Dynamic Bargaining for Relay-Based Cooperative Spectrum Sharing

Yang Yan, Jianwei Huang, *Senior Member, IEEE,* and Jing Wang, *Member, IEEE*

***Abstract*—Cooperative spectrum sharing can effectively improve the spectrum usage by allowing secondary users (SUs) to dynamically and opportunistically share the licensed bands with primary users (PUs). In return, an SU will relay a PU's traffic to improve the PU's effective data rate. In this paper, we study how one PU and one SU achieve an efficient spectrum sharing through dynamic noncooperative bargaining. The key challenge is that the PU does not have complete information of the SU's energy cost. We model the dynamic bargaining with incomplete information as a dynamic Bayesian game, and investigate the equilibria under both single-slot and multi-slot bargaining models. Theoretical analysis and numerical results indicate that our proposed scheme can lead to a win-win situation, where both the PU and the SU obtain data rate improvements via the bargaining-based CSS mechanism. Furthermore, the SU can take advantage of the incomplete information to improve its bargaining power (also called reputation effect), and thus gain a higher data rate.**

***Index Terms*—Cooperative spectrum sharing, bargaining, dynamic Bayesian game, incomplete information, sequential equilibrium, reputation effect.**

I. Introduction

Cooperative spectrum sharing (CSS) can effectively improve spectrum efficiency, and thus can alleviate the network pressure caused by the rapid increase of wireless data communication. One mechanism to realize CSS is *cooperative communications* [4] between the primary licensed users (PUs) and secondary unlicensed users (SUs): PUs with poor direct channel conditions can achieve higher data rates by using SUs as cooperative relays. Cooperative communications can achieve the benefit of multiple-input-multiple-output (MIMO) communications by using relay nodes as virtual antennas, and has also been extensively studied in the recent literatures (*e.g.,* [5]–[10]) and has become part of the next generation communication standards (*e.g.,* IEEE 802.16J standard [11]–[13]).

Different from the traditional cooperative communications where only relays help, CSS requires PUs to *compensate* SUs by allocating network resources for the SUs' own communications. Because of this, CSS leads to a *win-win* situation for PUs and SUs. The key question for designing a relay-based CSS mechanism is how the PUs and SUs agree on the cooperative transmissions and resource splitting. In this paper, we will focus on a noncooperative bargaining-based CSS mechanism under incomplete network information.

The concept of relay-based CSS has only been proposed in recent literatures (*e.g.,* see [14]–[20] and the references therein). Han *et al.* in [17] analyzed the achievable rates of CSS with one PU pair and one SU pair. Simeone *et al.* in [18] considered the case where PU optimizes the resource allocation based on the known channel state information and transmission power, in a network with one PU and many SUs. Zhang *et al.* in [19] considered a similar CSS between multiple SUs and one PU. Both primary and secondary users target at maximizing their utility functions. Wang *et al.* in [20] studied a CSS between one PU and multiple SUs, where SUs decide their power levels for relaying PU's traffic to achieve proportional access time to the channel. All results in [14]–[20] considered complete network information, where the PU and SUs know all the channels and important network parameters such as SUs' energy cost. The only recent work focusing on incomplete network information environment are [21] and [22], where the authors proposed contract-based CSS mechanisms in networks with multiple SUs (with one or multiple PUs). However, [21] and [22] only considered the spectrum sharing in a static setting (*i.e.,* one time slot).

In this paper, we study a dynamic bargaining-based CSS between one PU and one SU[1] with incomplete information[2] under several different interaction scenarios. In particular, we assume that the SU's energy cost for transmission is unknown to the PU. In a dynamic network environment, the PU and the SU must timely adjust their own strategies in order to achieve a satisfactory resource splitting. The dynamic and incomplete information model better captures the reality of wireless communications, but so far has received little attention in the literature due to the difficulty of analysis.

The main contributions of this paper are summarized as follows:

Manuscript received August 25, 2011; revised March 17, 2012.

Y. Yan and J. Wang are with the State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: yanyang07@mails.tsinghua.edu.cn, wangj@tsinghua.edu.cn).

J. Huang is with the Network Communication and Economics Lab, Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: jwhuang@ie.cuhk.edu.hk).

This work was supported by the General Research Funds (Project No. 412710 and 412511) established under the University Grant Committee of the Hong Kong Special Administrative Region, the CSC Postgraduate Scholarship, the National Basic Research Program of China (2012CB316000), the National S&T Major Project (2010ZX03003-002), the National NSF of China (60832008), China's 863 Project (2009AA011501), the Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation, Tsinghua Research Funding (2010THZ02-3), and the International Science Technology Cooperation Program (2010DFB10410).

Part of the results have appeared in *GameNets 2011* [1], *WiCON 2011* [2], and *GLOBECOM 2011* [3].

Digital Object Identifier 10.1109/JSAC.2013.1308.11

[1]In Section II-C, we will discuss further the motivation behind the study of the one-to-one bargaining and its connection to a more general many-to-many bargaining.

[2]Incomplete information means that not all players know the utility function and strategy space of each player in the game.

0733-8716/13/$31.00 © 2013 IEEE

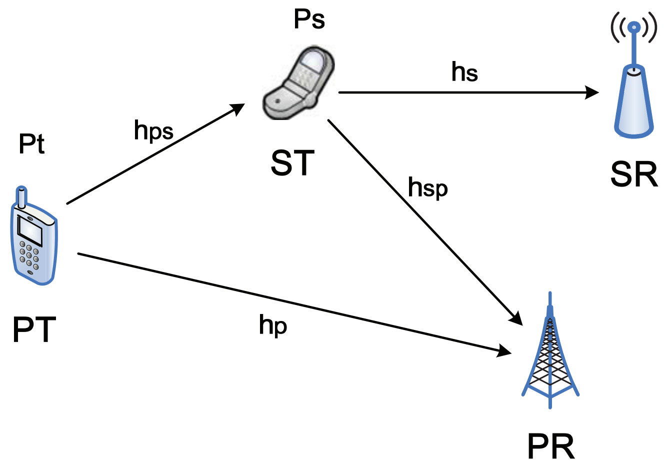


Fig. 1. The cooperation between one PU (PT-PR) and one SU (ST-SR).

- *New network model:* This is the first paper that studies cooperative spectrum sharing under a *dynamic* and *incomplete* information network environment based on *noncooperative bargaining theory*.
- *Dynamic bargaining with incomplete information:* During the dynamic bargaining scenario, the PU and SU must adjust their beliefs about incomplete information and strategies accordingly, so as to maximize their own payoffs. We model such interaction as a dynamic Bayesian game, which is challenging to analyze.
- *Multiple system scenarios and sequential equilibrium:* We investigate the bargaining interactions under two different system scenarios: single-slot and multi-slot. In both scenarios, we derive the sequential equilibrium (SE) for both scenarios, and evaluate their properties through numerical examples.
- *The analysis of SEs:* Theoretical analysis and numerical results indicate that the proposed mechanism leads to a win-win situation. Moreover, the SU can gain a higher data rate by taking advantage of the incomplete information. In particular, the SU can benefit from the multi-stage bargaining and the reputation effect in the multi-slot bargaining.

The rest of the paper is organized as follows. We introduce the system model and methodology in Section II. In Section III, we analyze the equilibria of the single-slot bargaining game. In Section IV, we extend the analysis to the multi-slot bargaining game. We finally conclude in Section V.

II. System Model and Methodology

A. Cooperative Spectrum Sharing Model

We consider a time-slotted system with the network model as in Fig. 1. Here, PT and PR represent PU's transmitter and receiver, and ST and SR represent SU's transmitter and receiver. Let h_p, h_s, h_{ps}, and h_{sp} denote the channel gains of the links PT-PR, ST-SR, PT-ST, and ST-PR, respectively. For simplicity, we assume that the channel gains

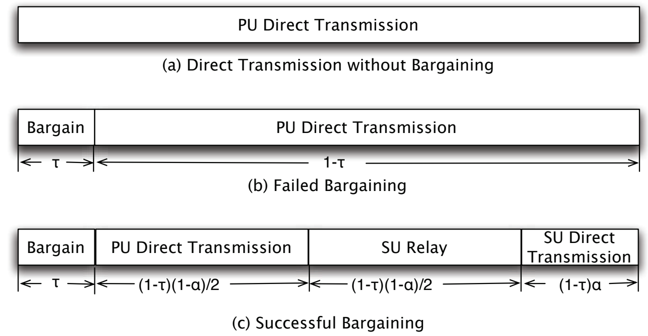


Fig. 2. Three possibilities of the bargaining process in a single time slot.

remain fixed across time slots.[3] We further assume that both the PU and the SU know the channel gains of all links through a proper feedback mechanism. The PU and SU transmit with fixed power levels of P_t and P_s, respectively.

In this paper, we assume that the SU is an *energy-constrained* device (*e.g.*, wireless sensor or mobile device). We let C denote the SU's energy cost, which is related to its current battery status. The actual value of C is known only by the SU but not by the PU. However, PU knows the distribution of C, and hence the information is asymmetric and incomplete.[4]

Spectrum bargaining process can happen within a single time slot or over multiple time slots. Fig. 2 illustrates three possibilities of the bargaining result within a single time slot. Without loss of generality, we normalize the time slot length T to 1, and denote $\tau \ll 1$ as the overhead due to bargaining.

- Fig. 2(a): If PU's direct channel condition h_p is good enough, it will choose direct transmission during the whole time slot and achieve a data rate (measured in bits/s/Hz) $R_{dir} = \log_2(1 + P_t h_p)$. In this case, the SU cannot transmit and thus achieves a zero utility.
- Fig. 2(b): If PU believes that cooperation may be beneficial, it can offer α fraction of the time slot for SU's own transmission. If SU rejects the offer, PU proceeds with direct transmission in the remaining time without the cooperation of the SU.
- Fig. 2(c): If SU accepts PU's offer α during the bargaining phase, then PU and SU transmit in the amplified and forward (AF) relay mode. The PU achieves a data rate (measured in bits/s/Hz) [4]
$$R_r = \frac{1}{2}\log_2\left(1 + P_t h_p + \frac{P_t P_s h_{ps} h_{sp}}{P_t h_{ps} + P_s h_{sp} + 1}\right), \quad (1)$$
and the normalized data rate (per unit time) is $(1 - \tau)(1 - \alpha)R_r$. The SU achieves a data rate (measured in bits/s/Hz)
$$R_s = \log_2(1 + P_s h_s), \quad (2)$$

[3]This assumption is only relevant when we discuss multi-slot bargaining in Section IV. This applies to the case where nodes in the network remain fixed for a relatively long time, and each time slot is much longer than the coherence time of the fast fading process (so that we only need to consider the average channel condition).

[4]For simplicity, we assume that an SU's energy cost C is fixed (but unknown to the PU). This is reasonable when the SU's energy does not change significantly during the time of interests.

TABLE I
A SUMMARY OF BARGAINING MODELS AND KEY RESULTS

Bargaining Models		Key Results	Sections in This Paper
One-Slot	**One-Stage**	A unique sequential equilibrium	III-A
	Multi-Stage	Two (potential) sequential equilibria	III-B
Multi-Slot One-Stage		A unique sequential equilibrium with reputation effect	IV

and the normalized data rate (per unit time) is $(1-\tau)\alpha R_s$.

Apparently, a larger α means a higher data rate (thus a larger benefit) for SU but a lower data rate for PU (due to reduction of transmission time). PU and SU will therefore bargain with each other to determine the value of α that is acceptable by both sides, by considering the network environment, opponent's history actions, and potential strategies.

In this paper, we first investigate how the PU and SU bargain with each other within a *single* time slot (*i.e.*, one-slot bargaining game). Note that the bargaining period with length τ in Fig. 2 (b) and (c) can include multiple stages, *i.e.*, PU and SU can bargain multiple times within the same time slot. Thus we will study both the *one-slot one-stage* and *one-slot multi-stage* bargaining game. Then we analyze how the PU and SU bargain over a *finite number* of time slots (*i.e.*, multi-slot bargaining game). We summarize various models and the key results in Table I.

B. Sequential Equilibrium

The bargaining process described above is a dynamic Bayesian game, which involves the PU's and SU's dynamic decision-making and belief updates. The *sequential equilibrium* (SE) is a commonly used equilibrium concept for the dynamic Bayesian game, and is defined as a *strategy profile* and *belief system* which satisfy the following three basic requirements [23]:

Requirement 1: The player taking the action must have a belief (probability distribution) about the incomplete information, reflecting what that player believes about everything has happened so far.

Requirement 2: The action taken by a player must be optimal given the player's belief and the other players' subsequent strategies.

Requirement 3: A player's belief is determined by the Bayes' rule whenever it applies and the players' hypothesized equilibrium strategies.

Requirement 2 shows that a player's strategy at a given information set is dependent on its belief. However, a player's belief is derived from the players' strategies higher up in the game tree according to Requirement 3. Requirements 2 and 3 together show that players' beliefs at different stages of the game are related to each other, and a single *backward induction* through the game tree typically will not be enough to compute an SE [24]. Requirement 3 requires the *consistency* between one player's belief and players' equilibrium strategies. We will refer back to Requirements 1-3 when deriving the SEs in later sections.

C. Connection with Many-to-Many Bargaining

In this paper, we analyze the bargaining-based CSS between one PU and one SU. However, enlightened by Antoni's work [25] [26], the general CSS with multiple PUs and multiple SUs could be decomposed into multiple pairs of bilateral bargaining game between one PU and one SU. In Fig. 3, we use a black solid square pair to represent one PU and a red solid dot pair to represent one SU. An SU will only bargain with PUs nearby, as it can provide good relay services to these PUs. For example, SU_1 in Fig. 3(a) is likely to bargain with PU_4 and PU_1 as they are close by, and is less likely to bargain with PU_2 and PU_3 as they are far away. A PU will choose an SU to bargain according to a similar principle. This means that we can remodel the network as an undirected graph (see Fig. 3(b)): square nodes represent PUs, and circle nodes represent SUs, and a link between nodes represents that the probability of bargaining is above certain thresholds.[5]

Let us consider a concrete example in Fig. 3(a). Assume that SU_1 bargains with PU_1, indicated by *ellipse a*. If the bargaining fails, SU_1 could *randomly* select another PU (*e.g.*, PU_4) for bargaining (*ellipse c*). Meanwhile, PU_1 may decide to bargain with SU_3 (*ellipse b*). We can also consider a variation of this random selection model, where each node (PU or SU) optimizes its choice of bargaining partner based on the network information such as nodes' locations, channel conditions, and the potential bargaining opponents' information [26]. In either case, it is clear that the study of the bargaining-based CSS between one PU and one SU serves as a basis for the general case involving multiple network entities, and therefore is of great interest to us.

III. ONE-SLOT BARGAINING GAME

In this section, we consider the single time slot bargaining game. As a preliminary result, we first study the one-slot one-stage bargaining game, which will help us to study the more general model of one-slot multi-stage bargaining.

A. One-Slot One-Stage Bargaining Game

In this case, there is *at most* one stage of bargaining in a time slot. The proportion of the slot after bargaining is $\delta = (1-\tau) < 1$, where τ is the duration of one-stage bargaining. PU needs to decide: (i) whether to bargain, and (ii) the optimal offer α if it decides to bargain. The SU should decide whether to accept α (if the PU offers one).

The SU's utility function U_s of cooperation is the difference between the achievable data rate and the energy cost

$$U_s(\alpha) = \delta\left(\alpha R_s - \frac{1+\alpha}{2} P_s C\right), \tag{3}$$

where R_s is given in (2). $\frac{1+\alpha}{2} P_s C$ denotes all SU's energy cost, including the costs in relaying PU's data and transmitting

[5]We may further consider a weighted graph, where the weights represent the probabilities of bargaining.

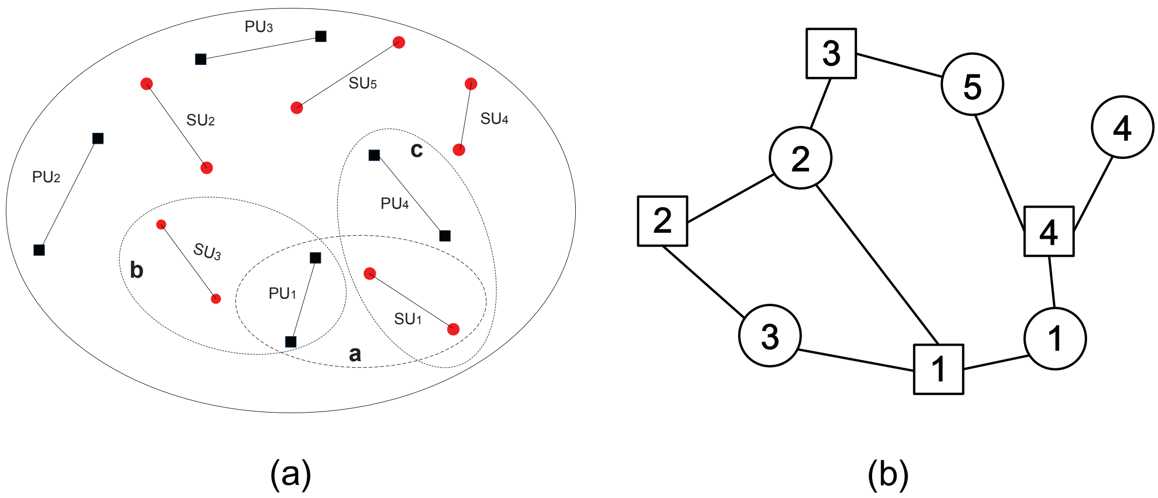


Fig. 3. (a) A network with multiple PUs and multiple SUs; (b) The corresponding undirected graph.

SU's own data in this mechanism. We can think C as data rate per *watt* the SU can get if it does not relay for PU. Therefore, $U_s(\alpha)$ is SU's *data rate increase* by accepting α. Note that the SU can always achieve a zero utility by not participating in the cooperative communication. Given PU's offer α, the optimal decision for SU is obvious: accept α if and only if $U_s(\alpha) > 0$.

Without bargaining, the PU can always achieve R_{dir} through the direct transmission as in Fig. 2(a). In that case, its *data rate increase* is zero. Without any prior knowledge, PU assumes that C follows a uniform distribution in $[K_1, K_2]$, where $0 \leq K_1 < K_2$.[6] If the SU rejects the offer as in Fig. 2(b), PU can only directly transmit in the remaining δ time and achieve a *negative* data rate improvement $(\delta R_{dir} - R_{dir}) < 0$. If SU accepts the offer as in Fig. 2(c), PU receives a data rate increase of $\delta(1-\alpha)R_r - R_{dir}$, which can be either positive or negative. Here, R_r is given in (1). Therefore, if the PU decides to bargain with the SU, it will choose α to maximize the ***PU's utility (expected data rate increase)***

$$U_p(\alpha) = (\delta R_{dir} - R_{dir})\,\text{Prob}\,(U_s(\alpha) \leq 0) + (\delta(1-\alpha)R_r - R_{dir})\,\text{Prob}\,(U_s(\alpha) > 0), \quad (4)$$

where $\text{Prob}\,(U_s(\alpha) \leq 0) = \text{Prob}\,(\text{SU rejects offer})$ and $\text{Prob}\,(U_s(\alpha) > 0) = \text{Prob}\,(\text{SU accepts offer})$. Let us denote the optimal value of (4) as α^*. The optimal choice α^* that maximizes (4) is given in the following theorem:

Theorem 1: When $K_1 > R_s/P_s$, $U_p(\alpha) < 0$ for any given $\alpha \in [0, 1]$. When $K_1 \leq R_s/P_s$, the optimal α^* is shown in (5) at the top of the next page, where $\bar{\alpha}_p = 2\sqrt{\frac{(R_s/P_s)(2R_r - R_{dir})}{2R_r(2R_s/P_s - K_1)}} - \frac{1}{2}$.

Proof Sketch (Details in [27]): We divide the discussion based on different relationships between K_1, K_2, $\frac{R_s}{P_s}$, and $\frac{2R_s}{P_s}$. In each case, $U_p(\alpha)$ is a piece-wise function and the optimal α^* can be found by considering the boundary values. ■

[6]To simplify the analysis, we assume that the SU's energy cost C follows a uniform distribution. The analysis for other distributions will be technically more involved but offers essentially the same engineering insights.

When K_1 is larger than R_s/P_s, the SU will not accept any offer α from the PU. In this case, the PU knows that the bargaining will fail and thus has to choose direct transmission. When $K_1 \leq R_s/P_s$, PU will choose α^* in (5) to achieve the best tradeoff of data rate increase and performance loss. Note that PU will compare $U_p(\alpha^*)$ with zero and decides whether it is worth bargaining or not. If $U_p(\alpha^*) < 0$, it will simply choose direct transmission. The one-stage bargaining game is a subgame for the multi-stage bargaining game, and we will use Theorem 1 in later analysis.

B. One-Slot Multi-Stage Bargaining Game

Now we consider the case where the bargaining within a time slot can happen over *more* than one stage. For the ease of illustration, we will focus on the two-stage bargaining case. The more general multi-stage model can be similarly analyzed.

1) Utility Functions and Game Tree: Similar to the one-stage game, the utility functions are PU's *expected data rate increase* and SU's *data rate increase*. We assume that PU's belief about C *at the beginning of first-stage of bargaining* is a uniform distribution over $[0, K]$. We denote δ_1 and δ_2 as the proportions of the time slot after bargaining in the first and second stage. By setting different values of δ_1 and δ_2, we can model different bargaining overheads. Compared with the one-stage bargaining, the key characteristic here is that PU's belief about C at the beginning of *second-stage* may no longer be uniform over $[0, K]$.

Fig. 4 illustrates the sequential decisions and possible scenarios of this one-slot two-stage bargaining game. PU and SU make decisions alternatively at the non-leaf nodes. PU first makes the decision on whether to bargain. If it selects direct transmission (**D**), the game ends. Otherwise, PU offers α_1 to SU. If SU accepts this offer, then the game ends. If SU rejects the offer, then PU makes a second offer α_2 to SU. Finally, SU either accepts or rejects α_2. The game ends in both cases. Every possible ending of the game is denoted by a black solid square together with the corresponding utilities of PU (upper value) and SU (lower value).

The two-stage bargaining game is more complex than the one-stage bargaining model for two reasons: (i) The SU may

$$\alpha^* = \min\left(\max\left(\bar{\alpha}_p, \frac{K_1}{2R_s/P_s - K_1}\right), \min\left(\left|\frac{K_2}{2R_s/P_s - K_2}\right|, 1\right)\right) \tag{5}$$

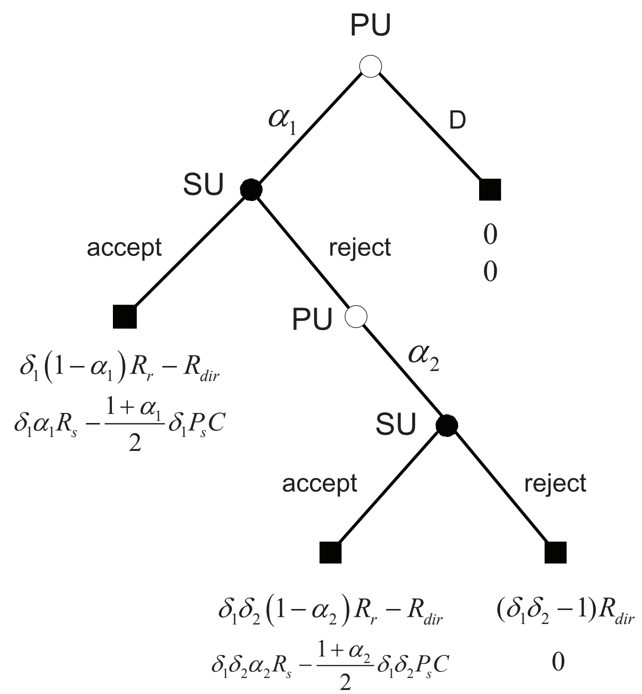


Fig. 4. Game tree of one-slot two-stage bargaining.

reject α_1 even though its utility is positive, if it believes that α_2 is much better; (ii) The PU needs to *update* its belief on C at the end of the first stage (*i.e.,* no longer uniform distribution over $[0, K]$) by taking the SU's strategic decision into consideration.

Based on Requirements 1-3, we further introduce the following notations:

- PU's strategy: whether to bargain, the first-stage offer α_1 if decides to bargain, and the second-stage offer $\alpha_2(\alpha_1)$ (*i.e.,* as a function of α_1) if SU rejects α_1.
- PU's belief: $\mu_1(C)$ (*i.e.,* uniform distribution in $[0, K]$) denotes PU's belief on SU's energy cost C at the beginning of the first stage, and $\mu_2(C|\alpha_1)$ denotes PU's *updated* belief about C at the beginning of the second stage if SU rejects α_1.
- SU's strategy: $[\mathcal{A}_1(\alpha_1|C), \mathcal{A}_2(\alpha_2|C, \alpha_1)]$. $\mathcal{A}_1(\alpha_1|C) = 1$ if SU accepts α_1 when its energy cost is C, and $\mathcal{A}_1(\alpha_1|C) = 0$ otherwise. Likewise, $\mathcal{A}_2(\alpha_2|C, \alpha_1) = 1$ if SU accepts α_2 (after rejecting α_1) when its energy cost is C, and $\mathcal{A}_2(\alpha_2|C, \alpha_1) = 0$ otherwise.
- SU's belief: since SU knows its own energy cost C, its belief is a singleton set (*i.e.,* no uncertainty).

2) Analysis of the Second Stage: We will start our analysis from the second stage. Since this is the last stage of the game, the analysis is similar to the one-stage game case. The PU needs to optimize the choice of α_2. More specifically, we can apply Requirement 2 to solve SU's strategy $\mathcal{A}_2(\alpha_2|C, \alpha_1)$. Since this is the *last* move of the game, the optimal strategy for SU in the second stage is to accept α_2 if and only if the SU's utility $\delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C > 0$. Such decision is independent of α_1.

Given SU's optimal strategy in the second stage $\mathcal{A}_2(\alpha_2|C, \alpha_1)$, we can apply Requirement 2 to compute the PU's optimal strategy in the second stage. The PU will calculate the optimal α_2 that maximizes PU's expected utility function U_p, given PU's *updated* belief $\mu_2(C|\alpha_1)$ and SU's *subsequent* strategy $\mathcal{A}_2(\alpha_2|C, \alpha_1)$. *The tricky part is how to compute the belief $\mu_2(C|\alpha_1)$, which depends on the interaction in the first stage.* In particular, we need to understand the SU's equilibrium strategy in the first stage in order to update PU's belief in the second stage.

3) Analysis of the First Stage: We start our analysis given *arbitrary* first and second stage offers α_1 and α_2. We further assume that K (upper bound of C) is reasonably large (*i.e.,*$K > \frac{R_s}{P_s}$).[7] Define

$$C^*(\alpha_1, \alpha_2) = \frac{2R_s(\alpha_1 - \delta_2\alpha_2)}{P_s((1+\alpha_1) - \delta_2(1+\alpha_2))}. \tag{6}$$

The following lemma provides SU's equilibrium strategy in the first stage for *given* α_1 and α_2.

Lemma 1: SU *rejects* α_1 in the first stage if one of the following three is true: (i) $C \in [\frac{2\alpha_1 R_s}{P_s(1+\alpha_1)}, K]$ and $\alpha_1 > \alpha_2$, (ii) $C \in [C^*(\alpha_1, \alpha_2), K]$ and $\delta_2\alpha_2 < \alpha_1 \le \alpha_2$, or (iii) $C \in [0, K]$ and $\alpha_1 \le \delta_2\alpha_2$. Otherwise, SU *accepts* α_1.

The detailed proof of Lemma 1 can be found in Appendix A.

4) Two Types of Sequential Equilibria for the One-Slot Two-Stage Bargaining: With Lemma 1, we can derive two types of SEs for the two-stage bargaining game by checking the consistence between the two stages.

For the first type of SE, α_2^* is slightly better than α_1^* (*i.e.,* $\delta_2\alpha_2^* < \alpha_1^* \le \alpha_2^*$). An SU with a small energy cost will accept α_1 in the first stage so that it can start to benefit immediately. An SU with a medium or large energy cost will wait for the second stage hoping for a better offer. In the second stage, only an SU with a medium energy cost will accept α_2, and an SU with a high energy cost has to reject α_2. Note that the SU does not know the value of α_2 in the first stage, and thus it needs to make the above decisions by *anticipating* the value of α_2. The PU needs to decide α_1 and α_2 by taking the SU's anticipation into consideration. An SE exists if the SU's anticipation is *consistent* with what the PU offers.

The first type of SE is summarized in the following theorem, with the proof in Appendix B.

Theorem 2: Given the first stage offer α_1, the beliefs and strategies for PU and SU are:

- $\alpha_2^*(\alpha_1)$: the solution of the fixed point equation (7) (in terms of α_2) at the top of the next page, where

[7] This assumption is made so as to have valid intervals in Lemma 1.

$$\alpha_2 = \min\left(\max\left(\alpha_p^*(K_1(\alpha_1,\alpha_2)), \frac{K_1(\alpha_1,\alpha_2)}{2R_s/P_s - K_1(\alpha_1,\alpha_2)}\right),\ \min\left(\left|\frac{K}{2R_s/P_s - K}\right|, 1\right)\right) \quad (7)$$

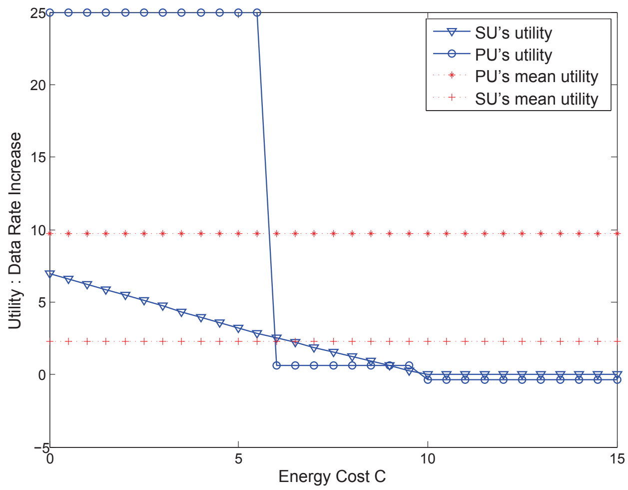


Fig. 5. First type SE (Theorem 2) in the one-slot two-stage bargaining.

$\alpha_p^*(K_1(\alpha_1,\alpha_2)) = \sqrt{\frac{2R_s(2R_r - R_{dir})}{P_s R_r (2R_s/P_s - K_1(\alpha_1,\alpha_2))}} - \frac{1}{2}$ and $K_1(\alpha_1,\alpha_2) = \frac{2R_s(\alpha_1 - \delta_2\alpha_2)}{P_s((1+\alpha_1) - \delta_2(1+\alpha_2))}$.

- $\mu_1(C)$: PU believes C is uniformly distributed in $[0, K]$.
- $\mu_2(C|\alpha_1)$: PU updates its belief on C as uniform distribution in $[C^*(\alpha_1, \alpha_2^*(\alpha_1)), K]$.
- $\mathcal{A}_1(\alpha_1|C)$: SU rejects α_1 if $C \in [C^*(\alpha_1, \alpha_2^*(\alpha_1)), K]$.
- $\mathcal{A}_2(\alpha_2|C,\alpha_1)$: SU accepts α_2 if and only if $\delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C > 0$.

Finally, PU chooses its first-stage offer α_1^* as follows,

$$\alpha_1^* = \arg\max_{\alpha_1\in[0,1]} \left((\delta_1(1-\alpha_1)R_r - R_{dir}) P_1 + (\delta_1\delta_2(1-\alpha_2^*(\alpha_1))R_r - R_{dir}) P_2 + (\delta_1\delta_2 - 1) R_{dir} P_3 \right), \quad (8)$$

where $P_1 = \frac{1}{K}K_1(\alpha_1)$, $P_2 = \frac{1}{K}\left(\frac{2R_s\alpha_2}{P_s(1+\alpha_2)} - K_1(\alpha_1)\right)$, and $P_3 = \frac{1}{K}\left(K - \frac{2R_s\alpha_2}{P_s(1+\alpha_2)}\right)$. PU will choose direct transmission if $U_p(\alpha_1^*) < 0$. *The above beliefs and strategies constitute an SE if and only if* $\delta_2\alpha_2^*(\alpha_1^*) < \alpha_1^* \le \alpha_2^*$.

Next we examine the second type of SE, where α_2^* is much larger than α_1^* (*i.e.*, $\alpha_1^* \le \delta_2\alpha_2^*$).

Theorem 3: The following beliefs and strategies constitute *infinitely many* SEs.

- α_2^*: a constant independent of α_1:
$$\alpha_2^* = \min\left(\max\left(\alpha_p^*, 0\right), \min\left(\left|\frac{K}{2R_s/P_s - K}\right|, 1\right)\right),$$
where $\alpha_p^* = \sqrt{\frac{2R_r - R_{dir}}{R_r}} - \frac{1}{2}$.
- α_1^*: any value satisfying $\alpha_1^* \le \delta_2\alpha_2^*$.
- $\mu_1(C) = \mu_2(C|\alpha_1)$: PU believes C is uniformly distributed in $[0, K]$ in both stages.
- $\mathcal{A}_1(\alpha_1|C)$: SU will never accept α_1.
- $\mathcal{A}_2(\alpha_2|C,\alpha_1)$: SU will accept α_2 if and only if $\delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C > 0$.

The proof of Theorem 3 is similar to the one of Theorem 2 and can be found in [27]. We want to emphasize that SE given in Theorem 2 is just a *potential* equilibrium result. Whether it exists in a particular game depends on the game parameters. Theorems 2 and 3 illustrates how the PU and SU should interact optimally in the one-slot multi-stage bargaining. Next, we will provide more insights of the SEs through some simulation results.

C. Simulation Results

We simulate and compare the data increases of the PU and SU at the equilibria in the two-stage game. We set $P_t = P_s = 1$, $K = 1.5R_s/P_s$, and $\delta_1 = \delta_2 = \delta = 0.8$. The PU's direct transmission rate $R_{dir} = 1$ and relay rate $R_r = 250$. In this case, SU's cooperative relay can bring a significant improvement to PU's data rate. The SU's own data rate $R_s = 10$.

Figs. 5 and 6 show the equilibrium data rate increases of PU with different energy cost C. Two figures illustrate two types of SEs of the same game. The *mean utility* (dotted curves) denotes the value averaged over all possible values of C. Since both PU and SU obtain positive mean utilities, CSS leads to a *win-win* situation comparing with no bargaining.

Fig. 5 corresponds to the first type of SE in Theorem 2. Let us focus on the solid curves: PU's utility (circle curve) and SU's utility (triangle curve). We can observe three regions depending on the value of the SU's energy cost C: (i) Small C (*e.g.*, $C \le 6$ in Fig. 5): SU accepts α_1, and thus both the PU and SU receive significant data rate increases; (ii) Medium C (*e.g.*, $C \in [7, 9]$ in Fig. 5): SU rejects α_1 but accepts α_2. Compared with the small C case, PU's utility dramatically decreases, since α_2 is larger than α_1 and more time is wasted in the bargaining. SU's utility decreases smoothly between these two regions, since a larger offer α_2 mitigates the negative effect of additional bargaining overhead; (iii) Large C (*e.g.*, $C \ge 10$ in Fig. 5): SU rejects both offers and PU receives a negative utility (*i.e.*, experiences data rate loss).

Fig. 6 corresponds to the second type of SE in Theorem 3. In this SE, the SU never accepts the first stage offer α_1, as it expects the second stage offer α_2 to be much better. As a result, the two-stage game becomes similar to a one-stage game. We can observe two regions based on the value of C, following a similar argument as for Fig. 5. By comparing Figs. 5 and 6, we notice that the PU's expected utility and SU's utility are both higher in Fig. 5 than in Fig. 6. The key reason is that the SE in Fig. 6 always wastes the first stage bargaining opportunity. In other words, the SE in Theorem 2 *Pareto dominates* the SE in Theorem 3.

As a comparison, we simulate the equilibrium result of the one-slot one-stage bargaining game discussed in Section III-A, as shown in Fig. 7. The system parameters are the same as the ones in Figs. 5 and 6. From the comparison of Fig. 7 and Fig. 5, we can see that at the first type SE in the two-stage bargaining the SU is better off (data rate increase changes from 2.14 to 2.44, *i.e.*, 14% improvement) and PU is worse off (data rate increase changes from 10.5 to 9.8, *i.e.*, 6.7% reduction).

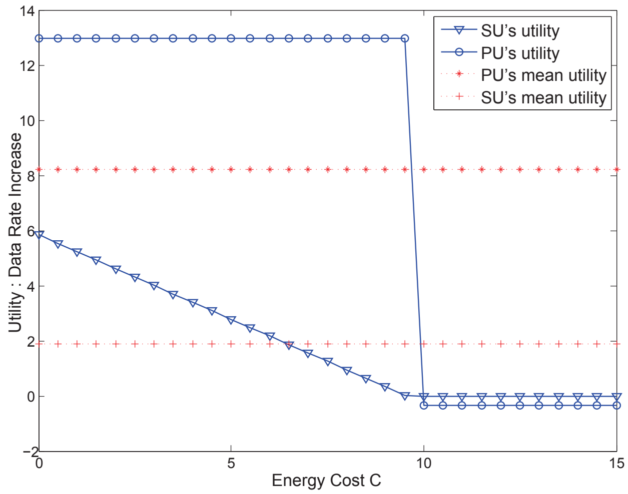


Fig. 6. Second type SE (Theorem 3) in the one-slot two-stage bargaining.

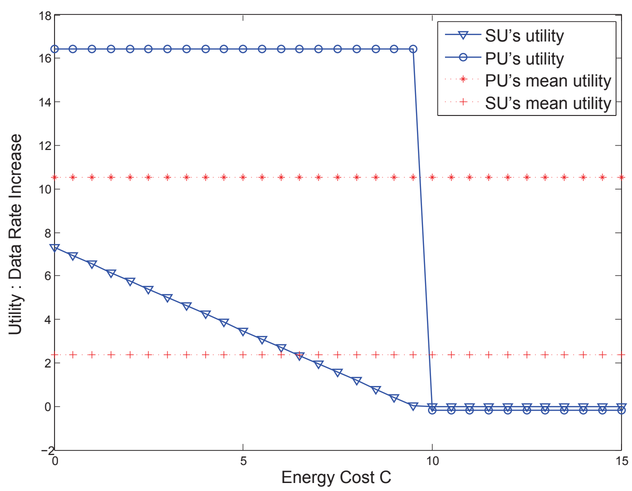


Fig. 7. Equilibrium in the one-slot one-stage bargaining.

As the total transmission time is limited, SU gains more transmission time means that the PU loses more transmission time, and thus the two players have opposite performance changes when moving from one-stage to two-stage bargaining. In particular, SU is better off in the two-stage bargaining, as it has more choices to negotiate with the PU and has relatively more bargaining power than in the one-stage bargaining.

From the comparison between Fig. 7 and Fig. 6, we can see that the basic shape of the curves are the same, since the second type of SE of the two-stage bargaining in Fig. 6 is similar to one-stage game due to the waste of the first-stage bargaining opportunity. The SU's data rate increase remains roughly the same in both figures. However, the PU's data rate increase is better in Fig. 7 with only one-stage bargaining. This confirms the same intuition that more bargaining reduces the bargaining power of the PU in our model. Figs. 5, 6, and 7 also show that the SU's data rate increase decreases with the energy cost C.

In this section, we have investigated how PU and SU conduct the dynamic bargaining in a single time slot. In Section IV, we extend the analysis to the multi-slot bargaining, where one PU and one SU sequentially bargain over a *finite number* of time slots.

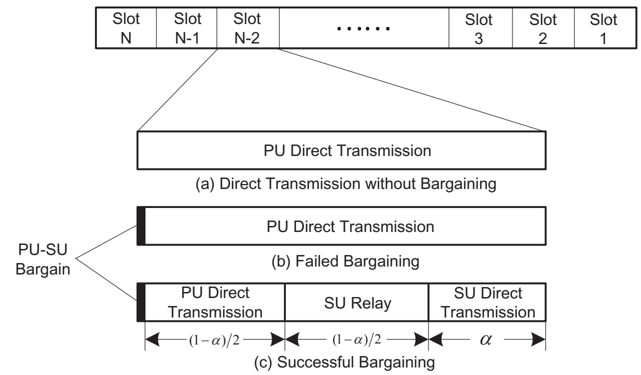


Fig. 8. The slotted system model with three possible bargaining results in a single time slot.

IV. Multi-Slot Bargaining Game

Comparing with the model in Section III, the model with multiple time slots is more interesting, since the PU and SU typically interact over many time slots in practice. The analysis, however, is much more complicated, and the equilibrium results are very different from the ones in Section III.

A. System Model

In this section, we assume that the SU's energy cost C follows a *binary* distribution with two values ($C_h > C_l$).[8] A larger C means that the SU's transmission is more costly. Accordingly, there are two types of SU based on C's value:

- **High type SU**: $C = C_h$,
- **Low type SU**: $C = C_l$.

The PU has a *belief* on C's distribution in each time slot $n = N, ..., 1$, *i.e.*, $\Pr(C = C_h) = q_n$ and $\Pr(C = C_l) = 1 - q_n$. PU's belief will be updated over time depending on the bargaining history in previous time slots.

The bargaining consists of N consecutive time slots as in Fig. 8. To make the discussions easier to follow later on, *we index time backwards, i.e., the bargaining starts with time slot N and ends in time slot 1.* We also normalize each time slot length to 1. For ease of discussion, the overhead due to bargaining is assumed to be negligible. The model where bargaining overhead is considered can be analyzed analogously. The bargaining process within one time slot is the same as in Fig. 2. We further assume that there is *at most* one stage of bargaining in each time slot, *i.e.*, this is a *multi-slot one-stage* game.

Furthermore, we assume that the SU is a *non-myopic* player, whose utility is its total utilities in all N time slots. For simplicity, we assume that the PU is *myopic*, *i.e.*, it only maximizes its utility in the current single time slot. PU and SU will bargain over the value of α in each time slot, by considering the bargaining results in previous time slots.

[8] For simplicity, here we consider a binary distribution of energy cost C. As shown in [28], the *continuous* distribution of C technically complicates the analysis but offers the same game theoretical insights about equilibrium outcomes.

B. Sequential Equilibrium Analysis

As a baseline, we will first consider the one-slot bargaining game (*i.e.*, $N = 1$). ***The SU's single-slot utility*** U_s after accepting an offer α is

$$U_s(\alpha) = \alpha R_s - \frac{1+\alpha}{2} P_s C. \tag{9}$$

Similar to (3), $U_s(\alpha)$ can also be viewed as SU's *data rate increase* by accepting α. Given PU's offer α, it is optimal for SU to accept the offer if and only if $U_s(\alpha) > 0$.

The PU's single-slot utility U_p is its *achievable data rate*. Without SU's relay, the PU can achieve a data rate R_{dir}. If PU's offer α is accepted by SU, then the PU's effective data rate is $(1-\alpha)R_r$, where R_r is given in (1). Knowing the SU's potential strategy, PU makes the optimal decision by maximizing its utility

$$U_p(\alpha) = \max\left\{R_{dir}, (1-\alpha)R_r\right\}. \tag{10}$$

Now let us consider the multi-slot case, the PU's belief is the probability assessment q_n about the High type SU in time slot n, with an initial value $q_N = \eta$ in the first time slot indexed by N. As the game progresses, both users observe all prior moves, which enable the PU to *update* its belief about the SU's type. The SU's belief in all time slots is deterministic since it knows its own type.

Intuitively, an SU will only choose to cooperate and serve as a relay node if it can get a positive total utility in N time slots. This not only depends on its energy cost (either C_h or C_l), but also on its achievable average data rate per unit power R_s/P_s. Next we will discuss three different cases based on the relationship between the energy cost and R_s/P_s. As we will see, in most cases we can decompose a multi-slot bargaining problem into multiple single-slot bargaining problems, and thus the result is not difficult to get. There is only one subcase of Case 3 that does not allow such decomposition, and we will discuss this subcase in Section IV-C. All cases are summarized in Fig. 9.

1) ***Case 1****: $R_s/P_s \in (0, C_l]$:* In this case, we have SU's single time slot utility $U_s(\alpha) \le 0$ for any $\alpha \in$ [0,1]. This means that even the low energy cost C_l is too costly for the SU to achieve a positive data rate increase in one time slot. Therefore, SU will reject any offer from PU in *all* time slots. PU thus will choose direct transmission without cooperation in each time slot (Fig. 8(a)).

2) ***Case 2****: $R_s/P_s \in (C_l, C_h]$:* In this case, a High type SU will reject any offer, since $U_s(\alpha) < 0$ for any $\alpha \in$ [0,1]. However, a Low type SU may accept a large enough α. It is easy to see that U_s is a linear *increasing* function of parameter α. The PU's single-slot utility $U_p(\alpha)$ can be simplified as $(1-\alpha)R_r$, and thus is a linear *decreasing* function of α. Therefore, in order to attract the help from a Low type SU, PU needs to provide an offer α that makes $U_s(\alpha)$ *slightly* larger than zero so that the SU will accept it.

The PU can first compute the *threshold* offer $\alpha_l' = \frac{1}{\frac{2R_s}{C_l P_s}-1}$, which leads to a zero single-slot utility for the Low type SU. Then, the PU's optimal offer can be $\alpha_l = \frac{1}{\frac{2R_s}{C_l P_s}-1} + \varepsilon$, where ε is an arbitrarily small positive number. Substitute α_l into (9), we get the SU's utility when accepting α_l, *i.e.*,

$$U_s(\alpha_l) = \left(R_s - \frac{1}{2}P_s C_l\right)\alpha_l - \frac{1}{2}P_s C_l = \left(R_s - \frac{1}{2}P_s C_l\right)\varepsilon.$$

For ease of discussion, we define

- $\Delta R_{sl} = \left(R_s - \frac{1}{2}P_s C_l\right)\varepsilon > 0$: the Low type SU's single-slot utility if accepting α_l,
- $R_l = (1-\alpha_l)R_r$: the PU's single-slot data rate if SU accepts α_l.

If there is no relay, PU's utility will be R_{dir}. If an SU relays for PU after accepting α_l, PU's utility will be R_l. Therefore, PU only has *two* options: no offer or offer α_l.[9] PU makes the decision based on the relationship of R_{dir} and R_l. Theorem 4 summarizes the SE of this case.

Theorem 4: Consider a multi-slot bargaining game where $R_s/P_s \in (C_l, C_h]$. If $R_{dir} \ge R_l$, PU always chooses direct transmission only regardless of SU's type. If $R_{dir} < R_l$, PU always offers α_l to SU. A High type SU rejects the offer α_l, and a Low type SU accepts the offer α_l.

Proof Sketch (Details in [27]): Take the case $R_{dir} \ge R_l$ for instance, it is easy to find that in a single time slot, PU's strategy of choosing direct transmission *dominates* the cooperative transmission strategy in an *expected* sense. Since such fact is independent of the PU's belief, it is easy to see that the PU and hence SU (Hiph or Low type) both have no incentive to deviate from their single-slot strategies in multi-slot case. It means that the multi-slot game can be decomposed into N independent single-slot game. ■

3) ***Case 3****: $R_s/P_s \in (C_h, \infty)$:* This is the most interesting case, as both the High and Low type SU may accept a large enough offer α. In this case, any type SU's single-slot utility function is the linear increasing function of parameter α. Similar to the discussions in Case 2, to attract the help from a High type SU, PU needs to provide the optimal offer $\alpha_h = \alpha_h' + \varepsilon = \frac{1}{\frac{2R_s}{C_h P_s}-1} + \varepsilon$, where ε is the same as in α_l. Let us define

- $\Delta R_{sh} = \left(R_s - \frac{1}{2}P_s C_h\right)\varepsilon > 0$: the High type SU's single-slot utility if accepting α_h,
- $R_h = (1-\alpha_h)R_r$: the PU's single-slot data rate if SU accepts α_h.

As $\alpha_h' > \alpha_l'$, we have $\alpha_h > \alpha_l$ and hence $R_h < R_l$. Notations α_l, α_h, R_{dir}, R_l, R_h, ΔR_{sl}, ΔR_{sh} have been defined. Here, we further define

- $R_{sh} = \alpha_l R_s - \frac{1+\alpha_l}{2} P_s C_h$: the High type SU's single-slot utility if it accepts α_l,
- $R_{sl} = \alpha_h R_s - \frac{1+\alpha_h}{2} P_s C_l$: the Low type SU's single-slot utility if it accepts α_h.

When ε approaches zero in the definitions of α_l and α_h, it is easy to see that $R_{sl} > 0$. We will show that $R_{sh} < 0$. Define $\Delta_C = C_h - C_l$, and we have $\Delta_C > 0$. Substitute α_l into R_{sh}

[9] Since the PU's utility in the relay mode is a linear decreasing function in α and SU's utility U_s is a linear increasing function in α, a larger offer α means a decrease of PU's utility but an increase of SU's utility (Low type SU). Anticipating this, the PU has no incentive to choose any offer larger than α_l.

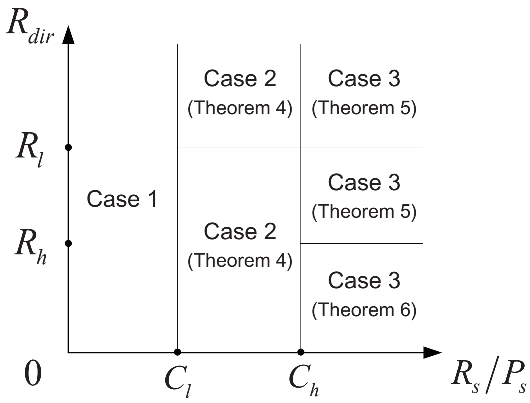


Fig. 9. Equilibrium outcomes in different regions.

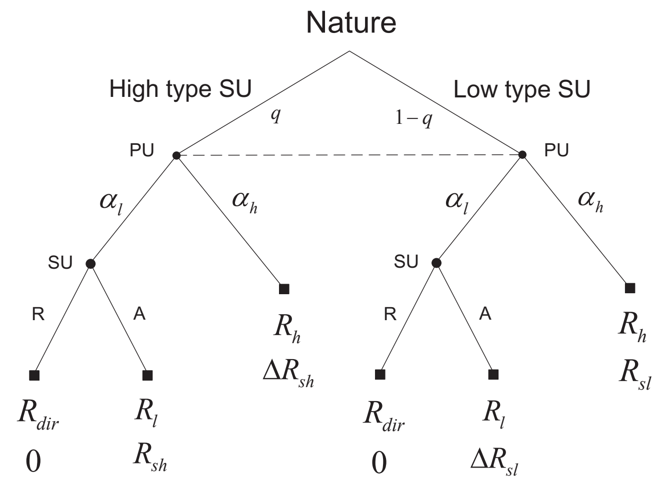


Fig. 10. Game tree of the single-slot bargaining with $R_l > R_h > R_{dir}$ and $C_l < C_h < R_s/P_s$.

to get

$$\begin{aligned} R_{sh} &= \alpha_l R_s - \frac{1+\alpha_l}{2} P_s (C_l + \Delta_C) \\ &= \left(R_s - \frac{1}{2} P_s C_h \right) \varepsilon - \frac{1+\alpha_l'}{2} P_s \Delta_C. \end{aligned} \tag{11}$$

When ε approaches zero, (11) will become

$$\begin{aligned} \lim_{\varepsilon \to 0} R_{sh} &= \lim_{\varepsilon \to 0} \left(\left(R_s - \frac{1}{2} P_s C_h \right) \varepsilon - \frac{1+\alpha_l'}{2} P_s \Delta_C \right) \\ &= -\frac{1+\alpha_l'}{2} P_s \Delta_C < 0. \end{aligned} \tag{12}$$

Any other offer between α_l and α_h will not change the decision of the High type SU but only increase the Low type SU's utility, thus PU will not offer it. Therefore, PU has *three* options: no offer, offer α_h, or offer α_l, with the corresponding utilities R_{dir}, R_h, and R_l. We have three subcases here, and the first two easy cases are summarized in Theorem 5.

Theorem 5: Consider a multi-slot bargaining game with $R_s/P_s \in (C_h, \infty)$. If $R_{dir} \in [R_l, \infty)$, PU always chooses direct transmission only regardless of the SU's type. If $R_{dir} \in [R_h, R_l)$, PU always offers α_l to SU. A High type SU rejects the offer, and a Low type SU accepts the offer.

The proof of Theorem 5 shares the same idea with the one of Theorem 4. Hence, we omit it here. In the two subcases of Theorem 5, PU's decision does not depend on its belief about the SU's type. The analysis of the multi-slot game can also be decomposed into N independent single-slot games.

The only remaining subcase is $R_{dir} \in (0, R_h)$, where the multi-slot game cannot be decomposed. The PU and SU need to make decisions by considering both the history and the future decisions of each other. In particular, a Low type SU may reject an offer α_l even if it leads a positive single-slot utility, in order to create a *reputation* indicating that it is a High type SU and induce the PU to provide a high offer in future time slots. We will discuss this reputation effect in details next.

C. Sequential Bargaining with Reputation Effect

In Section IV-B, we have discussed Case 1, Case 2 (Theorem 4), and part of Case 3 (Theorem 5) as shown in Fig. 9. In this subsection, we will consider the last subcase of Case 3, where $R_s/P_s \in (C_h, \infty)$ and $R_{dir} \in (0, R_h)$. In this subcase, PU has *three* options: no offer, offer α_h, or offer α_l. Since R_{dir} is small, PU will never choose "no offer". Thus, we will discuss whether the PU will provide α_h or α_l in each time slot $n = N, \cdots, 1$.

1) Single-Slot Game: Before discussing the SE of the multi-slot game, let us first consider the game tree of the *single-slot* game as in Fig. 10. *Nature* moves first and determines the SU's type.[10] PU and SU make decisions alternatively at the non-leaf nodes. The dotted line connecting two nodes indicates that the PU does not know the SU's type. Each possible game result is denoted by a leaf node (a black solid square) together with the corresponding PU's utility (upper value) and SU's utility (lower value). PU's belief (on SU's type) is Pr$(C = C_h) = q$. Notations α_l, α_h, R_{dir}, R_l, R_h, ΔR_{sl}, ΔR_{sh}, R_{sh}, and R_{sl} have been defined before.

Recall that we have $R_{sh} < 0$ and $R_{sl} > 0$, thus a High type SU will not accept a low offer, while a Low type SU has the incentive to accept a high offer. In Fig. 10, PU first decides to offer α_h or α_l. Then SU makes the acceptance (**A**) or rejection (**R**) decision. If PU offers α_h, an SU always accepts it regardless of its type since $\Delta R_{sh} > 0$ and $R_{sl} > 0$. Hence there is only one leaf node following α_h. If PU offers α_l, a High type SU will reject since $R_{sh} < 0$, and a Low type SU will accept since $\Delta R_{sl} > 0$.

By considering the SU's response, PU's *expected* utility if offering α_l is

$$U_p^{\alpha_l} = q R_{dir} + (1-q) R_l. \tag{13}$$

PU's utility if offering α_h is R_h. Thus PU will offer α_l if $U_p^{\alpha_l} > R_h$, or equivalently $q < (R_l - R_h)/(R_l - R_{dir})$ (*i.e.*, the probability of being a High type SU is low).

2) Multi-Slot Game: Now let us consider the multi-slot game, where the PU's belief might change over time (*i.e.*, q_n for time slot n instead of a fixed value q) based on the game history. The SU's behavior may also change depending on the game history and its anticipation of the future. In particular, a

[10]With Harsanyi's transformation, the game with incomplete information can be transformed into a game with complete but imperfect information by adding *Nature* as a third player into the game. See [29] for details.

Low type SU may choose to reject α_l in a particular time slot even if the offer brings a positive utility. The purpose of such a strategy is for this Low type SU to create a *reputation* of a High type SU and induce the PU to offer α_h in the future.

An SE of the multi-slot bargaining includes the following components in each time slot $n = N, \ldots, 1$: (i) The update of PU's belief q_n in time slot n, (ii) PU's strategy (offer α_l or α_h) in time slot n, and (iii) SU's strategy (accept or reject) in time slot n. We summarize the SE in the following theorem. Here, we define the parameter $d = \frac{R_l - R_h}{R_l - R_{dir}} \in (0, 1)$.

Theorem 6: The SE of the multi-slot bargaining game is given from (a) to (l):

- PU's Belief Updates:[11] (a) if $q_{n+1} = 0$, $q_n = 0$; (b) if $q_{n+1} > 0$ and SU accepts α_h in time slot $n+1$, $q_n = q_{n+1}$; (c) if $q_{n+1} > 0$ and SU accepts α_l in time slot $n+1$, $q_n = 0$; (d) if $q_{n+1} > 0$ and SU rejects α_l in time slot $n+1$, $q_n = \max(d^n, q_{n+1})$.
- PU's Strategy: (e) if $q_n < d^n$, offers α_l in time slot n; (f) if $q_n > d^n$, offers α_h in time slot n; (g) if $q_n = d^n$, offers α_h with probability $\frac{\Delta R_{sl}}{R_{sl} - \Delta R_{sl}}$ and offers α_l with probability 1-$\frac{\Delta R_{sl}}{R_{sl} - \Delta R_{sl}}$ in time slot n.[12]
- The High type SU's Strategy: (h) always accepts α_h and rejects α_l.
- The Low type SU's Strategy: (i) always accepts α_h; (j) if $n = 1$ (the last time slot), accepts α_l; (k) if $n > 1$ and $q_n \geq d^{n-1}$, rejects α_l; (l) if $n > 1$ and $q_n < d^{n-1}$, rejects α_l with probability $y_n = \frac{(1-d^{n-1})q_n}{d^{n-1}(1-q_n)}$ and accepts α_l with $1 - \frac{(1-d^{n-1})q_n}{d^{n-1}(1-q_n)}$.

The proof of Theorem 6 can be found in Appendix C. The bargaining over multiple time slots is a dynamic and repeated interaction between PU and SU. The *reputation effect* emerges from the bargaining interaction, which we will discuss in details next.

D. Reputation Effect Analysis

The Low type SU's reputation is the PU's belief Pr(High type)=q_n about the SU in time slot n. The reputation effect refers to the fact that a Low type SU has an incentive to reject the low offer α_l, in order to create a reputation of a High type SU so as to get higher utility (see (k) and (l) in Theorem 6). The incentive of doing so becomes higher when the bargaining lasts longer, *i.e.*, reputation effect is more likely to happen in long relationships than in short ones [29]. Therefore, we discuss such an effect when N is sufficiently large.

When the SU does not have a strong reputation yet (*i.e.*, $q_n < d^n$), it will choose to reject the low offer α_l with a certain probability to increase its reputation as in (l). The *balancing strategy* y_n^* (see Appendix C) indicates that the Low type SU's *willingness* to sustain a reputation of a High type SU. Fig. 11 shows how y_n^* evolves as the game proceeds. We choose three different values of $d = 0.4, 0.6, 0.7$. Note that the balancing strategy y_n^* with different values of d begins to decrease when $n = 5, 8, 11$, respectively. Prior to these time slots, the Low type SU *always* rejects α_l as in (k) and does not need to randomize its action, *i.e.*, it has a deterministic strategy with $y_n^* = 1$. As the game approaches to the end (*i.e.*, the time slot index decrease from 20 to 1), y_n^* monotonically decreases and becomes zero in the last time slot ($n = 1$). This is intuitive as the SU puts a higher priority on getting a positive utility in the current time slot (by accepting α_l) over the potential higher future utility (by rejecting α_l and sustaining the reputation) when the game is closer to the end.

[11]Recall that we index time backwards, and thus we will compute q_n based on q_{n+1} since time slot n is after time slot $n+1$.

[12]Since ε is an arbitrary small positive, ΔR_{sl} is arbitrarily small. Therefore, it is valid that $R_{sl} > 2\Delta R_{sl}$.

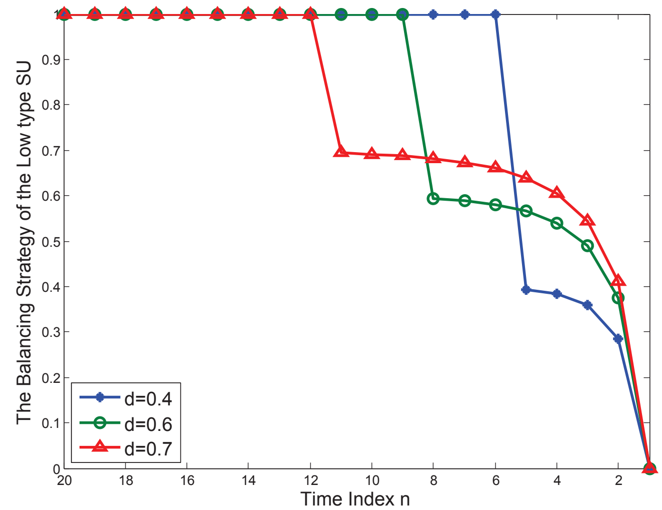


Fig. 11. The balancing strategy of the Low type SU for different d.

The Low type SU can obtain a higher total utility by taking advantage of the reputation effect. Since $d \in (0, 1)$ and N is large enough, then d^N can be arbitrarily small for a large enough N, and thus it is easy to satisfy the condition $q_N = \eta > d^{N-1}$ even if the initial belief η (in the first time slot N) is small. From (k), the Low type SU will reject α_l during the initial time slot $n = N$. Anticipating this, the PU will offer α_h as in (f). As a result, a Low type SU receives a high offer α_h and gets the high utility R_{sl} since the first time slot, *i.e.*, the Low type SU can benefit from incomplete information and thus gain more from the cooperative communication. As mentioned in Section II-C, it indicates that in this one-sided incomplete information bargaining game where PU has more bargaining power, the less *bargaining power* of SU can be increased (*i.e.*, SU can obtain higher bargaining utility). The reputation effect described here has been quite powerful. However, if there are several players who all have incomplete information within a game, they will all have incentives to maintain reputations to each other. But, it seems unlikely that the reputation effect could lead to a strong influence to the game [29].

For simplicity, we have assumed that there are only two types of SU. For the general case where there are multiple different types, Fudenberg *et al.* in [30] obtained the upper and lower bounds on the long-run player's utility that holds in any Nash equilibrium of the game. Milgrom *et al.* in [28] considered the case with the *continua* of types of players based on the one-sided incomplete information model. They pointed out that the difference between two and multiple types assumptions would not impair the conclusion related to the reputation effect as in [31], which shares essentially the same game theoretical insights as our paper's.

In this section, we have assumed that PU is a *myopic* player with the purpose of simplifying the analysis. However, through a much more involved analysis, we can show that the results of each subcase in Section IV also hold for the case of *non-myopic* PU assumption. We omit the proof due to the page limit, and details can be found in our online technical report [27].

V. Conclusions

This paper proposes a novel noncooperative bargaining-based cooperative spectrum sharing mechanism between one PU and one SU under two different network scenarios. The general CSS case with multiple PUs and multiple SUs can be modeled as probabilistic bargaining over an undirected graph, and can be decomposed into multiple pairs of bilateral bargaining game between one PU and one SU studied in this paper. We model the dynamic bargaining process under incomplete information as a dynamic Bayesian game, and characterize the corresponding sequential equilibria under different system parameters. Theoretical analysis and numerical results indicate that both the PU and the SU could obtain performance improvements (data rate increases) via the relay-based CSS mechanism. Thus, our proposed scheme leads to a win-win situation.

Appendix

A. Proof of Lemma 1

The SU may achieve different utilities depending on its choices in both stages: (i) if SU accepts α_1: $U_s = \delta_1\alpha_1 R_s - \frac{1+\alpha_1}{2}\delta_1 P_s C$; (ii) if SU accepts α_2 (and rejects α_1): $U_s = \delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C$; (iii) if SU rejects both offers: $U_s = 0$. SU therefore prefers accepting α_1 to accepting α_2 if $\delta_1\alpha_1 R_s - \frac{1+\alpha_1}{2}\delta_1 P_s C > \delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C$ (call it condition (a)), and SU prefers accepting α_1 to rejecting both offers if $\delta_1\alpha_1 R_s - \frac{1+\alpha_1}{2}\delta_1 P_s C > 0$ (call it condition (b)).

Thus, we can derive the condition under which SU will accept α_1 given (α_1, α_2). When $\alpha_1 > \delta_2\alpha_2$, from condition (a) we have $0 < C < C^*(\alpha_1, \alpha_2)$. Considering condition (b), SU will accept α_1 if $0 < C < \min\left(C^*(\alpha_1, \alpha_2), \frac{2\alpha_1 R_s}{P_s(1+\alpha_1)}\right)$, which can also be written equivalently as $0 < C < \frac{2\alpha_1 R_s}{P_s(1+\alpha_1)}$ when $\alpha_1 > \alpha_2$, and $0 < C < C^*(\alpha_1, \alpha_2)$ when $\delta_2\alpha_2 < \alpha_1 \le \alpha_2$. When $\alpha_1 \le \delta_2\alpha_2$, conditions (a) and (b) cannot be satisfied simultaneously, which means that SU will never accept α_1 for any $C \in [0, K]$. Thus we have completed the proof. ■

B. Proof of Theorem 2

With Lemma 1, we can obtain three possible belief updates for PU in the second stage. When $C \in [C^*(\alpha_1, \alpha_2), K]$ and $\delta_2\alpha_2 < \alpha_1 \le \alpha_2$, PU will update its belief at the beginning of the second stage as that C is uniformly distributed in $[K_1, K]$ with $K_1 = C^*(\alpha_1, \alpha_2)$ after observing SU's rejection. Note that $C^*(\alpha_1, \alpha_2) \le \frac{2\alpha_1 R_s}{P_s(1+\alpha_1)} \le \frac{2\alpha_2 R_s}{P_s(1+\alpha_2)} \le \frac{R_s}{P_s}$ when $\delta_2\alpha_2 < \alpha_1 \le \alpha_2$, thus according to Theorem 1, the optimal offer α_2^* is (14) at the top of the next page, where $\alpha_p^*(K_1) = \sqrt{\frac{2R_s(2R_r - R_{dir})}{P_s R_r(2R_s/P_s - K_1)}} - \frac{1}{2}$. Combining (14) with the fact that $K_1 = \frac{2R_s(\alpha_1 - \delta_2\alpha_2)}{P_s((1+\alpha_1) - \delta_2(1+\alpha_2))}$, we can eliminate K_1 and express α_2 as a function of α_1 (*i.e.*, $\alpha_2^*(\alpha_1)$) by solving a fixed point equation. Based on $\alpha_2^*(\alpha_1)$, we can further express K_1 as a function of α_1 only, *i.e.*, $K_1(\alpha_1)$.

Now we have reduced the game to a single-stage optimization problem (only optimizing over α_1) for PU: given α_1, we have specified $[\mu_1(C), \mu_2(C|\alpha_1)]$, $[\mathcal{A}_1(\alpha_1|C), \mathcal{A}_2(\alpha_2|C, \alpha_1)]$ and PU's $\alpha_2^*(\alpha_1)$:

- $\alpha_2^*(\alpha_1)$: calculated above by solving a fixed point equation.
- $\mu_1(C)$: PU believes C is uniformly distributed over $[0, K]$.
- $\mu_2(C|\alpha_1)$: PU updates its belief on C as uniformly distributed in $[C^*(\alpha_1, \alpha_2^*(\alpha_1)), K]$.
- $\mathcal{A}_1(\alpha_1|C)$: SU will reject α_1 if $C \in [C^*(\alpha_1, \alpha_2^*(\alpha_1)), K]$.
- $\mathcal{A}_2(\alpha_2|C, \alpha_1)$: SU will accept α_2 if and only if $\delta_1\delta_2\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1\delta_2 P_s C > 0$.

Thus, PU should choose the first-stage offer α_1^* to maximize

$$U_p(\alpha_1) = \left((\delta_1(1-\alpha_1)R_r - R_{dir}\right) P_1 + \left(\delta_1\delta_2(1-\alpha_2^*(\alpha_1))R_r - R_{dir}\right) P_2 + (\delta_1\delta_2 - 1)R_{dir}P_3), \quad (15)$$

where

- $P_1 = \text{Prob}(\text{SU accepts } \alpha_1)$,
- $P_2 = \text{Prob}(\text{SU accepts } \alpha_2)$,
- $P_3 = \text{Prob}(\text{SU rejects } \alpha_1, \alpha_2)$.

The probabilities are consistent with SU's subsequent strategies, *i.e.*,

- $P_1 = \text{Prob}\left(C \in [0, K_1(\alpha_1)]\right)$,
- $P_2 = \text{Prob}\left(U_2(\alpha_2^*) > 0, C \in [K_1(\alpha_1), K]\right)$,
- $P_3 = 1 - P_1 - P_2$.

In the second stage, SU will accept α_2 if

$$\delta_2\left(\delta_1\alpha_2 R_s - \frac{1+\alpha_2}{2}\delta_1 P_s C\right) > 0.$$

Thus, the three probabilities can be further given as $P_1 = \frac{1}{K}K_1(\alpha_1)$, $P_2 = \frac{1}{K}\left(\frac{2R_s\alpha_2}{P_s(1+\alpha_2)} - K_1(\alpha_1)\right)$, and $P_3 = \frac{1}{K}\left(K - \frac{2R_s\alpha_2}{P_s(1+\alpha_2)}\right)$, respectively. Substituting the three probabilities into (15), α_1^* can be calculated by maximizing (15). If $U_1(\alpha_1^*)$ is less than zero, then PU will choose direct transmission. Finally, we must also verify whether $\delta_2\alpha_2^*(\alpha_1^*) < \alpha_1^* \le \alpha_2^*$ is true. If yes, then the derived beliefs and strategies constitute an SE. ■

C. Proof of Theorem 6

First, let us verify the High type SU's equilibrium strategy, assuming that PU follows (a)-(g). First consider PU's offer α_l. If the High type SU accepts α_l, it will get a negative utility R_{sh} in the current time slot. Furthermore, PU will offer α_l in the rest of the time slots. The overall SU utility will be negative. This means that the High type SU should reject α_l and thus get a zero utility in the current time slot (and possible positive utility in future time slots as PU might offer α_h). Next consider the case where PU offers α_h. Accepting this offer leads to a positive SU's utility in the current time slot, and

$$\min\left(\max\left(\alpha_p^*(K_1), \frac{K_1}{2R_s/P_s - K_1}\right), \min\left(\left|\frac{K}{2R_s/P_s - K}\right|, 1\right)\right) \tag{14}$$

$$\begin{aligned} q_1^* &= \text{Prob(SU is High type} \mid \text{SU rejects } \alpha_l) \\ &= \frac{\text{Prob(High type SU)Prob(SU rejects } \alpha_l \mid \text{High type SU)}}{\text{Prob(High type SU)Prob(SU rejects } \alpha_l \mid \text{High type SU)} + \text{Prob(Low type SU)Prob(SU rejects } \alpha_l \mid \text{Low type SU)}} \end{aligned} \tag{18}$$

$$\begin{aligned} U_p(\alpha_l) &= q_n R_{dir} + y_n(1-q_n)R_{dir} + (1-y_n)(1-q_n)R_l = \left(q_n + \frac{q_n(1-d^{n-1})}{d^{n-1}}\right) R_{dir} + \frac{(d^{n-1}-q_n)}{(1-q_n)d^{n-1}}(1-q_n) R_l \\ &= \frac{1}{d^{n-1}}\left[q_n R_{dir} + (d^{n-1}-q_n)R_l\right] \end{aligned} \tag{22}$$

does not decrease the SU's possible utility in the future time slots. Thus the High type SU should accept α_h.

Next, we verify PU's equilibrium strategy (e)-(g). Define q_n^* as *the limiting belief*, which can be interpreted as a decision threshold to determine which offer the PU should provide (α_l or α_h). It is easy to analyze the last time slot. PU in the last time slot, *i.e.*, $\mathbb{P}_1$[13], is indifferent between offering α_l and α_h if

$$q_1 R_{dir} + y_1(1-q_1)R_{dir} + (1-y_1)(1-q_1)R_l = R_h. \tag{16}$$

The *LHS* of (16) is PU's utility if offering α_l, and the *RHS* is the utility if offering α_h. In (16), y_1 denotes the probability that the Low type SU rejects α_l. If PU offers α_l in the last time slot, the Low type SU will definitely accept (*i.e.*, $y_1 = 0$). Given that $y_1 = 0$, we solve (16) and get the limiting belief $q_1^* = \frac{R_l - R_h}{R_l - R_{dir}} = d$ for $n = 1$. If $\mathbb{P}_1$'s utility if offering α_l is larger than R_h, then we must have $q_1 < \frac{R_l - R_h}{R_l - R_{dir}} = d$ and $\mathbb{P}_1$ will choose to offer α_l. On the other hand, if $q_1 > d$, $\mathbb{P}_1$ will offer α_h. If $q_1 = d$, $\mathbb{P}_1$ will be indifferent and randomize between α_l and α_h with a probability that we will discuss later.

When $n = 2$ (the second to last time slot), q_2^* has to satisfy

$$q_2 R_{dir} + y_2(1-q_2)R_{dir} + (1-y_2)(1-q_2)R_l = R_h, \tag{17}$$

where y_2 are defined similarly to y_1. Because time slot 2 is followed by time slot 1, the history action in time slot 2 may have influence on the last time slot. It implies that y_2 is an unknown variable. Thus there should be one more restriction for solving the two dependent variables (q_2 and y_2) in (17).

Recall Requirement 3, given q_1^*, which is a *consistent* belief, q_1^* should satisfy the Bayes' rule and is derived from q_2 and y_2 as (18). Note that the Bayesian process only applies if PU provides α_l and SU rejects it. Since the High type SU always rejects α_l, an SU's accepting α_l in time slot 2 will reveal that it belongs to the Low type and thus result in $q_1 = 0$, which contradicts $q_1^* = d$. If PU provides α_h, any type of SU will accept it, from (b) we get $q_2 = q_1^* = d$, which we will verify later.

[13]For clear illustrations, we interchangeably use different notations, $\mathbb{P}_1$, $\mathbb{P}_2$,..., $\mathbb{P}_N$ to denote the PU in different time slots. However, these notations all indicate the *unique* PU in the multi-slot bargaining game.

With (18), we have $q_1^* = \frac{q_2 \times 1}{q_2 \times 1 + y_2(1-q_2)} = d$. Further, y_2 can be given as $y_2 = \frac{q_2(1-d)}{(1-q_2)d}$. Substitute y_2 into (17), and we obtain $q_2^* = \left(\frac{R_l - R_h}{R_l - R_{dir}}\right)^2 = d^2$. Combine q_2^* and y_2, then we get y_2^*, *i.e.*,

$$y_2^* = \frac{d^2(1-d)}{(1-d^2)d} = \frac{d}{1+d}. \tag{19}$$

Here y_2^* is the probability that the Low type SU rejects α_l and makes $\mathbb{P}_2$ indifferent between offering α_l and α_h. Following [32], we call y_2^* as *the balancing strategy* of the Low type SU.

For the general case where $n \geq 3$, the limiting belief q_n^* and the balancing strategy y_n^* can be similarly derived from

$$q_n R_{dir} + y_n(1-q_n)R_{dir} + (1-y_n)(1-q_n)R_l = R_h. \tag{20}$$

By using induction, we can *conjecture* that $q_{n-1}^* = d^{n-1}$ in time slot $n-1$. Again, we apply the Bayes' rule to q_{n-1}^* and get the restriction between q_n and y_n,

$$q_{n-1}^* = \frac{q_n}{q_n + y_n(1-q_n)} = d^{n-1}. \tag{21}$$

From (21), we get $y_n = \frac{q_n(1-d^{n-1})}{(1-q_n)d^{n-1}}$. Substitute y_n into (20), and we get $q_n^* = d^n$. Besides, we get the *balancing strategy* $y_n^* = y_n(q_n^*) = \frac{d-d^n}{1-d^n}$. Substitute y_n into the *LHS* of (20), and we get $\mathbb{P}_n$'s utility if offering α_l as (22). $\mathbb{P}_n$'s utility if offering α_h is $U_p(\alpha_h) = R_h$. If $U_p(\alpha_l) > U_p(\alpha_h)$, we have $q_n < d^n$ and $\mathbb{P}_n$ will offer α_l. If $U_p(\alpha_l) < U_p(\alpha_h)$, we have $q_n > q_n^*$ and $\mathbb{P}_n$ will offer α_h. If $U_p(\alpha_l) = U_p(\alpha_h)$, *i.e.*, $q_n = d^n$, $\mathbb{P}_n$ will be indifferent and randomize to choose α_l and α_h. For the mixed strategy in (g), we will discuss later.

Next let us consider the Low type SU's equilibrium strategy (i)-(l). It is obvious that the Low type SU always accepts α_h if PU offers it. The probability that the Low type SU rejects α_l in time slot n is y_n. When $n > 1$ and $q_n \geq d^{n-1}$, y_n equals 1, *i.e.*, the Low type SU will always reject α_l. When $n > 1$ and $q_n < d^{n-1}$, $y_n = \frac{(1-d^{n-1})q_n}{d^{n-1}(1-q_n)} \in (0,1)$. Thus, we get results in (i)-(l).

Finally, let us verify PU's belief update in (a)-(d). For offer α_h, any type of SU will definitely accept it. It indicates that there is no information to help update the PU's belief. Thus, we have $q_n = q_{n+1}$. Next we verify (a) and (c) from two aspects:

- If there is an offer α_l in time slot $n+1$ and the SU accepts it, then PU immediately knows that the SU is a Low type. Therefore, PU will update its belief as $q_n = 0$.
- If $q_{n+1} = 0$, then PU knows that the SU is a Low type for time slot $n+1$ and the whole following time slots. Therefore, PU will always set its belief as zero.

When $q_{n+1} > 0$ and SU rejects α_l in time slot $n+1$, with the Bayes' rule we have (23) at the top of the next page. According to (k) and (l), (23) could be further divided into two cases:

- When $q_{n+1} \geq d^n$, the Low type SU will always reject α_l. Thus, we have

$$q_n = \frac{q_{n+1} \times 1}{q_{n+1} + (1 - q_{n+1}) \times 1} = q_{n+1}. \quad (24)$$

- When $q_{n+1} < d^n$, the Low type SU will reject α_l with probability $\frac{(1-d^n)q_{n+1}}{d^n(1-q_{n+1})}$. We have

$$q_n = \frac{q_{n+1}}{q_{n+1} + (1 - q_{n+1}) \times \left(\frac{(1-d^n)q_{n+1}}{d^n(1-q_{n+1})}\right)} = d^n. \quad (25)$$

From (24) and (25), we have $q_n = \max(d^n, q_{n+1}) = \max(q_n^*, q_{n+1})$.

Now we consider (g) and focus on the *randomization (mixed)* strategy. When $q_n = d^n = q_n^*$, $\mathbb{P}_n$ is indifferent between offering α_l and α_h. In general, all mixed strategies on choosing α_l and α_h might be applicable. However, Requirement 3 will lead to a specific mixed strategy.

Reconsider $\mathbb{P}_n$'s utility in (20), denoted as $U_{\mathbb{P}_n}(q_n^*)$, is a constant R_h if the *balancing strategy* y_n^* applies. It is also subject to q_n^*, and independent of whether $\mathbb{P}_n$ chooses α_l or α_h, or randomizes on both alternatives. This follows from substituting q_n^* and y_n^* into the condition (20). Thus we can get

$$U_{\mathbb{P}_n}(x_n,\ y_n^*/q_n^*) = R_h, \text{ for all possible } x_n \text{ in } X_n, \quad (26)$$

where X_n is the set of all possible $\mathbb{P}_n$'s strategies, including the mixed strategies. The mixed strategy pair $(x_n, 1 - x_n)$ means choosing α_h with probability x_n and α_l with probability $1-x_n$. From (26), we can observe that all x_n in X_n are the *best responses* to y_n^* since R_n is a constant. However, when $q_n = q_n^*$, the Low type SU will only be expected to choose *the balancing strategy* $y_n^* = y_n(q_n^*)$. Since y_n^* is derived based on the sequential equilibrium requirements, thus as an equilibrium strategy, it should maximize the Low type SU's total utility. If we assume $(x^*, 1-x^*)$ from the set of X_n as $\mathbb{P}_n$'s *optimal* mixed strategy, then it implies that y_n^* is the *best response* to x^*. In other words, it indicates that the strategy pair (x^*, y_n^*) composes a *Nash equilibrium* in time slot n.

To get x^*, we *conjecture* that there exists a strategy profile $\mathcal{X}$ for the PU in each time slot such that its actual belief is equal to the limiting belief,

$$\mathcal{X} = (x_N, ..., x_n, ..., x_1)\,,\ x_N = \cdots = x_1 = x^* > 0, \quad (27)$$

where x_n is the PU's mixed strategy when $q_n = q_n^*$, *which makes the Low type SU is indifferent between rejecting and accepting* α_l. Since accepting α_l will lead to the utility $\Delta R_{sl} > 0$, $\mathcal{X}$ in (27) implies the Low type SU's *total* utility is $n\Delta R_{sl}$, irrespective of what strategy $\tilde{y}_n$ (including mixed strategy) the Low type SU will choose.[14] Therefore, any strategy $\tilde{y}_n$ can maximize the Low type SU's utility, including the equilibrium strategy y_n^*. Thus, it means that y_n^* is the best response to x^*. If we can verify that the strategy in (27) exists and obtain the value of x^*, then we finish the proof of (g). In fact, we can get $x^* = \frac{\Delta R_{sl}}{R_{sl} - \Delta R_{sl}}$. The detailed calculation of x^* can be found in [27]. We only sketch it here: x^* can be solved in a backward induction way by establishing the balancing condition for the Low type SU, *i.e.*, we first analyze the last time slot ($n = 1$), and then the next to last time slot ($n = 2$) and so forth ($n \geq 3$). The key point in each case is that $\mathcal{X}$ in (27) makes the Low type SU indifferent between rejecting and accepting α_l.

Thus, we have completed the proof. ■

REFERENCES

[1] Y. Yan, J. Huang, X. Zhong, and J. Wang, "Dynamic spectrum negotiation with asymmetric information," in *Proc. Int. Conf. Game Theory Networks (GameNets)*, Apr. 2011.

[2] Y. Yan, J. Huang, and J. Wang, "Dynamic Bayesian spectrum bargaining with non-myopic users," in *Proc. Int. ICST Conf. Wireless Internet (WiCON)*, Oct. 2011.

[3] Y. Yan, J. Huang, X. Zhong, M. Zhao, and J. Wang, "Sequential bargaining in cooperative spectrum sharing: Incomplete information with reputation effect," in *Proc. IEEE GLOBECOM*, Nov. 2011.

[4] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outrage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.

[6] S. A. Astaneh and S. Gazor, "Resource allocation and relay selection for collaborative communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6126–6133, Dec. 2009.

[7] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009.

[8] P. A. Parker, P. Mitran, D. W. Bliss, and V. Tarokh, "On bounds and algorithms for frequency synchronization for collaborative communication systems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3742–3752, Aug. 2008.

[9] J. Huang, Z. Han, M. Chiang, and H. V. Poor, "Auction-based resource allocation for cooperative communications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1226–1237, Sep. 2008.

[10] W. Saad, Z. Han, T. Başar, M. Debbah, and A. Hjørungnes, "Network formation games among relay stations in next generation wireless networks," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2528–2542, Sep. 2011.

[11] V. Genc, S. Murphy, Y. Yang, and J. Murphy, "IEEE 802.16J relay-based wireless access networks: An overview," *IEEE Wireless Commun. Mag.*, vol. 15, pp. 56–63, Oct. 2008.

[12] D. Niyato, E. Hossain, D. I. Kim, and Z. Han, "Relay-centric radio resource management and network planning in IEEE 802.16J mobile multihop relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6115–6125, Dec. 2009.

[13] D. Niyato and E. Hossain, "Integration of IEEE 802.11 WLANs with IEEE 802.16-based multihop infrastructure mesh/relay networks: A game-theoretic approach to radio resource management," *IEEE Netw. Mag.*, vol. 21, no. 3, pp. 6–14, May-June 2007.

[14] Y. Han, A. Pandharipande, and S. H. Ting,"Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4945–4950, Oct. 2009.

[15] Y. Han, S. H. Ting, and A. Pandharipande, "Cooperative spectrum sharing protocol with secondary user selection," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2914–2923, Sep. 2010.

[14]Note that $\tilde{y}_n$ is different from y_n in (l) of Theorem 6. y_n is derived from the Bayes' rule and thus is an *equilibrium strategy*. However, $\tilde{y}_n$ is not necessarily an equilibrium strategy, but definitely includes y_n (or y_n^* when $q_n = q_n^*$).

$$q_n = \frac{q_{n+1} \times 1}{q_{n+1} + (1 - q_{n+1}) \times \text{Prob}(\text{SU rejects } \alpha_l \mid \text{Low type SU})} \tag{23}$$

[16] Q. Li, S. H. Ting, A. Pandharipande, and Y. Han, "Cognitive spectrum sharing with two-way relaying systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1233–1240, Mar. 2011.

[17] Y. Han, A. Pandharipande, and S. H. Ting, "Cooperative spectrum sharing via controlled amplify-and-forward relaying," in *Proc. IEEE PIMRC*, Sep. 2008.

[18] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[19] J. Zhang and Q. Zhang, "Stackelberg game for utility-based cooperative cognitive radio networks," in *Proc. MobiHoc*, May 2009, pp. 23–32.

[20] H. Wang, L. Gao, X. Gan, X. Wang, and E. Hossain, "Cooperative spectrum sharing in cognitive radio networks: A game-theoretic approach," in *Proc. IEEE ICC*, 2010.

[21] L. Duan, L. Gao, and J. Huang, "Contract-based cooperative spectrum sharing," in *Proc. IEEE DySPAN*, 2011.

[22] L. Gao, L. Duan, and J. Huang, "Matching-based cooperative spectrum sharing," working paper [Onlne]. Available: http://jianwei.ie.cuhk.edu.hk/publication/MatchingCSS.pdf

[23] D. M. Kreps and R. Wilson, "Sequential equilibria," *Econometrica*, vol. 50, pp. 863–894, 1982.

[24] R. Gibbons, *Game Theory for Applied Economists*. Princeton University Press, 1992.

[25] C. Antoni, "Bargaining power in communication networks," *Mathematical Social Sciences*, vol. 41, pp. 69–87, 2001.

[26] C. Antoni, "On bargaining partner selection when communication is restricted," *Int. J. Game Theory*, vol. 30, no. 4, pp. 503–515, 2001.

[27] Y. Yan, J. Huang, and J. Wang, "Dynamic bargaining for relay-based cooperative spectrum sharing," Tech. Report [Online]. Available: http://jianwei.ie.cuhk.edu.hk/publication/BargainReputationTechReport.pdf.

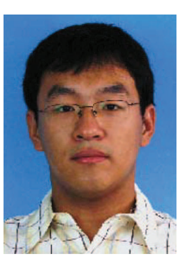
[28] P. Milgrom and J. Roberts, "Predation, reputation, and entry deterrence," *J. Economic Theory*, vol. 27, pp. 280–312, 1982.

[29] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: The MIT Press, 1991.

[30] D. Fudenberg and D. Levine, "Maintaining a reputation when strategies are imperfectly observed," *Review Economic Studies*, vol. 59, no. 3, pp. 561–579, 1992.

[31] D. M. Kreps and R. Wilson, "Reputation and imperfect information," *J. Economic Theory*, vol. 27, pp. 253–279, 1982.

[32] Manfred J. Holler, "The Kreps-Wilson monopoly-entrant game and cautiously rationalizable sequential equilibria," *Quality Quantity*, vol. 25, pp. 69–83, 1991.



Yang Yan received the B.Eng. degree in the Department of Electronics and Information Engineering, Harbin Institute of Technology (Harbin, Heilong Jiang, China) in 2007. He is working towards his Ph.D. degree in the Department of Electronic Engineering, Tsinghua University (Beijing, China). He is currently a Visiting Assistant in Research at Yale University (New Haven, CT, USA).

His research interests are in the areas of resource allocation and game theoretical analysis of wireless networks, with current emphasis on cooperative communications and cognitive radio networks. He is the recipient of the International ICST Conference on Wireless Internet (WiCON) Best Paper Award 2011.

Jianwei Huang (S'01-M'06-SM'11) is an Assistant Professor in the Department of Information Engineering at the Chinese University of Hong Kong. He received his B.S. in electrical engineering from Southeast University (Nanjing, Jiangsu, China) in 2000, and the M.S. and Ph.D. in electrical and computer engineering from Northwestern University (Evanston, IL, USA) in 2003 and 2005, respectively. He worked as a Postdoc Research Associate in the Department of Electrical Engineering at Princeton University during 2005–2007. He was a visiting scholar in the School of Computer and Communication Sciences at Ecole Polytechnique Federale De Lausanne (EPFL) during the Summer Research Institute in June 2009, and was a visiting scholar in the Department of Electrical Engineering and Computer Sciences at the University of California-Berkeley in August 2010. He is a Guest Professor of Nanjing University of Posts and Telecommunications.

Dr. Huang currently leads the Network Communications and Economics Lab (ncel.ie.cuhk.edu.hk), with the main research focus on nonlinear optimization and game theoretical analysis of communication networks, especially on network economics, cognitive radio networks, and smart grid. He is the recipient of the IEEE SmartGridComm Best Paper Award in 2012, the IEEE Marconi Prize Paper Award in Wireless Communications in 2011, the International Conference on Wireless Internet Best Paper Award in 2011, the IEEE GLOBECOM Best Paper Award in 2010, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2009, and the Asia-Pacific Conference on Communications Best Paper Award in 2009.

Dr. Huang has served as an Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Cognitive Radio Series, as Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, as Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS special issue on "Economics of Communication Networks and Systems," as Lead Guest Editor of the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS special issue on "Game Theory in Communication Systems." as Lead Guest Editor of *IEEE Communications Magazine's* Feature Topic on "Communications Network Economics," and as Guest Editor of several other journals including (Wiley) *Wireless Communications and Mobile Computing*, the *Journal of Advances in Multimedia*, and the *Journal of Communications*.

Dr. Huang serves as Chair of IEEE MMTC (Multimedia Communications Technical Committee) as the TPC Co-Chair of the IEEE GLOBEBOM Selected Areas in Communications Symposium (Game Theory for Communications Track) 2013, as the TPC Co-Chair of IEEE WiOpt (International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks) 2012, as the Publicity Co-Chair of the IEEE Communications Theory Workshop 2012, as the TPC Co-Chair of the IEEE ICCC Communication Theory and Security Symposium 2012, as the Student Activities Co-Chair of IEEE WiOpt 2011, as the TPC Co-Chair of the IEEE GlOBECOM Wireless Communications Symposium 2010, as the TPC Co-Chair of the IWCMC (International Wireless Communications and Mobile Computing) Mobile Computing Symposium 2010, and as the TPC Co-Chair of GameNets (International Conference on Game Theory for Networks) 2009. He is also a TPC member of leading conferences such as INFOCOM, MobiHoc, ICC, GLBOECOM, DySPAN, WiOpt, NetEcon, and WCNC.

Jing Wang received B.S. and M.S. degrees in the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 1983 and 1986, respectively. He has been on the faculty of Tsinghua University since 1986. He is currently a professor and the vice director of the Tsinghua National Laboratory for Information Science and Technology.

His research interests are in wireless communications, including channel modeling, multi-user detection, MIMO, and DWCS. He has published more than 100 conference and journal papers. He is a member of the Radio Communication Committee of the Chinese Institute of Communications and a senior member of the Chinese Institute of Electronics.