{R}_{d_P}^- = I - d_P (d_P^H d_P)^{-1} d_P^H \) is the orthogonal projection operator which projects vectors onto the complementary subspace of \( \langle d_P \rangle \). In this way, the following result holds true

\[
E_{d_S,d_P} y_S(t) = h_{SS} U_S x_S(t) + E_{d_S,d_P} w.
\]  

(5)

Along the same analysis, the SU can extract the PU’s signal by using \( E_{d_P,d_S} \) which is the operator that obliquely projects vectors onto \( \langle d_P \rangle \) along \( \langle d_S \rangle \).

To avoid interference from SUs to PU, polarization zero-forcing is used. The signal received by PR is given by

\[
y_P(t) = \{ h_P P D_{PP} U_P x_P + h_{SP} D_{SP} U_S x_S \} |_V + w,
\]  

(6)

where \( \{x(t)\}_{|V} \) is the V-polarized component of \( x(t) \), and \( w \) is the 1-D additive noise with zero mean and variance \( \sigma^2 \). If \( D_{SP} U_S = 0 \) is achieved, SU’s signal is zero-forced at PR in the polarization domain. Therefore, after getting \( D_{SP} \), the SU can set its polarization state accordingly.

D. Multi-timescale Cooperation Framework

As shown in Fig. 2(a), time is divided into superframes. A PU keeps cooperating with the same SUs for one superframe with length \( T + N(T + \omega) \), i.e., the PU changes its cooperators on a superframe basis. Each superframe is further divided into two parts, the first part is the multi-user coordination process with duration \( \Omega \), i.e., the PU selects the most effective SUs for cooperation. The second part is the cooperation process with length \( N(T + \omega) \), and this process is composed of \( N \) consecutive frames. Constant power is used by a PU to cooperate with SUs within each superframe.

Since XPD changes much slower as compared with spatial multi-path fading, the duration of each superframe is chosen such that XPDs remain constants\(^3\). The length of \( T + \omega \) is determined by Doppler spread w.r.t. licensed spectrum and mobility, e.g., in duration \( T + \omega \), channel suffers flat-fading, and channel coefficients are constants.

For the network scenario under consideration, we investigate two ST-SR pairs simultaneously cooperate with one PT-PR pair during one cooperation period. Each frame in one superframe is divided into two parts, as shown in Fig. 2(b). The first part with duration \( \omega \) is used to estimate channel coefficients of each frame, while the second part is used for two-phase cooperation. Consider one specific frame within the \( m \)th superframe. In the first cooperation phase with duration \( \beta_m T \), say phase A, PT and ST_A transmit to PR and SR_A. At the same time, SR_A and ST_B receive signals from both ST_A and PT. For PR and ST_B, the signal from ST_A is considered as interference. However, by using oblique projection polarization filtering, ST_A’s signal can be nullled at PR, and ST_B

\(^3\)In practice, typical XPD values fall between 8 and 10 dBs in NLOS or LOS cases, respectively, and are surprisingly independent of other environment characteristics and of link distance [21]. In addition, different XPDs can be easily modeled with log-normal distributions (or normal when expressed in dB) [22].
can extract primary signal from mixed signals, while $S_{UA}$ can separate its own desired signal and primary signal. In the second phase with duration $(1 - \beta_m)T$, say phase $B$, $ST_B$ cooperatively relays PT’s information received in phase $A$ while concurrently sends its own signals to $SR_B$. Meanwhile, $SR_A$ forwards the primary information to $PR$. Based on the cooperation framework, the throughput achieved within each phase can be calculated as follows:

**Phase A:** In phase $A$ of the $n$th frame, $PT$ sends $x_{P}\{n\}$ to $PR$ with power $P_{P,n}$. Meanwhile, $STA_A$ sends $x_{SA}\{n\} = U_{STA_A,n}x_{STA_A}\{n\}$ to $STB_A$ and $SR_A$ with power $P_{STA_A,n}$. By using polarization filtering and polarization zero-forcing, the achievable rate of primary and secondary links in the $n$th frame are

$$R_{PS,n,m} = S \left( \frac{\min_{i \in \{A,B\}} \{ \sin^2 \gamma_{P,n,m} |h_{P,n,m}^i|^2 \} P_{P,n} }{\sigma^2} \right)$$

and

$$R_{SA,n,m} = S \left( \frac{\sin^2 \gamma_{P,n,m} |h_{A,n,m}^i|^2 P_{STA_A,n} }{\sigma^2} \right),$$

where $S(x) = \log_2(1 + x)$ is the Shannon capacity, $h_{pq,n,m}$ is the channel coefficient between nodes $p$ and $q$ in the $n$th superframe, and $\gamma_{P,n,m} (i \in \{A,B\})$ is determined by the angle of polarization vectors between PT and SUs, e.g., $\gamma_{PA,m} = \arccos(d_{PA}^m)$ [23].

**Phase B:** $SR_A$ forwards to $PR$ with $P_{STA_A,n}$, and $ST_B$ forwards to $PR$ with $P_{STB,n}$ and concurrently sends to $SR_B$ with $P_{STBS,n}$. $PR$ uses maximal ratio combining to combine signals received from different paths. Therefore, $PU$ and $SU_B$’s achievable rate are given by

$$R_{SP,n,m} = S \left( \frac{|h_{AP,n,m}^i|^2 P_{STA_A,n}}{\sigma^2} + \frac{|h_{BP,n,m}^i|^2 P_{STB,n}}{\sigma^2} \right),$$

and

$$R_{SB,n,m} = S \left( \frac{|h_{BB,n,m}^i|^2 P_{STBS,n}}{\sigma^2} \right).$$

The achievable primary and secondary throughput in the $n$th frame can be obtained by $T_{PS,n} = \min\{\beta_m T R_{PS,n,m}, (1 - \beta_m)T R_{SB,n,m}\}$, $T_{SA,n} = \beta_m T R_{SA,n,m}$, and $T_{SB,n} = (1 - \beta_m)T R_{SB,n,m}$, respectively. Energy consumptions of PT, SU$A$, and SU$B$ in the $n$th frame of the $n$th superframe can be calculated as $\beta_m T P_{P,n} + \beta_m T P_{SA,n,m}$ and $(1 - \beta_m)T (P_{STB,n} + P_{STBS,n})$, respectively. Denote the residual energy of PT, $STA_A$, and $STB_B$ at the beginning of the $n$th frame in the $n$th superframe as $X_{P,n,m} \in [0, X_{PM}]$, $X_{STA_A,n,m} \in [0, X_{STAM}]$, and $X_{STB,n,m} \in [0, X_{STBM}]$, respectively, where $X_{PM}$, $X_{STAM}$, and $X_{STBM}$ are battery capacities. Then, residual energies of PT, $STA_A$, and $STB_B$ in the $(n+1)$th frame are (7), where $[X]^+$ equals to $X$ if $X > 0$ and 0 otherwise. Since the maximum volume of data that can be delivered to PR is $\beta_m T R_{PS,n,m}$, the relaying power $P_{STBP,n}$ used by $STB_B$ is further bounded by $P_{STBPM,n}$, which is given as

$$\left( 1 + \frac{\min_{i \in \{A,B\}} \sin^2 \gamma_{P,n,m} |h_{P,n,m}^i|^2 P_{P,n} }{\sigma^2} \right) \frac{N_0}{|h_{BB,n,m}^i|^2} 2^{\frac{N_0}{|h_{BB,n,m}^i|^2} - 1}.$$
the allocation policies of STA and STB are given by the energy levels and CSI of the PU and SUs, the resource

B. Resource Allocation Problem Formulation

The state variables are defined w.r.t. the residual energy of PT, STA, and STB, i.e., $X_{P,m}$, $X_{STA,m}$, and $X_{STB,m}$, respectively, which evolve according to (7). Denote the CSI at the beginning of the $n$th frame of the $m$th superframe as $S_m = (\gamma_m, h_m)$. The objective of the resource allocation is to maximize the sum of the weighted sum throughput over the $M$ superframes. Then we have the following optimization problem:

$$\begin{align*}
\text{(P1):} \quad \max_{P_P \in P_P} & \quad \sum_{PST \in P_{ST}} \lim_{M \to \infty} \left\{ E \left[ \sum_{m=0}^{M-1} \eta^m \right] \right. \\
& \left. \sum_{n=0}^{N-1} T_{WSum}(S_m, P_{P,m}, P_{ST}, P_{STBS}) \right\}, \quad (11)
\end{align*}$$

where $T_{WSum}(\cdot)$ represents the weighted sum throughput of the $n$th frame in the $m$th superframe defined in Section II. The discount factor $\eta$ is used to emphasize the short-term reward since the system statistics are more likely to change in a distant future. Note that the number of superframes is considered to be sufficiently large in our work, i.e., an approximately infinite horizon problem ($M \to \infty$) is investigated.

IV. Optimal Resource Allocation Policy

According to Markov decision process (MDP), for the multi-timescale resource allocation problem P1, optimal resource allocation policies in the forms of (8), (9), and (10) exist for PT, STA, and STB, respectively [24]. However, finding the optimal policies is computationally prohibitive because of the multi-timescale framework. In order to reduce the computational complexity, we consider a transformation of problem P1, based on which the optimal frame-level resource allocation policy can be obtained explicitly, while the superframe-level resource allocation can be achieved by existing policy iteration algorithms with bounded performance guarantee.

Denote the value function of problem P1 as $\{V(s, a)\}_{s \in S, a \in A}$, where $X_{P,0} = X_P, X_{STA,0} = X_{STA},$ and $X_{STB,0} = X_{STA}$ are the initial battery energy, while $S$ represents the initial value of CSI. Note that the equality in (12) holds since the expectation of summation equals the summation of expectations. By approximating the CSI and the selected SUs as independent among superframes, the value function is given in (13), where the expectation is performed w.r.t. the randomness in the available battery energy of newly selected SUs in the next superframe ($X_{STA}$ and $X_{STB}$) and the CSI ($S'$), while the battery energy of the PU under consideration (from $X_P$ to $X_P'$) evolves according to (7). In (13), the resource allocation vectors within a superframe are simplified by removing subscript $m$ (w.r.t. a tagged superframe) and are given by $P_{STA} = \{P_{STA,m} | m \in \{0,1,2,\ldots,N-1\}\}$, $P_{STBP} = \{P_{STBP,m} | m \in \{0,1,2,\ldots,N-1\}\}$, $P_{STBS} = \{P_{STBS,m} | m \in \{0,1,2,\ldots,N-1\}\}$, and let $P_P, P_{STA}, P_{STBP},$ and $P_{STBS}$ be the class of all admissible policies, respectively.

A quantitative performance degradation of the approximation can be evaluated based on Müller’s work [24]. However, different from traditional problems, the frame-level resource
allocation is performed w.r.t. a finite number of frames (within one superframe). Specifically, based on the specific structure of the framework, an optimal frame-level resource allocation policy can be explicitly derived, to be discussed as follows. Denote the objective functions of STA and STB as $T_{\text{WSum}, A}()$ and $T_{\text{WSum}, B}()$, respectively. By separating the terms in $T_{\text{WSum}, A}()$ w.r.t. STA and STB, we have

\[
T_{\text{WSum}, A}(\mathbf{S}^{(n)}, P_{\text{STA}}^{(n)}) = \zeta_{\text{TSTA}}(\mathbf{S}^{(n)}, P_{\text{STA}}^{(n)}),
\]

\[
T_{\text{WSum}, B}(\mathbf{S}^{(n)}, P_{\text{STB}}, P_{\text{STB}}^{(n)}) = \zeta_{\text{TSTB}}(\mathbf{S}^{(n)}, P_{\text{STB}}^{(n)}) + (1 - \zeta) T_{P}(\mathbf{S}^{(n)}, P_{\text{STB}}^{(n)}).
\]  

Since the resource allocation of STA can be considered as a special case of STB (i.e., one-dimensional power allocation instead of two-dimensional power allocation), we derive the optimal frame-level resource allocation policy for STB first, and then extend the policy to that of STA.

A. Frame-Level Resource Allocation Policy for STB

Given a known $P_{P}$ which is constant during one superframe, the optimal frame-level resource allocation policy for STB within a superframe is given in (14), where the CSI at the frame level evolves according to conditional probability $P_{R}(\mathbf{S}^{(n+1)}|\mathbf{S}^{(n)})$ based on the fading profile. Denote the value function of problem P3-B within frames $\{n, n+1, \cdots, N-1\}$ in (15), where $X_{\text{STB}} = X_{\text{STB}}$ and $\mathbf{S}^{(n)} = \mathbf{S}$. For $n \in [0, N-2]$, the dynamic programming equation of the value function is given in (16), where the conditional expectation is performed w.r.t. the conditional probability $P_{R}(\mathbf{S}|\mathbf{S})$ according to the fading profile. For $n = N - 1$, we have (17).

For notational simplicity, in the following analysis, we consider the $m$th frame as the tagged frame and denote $P_{\text{STB}}^{(n)} = P_{1}$ and $P_{\text{STB}}^{(n)} = P_{2}$, respectively. An illustration of the domain of $(P_{1}, P_{2})$ is given by Fig. 3. Define a resource allocation policy as follows:

**Definition 1**: (Frame-Level Resource Allocation Policy of STB)

\[
(\tilde{P}_{1}, \tilde{P}_{2}) = \begin{cases} (P_{1}^{*}, P_{2}^{*}), & \text{if } (P_{1}^{*}, P_{2}^{*}) \in D \\ (P_{1}^{1B}, P_{2}^{1B}), & \text{otherwise} \end{cases}
\]  

where $D$ is the domain of $(P_{1}, P_{2})$, as shown in Fig. 3. Denote $Q_{n}(\cdot)$ as the objective function to be optimized, i.e.,

\[
Q_{n}(\mathbf{S}, P_{P}, X_{\text{STB}}, P_{1}, P_{2}) = T_{\text{WSum}, B}(\mathbf{S}, P_{P}, P_{1}, P_{2}) + E[V_{n+1}(\mathbf{S}', P_{P}, X_{\text{STB}} - (1 - \beta) T(P_{1} + P_{2})|\mathbf{S})],
\]

where $Q_{N-1} = T_{\text{WSum}, B}(\mathbf{S}, P_{P}, P_{1}, P_{2})$. Let $Q'(\cdot)$ and $Q''(\cdot)$ represent the two lines w.r.t. boundary 1 and boundary 2, respectively, given by

\[
Q'(\mathbf{S}, P_{P}, X_{\text{STB}}, P_{1}) = Q_{n}(\mathbf{S}, P_{P}, X_{\text{STB}}, P_{1}, \min\{P_{\text{STBM}} - \frac{X_{\text{STB}}}{1 - \beta}, P_{1}\}),
\]

\[
P_{1} \in \left[0, \min\{P_{\text{STBM}} - \frac{X_{\text{STB}}}{1 - \beta}, P_{1}\}\right],
\]

\[
Q''(\mathbf{S}, P_{P}, X_{\text{STB}}, P_{2}) = Q_{n}(\mathbf{S}, P_{P}, X_{\text{STB}}, P_{\text{STBM}}^{(n)}, P_{2}),
\]

\[
P_{2} \in \left[0, \min\{P_{\text{STBM}} - \frac{X_{\text{STB}}}{1 - \beta}, P_{2}\}\right].
\]  

Note that boundary 2 exists only when $\min\{P_{\text{STBM}} - \frac{X_{\text{STB}}}{1 - \beta}\} \geq P_{\text{STBM}}^{(n)}$. Accordingly, in (19), $(P_{1}^{*}, P_{2}^{*})$ is an arbitrary point in the optimal set of $Q_{n}(\cdot)$ without considering energy constraints, i.e.,
For the effect of applying the threshold policy, we have the following theorem:

**Theorem 1:** The resource allocation policy given by Definition 1 is optimal for problem P3-B.

**Proof:** The proof is completed by induction. For $n = N - 1$, the optimality of the resource allocation policy is straightforward based on the concavity of $T_{WSum,B}$. Suppose the theorem holds for $n = k + 1$, based on Lemma 4, $V_{k+1}(S,P,X)$ is concave w.r.t. $X$. It follows that $Q_k(S,P_P,P_1,P_2)$ is concave w.r.t. $P_1$ and $P_2$ with the optimal policy for the $k$th frame being given by Definition 1. After applying the policy, $V_k(S,P,P,X)$ is also concave w.r.t. $X$ according to Lemma 4. This completes the proof.

Based on Definition 1, the values of $Q^*, (P^*_1,P^*_2), Q^*_1,$ and $(P^*_1,P^*_2)$ need to be calculated. The calculations of $Q^*$ and $(P^*_1,P^*_2)$ are straightforward since they are not dependent on the residual energy of STB. On the other hand, as the boundary of the domain of $(P_1,P_2)$ is related to the residual energy of STB, the values of $Q^*_2$ and $(P^*_1,P^*_2)$ are functions of $X_{STB}$.

However, by further investigating the problem, we can find simplified forms of the functions.

Let us consider boundary 2 first. Since the value function is concave w.r.t. $P_1$ and $P_2$, it is concave w.r.t. the restriction to any lines [27]. For boundary 2, the optimal value of $Q^*_2(\cdot)$ without considering the energy and transmission power limitation is given by $Q'^* = \max_{P_1 \geq 0}\{Q'(S,P_P,X_{STB},P_2)\}$, which can be achieved by an arbitrary optimal point $(\tilde{P}_{1B_2},\tilde{P}_{2B_2}) \in \{(P_1,P_2)|Q_n(S,P_P,X_{STB},P_1,P_2) = Q'^*\}$. For a concave function, since any local optimum is also global optimum and the optimal set is a convex set, the optimal point on boundary 2 can be calculated using (24).

For boundary 1, if $P_{STBM} \leq X_{STB}/(1-\beta)$, since the line equation of boundary 1 is independent of $X_{STB}$, the optimal point $(P^*_{1B_1},P^*_{2B_1})$ can be obtained in the same way as (24), given by (25), where $Q' = \max_{P_1 \geq 0}\{Q'(S,P_P,X_{STB},P_1)\}$ and $(P^*_{1B_1},P^*_{2B_1}) \in \{(P_1,P_2)|Q_n(S,P_P,X_{STB},P_1,P_2) = Q'^*\}$.

If $P_{STBM} > X_{STB}/(1-\beta)$, since the resource allocation on this boundary saturates the residual energy, there is no energy for subsequent frames. Therefore, the opti-
with \( P_1 \in \left[ 0, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} \right] \) and \( P_2 = \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_1 \) or \( P_1 = P_{STBM}^{(n)} \),
\[
P_2 \in \left[ 0, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_{STBM}^{(n)} \right],
\]
where \( Q_2^* = \max_{i \in \{1,2\}} \{ Q_i^* \} \) with \( Q_1^* = \max_{P_1 \in \left[ 0, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} \right]} \{ \mathcal{Q}(S, P, X_{STB}, P_1) \}, \)
\[
Q_2^* = \max_{P_2 \in \left[ 0, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_{STBM}^{(n)} \right]} \{ \mathcal{Q}''(S, P, X_{STB}, P_2) \}.
\]

\[ (P_{1B_2}^*, P_{2B_2}^*) = \begin{cases} 
(P_{STBM}^{(n)}, P_{STBM}^{(n)}) & \text{if } P_{2B_2} \in \left[ 0, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_{STBM}^{(n)} \right] \\
(P_{STBM}^{(n)}, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_{STBM}^{(n)}) & \text{otherwise}.
\end{cases} \]

(24)

\[ (P_{1B_1}^*, P_{2B_1}^*) = \begin{cases} 
(P_{1B_1}, P_{2B_1}) & \text{if } P_{1B_1} \in \left[ 0, P_{STBM}^{(n)} \right] \\
(P_{STBM}^{(n)}, \min\{ P_{STBM}, \frac{X_{STB}}{(1-\beta)T} \} - P_{STBM}^{(n)}) & \text{otherwise}.
\end{cases} \]

(25)

\[ (\tilde{V}_n(S, X_{STA}) = \max_{P_{STA} \in \left[ 0, \min\{ P_{STAM}, \frac{X_{STA}}{(1-\beta)T} \} \right]} \{ T_{WSum,A}(S, P_{STA}^{(n)}) + E \left[ \tilde{V}_{n+1}(S', X_{STA} - \beta TP_{STA}^{(n)} | S) \right] \}], \]

(31)

\[ \tilde{V}_{n-1}(S, X_{STA}) = \max_{P_{STA}^{(n-1)} \in \left[ 0, \min\{ P_{STAM}, \frac{X_{STA}}{(1-\beta)T} \} \right]} \{ T_{WSum,A}(S, P_{STA}^{(n-1)}) \}. \]

(32)

\[ \tilde{P}_1 = \begin{cases} 
P_1^* & \text{if } P_1^* \in \left[ 0, \min\{ P_{STAM}, \frac{X_{STA}}{(1-\beta)T} \} \right] \\
\min\{ P_{STAM}, \frac{X_{STA}}{(1-\beta)T} \} & \text{otherwise},
\end{cases} \]

(29)

where \( P_* \in \arg \max_{P_1 \in [0, P_{STAM}]} \tilde{Q}_n(S, X_{STA}, P_1) \), and \( \tilde{Q}_n(\cdot) \) is given by
\[
\tilde{Q}_n(S, X_{STA}, P_1) = T_{WSum,A}(S, P_1) + E \left[ \tilde{V}_{n+1}(S', X_{STA} - \beta TP_1 | S) \right],
\]

(30)

where \( \tilde{Q}_{n-1}(S, X_{STA}, P_1) = T_{WSum,A}(S, P_1) \). For \( n \in [0, N-2] \), we have (31). For \( n = N-1 \), we have (32).

C. Superframe-level Resource Allocation Policy

Unlike the resource allocation policies for SUs in the frame level, the resource allocation for the PU cannot be analyzed by using the above analytical method, i.e., it is impossible to get a close-form solution. Here, we propose to use the heuristic online methods in [24] to get the superframe-level resource allocation policy for the PU. The parallel rollout scheme which is based on the decision rule/policy improvement principle in the policy iteration algorithm in [28] can be adopted here. Since the value of any decision rule pair is a lower bound to the optimal value, the parallel rollout iteration is a lower bound scheme. The upper bound based iteration can also be used for the PU, for example, hindsight approach in [29]. The number
\begin{algorithm}[H]
\caption{Backward Iteration Algorithm of ST\textsubscript{A}}
\textbf{Input:} \(X_{STAM}, X_{STA}, P_{STAM}, \{b_{AA}\}\), transition probability matrix of channel state, steady-state probability of each state, \(N_0,\beta, \gamma, \eta, T, N\).
\textbf{Output:} Optimal \(P^{(q)}_{STA}\).
\begin{algorithmic}[1]
\State 1: If \(n = N\) then
\State 2: If \(X_{N}^N \geq P_{STAM}^\beta T\) then
\State 3: \(P_{STA}^N = P_{STAM}^\beta T\);
\State 4: else\;
\State 5: \(P_{STA}^N = \max \left(0, \frac{X_{N}^N}{\beta T} \right)\);
\State 6: end if\;
\State 7: end if\;
\State 8: for \(n = N - 1\) to 1 do
\State 9: for Channel State Transition \(h_{AA}^{(n)}(i) \rightarrow h_{AA}^{(n+1)}(j)\) do
\State 10: Calculate optimal \(P_{STA}^{(q)}\) according to (31)\;
\State 11: end for\;
\State 12: end for\;
\State 13: return \;
\end{algorithmic}
\end{algorithm}

Algorithm 2 Backward Iteration Algorithm of ST\textsubscript{B}
\textbf{Input:} \(X_{STBM}, X_{STB}, P_{STBM}, \{b_{BB}\}, \{b_{PA}\}, \{b_{PB}\}, \{b_{PS}\}\), transition probability matrix, steady-state probability matrix, \(N_0,\beta, \gamma, \eta, T, N\).
\textbf{Output:} Optimal \(P^{(q)}_{STBM}, P^{(q)}_{STBS}\).
\begin{algorithmic}[1]
\State 1: If \(n = N\) then
\State 2: If \(X_{N}^{(n)} \geq P_{STBM}^\beta T\) then
\State 3: \(P_{STBP} = \min \left(P_{STBM}, \eta \cdot P_{STBPM}^{\gamma} \right)\) and \(P^{(q)}_{STBS} = P_{STBM} - P_{STBP}^\eta\);
\State 4: else\;
\State 5: \(P_{STBP}^n = 0\) and \(P^{(q)}_{STBS} = P_{STBM}\);
\State 6: end if\;
\State 7: end if\;
\State 8: for \(n = N - 1\) to 1 do
\State 9: for Channel State Transition \(h_{BB}^{(n)}(i) \rightarrow h_{BB}^{(n+1)}(j)\) do
\State 10: Calculate optimal \(P_{STB}^{(q)}\) and \(P^{(q)}_{STBS}\) using (16)\;
\State 11: end for\;
\State 12: end for\;
\State 13: return \;
\end{algorithmic}

of iterations can be determined by many factors such as the time required to obtain the optimal solution. If the parallel rollout is used for the PU, the lower bound of the iteration approach is given by \(|Q^*(S, P_{P}) - Q(S, P_{P})| \leq \frac{2\epsilon}{\pi} P_{P}^\gamma\), where \(Q^*(S, P_{P}) = \max_{P_{P}} \{T_{WSum}\}\) is the analytically optimized value of the PU, \(Q(S, P_{P})\) is the iteration solution, \(\epsilon\) is defined as \(\sup_{P_{P}} |Q^*(S, P_{P}) - U(S, P_{P})| \leq \epsilon\), and \(U(S, P_{P})\) is a bounded and measurable function which is determined by iteration times. Due to space limitations, the detailed iteration algorithms are omitted.

V. APPROXIMATION ALGORITHM AND NUMERICAL RESULTS

To simplify the calculation and interpretation of the proposed policies, two modified backward iteration based approximation algorithms for ST\textsubscript{A} and ST\textsubscript{B} policies are given in Algorithm 1 and Algorithm 2, respectively. For simplicity of optimization, the calculation of Shannon capacity in the iteration process is approximated by linear fitting. The power policy obtained by the approximation backward iteration is used to calculate the real reward functions without approximation. Note that step 10 in Algorithm 1 and step 20 in Algorithm 2 involve two convex optimization problems which can be readily solved by sophisticated algorithms [27], according to the stochastic inventory theory [30], [31].

The simulations are done using Matlab for ST\textsubscript{A} and ST\textsubscript{B} separately. Two Markov channel models are adopted in our simulations. Specifically, in the first scenario, each user is uniformly distributed in the area of 1000 \(\times\) 1000 \(m^2\) with low speed, e.g., \(v = \{1, 2\}\) \(m/s\). To investigate the performance in the high speed environment, we simply model the channel state as a two state ON-OFF channel in the second scenario, as discussed in [32] where the velocity is up to 30 \(m/s\). In the first scenario, each channel has nine states in terms of SNR listed as \(\text{SNR} = \{1, 6.02, 7.78, 9.03, 10.79, 17.04, 18.80, 24.05, 24.56\}\) in \(dB\), and the path loss exponent is 3. In the second scenario, the off state is with a SNR of 0 \(dB\) and the on state is with a SNR of 14.77 \(dB\), respectively. In our setting, \(T = 100\ m/s\), \(\zeta = 0.4\) and \(\eta = 1\). The bandwidth in our system is 1 \(MHz\). The associated channel state transition probability matrix and the steady-state probability are obtained by using these parameters described in [32], [33]. The reference path loss and the reference distance are 5.105 and 100 \(m\).

In the simulation of ST\textsubscript{A} power policy, we consider both on-off Markov channel and 9-state Markov channel. The maximum transmit power is \(P_{STAM} = 5 \times 10^{-3}\ W\), and the initial energy of ST\textsubscript{A} is \(X_{STA} = 1.2 \times 10^{-3}\ W \cdot s\). We compare the result obtained using Algorithm 1 with the greedy and random power policies. In the greedy one, ST\textsubscript{A} always transmits at its maximum power. In the random one, ST\textsubscript{A} randomly chooses a power not greater than \(P_{STAM}\). The throughput performance of each power policy under different conditions is shown in Fig. 4(a) and Fig. 4(b).

Specifically, \(T_{WSum}\) of different frame numbers in one superframe under the on-off Markov channel is shown in Fig. 4(a). Since \(X_{STA}\) is only enough for less than \(N_{STAM}\) 5 times transmission at the power level of \(P_{STAM}\), it can be seen that MDP based policy has the same throughput with that of greedy policy when \(N = 4\). This shows MDP policy can achieve the same performance when the energy is enough for one superframe transmission. We can find that the throughput of MDP based power policy increases with the increasing of \(N\). The rationale behind this result is that with more number of frames, the probability that ST\textsubscript{A} transmits at the good channel condition increases. The MDP policy in this two state on-off model can be interpreted as ST\textsubscript{A} aims to transmit as much as
possible when the channel is good, while ST_A tries to save energy for future good channels when the current channel is bad. In this regard, the optimal performance for MDP happens when ST_A transmits using $P_{STAM}$ only when the channel is good, while keeping idle when the channel is bad. This optimal performance requires a sufficiently large $N$ to enable ST_A to consume its energy only in the on state. It is evident that the throughput of greedy based power policy is a constant after $N = 5$. The throughput using random power policy increases with increasing $N$, because the probability that ST_A can transmit in the good channel condition increases when $N$ is large. It is noted that the MDP and greedy policies achieve the same throughput when $N$ is 5 and 6 as shown in Fig. 4(a). The reason for this result comes from that channel conditions for transmissions using MDP are most likely the same as that of using the greedy policy, as the maximum number of frames in one superframe is near $N_{STAM}$.

Performance comparison of ST_A using MDP, greedy and random power policies in the 9-state Markov channel model with different velocities and initial channel states is shown in Fig. 4(b). Given a velocity, it is straightforward that all these three policies achieve better performance when the initial channel state is under a better SNR condition. When the initial channel condition is not sufficiently good, for a given initial channel state, it can be seen from the figure that the throughput increases with the increasing of velocity. Since the channel state in current frame will more likely to transit to the next channel state. In this context, ST_A can transit from a low SNR to a high one with a higher probability at a higher velocity.

The maximum power is $P_{STBM} = 8 \times 10^{-3}$ W, and the initial energy of ST_B is $X_{STB} = 1.2 \times 10^{-3}$ W · s. PU’s maximum power is $P_P = 2 \times 10^{-3}$ W. The optimal MDP policy using Algorithm 2 for ST_B is that ST_B aims to transmit on the channel with higher throughput at each frame, e.g., when $T_P^{(n)} \geq T_{SB}^{(n)}$ in (18), ST_B will forward the PU’s data using $P_{STBPM}$ in the $n$th frame, otherwise ST_B transmits its own data with $P_{STBM}$. In this simulation, all of $\{h_{BB}\}, \{h_{PA}\}, \{h_{PB}\}, \{h_{BP}\}$ are modeled based on the on-off Markov channel models for computational simplicity. It can be seen from Fig. 4(c) that MDP policy can get higher throughput with an increasing $N$. Since $X_{STB}$ can only guarantee 3 transmissions when greedy policy is used,
the throughput of greedy policy is a constant for $N \geq 3$.

To investigate the superframe level policy, numerical result of optimal $P_{P}$ is given in Fig. 4(d). XPD of each superframe is assumed to be independent and identically distributed. The PU transmits using a constant power for all superframes. In this context, the optimization problem in the superframe level is to find the optimal $P_{P}$ to reach the maximum $\bar{T}_{WSum}$. The initial energy level of the PU is $X_P = 1 \times 10^{-3}$ W·s. The channel model of spatial fading in each superframe for $ST_{A}$ is according to the on-off model, and that of $ST_{B}$ is the 9-state SNR model with $v = 1$ ms, respectively. In each superframe, there are $N = 7$ frames, and $M = 4$ superframes for the PU are considered. It can be seen that optimal $P_{P}$ that can reach the maximum $\bar{T}_{WSum}$ is achieved at $10^{-4}$ W. The rationale behind this result is that, with a low $P_{P}$, although the PU still has energy after four superframes cooperation, the PU contributes little to the weighted sum throughput of each superframe. With the increasing of $P_{P}$, although the weighted sum throughput of each superframe increases, the PU depletes its energy which terminate the cooperative communications between the PU and SUs in later superframes. Therefore, we can conclude that the optimal value of $P_{P}$ is $6 \times 10^{-4}$. To further reduce the complexity of computing the optimal $P_{P}$, the heuristic online method in Section IV-C can be used, and the performance bound is given mathematically.

VI. CONCLUSIONS

In this work, we have proposed a novel polarization enabled two-phase cooperation framework for CCRN. By utilizing ODPA at SU transceivers, SUs can simultaneously relay the PU’s data and transmit their own information without mutual interference. We have modeled this system as a two time-scales cooperation framework by taking both the spatial and polarization domains into consideration. To evaluate the effectiveness of the proposed framework, a sum throughput maximization problem with throughput and power constraints has been formulated. The optimization problem has been analyzed and solved by using MDP. In practice, CSI is imperfect; our future work will address the problem of imperfect CSI. In addition, we will fully and deeply investigate the properties of depolarization in wireless channels to model the CCRN cooperation framework.

REFERENCES


**Bin Cao** received his M.A.Sc. degree in information and communication engineering from the Harbin Institute of Technology Shenzhen Graduate School, China. He is now working on his Ph.D. degree at Harbin Institute of Technology Shenzhen Graduate School. Financially supported by the Chinese Scholarship Council, he was a visiting student at the BBCR group from 2010 to 2012 at the University of Waterloo, Waterloo, Ontario, Canada. His research interests include signal processing for wireless communications, cognitive radio networking, and cross-layer optimization.

**Hao Liang** received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Canada, in January 2013. Since February 2013, he has been a postdoctoral research fellow at the Broadband Communications Research (BBCR) Lab and Power and Energy Systems Group, University of Waterloo. His research interests are in the areas of smart grid, wireless communications, and wireless networking. He is a recipient of the Best Student Paper Award from IEEE VTC Fall-2010, Ottawa, ON, Canada. He has been serving as the System Administrator of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY since 2009.

**Qinyu Zhang** (M’08–SM’12) received his Ph.D. degree from the University of Tokushima, Japan. From 1999 to 2003, he was an assistant professor at the University of Tokushima. He is a full professor of the Harbin Institute of Technology Shenzhen Graduate School. His research interests include wireless communications, cognitive radio, signal processing, and biomedical engineering. He has published more than 200 technical papers and 43 issued or pending patents, and was awarded as one of the Chinese New Century Excellent Talents in University. He is the founding chair of the IEEE Communications Society Shenzhen Chapter.

**Jon W. Mark** (M’62–SM’80–F’88–LF’03) received his Ph.D. degree in electrical engineering from McMaster University. In September 1970, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, where he is currently a Distinguished Professor Emeritus. His current research interests are in broadband wireless communications, resource and mobility management, and cross domain interworking. He is a co-author of the text *Wireless Communications and Networking*, Prentice-Hall, 2003. He is a Life Fellow of the IEEE and a Fellow of the Canadian Academy of Engineering.