Dynamic Spectrum Sharing Auction with Time-Evolving Channel Qualities

Mehrdad Khaledi, *Student Member, IEEE,* Alhussein A. Abouzeid, *Senior Member, IEEE,*

***Abstract*—Spectrum auction is considered a suitable approach to efficiently allocate spectrum among unlicensed users. In a typical spectrum auction, Secondary Users (SUs) bid to buy spectrum bands from a Primary Owner (PO) who acts as the auctioneer. Existing spectrum auctions assume that SUs have static and known values for the channels. However, in many real world settings, the SUs do not know the exact value of channel access at first, but they learn it and adapt it over time. In this paper, we study spectrum auctions in a dynamic setting where SUs can change their valuations based on their experiences with the channel quality. We propose ADAPTIVE, a dynAmic inDex Auction for sPectrum sharing with TIme-evolving ValuEs that maximizes the social welfare of the SUs. ADAPTIVE is based on multi-armed bandit models where for each user an allocation index is independently calculated in polynomial time. Then we generalize ADAPTIVE to Multi-ADAPTIVE that auctions multiple channels at each time. We provide a sufficient condition under which Multi-ADAPTIVE achieves the maximum social welfare. Both ADAPTIVE and Multi-ADAPTIVE have some desired economic properties that are formally proven in the analysis. Also, we provide a numerical performance comparison between our proposed mechanisms and the well known static auctions, namely the Vickrey second price auction and the VCG mechanism.**

***Keywords*—Cognitive Radio Networks, Spectrum Sharing, Game Theory, Auction, Multi-armed Bandit.**

I. Introduction

WITH the ever-increasing demand for wireless communications, the wireless spectrum is becoming overcrowded. Currently, the Federal Communications Commission (FCC) allocates the spectrum for a long period of time through auctions among giant wireless operators. However, this static allocation results in inefficient use of the wireless spectrum. According to the measurements by the FCC's Spectrum Policy Task Force, most of the allocated spectrum is under-utilized [1]. Dynamic spectrum sharing has been proposed recently to improve the spectrum utilization [2]. Dynamic spectrum sharing enables new methods of cooperation and competition where a Primary Owner (PO) can re-allocate its idle spectrum bands to unlicensed or Secondary Users (SUs). Therefore, designing mechanisms that provide incentives for both PO and SUs is imperative.

We focus on auction-based mechanisms as they are very well-suited to the spectrum sharing problem, compared to the other possible market mechanisms. For instance, in pricing mechanisms the seller is assumed to have prior knowledge about the value of items to the potential buyers. However, in an auction the seller gets this information through bidding and the prior knowledge is not necessary. Also, auction mechanisms are more practicable compared to other market mechanisms (e.g. bargaining games [3]), since they incur less communication overhead.

In a simple spectrum auction, SUs bid to buy spectrum bands from a PO that sells its idle bands for a profit. An underlying assumption in existing spectrum auctions is that SUs know the exact value of channel access, and they bid accordingly (see section II for a review of prior work). However, in real world scenarios, the value of obtaining channel access is not exactly known to the SUs a priori, but they learn it over time. In fact, SUs revise their estimates of values for channel access based upon what they experience.

In this paper, we study spectrum auctions with dynamically evolving values. The setting allows SUs to learn their valuations based on their recent experiences of the channel quality. In this context, an SU's experience is estimated as a function of the channel quality or Signal to Noise Ratio (SNR) of the channel. We propose ADAPTIVE, a dynAmic inDex Auction for sPectrum sharing with TIme-evolving ValuEs. To the best of our knowledge, ADAPTIVE is the first spectrum auction that considers dynamically evolving values. ADAPTIVE is technically a repeated auction of a channel in which SUs learn their values over time. The proposed auction results in efficient allocation that maximizes the expected discounted social welfare.

The challenge presented by dynamic valuations is that the allocation needs to consider evolution of values that requires taking into account the consequences of current allocation by looking at future values. With non-deterministic evolution of values, the channel allocation is a stochastic dynamic programming problem. However, dynamic programming, which utilizes standard techniques such as backwards induction, is computationally intensive. The novelty of this work is that, by integrating multi-armed bandit models [4] in our auction models, we develop index-type allocation policies in our proposed auction where dynamic allocation indices are computed via forward induction in polynomial time.

Every auction is determined by a pair of functions (or rules); the allocation function and the payment function. We cast the allocation part of ADAPTIVE into an infinite horizon

This material is based upon work supported by the National Science Foundation under grant numbers 1422153 and 1456887, and a Tekes FiDiPro Fellow award.

M. Khaledi is with the Department of Electrical and Computer Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180 USA (e-mail: khalem@rpi.edu).

A. A. Abouzeid is with the Department of Electrical and Computer Engineering, Rensselaer Polytechnic Institute, Troy NY 12180 USA, and also has a visiting appointment with the Department of Communications Engineering, University of Oulu, Oulu 90014, Finland (e-mail: abouzeid@ecse.rpi.edu).

multi-armed bandit problem [4]. The main idea is to allocate the channel based on dynamic allocation indices of SUs that are computed independent of each other. This allocation rule results in an efficient allocation that maximizes the expected discounted social welfare. To obtain the payment function, we take into account the externality that the winning SU imposes on the other SUs, which is the surplus that other SUs could have achieved in the absence of the winning SU.

We then generalize ADAPTIVE to Multi-ADAPTIVE where multiple channels are auctioned at each time. The channel allocation in Multi-ADAPTIVE is inspired by multi-armed bandit models with multiple plays [4]–[6]. Up to our knowledge, there is no optimal index policy for the general multi-armed bandit problem with multiple plays. Thus, we specify sufficient conditions under which the generalized index policy of ADAPTIVE yields the maximum discounted social welfare. We adopt the model in [7] that determines conditions for optimality of index policies in multi-armed bandits with multiple plays. Then, we derive payments and allocation indices for the Multi-ADAPTIVE auction.

Both ADAPTIVE and Multi-ADAPTIVE run in polynomial time and have desired economic properties (namely, periodic ex post incentive compatibility, periodic ex post individual rationality and no positive transfers) that are formally proven in the analysis. Furthermore, we provide numerical results to show the effect of dynamic values. Our proposed auctions are compared with well-known static auctions (the second price auction and the VCG mechanism [8]) in terms of the PO's revenue, social welfare, average payments and average utilities of SUs.

The main contributions of this paper can be summarized as follows. We consider a dynamic spectrum auction setting where SUs can use their experiences with the channel quality to revise their valuations. This model allows learning of valuations over time and it is more realistic compared to prior work. We propose an efficient auction mechanism called ADAPTIVE that maximizes the expected discounted social welfare. To the best of our knowledge, ADAPTIVE is the first spectrum auction that enables dynamically evolving valuations. We then propose Multi-ADAPTIVE that extends ADAPTIVE to the case of auctioning multiple channels at each time. We provide sufficient conditions under which Multi-ADAPTIVE achieves the maximum social welfare. The payments and index-based allocation rules are determined in a way to ensure certain economic properties. We formally prove the economic properties of both ADAPTIVE and Multi-ADAPTIVE in the analysis. Furthermore, we provide a numerical performance comparison between our proposed auctions and well known static auctions.

The remainder of the paper is organized as follows. In Section II, we provide a brief review and discussion of related work. Section III describes the system model that is the basis of our proposed mechanism. In Section IV, we propose the ADAPTIVE auction and prove its economic properties. Section V presents the Multi-ADAPTIVE auction and its economic properties. Numerical results are reported and discussed in Section VI. Finally, Section VII concludes the paper and outlines possible avenues for future work.

II. Related Work

Several auction mechanisms have been proposed recently for wireless spectrum sharing in various settings [9]–[27]. In this section, we briefly overview the most relevant studies.

Heterogeneous spectrum sharing methods have been studied recently. In [9], the authors present a pricing method in which the PO offers channels of different qualities. In the model, the PO has prior knowledge about SUs' values for accessing the channels. Thus, the PO acts monopolistically and determines channel prices such that its revenue is maximized. A truthful auction mechanism for sharing variable bandwidth spectra is presented in [27]. The key assumption is that SUs bid their valuation functions to the PO so that the PO can evaluate SUs' values for any bandwidth. The authors in [21], propose an efficient heterogenous spectrum sharing auction in which the SUs can submit channel-specific bids depending on channel characteristics. The model is extended to a reserve price auction in [24] where the PO imposes reserve prices on the available channels.

Spectrum sharing in presence of multiple POs is studied in [10]. In the model, POs compete with each other by gradually raising the prices. The authors show that the algorithm converges to an equilibrium point where no SU and PO deviates. In a similar competitive environment, Niyato et al. [13] utilize the noncooperative game theory to model the dynamics of spectrum pricing. The authors in [11], study the competition between two POs that offer channels on different frequency bands. They show that the equilibrium price and its uniqueness is dependant on the SUs' geographical density and the spectrum propagation characteristics.

Spectrum double auctions provide a framework in which the POs can request their asking prices and SUs can submit their bids. Zhou et al. [15], proposed a double auction framework that enables spectrum reuse. The framework takes any reusability-driven spectrum allocation method, and implements its own winner determination and payment scheme. Another truthful double auction, called TAHES, is proposed in [14] that considers heterogeneous spectrum bands. The main assumption in double auctions is that, there must be an external third party with complete information to run the auction.

Spectrum auctions have also been studied in dynamic settings. For instance, [18], [19] present online spectrum auctions that allow SUs to join and leave the auction at different times. The authors in [25] present a repeated second price auction in which SUs can choose to enter the auction or stay out and monitor the results. Learning algorithms are utilized by SUs to optimize their decisions. Similarly in [12], a sequential second price auction is utilized for power and bandwidth allocation where one resource unit is auctioned at each time step. Thus, a static auction is performed repeatedly and auction results are observed by SUs. In this paper, however, we propose a dynamic auction that allocates dynamically with the objective of social welfare maximization.

Despite all the prior work, the problem of designing a spectrum auction with dynamically evolving values for SUs has not been addressed. The model used in this paper is inspired by [28] where the authors present a revenue maximizing

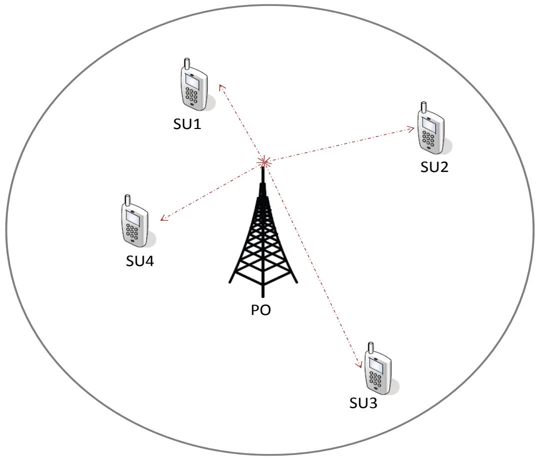

Fig. 1. A cognitive radio network with one primary owner and four secondary users.

dynamic auction. We specialize their model to the context of spectrum sharing in cognitive radio networks, and we present a social welfare maximizing dynamic auction with less restrictive assumptions. We also generalize the model to the case of auctioning multiple items and extend our preliminary work [26] on single channel dynamic spectrum sharing.

III. System Model

Here, we study the problem of auctioning a single channel where SUs' valuations dynamically evolve over time based on their experiences. In section V we extend the model to the case of auctioning multiple channels. We consider a cognitive radio network with one PO (a base station or an access point) who is willing to auction its idle channel to the SUs. An example of cognitive radio network is depicted in Fig. 1.

The spectrum sharing process is modeled by an auction in which PO acts as the auctioneer, and SUs are the bidders. The objective is to maximize social welfare while satisfying desired economic properties (such as incentive compatibility and individual rationality). There are k SUs competing with each other at each time step for an infinite horizon. The type of SU i is denoted by θ_i which is a real number reflecting how much SU i values channel access. It also captures the urgency for channel access, the more urgent the channel access to SU i, the higher the monetary value θ_i. For instance, SUs can set their types based upon their application types. For delay sensitive multimedia applications they have a different urgency than delay tolerant services. It should be noted that the type of SUs can also be related to physical measures such as the SU's queue backlog.

SUs gain experience over time dealing with the channel. We denote SU i's experience at time t by $e_{i,t} \in \xi_i$ where ξ_i can be a potentially arbitrary set. It should be noted that SU's experience evolves only when it gets the channel, otherwise its experience does not change. In our model, we consider SU's experience as the channel quality (or SNR) of the channel. An SU's experience at the instants that it gets the channel evolves in a Markovian model (i.e. the process is semi-Markov), which is known to the PO. That is, the probability that the next experience is $e_{i,t+1}$ is $P(e_{i,t+1}|e_{i,t})$, only depends on the current experience.

The channel evolution model follows the widely used Finite State Markov Channel (FSMC) in [29], in which SNR values of the channel are mapped into a finite number of discrete values, and the channel evolution is based on Markov transition probabilities. While most current work (e.g. [30], [31]) assume that the *instantaneous* channel state information is known, our model only assumes knowledge of channel statistics. Form the practical perspective, obtaining such a probabilistic model of channels is a feasible task. Since the PO collects the information (via submitted bids) at each time step, after some steps of running the mechanism, the PO can reasonably build a probabilistic model of the channel state.

SU's valuation for the channel is a stationary function of its type and experience. Without loss of generality, we assume that all SUs use the same valuation function. It is worth noting that even though the function is shared, the parameters of the function are private.[1] This actually makes sense practically, since SUs value a channel based on its capacity. So, the valuation functions can be similar but with different parameters, as SUs have different experiences and monetary preferences. We define SU i's valuation for the channel at time t as:

$$v(\theta_i, e_{i,t}) = \theta_i \; B \log(1 + e_{i,t}) \tag{1}$$

where B is the channel bandwidth. The function v takes into account both the channel quality experienced by SUs and SU's monetary value that reflects the urgency for channel access. It is worth noting that the model presented in this paper works with any other stationary valuation function, and equation (1) is one example of such a function that we use for evaluations. The expected (next) value of SU i for the channel at time t equals to $\delta^{t-1} v(\theta_i, e_{i,t})$ where $0 < \delta < 1$ is the common discount factor.

We assume that bidders are rational, which is an inherent assumption in designing truthful auction mechanisms. That implies that bidders act solely with the purpose of maximizing their own utilities. Also, by using the Revelation Principle[2] [8], we focus on direct mechanisms where bidders are willing to reveal their private information to the auctioneer. If an SU gets the channel, its utility is the difference between its valuation for the channel and the price it has to pay. The objective of our auction is to maximize the social welfare which is the total utility of all the users, including the PO. As the PO's utility is the sum of the received payments and an SU's utility is the difference between its valuation and its payment, the social welfare is equal to sum of the winning bidders' valuations [8].

IV. The ADAPTIVE Mechanism

In this section, we propose ADAPTIVE, a dynAmic inDex Auction for sPectrum sharing with TIme-evolving ValuEs.

[1]This scheme should not be confused with common value auctions where in the latter case, the item is of the same value for all bidders but they do not know what this value is [8].

[2]The principle states that an outcome of any indirect mechanism can be obtained by a (truthful) direct mechanism.

ADAPTIVE is a repeated auction for an infinite horizon that allows SUs to dynamically learn and revise their values. The proposed auction has certain guaranteed economic properties that will be proven in this section.

At every time step, each SU i reports its bid, which is $(\theta_i, e_{i,t})$, to the PO who takes the role of the auctioneer and runs ADAPTIVE. The auction should determine two output functions; the allocation and the payment. Channel allocation is denoted by $Q \in \mathbb{Q}$, where $\mathbb{Q}$ is the set of all possible allocation rules. Q contains $q_{i,t} \in \{0, 1\}$ that determine which SU gets the channel at every time step. That is, $q_{i,t} = 1$ indicates that the SU i has obtained the right to access the channel at time t and $q_{i,t} = 0$ otherwise. Similarly, $p_{i,t}$ represents the payment of SU i at time t.

After the channel is allocated at each round, the winner gets the chance to update its experience with the channel for the next round. As mentioned earlier, the PO knows the Markov probability model for the evolution of experiences. Therefore, with SUs' reports, the PO can compute expected future values and make decisions accordingly.

The objective is to find an efficient allocation scheme that satisfies desired economic properties. An efficient allocation rule maximizes the winning bidders' valuations. We can formally define the expected future social welfare at time t as:

$$S(\theta, e_t) = \max_{Q\in\mathbb{Q}} \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{i} \delta^{t'-t}\, q_{i,t'}\, v(\theta_i, e_{i,t'})\middle|\theta, e_t\right] \quad (2)$$

where $0 < \delta < 1$ is the common discount factor, θ and e_t are vectors of SUs' types and experiences at time t, respectively. To achieve efficiency, we need to find an allocation scheme $Q \in \mathbb{Q}$ that maximizes the equation (2). In the next subsection, we present a method to achieve the efficient allocation.

A. Efficient Allocation Policy of ADAPTIVE

The channel allocation component of ADAPTIVE is essentially a sequential resource allocation problem over an infinite time horizon. The PO needs to take into account the evolution of values when making decisions. With non-deterministic evolution of values, the channel allocation is a stochastic dynamic programming problem that utilizes backward induction to find the optimal decision rule. Although backward induction is a powerful method, it suffers from high computational complexity. Inspired by multi-armed bandit models [4], we use index-type allocation policies where dynamic allocation indices can be computed via forward induction in polynomial time.

Multi-armed bandit problems refer to a class of sequential resource allocation problems that are concerned with the dilemma of making decisions that bring high current gains or making decisions that sacrifice immediate payoffs with the hope of better future rewards [4]. In a multi-armed bandit problem, there is an operator with a collection of independent single-armed bandits. At each time step, the operator chooses to operate exactly one of the machines. The chosen machine generates a reward and updates its state. All other machines retain their current states till the next time step. The objective of the operator is to maximize the sum of rewards.

In [32], Gittins and Jones presented an index policy for the operator to obtain an optimal solution (that yields maximum sum of rewards). They introduced a *dynamic allocation index* which can be computed independently for each bandit. At each time step, the operator needs to choose the highest index bandit to achieve the optimal solution. Dynamic allocation index (also called *Gittins index*) reduces the complexity of the problem exponentially, since instead of finding the solution of a multi-armed bandit problem, the operator is required to determine Gittins indices for some single-armed bandit problems.

The channel allocation problem in ADAPTIVE can be transformed into a classical multi-armed bandit problem. Each arm in the bandit model can be thought of as an SU and rewards generated by pulling arms resemble SUs' valuations. The operator chooses a machine in the multi-armed bandit model just like the PO chooses an SU to allocate the channel to it. State changes in the bandit model is similar to experience updates of the winning SU in the ADAPTIVE mechanism. With the transformation, we see that the optimal solution to the multi-armed bandit problem is equivalent to the efficient channel allocation in ADAPTIVE.

Therefore, we can use the Gittins index policy to solve the efficient allocation problem. According to this policy, the PO gives the channel to the SU with the highest index. The Gittins index of SU i at time t (conditioned on its current experience and type) is the maximum expected discounted value per unit of expected discounted time [32].

$$G(\theta_i, e_{i,t}) = \max_{\tau} \mathbb{E}\left[\frac{\sum_{t'=t}^{\tau} \delta^{t'-t}\, v(\theta_i, e_{i,t'})}{\sum_{t'=t}^{\tau} \delta^{t'-t}}\middle|\theta_i, e_{i,t}\right] \quad (3)$$

where the maximization is taken over all the stopping times τ. An important feature of the Gittins index policy is that the index of SU i can be computed independently and without any information about other SUs. Also, it is worth noting that, the index of SU i will not change if it does not get the channel. There are several polynomial time algorithms to find the indices. For instance, Sonin in [33] proposes an algorithm to solve equation (3) in $n^3 + O(n^2)$ operations.

In addition to the allocation policy, we need to specify the payment function of the mechanism, i.e. the price the winning SU has to pay. In the next subsection, we propose the payment rule of the ADAPTIVE.

B. The Payment Rule of ADAPTIVE

We propose the payment rule of the ADAPTIVE mechanism in this subsection and we discuss and prove its economic properties in the next subsection. We specify payments such that under the efficient allocation policy, each SU's utility coincides with its marginal contribution to the social welfare [28].

Let $m_{i,t}$ denote SU i's marginal contribution to the social welfare at time t:

$$m_{i,t} = S(\theta, e_t) - S_{-i}(\theta, e_t) - \delta\mathbb{E}\Big[S(\theta, e_{t+1}) - S_{-i}(\theta, e_{t+1})\Big] \tag{4}$$

where $S_{-i}(\theta, e_t)$ is the expected future social welfare without SU i. When SU i gets the channel at time t, we have:

$$S(\theta, e_t) = v(\theta_i, e_{i,t}) + \delta\mathbb{E}\Big[S(\theta, e_{t+1})\Big] \tag{5}$$

Also, SU's experience will not change without getting the channel, i.e. $S_{-i}(\theta, e_t) = S_{-i}(\theta, e_{t+1})$. Therefore, using (5), we can simplify (4) as:

$$m_{i,t} = v(\theta_i, e_{i,t}) - (1-\delta)\, S_{-i}(\theta, e_t) \tag{6}$$

Now, we want that an SU's immediate utility coincides with its marginal contribution. That is,

$$m_{i,t} = v(\theta_i, e_{i,t}) - p_{i,t}. \tag{7}$$

As a result of combining equations (6) and (7), the winning SU i at time t pays

$$p_{i,t} = (1-\delta)\, S_{-i}(\theta, e_t). \tag{8}$$

If SU i does not get the channel at time t, then $p_{i,t} = 0$. Also, It should be noted that SU i has no control over its payment and its valuation is excluded in (8). This property prevents SUs from manipulating their payments to gain some profit. We will discuss and the prove economic properties of ADAPTIVE in the next subsection.

C. Economic Properties

An auction is required to satisfy certain economic properties such as incentive compatibility and individual rationality. In this subsection, we define these properties and prove that the ADAPTIVE mechanism satisfies them.

An auction is called *ex post incentive compatible* if truth-telling is always the best strategy for bidders, regardless of the history and current state (i.e. type and experience) of other bidders [28]. We should note here that SUs observe their history, which includes their past states, reports (bids) and allocations. Also, it is worth noting that, in dynamic settings, we have the notion of *periodic* ex post incentive compatibility. That is, the mechanism is ex post incentive compatible with respect to the information received in time t, but it is not ex post with respect to the information arriving after time t. In other words, a bidder may get some information in the future that it would regret its report at time t.

Before we formally define the properties, we need a few definitions. A *reporting strategy* for bidder i, denoted by R_i, provides a mapping from its state (type and experience) to a report. We denote the (joint) truth-telling strategy by T in which all the bidders report truthfully. The expected future utility of bidder i at time t under (joint) reporting strategy R is defined as:

$$U_{i,t}^{R} = \mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}\,(q_{i,t'}^{R}\; v(\theta_i, e_{i,t'}) - p_{i,t'})\right] \tag{9}$$

where $q_{i,t}^{R}$ is the allocation induced by R. Now, we can define the economic properties.

- **Periodic Ex Post Incentive Compatibility**; An auction is periodic ex post incentive compatible if for every bidder i and at any time t, truth-telling is the best response to the truthfulness of the other bidders. That is $U_{i,t}^{T} \geq U_{i,t}^{R_i,T_{-i}}$, where $U_{i,t}^{R_i,T_{-i}}$ is the utility of bidder i when i uses an arbitrary report strategy R_i, while all the other bidders use the truth-telling strategy T.
- **Periodic Ex Post Individual Rationality**; An auction is periodic ex post individually rational if for every bidder i and at any time t, we have $U_{i,t}^{T} \geq 0$. That means, bidders do not suffer as a result of participating in the auction.

Theorem 1: The ADAPTIVE mechanism is periodic ex post incentive compatible and periodic ex post individually rational.

Proof: Let $S^{R_i}(\theta, e_t)$ denote the expected future social welfare at time t when the bidder i uses the reporting strategy R_i and all other bidders report truthfully, defined as:

$$S^{R_i}(\theta, e_t) = \max_{Q\in\mathbb{Q}} \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{j} \delta^{t'-t}\; q_{j,t'}^{R_i}\; v(\theta_j, e_{j,t'})\Big|\theta, e_t\right]$$

where $\mathbb{Q}$ is the set of all possible allocation rules and $q_{j,t'}^{R_i}$ is the allocation induced by (R_i, T_{-i}) (when i uses the reporting strategy R_i and others report truthfully). The expected future social welfare at time t and without bidder i, can be defined similarly:

$$S_{-i}(\theta, e_t) = \max_{Q\in\mathbb{Q}_{-i}} \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{j\neq i} \delta^{t'-t}\; q_{j,t'}\; v(\theta_j, e_{j,t'})\Big|\theta, e_t\right]$$

where $\mathbb{Q}_{-i}$ is the set of allocation rules that disregard bidder i. Now, the marginal contribution of bidder i to the social welfare, at time t, will be:

$$\begin{aligned} m_{i,t} = {} & S^{R_i}(\theta, e_t) - S_{-i}(\theta, e_t) \\ & - \delta\mathbb{E}\Big[S^{R_i}(\theta, e_{t+1}) - S_{-i}(\theta, e_{t+1})\Big] \end{aligned} \tag{10}$$

If bidder i does not get the channel at time t, $m_{i,t} = p_{i,t} = 0$. However, if bidder i gets the channel at time t, then:

$$S^{R_i}(\theta, e_t) = v(\theta_i, e_{i,t}) + \delta\mathbb{E}\Big[S^{R_i}(\theta, e_{t+1})\Big]$$

Also, it should be noted that if bidder i gets the channel at time t, other bidders' state will not change (their experiences with the channel remain the same). That is:

$$S_{-i}(\theta, e_t) = S_{-i}(\theta, e_{t+1})$$

Therefore, the marginal contribution of bidder i (who gets the channel at time t) can be rewritten from (10) as

$$\begin{aligned} m_{i,t} &= v(\theta_i, e_{i,t}) - (1-\delta)\, S_{-i}(\theta, e_t) \\ &= v(\theta_i, e_{i,t}) - p_{i,t} \end{aligned} \tag{11}$$

where the second equality uses the payment rule, equation (8). The expected future utility of bidder i at time t when i uses reporting strategy R_i and others report truthfully using T_{-i} is defined as (for simplicity of notation we omit T_{-i}):

$$\begin{aligned} U_{i,t}^{R_i} &= \mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}\,(q_{i,t'}^{R_i}\, v(\theta_i, e_{i,t'}) - p_{i,t'})\right] \\ &= \mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}\, m_{i,t'}\right] \\ &= S^{R_i}(\theta, e_t) - S_{-i}(\theta, e_t) \end{aligned}$$

where the second equality follows from equation (11) and the third equality holds from definition of marginal contribution, equation (10), and noting that all the terms except for time t will cancel out. Clearly, S_{-i} is independent of bidder i's reports. Also, since the social welfare is defined with respect to the true states, S^{R_i} is maximized if bidder i reports truthfully. Therefore, the auction is periodic ex post incentive compatible. We can also see that $U_{i,t}^{T} \geq 0$ that is because $S^T \geq S_{-i}$. As a result, the ADAPTIVE mechanism is also periodic ex post individually rational. ■

It is worth noting that the ADAPTIVE mechanism has no positive transfers (another economic property), i.e. for all times t and every bidder i we have $p_{i,t} \geq 0$. This can be easily seen from the payment function, equation (8) which is sum of non-negative numbers multiplied by a factor $(1-\delta)$ in [0,1].

V. The Multi-ADAPTIVE Mechanism

The ADAPTIVE mechanism in section IV auctions a single channel at each time. In this section, we develop the Multi-ADAPTIVE mechanism that extends ADAPTIVE to the case of auctioning multiple channels. We determine the allocation policy, derive payments for winning SUs and prove the economic properties of Multi-ADAPTIVE.

The system model is similar to that of ADAPTIVE, but now the PO auctions multiple channels. We consider a frequency division multiplexing (FDM) scheme where the total available bandwidth is divided into m channels that are indexed by $j = 1, \cdots, m$. The channels can be of different qualities (heterogeneous channels) or identical. With heterogeneous channels, the valuation function (equation (1)) may be modified to include channel qualities. For instance, in the heterogeneous spectrum sharing models in [21], [24], the valuation function takes the maximum allowable transmission power as the measure for quality of channels. The model in this paper works with both identical and heterogeneous cases.

We assume that each channel can only be utilized by one SU at a time. Also, each SU can only use one channel at a time. SU i's valuation for channel j at time t is denoted by $v(\theta_i, e_{i,j,t})$. SU i's experience with channel j at time t is denoted by $e_{i,j,t}$. Similar to ADAPTIVE, SU's experience with a channel evolves only when it gets that channel, but now we have multiple Markovian models, one for each channel. That is, the probability that the next experience with channel j is $e_{i,j,t+1}$ is $P_j(e_{i,j,t+1}|e_{i,j,t})$, only depends on the current experience.

At each time step, SUs report $(\theta_i, e_{i,t})$ to the PO, where $e_{i,t}$ is SU i's vector of experiences containing $e_{i,j,t}$s. The PO runs the Multi-ADAPTIVE auction and determines the allocation rule and the payments. Channel allocation Q contains $q_{i,j,t} \in \{0,1\}$ that determine which SU gets which channel at every time step. The payment of SU i at time t is denoted by $p_{i,t}$. The objective is to find an efficient allocation scheme that satisfies desired economic properties. We can formally define the expected future social welfare at time t as:

$$S(\theta, e_t) = \max_{Q \in \mathbb{Q}} \mathbb{E}\left[\sum_{t'=t}^{\infty} \sum_{i} \sum_{j} \delta^{t'-t}\, q_{i,j,t'}\, v(\theta_i, e_{i,j,t'}) \Big| \theta, e_t\right] \tag{12}$$

where $0 < \delta < 1$ is the common discount factor, θ is the vector of SUs' types and e_t is the matrix of SUs' experiences over channels at time t. To achieve efficiency, we need to find an allocation scheme $Q \in \mathbb{Q}$ that maximizes the equation (12).

A. Efficient Allocation Policy of Multi-ADAPTIVE

The channel allocation in Multi-ADAPTIVE is inspired by multi-armed bandit models with multiple plays [4], [5]. In a multi-armed bandit model with multiple plays, there is an operator with a collection of independent single-armed bandits. At each time step, the operator operates m machines. The chosen machines generate rewards and update their states while other machines remain frozen till the next time step. The objective of the operator is to maximize the sum of rewards. Similar to ADAPTIVE, the channel allocation problem in Multi-ADAPTIVE can be transformed into a multi-armed bandit problem with multiple plays. Here, the PO chooses m SUs at each time step and they get a chance to update their experiences with the assigned channels.

Learning-based algorithms have been proposed in the literature that aim to find the minimum regret policy for the multi-armed bandit problem with multiple plays (see [5], [6]). However, up to our knowledge, there is no optimal index policy for this problem. One possible solution is to find the dynamic allocation indices (equation 3) and choose the SUs that have the m highest indices. But, this policy is not optimal in general. Pandelis and Teneketzis [7] specified a sufficient condition that guarantees the optimality of the policy that chooses the m highest indices at each time.

We first present the index-based allocation policy for Multi-ADAPTIVE, then we provide the sufficient condition for its optimality. Similar to ADAPTIVE, we define allocation indices (Gittins Indices) for SUs. However, in Multi-ADAPTIVE there are m Gittins indices for each SU, one for each channel. This is because SUs get different experiences with different channels. Thus, we define the Gittins index of SU i for channel j at time t as:

$$G(\theta_i, e_{i,j,t}) = \max_{\tau} \mathbb{E}\left[\frac{\sum_{t'=t}^{\tau} \delta^{t'-t}\, v(\theta_i, e_{i,j,t'})}{\sum_{t'=t}^{\tau} \delta^{t'-t}} \Big| \theta_i, e_{i,j,t}\right] \tag{13}$$

Since the Gittins index represents the maximum expected discounted value per unit of expected discounted time, we want to assign channels in a way to maximize the sum of Gittins indices. Thus, once the indices are computed, the PO's problem is to match channels with SUs to maximize the sum of Gittins indices. This problem can be cast into a maximum weight matching problem in a bipartite graph [34]. A *bipartite graph* is a graph whose vertices can be divided into two disjoint sets V_1 and V_2, such that every edge in the graph connects a vertex in V_1 to one in V_2. A graph can be weighted in which edges are associated with weights, usually a real number. In a bipartite graph, a *matching* is a subset of edges such that they do not share an endpoint. In other words, a matching is a subset of edges such that for each vertex, there is at most one edge in the matching that is incident upon this vertex.

The PO can build a bipartite graph $G(V_1, V_2)$ at each time by letting V_1 be the set of SUs and V_2 be the set of available channels. The weight of the edge ij (connecting SU i to channel j) represents the Gittins index of the SU i for the channel j. Therefore, the channel allocation problem in Multi-ADAPTIVE is a maximum weight matching problem that can be solved by the Kuhn-Munkres algorithm (also known as the Hungarian algorithm) in $O(N^3)$, where $N = \max$(number of SUs, number of channels) [35].

Using the optimality conditions in [7], we describe a sufficient condition that guarantees the efficiency[3] of the above mentioned allocation policy.

C1 For any realization of the problem, for any SUs i, j such that $i \neq j$, any channel x and any p, q such that

$$G(\theta_i, e_{i,x,p}) > G(\theta_j, e_{j,x,q}),$$

we have

$$G(\theta_i, e_{i,x,p})(1-\delta) > G(\theta_j, e_{j,x,q}).$$

Condition (C1) provides a situation where there is enough separation among the Gittins indices. The results in [7] states that when the Gittins indices are sufficiently separated, forward induction can be used to find an optimal policy. Thus, the Gittins index rule is optimal.

It should be noted that in the above-mentioned allocation scheme, we consider the general case where channels are heterogeneous. If channels are identical, an SU's experiences with channels are the same, thus there is only one Gittins index for each SU. As a result, the channel allocation simplifies to choosing the m highest indices.

B. The Payment Rule of Multi-ADAPTIVE

In this subsection, we propose the payment rule of Multi-ADAPTIVE that determines the price the winning SUs have to pay. We specify payments such that under the efficient allocation policy, each SU's utility coincides with its marginal contribution to the social welfare.

[3] efficiency refers to maximizing the expected future social welfare, equation (12)

Recall from equation (4) that $m_{i,t}$ denotes SU i's marginal contribution to the social welfare at time t:

$$m_{i,t} = \Big(S(\theta, e_t) - \delta \mathbb{E}\big[S(\theta, e_{t+1})\big]\Big) - \Big(S_{-i}(\theta, e_t) - \delta \mathbb{E}\big[S_{-i}(\theta, e_{t+1})\big]\Big) \quad (14)$$

where $S_{-i}(\theta, e_t)$ is the expected future social welfare without SU i. We can write the expected future social welfare (equation (12)) recursively as:

$$S(\theta, e_t) = \sum_{\substack{k,j \\ q_{k,j,t} \in Q^*}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) + \delta \mathbb{E}\Big[S(\theta, e_{t+1})\Big] \quad (15)$$

Similarly for $S_{-i}(\theta, e_t)$ we have:

$$S_{-i}(\theta, e_t) = \sum_{\substack{k,j \\ k \neq i \\ q_{k,j,t} \in Q^*_{-i}}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) + \delta \mathbb{E}\Big[S_{-i}(\theta, e_{t+1})\Big] \quad (16)$$

where Q^* and Q^*_{-i} are the allocation rules that maximize the expected social welfare with and without SU i, respectively. Now, using equations (15) and (16), the marginal contribution (equation 14) can be simplified as:

$$m_{i,t} = \sum_{\substack{k,j \\ q_{k,j,t} \in Q^*}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) - \sum_{\substack{k,j \\ k \neq i \\ q_{k,j,t} \in Q^*_{-i}}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) \quad (17)$$

Now, we want that SU's immediate utility coincide with its marginal contribution. That is:

$$m_{i,t} = \sum_{\substack{j \\ q_{i,j,t} \in Q^*}} q_{i,j,t}\ v(\theta_i, e_{i,j,t}) - p_{i,t} \quad (18)$$

As a result of combining equations (17) and (18), the winning SU i at time t pays

$$p_{i,t} = \sum_{\substack{k,j \\ k \neq i \\ q_{k,j,t} \in Q^*_{-i}}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) - \sum_{\substack{k,j \\ k \neq i \\ q_{k,j,t} \in Q^*}} q_{k,j,t}\ v(\theta_k, e_{k,j,t}) \quad (19)$$

It is worth noting that if SU i does not get any channel at time t, then $p_{i,t} = 0$. Also, it should be noted that SU i cannot manipulate its payment, since its valuation is excluded in (19). This is essential towards achieving desired economic properties that we prove in the next subsection.

C. Economic Properties

Here we prove the economic properties of Multi-ADAPTIVE, namely periodic ex post incentive compatibility and periodic ex post individual rationality that we defined them in section IV-C.

Theorem 2: The Multi-ADAPTIVE mechanism is periodic ex post incentive compatible and periodic ex post individually rational.

Proof: The proof is similar to the proof of Theorem 1. Let $S^{R_i}(\theta, e_t)$ denote the expected future social welfare at time t when bidder i uses the reporting strategy R_i and all other bidders report truthfully, defined as:

$$S^{R_i}(\theta, e_t) = \max_{Q\in\mathbb{Q}} \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{k,j}\delta^{t'-t}\, q^{R_i}_{k,j,t'}\, v(\theta_k, e_{k,j,t'})\Big|\theta, e_t\right]$$

where $\mathbb{Q}$ is the set of all possible allocation rules and $q^{R_i}_{k,j,t'}$ is the allocation induced by (R_i, T_{-i}) (when i uses the reporting strategy R_i and others report truthfully). The expected future social welfare without bidder i, can be defined similarly:

$$S_{-i}(\theta, e_t) = \max_{Q\in\mathbb{Q}_{-i}} \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{\substack{k,j\\k\neq i}}\delta^{t'-t}\, q_{k,j,t'}\, v(\theta_k, e_{k,j,t'})\Big|\theta, e_t\right]$$

where $\mathbb{Q}_{-i}$ is the set of allocation rules that disregard bidder i. Now, bidder i's marginal contribution to the social welfare, at time t, is:

$$\begin{aligned} m_{i,t} &= S^{R_i}(\theta, e_t) - S_{-i}(\theta, e_t) \\ &\quad - \delta\mathbb{E}\Big[S^{R_i}(\theta, e_{t+1}) - S_{-i}(\theta, e_{t+1})\Big] \end{aligned} \tag{20}$$

Using the recursive definitions of expected social welfare (i.e. equations (15) and (16)), $m_{i,t}$ can be simplified to:

$$m_{i,t} = \sum_{\substack{k,j\\q_{k,j,t}\in Q^*}} q_{k,j,t}\, v(\theta_k, e_{k,j,t}) - \sum_{\substack{k,j\\k\neq i\\q_{k,j,t}\in Q^*_{-i}}} q_{k,j,t}\, v(\theta_k, e_{k,j,t})$$

The above equation can be rewritten by using the payment formula (i.e. equation (19))

$$m_{i,t} = \sum_{\substack{j\\q_{i,j,t}\in Q^*}} q_{i,j,t}\, v(\theta_i, e_{i,j,t}) - p_{i,t} \tag{21}$$

The expected future utility of bidder i at time t when i uses the reporting strategy R_i and others report truthfully using T_{-i}, is defined as (for simplicity of notation we omit T_{-i}):

$$\begin{aligned} U^{R_i}_{i,t} &= \mathbb{E}\left[\sum_{t'=t}^{\infty}\sum_{j}\delta^{t'-t}\,(q^{R_i}_{i,j,t'}\, v(\theta_i, e_{i,j,t'}) - p_{i,t'})\right] \\ &= \mathbb{E}\left[\sum_{t'=t}^{\infty}\delta^{t'-t}\, m_{i,t'}\right] \\ &= S^{R_i}(\theta, e_t) - S_{-i}(\theta, e_t) \end{aligned}$$

where the second equality follows from equation (21) and the third equality holds from definition of marginal contribution, equation (20), and noting that all the terms except for time t will cancel out. Clearly, S_{-i} is independent of bidder i's reports. Thus, the bidder i wants to use a reporting strategy R_i that maximizes S^{R_i}. Since the social welfare is defined with respect to the true states, S^{R_i} is maximized if bidder i reports truthfully. Therefore, the auction is periodic ex post incentive compatible. We can also see that $U^T_{i,t} \geq 0$ that is because $S^T \geq S_{-i}$. As a result, the ADAPTIVE mechanism is also periodic ex post individually rational. ■

VI. Numerical Results

In this section, we provide numerical results to show the effect of dynamism on the performance of spectrum auctions. Results are presented in two subsections, for cases where a single channel or multiple channels are auctioned off, respectively.

A. Single channel

For the single channel case, we compare the performance of ADAPTIVE, the dynamic valuation auction, with the well-known Vickrey auction (also called second price auction) as the representative of static auctions.

We set the common discount factor, δ, to 0.7 and change the number of SUs from 3 to 21. Each setting is run 500 times in MATLAB to eliminate the effect of random initialization. Social welfare (sum of winning SUs' valuations), discounted social welfare, average payment of SUs, average utility of SUs, and revenue of the PO (sum of SUs' payments) are considered as performance metrics. In order to study the impact of the discount factor, we fix the number of SUs and compute the revenue of PO for different values of δ.

The bandwidth B is set to 1 MHz and SUs' initial experiences and types are randomly drawn from discrete uniform distributions. Thus, SUs' initial valuations can be computed. The ADAPTIVE mechanism computes the Gittins indices and chooses the maximum index SU as the winner and uses (8) for determining payments. We use Sonin's algorithm [33] to compute Gittins indices in polynomial time. The winner updates its experience and the mechanism repeats. On the other hand, the second price auction uses a myopic policy for allocation, that is the SU with the highest valuation gets the channel and pays the second highest value. The winner gets a chance to update its experience, and the second price auction continues.

In this setting, we define SUs' experiences as the Signal to Noise Ratio (SNR) of the channel they get. The winner updates its experience according to an Auto-Regressive (AR) model. If SU i is the winner, we have $e_{i,t+1} = e_{i,t} + z$, where z is a discrete random variable. We consider limited discrete values for SNR (ranging from -30db to 30db with increments of 1db) which provides a finite small sized state space. z follows a Binomial distribution with probability of 0.5 that is a good approximation of a Gaussian distributed random variable centered at 0. With this evolution model, we can easily build a Markov probability model for transitions between experiences, $P(e_{i,t+1}|e_{i,t})$.

In Fig. 2 the average social welfare is shown versus number of SUs. As can be seen, in both ADAPTIVE and second price auctions the social welfare increases with number of SUs. This is because when more SUs participate in the auction, there will be a wider range of valuations. Since SUs with high valuations are favored, it is more probable that the winner has a higher valuation compared to the case of having less SUs that consequently results in a higher social welfare. Fig. 2 also indicates that the ADAPTIVE mechanism results in a better social welfare than the second price auction. This happens due to the fact that in ADAPTIVE, we have a probabilistic model of

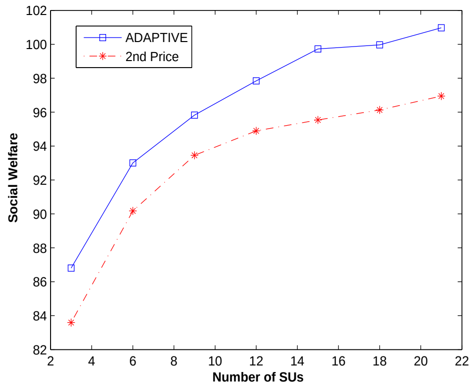

Fig. 2. Average social welfare versus the number of SUs.

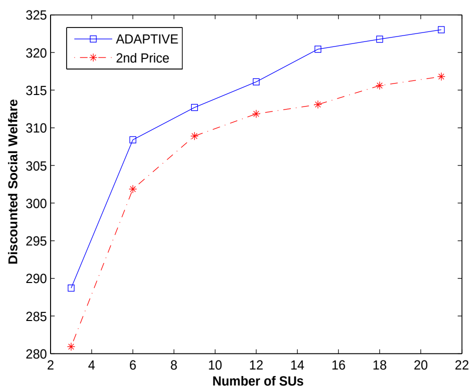

Fig. 3. Total discounted social welfare versus the number of SUs.

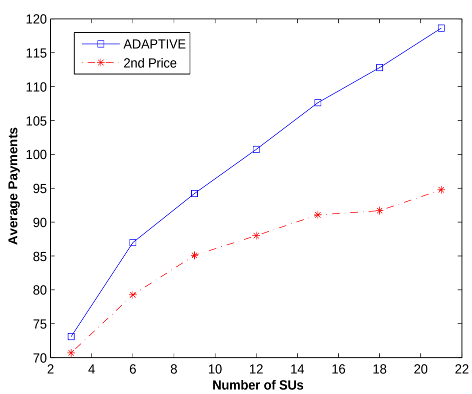

Fig. 4. Average payments versus the number of SUs.

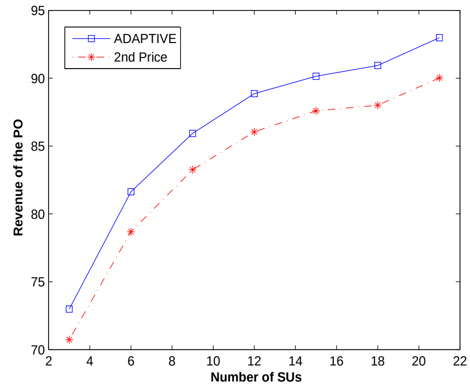

Fig. 5. Revenue of the PO versus the number of SUs.

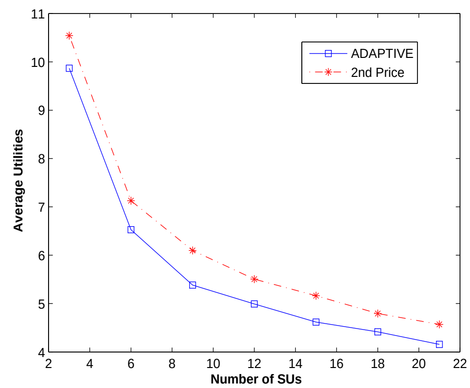

Fig. 6. Average utilities versus the number of SUs.

the future, because of the experience evolution model. Thus, the auctioneer takes into account the expected future values when allocating the channel. We observe a similar behavior in Fig. 3 that shows the total discounted social welfare of the two auctions versus number of SUs.

The average payments of SUs is depicted in Fig. 4. We observe that in both auctions as the number of SUs increases, payments increase. This is because with more SUs, channel access becomes more competitive. Thus, the winner causes more externality to the other SUs, and consequently he has to pay more. This figure also shows that the ADAPTIVE mechanism induces higher payments than the second price auction. In the payment function of the ADAPTIVE mechanism (equation 8), the PO takes into account the future expected values and in the eye of ADAPTIVE, the winning SU causes more externality than that of the second price auction (i.e. second highest valuation). Therefore, the payments in ADAPTIVE are higher than payments in second price auction.

High payments are favorable to the PO, as it leads to more revenue. In Fig. 5 the revenue of the PO versus number of SUs is shown. Since higher payments imply higher revenue, we observe a similar behavior in this figure as in Fig. 4. In both auctions the revenue of the PO increases with number of SUs. In addition, the proposed dynamic auction yields more revenue than the static second price.

From the SUs' point of view, however, increase in the number of competitors (other SUs) is not favorable. Fig. 6 shows the average utility of SUs versus the number of SUs. We see that in both mechanisms, average utilities decreases

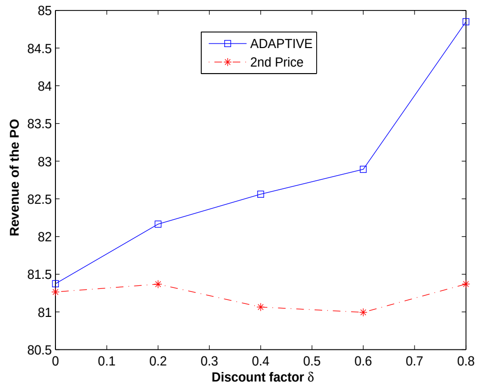

Fig. 7. Revenue of the PO versus δ, with 12 SUs.

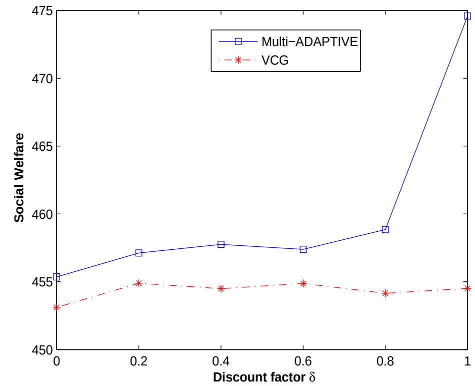

Fig. 8. Average social welfare versus δ, with 12 SUs and 6 channels.

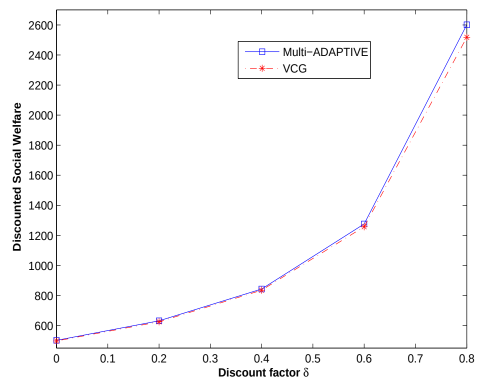

Fig. 9. Total discounted social welfare versus δ, with 12 SUs and 6 channels.

with the number of SUs. This is due to the increase in SU payments that consequently lowers utilities (since utility is the difference between valuation and payment).

In all the diagrams so far, we had a fixed discount factor, $\delta = 0.7$. Now, we fix the number of SUs at 12 and increase the discount factor. Fig. 7 shows the revenue of the PO versus the discount factor. The revenue from the second price auction remains almost constant, since it does not rely on the discount factor. However, we observe that the revenue from ADAPTIVE slightly increases with δ. As δ increases, the summation in the payment formula ($S_{-i}(\theta, e_t)$ in (8)) increases as well. This increase in summation leads to higher payments and higher revenue, even though the factor $(1-\delta)$ in the payment formula decreases. Fig. 7 implies that the more we weigh the future (as opposed to the current gains), the higher revenue we get from the ADAPTIVE mechanism.

B. Multiple channels

In this section we present numerical results to compare Multi-ADAPTIVE mechanism with the well-known VCG mechanism [8]. The setting and initial parameters are similar to the single channel case, except that we present the results for a fixed number of SUs, but variable discount factor (δ). This is because with variable SUs, comparing Multi-ADAPTIVE with VCG will result in the same conclusions as in the single channel case. Instead, we change the discount factor that may affect the efficiency of Multi-ADAPTIVE and show how the auction performs. Recall from Section V-A that the sufficient condition (C1) for efficiency of Multi-ADAPTIVE depends on δ.

In the following diagrams, there are 12 SUs competing for $m = 6$ channels. For simplicity we consider identical channels, thus, there is one allocation index for each SU. The Multi-ADAPTIVE mechanism computes the allocation indices (equation (13)) for each SU and chooses the m highest index SUs as the winners. Then, the winners pay the price according to the payment rule (19), update their experiences and the mechanism repeats. The VCG mechanism, however, takes the m highest valuation SUs as winners and charges them the $(m+1)$th valuation. We also allow the winning SUs to update their experiences for the next round.

Fig. 8 shows the average social welfare versus the discount factor (δ). As can be seen, the Multi-ADAPTIVE mechanism results in a better social welfare than the VCG auction. This is due to the fact that Multi-ADAPTIVE uses a look ahead approach for channel allocation (i.e. the future expected values are taken into account), which performs better than the myopic approach in the VCG auction. We also observe an increase in social welfare for Multi-ADAPTIVE when $\delta = 1$. When the discount factor is 1, we favor the future rewards as opposed the current gains, thus the dynamic allocation proceeds in a way that, over the long run, we get a larger social welfare.

Fig. 9 shows that the total discounted social welfare increases with the discount factor δ. Clearly, when δ increases, we have larger numbers being summed up in the total discounted social welfare. In order to clarify the performance of Multi-ADAPTIVE compared to VCG, we have depicted the difference between the total discounted social welfare of Multi-ADAPTIVE and that of VCG in Fig. 10. As can be seen, Multi-ADAPTIVE achieves a better discounted social welfare than VCG, and the difference goes even higher when δ increases.

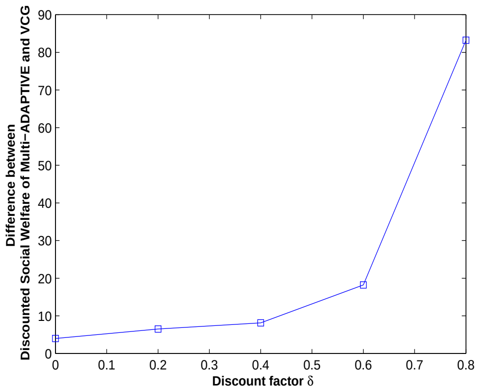

Fig. 10. Difference between the total discounted social welfare of Multi-ADAPTIVE and VCG versus δ, with 12 SUs and 6 channels.

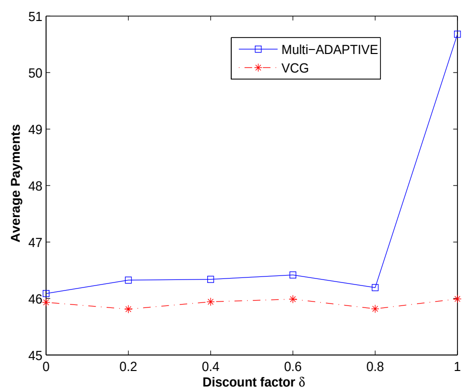

Fig. 11. Average payments versus δ, with 12 SUs and 6 channels.

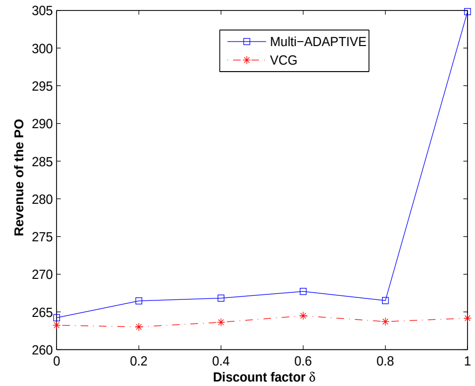

Fig. 12. Revenue of the PO versus δ, with 12 SUs and 6 channels.

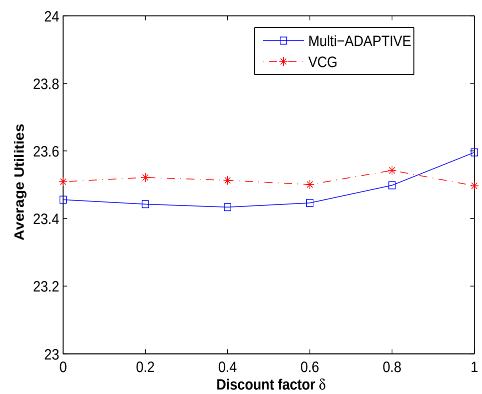

Fig. 13. Average utilities versus δ, with 12 SUs and 6 channels.

In these figures, we observe that Multi-ADAPTIVE performs better than VCG regardless of whether the sufficient condition (C1) for its efficiency holds. When $\delta = 0$ the condition (C1) holds and Multi-ADAPTIVE is efficient, but as we go from 0 to 1, there is less chance that condition (C1) holds. From figures 8 and 10, we see that even though Multi-ADAPTIVE may not achieve the efficient allocation, it outperforms the VCG auction.

In Fig. 11 the average payments of SUs is illustrated. This figure shows that Multi-ADAPTIVE induces higher payments than VCG. Recall that VCG ranks the SUs based on their valuations and the winners pay the $(m + 1)$th valuation. However, Multi-ADAPTIVE ranks the SUs based upon their dynamic allocation indices, and charges the winners according to (19) which equals the valuation of the SU with the $(m + 1)$th allocation index (when channels are identical). The dynamic allocation in Multi-ADAPTIVE, over the long run, leads to higher average valuations (as in Fig. 8), that consequently makes winning SUs cause more externality than that of VCG. Therefore, the average payments are higher in Multi-ADAPTIVE than VCG. For the same argument as in Fig. 8, we see an increase when $\delta = 1$.

The revenue of the PO is depicted in Fig. 12. Since higher payments imply higher revenue, we see that the Multi-ADAPTIVE mechanism leads to higher revenue for the PO than VCG. However, from the SUs' perspective, high payments are not favorable. Fig. 13 shows the average utilities of SUs. Since utility is the difference between valuation and payment, we see that VCG results in slightly better utilities for SUs than Multi-ADAPTIVE, except for $\delta = 1$ where the valuations are high enough to compensate the high payments of Multi-ADAPTIVE.

VII. Conclusion

In this paper, we study spectrum auctions in a realistic setting where SUs are allowed to revise their estimates of values for channel access, based upon what they experience over time. In this setting, we propose ADAPTIVE, a dynAmic inDex Auction for sPectrum sharing with TIme-evolving ValuEs. To the best of our knowledge, ADAPTIVE is the first spectrum auction that considers dynamically evolving values. Then we generalize ADAPTIVE to Multi-ADAPTIVE that auctions multiple channels at each time. We provide a sufficient condition under which Multi-ADAPTIVE achieves the maximum social welfare. Both ADAPTIVE and Multi-ADAPTIVE run in polynomial time and have desired economic properties that are formally proven in the analysis. Furthermore, we provide a numerical performance comparison between our proposed auctions and the well known static auctions (the second price auction and the VCG mechanism). In our model, we assumed that SUs' population is static, that is, SUs cannot leave or enter the auction at arbitrary rounds. A possible direction for future work is to extend the proposed model to a dynamic population model that allows a dynamic population of SUs with dynamic valuations. We also considered an infinite time horizon in our auction models. Another direction for future work is to consider a finite time horizon model which is suitable for modeling situations where the channel is available to the operator for a limited duration.

Acknowledgment

This material is based upon work supported by the National Science Foundation under grant numbers 1422153 and 1456887, and a Tekes FiDiPro Fellow award.

References

[1] FCC Spectrum Policy Task Force, *Report of the spectrum efficiency working group*, Available: http://www.fcc.gov/sptf/reports.html, Nov. 2002.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[3] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.

[4] A. Mahajan and D. Teneketzis, "Multi-armed bandit problems," in *Foundations and Applications of Sensor Management*. Springer US, 2008, pp. 121–151.

[5] V. Anantharam, P. Varaiya, and J. Walrand, "Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-part ii: Markovian rewards," *Automatic Control, IEEE Transactions on*, vol. 32, no. 11, pp. 977–982, Nov 1987.

[6] V. Anantharam, P. Varaiya, and J. Walrand, "Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-part i: I.i.d. rewards," *Automatic Control, IEEE Transactions on*, vol. 32, no. 11, pp. 968–976, Nov 1987.

[7] D. G. Pandelis and D. Teneketzis, "On the optimality of the gittins index rule for multi-armed bandits with multiple plays," *Computational Statistics*, vol. 50, no. 3, pp. 449–461, 1999. [Online]. Available: http://EconPapers.repec.org/RePEc:spr:compst:v:50:y:1999:i:3:p:449-461

[8] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. New York, NY, USA: Cambridge University Press, 2007.

[9] L. Gao, X. Wang, Y. Xu, and Q. Zhang, "Spectrum trading in cognitive radio networks: A contract-theoretic modeling approach," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 4, pp. 843–855, 2011.

[10] L. Gao, Y. Xu, and X. Wang, "MAP: Multiauctioneer progressive auction for dynamic spectrum access," *Mobile Computing, IEEE Transactions on*, vol. 10, no. 8, pp. 1144–1161, 2011.

[11] A. Min, X. Zhang, J. Choi, and K. Shin, "Exploiting spectrum heterogeneity in dynamic spectrum market," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 12, pp. 2020–2032, 2012.

[12] J. Bae, E. Beigman, R. Berry, M. Honig, and R. Vohra, "Sequential bandwidth and power auctions for distributed spectrum sharing," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 7, pp. 1193–1203, September 2008.

[13] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game-theoretic modeling approach," *Mobile Computing, IEEE Transactions on*, vol. 8, no. 8, pp. 1009–1022, 2009.

[14] X. Feng, Y. Chen, J. Zhang, Q. Zhang, and B. Li, "TAHES: a truthful double auction mechanism for heterogeneous spectrums," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 11, pp. 4038–4047, 2012.

[15] X. Zhou, A. Sala, and H. Zheng, "Towards large-scale economic-robust spectrum auctions," *15th Annual International Conference on Mobile Computing and Networking (MobiCom 2009)*, 2009.

[16] M. Dong, G. Sun, X. Wang, and Q. Zhang, "Combinatorial auction with time-frequency flexibility in cognitive radio networks," in *INFOCOM, 2012 Proceedings IEEE*, 2012, pp. 2282–2290.

[17] Y. Zhu, B. Li, and Z. Li, "Core-selecting combinatorial auction design for secondary spectrum markets," in *INFOCOM, 2013 Proceedings IEEE*, 2013.

[18] L. Deek, X. Zhou, K. Almeroth, and H. Zheng, "To preempt or not: Tackling bid and time-based cheating in online spectrum auctions," in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 2219–2227.

[19] Y. Yang, J. Wu, C. Long, and B. Li, "Online market clearing in dynamic spectrum auction," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, 2011, pp. 1–5.

[20] I. Kash, R. Murty, and D. Parkes, "Enabling spectrum sharing in secondary market auctions," *Mobile Computing, IEEE Transactions on*, 2013.

[21] M. Khaledi and A. Abouzeid, "Auction-based spectrum sharing in cognitive radio networks with heterogeneous channels," in *Information Theory and Applications Workshop (ITA), 2013*, 2013, pp. 1–8.

[22] Y. Zhang, C. Lee, D. Niyato, and P. Wang, "Auction approaches for resource allocation in wireless systems: A survey," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 3, pp. 1020–1041, Third 2013.

[23] Y. Zhang, D. Niyato, P. Wang, and E. Hossain, "Auction-based resource allocation in cognitive radio systems," *Communications Magazine, IEEE*, vol. 50, no. 11, pp. 108–120, November 2012.

[24] M. Khaledi and A. Abouzeid, "A reserve price auction for spectrum sharing with heterogeneous channels," in *Computer Communications and Networks (ICCCN), 2013 22nd International Conference on*, July 2013, pp. 1–7.

[25] Z. Han, R. Zheng, and H. Poor, "Repeated auctions with bayesian nonparametric learning for spectrum access in cognitive radio networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 3, pp. 890–900, March 2011.

[26] M. Khaledi and A. Abouzeid, "Adaptive: A dynamic index auction for spectrum sharing with time-evolving values," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on*, May 2014, pp. 513–520.

[27] T. Chen and S. Zhong, "Truthful auctions for continuous spectrum with variable bandwidths," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 2, pp. 1116–1128, February 2014.

[28] S. M. Kakade, I. Lobel, and H. Nazerzadeh, "An optimal dynamic mechanism for multi-armed bandit processes," *arXiv preprint arXiv:1001.4598*, 2010.

[29] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *Vehicular Technology, IEEE Transactions on*, vol. 44, no. 1, pp. 163–171, Feb 1995.

[30] N. Zhang, N. Lu, N. Cheng, J. Mark, and X. Shen, "Cooperative spectrum access towards secure information transfer for crns," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 11, pp. 2453–2464, November 2013.

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

[32] J. C. Gittins and D. M. Jones, "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics*, J. Gani, Ed. Amsterdam, NL: North-Holland, 1974, pp. 241–266.

[33] I. M. Sonin, "A generalized gittins index for a markov chain and its recursive calculation," *Statistics and Probability Letters*, vol. 78, no. 12, pp. 1526 – 1533, 2008.

[34] D. B. West, *Introduction to Graph Theory (2nd Edition)*, 2nd ed. Prentice Hall, Sep. 2000.

[35] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. pp. 32–38, 1957.