Indian Buffet Game With Negative Network Externality and Non-Bayesian Social Learning

Chunxiao Jiang, *Member, IEEE,* Yan Chen, *Senior Member, IEEE,* Yang Gao, *Student Member, IEEE,* and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—In a dynamic system, how to perform learning and make decisions are becoming more and more important for users. Although there are some works in social learning-related literature regarding how to construct belief for an uncertain system state, few studies have been conducted on incorporating social learning with decision making. Moreover, users may have multiple concurrent options on different objects/resources and their decisions usually negatively influence each other's utility, which makes the problem even more challenging. In this paper, we propose an Indian Buffet Game to study how users in a dynamic system learn about the uncertain system state and make multiple concurrent decisions by not only considering the current myopic utility, but also the influence of subsequent users' decisions. We analyze the proposed Indian Buffet Game under two different scenarios: 1) on customers requesting multiple dishes without budget constraint and 2) with budget constraint. For both cases, we design recursive best response algorithms to find the subgame perfect Nash equilibrium (NE) for customers and characterize special properties of the NE profile under homogeneous setting. Moreover, we introduce a non-Bayesian social learning algorithm for customers to learn the system state, and theoretically prove its convergence. Finally, we conduct simulations to validate the effectiveness and efficiency of the proposed algorithms.**

***Index Terms*—Decision making, game theory, Indian Buffet Game, negative network externality, non-Bayesian social learning.**

I. INTRODUCTION

IN A dynamic system, users are usually confronted with uncertainty about the system state when making decisions. For example, in the field of wireless communications, when choosing channels to access, users may not know exactly the channel capacity and quality. Besides, users have to consider others' decisions since a large number of users sharing a same channel will inevitably decrease the average data rate and increase the end-to-end delay. Such phenomenon is known as negative network externality [1], [2], i.e., the negative influence of other users' behaviors on one user's reward, due to which users tend to avoid making the same decisions with others to maximize their own utilities. Similar phenomenon can be found in our daily life such as selecting online cloud storage service and choosing WiFi access point [3]. Therefore, how users in a dynamic system learn the system state and make best decisions by considering the influence of others' decisions becomes an important research issue in many fields [4]–[7].

Although users in a dynamic system may only have limited knowledge about the uncertain system state, they can construct a probabilistic belief regarding the system state through social learning. In the social learning literature [8]–[13], different kinds of learning rules were studied where the essential objective is to learn the true system state eventually. In most of these existing works, the learning problem is typically formulated as a dynamic game with incomplete information and the main focus is to study whether users can learn the true system state at the equilibria. However, all of the previous works assumed that users' utility functions are independent of each other, i.e., they did not consider the concept of network externality, which is indeed a common phenomenon in dynamic systems and can influence users' utilities and decisions to a large extent.

To study the social learning problem with negative network externality, in [14]–[16], we have proposed a general framework called Chinese Restaurant Game. The concept is originated from Chinese Restaurant Process [17], which is used to model unknown distributions in the nonparametric learning methods in the field of machine learning. In the Chinese Restaurant Game, there are finite tables with different sizes and finite customers sequentially requesting tables to be seated. Since customers do not know the exact size of each table, which affects each other's utility, they have to learn the table sizes according to some external information. Moreover, when requesting one table, each customer should take into account the subsequent customers' decisions due to the limited dining space in each table, i.e., the negative network externality. Then, the Chinese Restaurant Game is extended to a dynamic population setting in [18], where customers arrive at and leave the restaurant with a Poisson process. With the general Chinese Restaurant Game theoretic framework, one

Manuscript received March 13, 2014; revised June 4, 2014; accepted August 2, 2014. Date of publication December 12, 2014; date of current version March 13, 2015. This work was supported in part by Project 61371079, Project 61273214, and Project 61271267, in part by the National Science Foundation of China, in part by the National Basic Research Program of China (973 Program) under Grant 2013CB329105, and in part by the Post-Doctoral Science Foundation funded project. This work was done during C. Jiang's visit at the University of Maryland. This paper was recommended by Associate Editor F.-Y. Wang.

C. Jiang is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: chx.jiang@gmail.com).

Y. Chen, Y. Gao, and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2014.2376880

is able to analyze the social learning and strategic decision making of rational users in a network.

One underlying assumption in the Chinese Restaurant Game is that each customer can only choose one table. However, in many real applications, users can have multiple concurrent selections, e.g., mobile terminals can access multiple channels at once and users can have multiple cloud storage services. To tackle such a challenge, similar to the Chinese Restaurant Game which introduced the strategic behaviors into the nonstrategic Chinese Restaurant Process, in this paper, we propose a new game called Indian Buffet Game, which introduces strategic behaviors into the nonstrategic Indian Buffet Process [19]. In the Indian Buffet Process, there exist an infinite number of dishes and each customer can order a number of dishes following Poisson process. Such a stochastic process defines a probability distribution for use as a prior in probabilistic models. By incorporating customers' rational behaviors, we extend the Indian Buffet Process into the Indian Buffet Game and study what is the customers' optimal decisions when choosing the dishes in a sequential manner. The proposed Indian Buffet game is an ideal framework to study the multiple dishes selection problem by integrating social learning and strategic decision making with negative network externality. While lots of works have been done regarding the simultaneous decision making problem [7], [20], the sequential decision making problem has not been well investigated in the literature, whereas a user has to consider both the previous users' decisions and predict the subsequent users' decisions. Furthermore, in practice, the sequential decision making scenarios are even more prevailing, especially in the field of wireless communication where synchronization among all users is quite difficult. Therefore, in this paper, we focus on sequential decision making analysis and try to provide some insights and results on this problem. We will discuss two cases: Indian Buffet Game without budget constraint and with budget constraint, where, "with budget constraint" means the number of dishes each customer can require is limited, and, "without budget constraint" means no limitation. The main contributions of this paper can be summarized as follows.

1) We propose a general framework, Indian Buffet Game, to study how users make multiple concurrent selections under uncertain system states. Specifically, such a framework can reveal how users learn the uncertain system state through social learning and make optimal decisions to maximize their own expected utilities by considering negative network externality.
2) In the learning stage of the Indian Buffet Game, we propose a non-Bayesian social learning algorithm for customers to learn the dish states. Moreover, we prove theoretically the convergence of the proposed non-Bayesian social learning algorithm to the true belief and show with simulations the fast convergence rate.
3) For the case without budget constraint, we show that the multiple concurrent dishes selection problem can be decoupled into a series of independent Indian Buffet Games. We then design a recursive best response algorithm to find the subgame perfect Nash equilibrium (NE) of the elementary Indian Buffet Game. We show that, under the homogeneous setting, the NE profile exhibits a threshold structure.
4) For the case with budget constraint, we design a recursive best response algorithm to find the corresponding subgame perfect NE. We then show that, under the homogeneous setting, the NE profile exhibits an equal-sharing property.

The rest of this paper is organized as follows. The system model is described in Section II. While, the Indian Buffet Game without and with budget constraint are discussed in details in Sections III and IV, respectively. In Section V, we give the theoretical proof of the convergence of the proposed non-Bayesian learning rule. Finally, we show simulation results in Section VI and draw the conclusion in Section VII.

II. System Model

A. Indian Buffet Game Formulation

Let us consider an Indian buffet restaurant which provides M dishes denoted by $r_1, r_2, \ldots, r_M$. There are N customers labeled with $1, 2, \ldots, N$ sequentially requesting dishes for a meal. Each dish can be shared among multiple customers and each customer can select multiple dishes. We assume that all N customers are rational in the sense that they will select dishes which can maximize their own utilities. In such a case, the multiple dishes selection problem can be formulated to be a noncooperative game, called Indian Buffet Game, as follows.

1) *Players:* N rational customers in the restaurant.
2) *Strategies:* Since each customer can request multiple dishes, the strategy set can be defined as

$$\mathcal{X} = \{\emptyset, \{r_1\}, \ldots, \{r_1, r_2\}, \ldots, \{r_1, r_2, \ldots, r_M\}\} \quad (1)$$

where each strategy is a combination of dishes and $\emptyset$ means no dish is requested. Obviously, customers' strategy set is finite with 2^M elements. We denote the strategy of customer i as $\mathbf{d}_i = (d_{i,1}, d_{i,2}, \ldots, d_{i,M})'$, where $d_{i,j} = 1$ represents customer i requests dish r_j and otherwise we have $d_{i,j} = 0$. The strategy profile of all customers can be denoted by a $M \times N$ matrix as follows[1]:

$$\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N) = \begin{bmatrix} d_{1,1} & d_{2,1} & \cdots & d_{N,1} \\ d_{1,2} & d_{2,2} & \cdots & d_{N,2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{1,M} & d_{2,M} & \cdots & d_{N,M} \end{bmatrix}. \quad (2)$$

3) *Utility Function:* The utility of each customer is determined by both the quality of the dish and the number of customers who share the same dish due to the negative network externality. The quality of one dish can be interpreted as the deliciousness or the size. Let $q_j \in Q$ denote the quality of dish r_j where Q is the quality space, and N_j denote the total number of customers requesting dish r_j. Then, we can define the utility function of customer i requesting dish r_j as

$$u_{i,j}\left(q_j, N_j\right) = g_{i,j}\left(q_j, N_j\right) - c_{i,j}\left(q_j, N_j\right) \quad (3)$$

where $g_{i,j}(\cdot)$ is the gain function and $c_{i,j}(\cdot)$ is the cost function. Note that the utility function is an increasing

[1] In this paper, the bold symbols represent vectors, the bold capital symbols represent matrixes, the subscript i denotes the customer index, subscript j denotes the dish index, and the superscript (t) denotes time slot index.

function in terms of q_j, and a decreasing function in terms of N_j, which can be regarded as the characteristic of negative network externality since the more customers request dish r_j, the less utility customer i can obtain. Note that if no dish is requested, the utility is zero for a customer.

We hereby define the dish state $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_M\}$, where $\theta_j \in \Theta$ denotes the state of dish r_j. The dish state can be interpreted as how good the ingredient or cooking is. Θ is the set of all possible states, which is assumed to be finite. The dish state keeps unchanged along with time until the whole Indian buffet restaurant is remodeled. The aforementioned quality of dish r_j, q_j, is assumed to be a random variable following the distribution $f_j(\cdot|\theta_j)$, which means that the state of the dish θ_j determines the distribution of the dish quality q_j. Note that the dish quality is the same for all users in a certain time slot, which follows distribution $f_j(\cdot|\theta_j)$ and there is a realization in each time slot. The dish state $\boldsymbol{\theta} \in \Theta^M$ is unknown to all customers, e.g., they do not know exactly whether the dish is delicious or not before requesting and tasting. Nevertheless, they may have received some advertisements or gathered some reviews about the restaurant. Such information can be regarded as some kinds of signals related to the true state of the restaurant. In such a case, customers can estimate $\boldsymbol{\theta}$ through those available information, i.e., the information they know in advance and/or gather from other customers.

In the Indian Buffet Game model, we divide the system time into time slots and assume that the dish quality q_j with $j = 1, 2, \ldots, M$ varies independently from time slot to time slot following the corresponding conditional distributions $f_j(\cdot|\theta_j)$. In each time slot, customers make sequential decisions on which dishes to request. There are mainly two issues to be addressed in the Indian Buffet Game. First, since the states are unknown, it is very important to design an effective social learning rule for customers to learn from others and their previous outcomes. Second, given customers' knowledge about the state, we should characterize the equilibrium that rational customers will adopt in each time slot. In this paper, to ensure fairness among customers, we assume that customers have different orders of selecting dishes at different time slots. In the sequel, the customer index $1, 2, \ldots, N$ means the dish request order of them, i.e., customer i means the ith customer. In such a case, it is sufficient for customers to only consider the expected utilities at current time slot. Moreover, although each customer can request more than one dish, the total number of requests is subject to the following budget constraint:

$$\sum_{j=1}^{M} d_{i,j} \leq L, \quad \forall\, i = 1, 2, \ldots, N. \tag{4}$$

A special case of (4) is that $L \geq M$, which is equivalent to the case without budget constraint where customers can request as many dishes as possible. In Sections III and IV, we will discuss the Indian Buffet Game under two scenarios: without budget constraint ($L \geq M$) and with budget constraint ($L < M$), respectively.

B. Time Slot Structure of Indian Buffet Game

Since the dish state $\boldsymbol{\theta} \in \Theta^M$ is unknown to customers, we introduce the concept of belief to describe their uncertainty about the state [22]. Let us denote the belief as $\mathbf{P}^{(t)} = \{\mathbf{p}_j^{(t)}, j = 1, 2, \ldots, M\}$, where $\mathbf{p}_j^{(t)} = \{p_j^{(t)}(\theta), \theta \in \Theta\}$ represents customers' estimation about the probability distribution regarding the state of dish r_j at time slot t. Since customers can obtain some prior information about the dish state, we assume that all customers start with a prior belief $p_{i,j}^{(0)}(\theta)$ for every state $\theta \in \Theta$. Note that the prior beliefs of all customers can be different, however, since all customers share their belief information with each other, they will have the same belief after the first belief updating. In this section, we will discuss the proposed social learning algorithm, i.e., how customers update their belief $\mathbf{P}^{(t)}$ at each time slot, and leave the convergence and performance analysis in Section V.

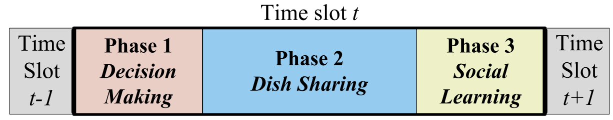


Fig. 1. Time slot structure of the Indian Buffet Game.

In Fig. 1, we illustrate the time slot structure of the proposed Indian Buffet Game. At each time slot $t \in \{1, 2 \ldots\}$, there are three phases: 1) decision-making phase; 2) dish sharing phase; and 3) social learning phase.

1) Phase 1—Decision Making: In this phase, customers sequentially make decisions on which dishes to request and broadcast their decisions to others, or simply put, everyone knows what others are getting. For customer i, his/her decision is to maximize his/her expected utility at current time slot, based on the belief at current time slot $\mathbf{P}^{(t)}$, the decisions of the previous $(i-1)$ customers $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{i-1}\}$, and his/her predictions of the subsequent $(N-i)$ customers' decisions.

2) Phase 2—Dish Sharing: In the second phase, each customer requests his/her desired dishes, receives a utility $u_{i,j}(q_j, N_j)$ according to the dish quality q_j and the number of customers N_j sharing the same dish as defined in (3). Notice that since N_j is known to all customers after the decision making phase, the customers who requested dish r_j at time slot t can infer the dish quality q_j from the received utility. Let us denote such inferred information as $s_{i,j}^{(t)} \in \mathcal{Q}, s_{i,j} \sim f_j(\cdot|\theta_j)$, which serves as the signal in the learning procedure. On the other hand, the customers who have not requested r_j at time slot t, cannot infer the dish quality q_j and thus have no inferred signal by themselves. Such an asymmetric structure, i.e., not every customer infers signals, makes the learning problem different from the traditional social learning settings and thus poses more challenges on learning the true state.

3) Phase 3—Social Learning: According to the observed/inferred signals in the second phase, customers can update the belief through the proposed non-Bayesian social learning rule. As illustrated in Fig. 2, there are mainly two steps in the proposed social learning rule. In the first step, each customer updates his/her local intermediate belief on θ_j, $\boldsymbol{\mu}_{i,j}^{(t)}$, and then reveals this intermediate belief to others. Since sharing the belief can help each customer to get more information and thus enhance the learning performance, we assume the customers have the incentive to do so. In the second step, each customer

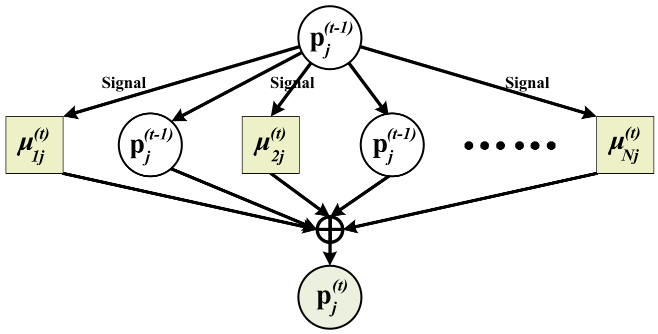


Fig. 2. Non-Bayesian learning rule for each dish.

combines his/her intermediate belief with other customers' intermediate beliefs in a linear manner.[2] Based on the Bayes' theorem [23], the customer is intermediate belief on the state of dish r_j, $\boldsymbol{\mu}_{i,j}^{(t)} = \{\mu_{i,j}^{(t)}(\theta), \theta \in \Theta\}$, can be calculated by

$$\mu_{i,j}^{(t)}(\theta) = \frac{f_j\left(s_{i,j}^{(t)}|\theta\right) p_j^{(t-1)}(\theta)}{\sum_{\Theta} f_j\left(s_{i,j}^{(t)}|\theta\right) p_j^{(t-1)}(\theta)}, \quad \forall\ \theta \in \Theta. \tag{5}$$

From (5), we can see that when customer i has requested r_j at time slot t, he/she will incorporate the corresponding signal into his/her intermediate belief $\boldsymbol{\mu}_{i,j}^{(t)}$. Otherwise, he/she will use the previous belief $\mathbf{p}_j^{(t-1)}$. Then, each customer linearly combines his/her intermediate belief with others customers' intermediate beliefs as follows:

$$p_j^{(t)}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left[d_{i,j}^{(t)}\mu_{i,j}^{(t)}(\theta) + \left(1 - d_{i,j}^{(t)}\right) p_j^{(t-1)}(\theta)\right]$$
$$\forall\ \theta \in \Theta,\ \text{and } j = 1, 2, \ldots, M \tag{6}$$

where $d_{i,j}^{(t)}$ is the strategy of customer i at time slot t.

One may have already noticed that, different from the Chinese Restaurant Game where each customer can only choose one table [16], customers in the Indian Buffet Game can request multiple dishes concurrently. Moreover, there are another two significant differences between those two games. One difference is that each customer in the Chinese Restaurant Game can only make decision once while customers in the Indian Buffet Game can make decisions repeatedly, i.e., one-shot game versus repeated game, due to which customers in the Indian Buffet Game can learn from their previous experiences and an effective learning rule that can guarantee convergence is required. The other difference is the learning rule. In the Chinese Restaurant Game, customers sequentially make decisions and then reveal their signals to subsequent customers, where the Bayesian social learning rule is used for customers to combine the signals from previous customers. However, in the Indian Buffet Game, instead of revealing signals, customers only reveal their intermediate belief and the non-Bayesian social learning rule is applied to generate the final belief.

[2]Note that the state learning processes, i.e., the belief update, of all dishes are independent.

III. Indian Buffet Game Without Budget Constraint

In this section, we study the Indian Buffet Game without budget constraint, which is corresponding to the case where $L \geq M$ in (4). When there is no budget constraint, customers should request all dishes that can give them positive expected utility to maximize their total expected utilities. We will first show that without budget constraint, the decision of whether to request one dish or not is independent from other dishes, i.e., the Indian Buffet Game that selects multiple concurrent dishes is decoupled into a series of elementary Indian Buffet Games that select a single dish. Although in elementary Indian Buffet Game each customer can only choose one dish, which is similar to the Chinese Restaurant Game model [14], the difference is that all customers cooperatively estimate the dish state in the Indian Buffet Game model, instead of sequential learning in the Chinese Restaurant Game model. We present a recursive algorithm that characterizes the subgame perfect equilibrium of the Indian Buffet Game without budget constraint. Finally, we also discuss the homogeneous case where customers have the same form of utility function to gain more insights.

To show the independence among different dishes, we first define the best response of a customer given other customers' actions. Let us define $\mathbf{n}_{-i} = \{n_{-i,1}, n_{-i,2}, \ldots, n_{-i,M}\}$ with

$$n_{-i,j} = \sum_{k \neq i} d_{k,j} \tag{7}$$

being the number of customers except customer i choosing r_j. Let $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$, where $\mathbf{p}_j = \{p_j(\theta), \theta \in \Theta\}$ is the customer's belief regarding the state of dish r_j at current time slot.[3] Given $\mathbf{P}$ and $\mathbf{n}_{-i}$, the best response of customer i, $\mathbf{d}_i^* = (d_{i,1}^*, d_{i,2}^*, \ldots, d_{i,M}^*)'$, can be written as

$$\mathbf{d}_i^* = \mathrm{BR}_i(\mathbf{P}, \mathbf{n}_{-i}) = \arg\max_{\mathbf{d}_i \in \{0,1\}^M} \sum_{j=1}^{M} d_{i,j} \cdot U_{i,j} \tag{8}$$

where $U_{i,j}$ is customer is expected utility of requesting dish r_j given belief $\mathbf{P}$, which can be calculated by

$$U_{i,j} = \sum_{\Theta}\sum_{Q} u_{i,j}\left(q_j, n_{-i,j} + d_{i,j}\right) f_j\left(q_j|\theta_j\right) p_j\left(\theta_j\right) \tag{9}$$

where Q is the quality/signal set and $q_j \in Q$.

Since there is no constraint for the optimization problem in (8) and $U_{i,j}$ is only related with $d_{i,j}$, we can rewrite (8) by

$$\mathbf{d}_i^* = \mathrm{BR}_i(\mathbf{P}, \mathbf{n}_{-i}) = \sum_{j=1}^{M} \arg\max_{\mathbf{d}_i \in \{0,1\}^M} d_{i,j} \cdot U_{i,j}. \tag{10}$$

From (10), we can see that the optimal decision on one dish is irrelevant to the decisions on others, which leads to the independence among different dishes. In such a case, we have

$$d_{i,j}^* = \arg\max_{d_{i,j} \in \{0,1\}} d_{i,j} \cdot U_{i,j}. \tag{11}$$

The independence property enables us to simplify the analysis by decoupling the origin Indian Buffet Game into M elementary Indian Buffet Games. The elementary Indian

[3]Since we discuss the Indian Buffet Game in one time slot, the superscript (t) is omitted in Sections III and IV.

Buffet Game is defined as the case when there is only one dish for the customers, and thus, the decision for each customer is binary, i.e., to request or not to request. In the remaining of this section, we will focus on the analysis of the elementary Indian Buffet Game and discard the dish index j for notational simplification. As a result, we can rewrite the best response of customer i based on the latest belief information $p(\theta)$ as follows:

$$d_i^* = \mathrm{BR}_i(\mathbf{p}, n_{-i}) = \arg\max_{d_i \in \{0,1\}} d_i \cdot U_i = \begin{cases} 1, & \text{if } U_i = \sum_\Theta \sum_Q u_i(q, n_{-i}+1) f(q|\theta) p(\theta) > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

A. Recursive Best Response Algorithm

In this section, we study how to solve the best response defined in (12) for each customer. From (12), we can see that customer i needs to know n_{-i} to calculate the expected utility U_i in order to decide whether to request the dish or not. However, since the customers make their decisions sequentially, customer i does not know the decisions of those who are after him/her and thus needs to predict the subsequent customers' decisions based on the belief and known information.

Let m_i denote the number of customers who will request the dish after customer i, then we can write the recursive form of m_i as

$$m_i = m_{i+1} + d_{i+1}. \tag{13}$$

Let $m_i|_{d_i=0}$ and $m_i|_{d_i=1}$ represent m_i under the condition of $d_i = 0$ and $d_i = 1$, respectively. Denoted by $n_i = \sum_{k=1}^{i-1} d_k$ is the number of customers choosing the dish before customer i. Then, the estimated number of customers choosing the dish excluding customer i can be written as follows:

$$\hat{n}_{-i}|_{d_i=0} = n_i + m_i|_{d_i=0} \tag{14}$$

$$\hat{n}_{-i}|_{d_i=1} = n_i + m_i|_{d_i=1}. \tag{15}$$

Note that $\hat{n}_{-i}|_{d_i=0}$ and $\hat{n}_{-i}|_{d_i=1}$ are different from n_{-i} since the values of $d_{i+1}, d_{i+2}, \ldots, d_N$ are estimated instead of true observations.

According to (15), we can compute the expected utility of customer i when $d_i = 1$ as

$$U_i|_{d_i=1} = \sum_\Theta \sum_Q u_i(q, n_i + m_i|_{d_i=1} + 1) f(q|\theta) p(\theta). \tag{16}$$

Since the utility of customer i is zero when $d_i = 0$, the best response of customer i can be obtained as

$$d_i^* = \begin{cases} 1, & \text{if } \; U_i|_{d_i=1} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

With (17), we can find the best response of customer i given belief $\mathbf{p}$, current observation n_i and predicted number of subsequent customers choosing the dish, $m_i|_{d_i=1}$. To predict $m_i|_{d_i=1}$, customer i needs to predict the decisions of all customers from $i+1$ to N. When it comes to customer N, since he/she knows exactly the decisions of all the previous customers, he/she can find the best response without making any prediction, i.e., $m_N = 0$. Along this line, it is intuitive to design a recursive algorithm to predict $m_i|_{d_i=1}$ by considering all possible decisions of customers from $i+1$ to N and sequentially updating $m_i = m_{i+1} + d_{i+1}$. In Algorithm 1, we show the recursive algorithm BR_EIBG($\mathbf{p}, n_i, i$) that describes how to predict $m_i|_{d_i=1}$ and find the best response d_i for customer i, given current belief $\mathbf{p}$ and observation n_i. Moreover, in order to obtain a correct prediction of m_i in the recursion procedure, we calculate and return $m_i|_{d_i=0}$ when the best response of customer i is 0. Note that in Algorithm 1, we have assumed that the functional form of the utilities of all users are known by each other. In case that the utility function forms are unknown, users can take expectation over the types of users based on some empirical user-type distribution, which is not the focus of this paper. Moreover, as for the complexity of Algorithm 1, it enumerates all possible combinations of the choices made by customers in a dynamic programming manner. Different from the exhaustive search with the complexity order of $O(2^N)$, Algorithm 1 is with order $O(N^2)$, where N is the total number of customers. In the following, we will prove that the action profile specified in BR_EIBG($\mathbf{p}, n_i, i$) is a subgame perfect NE for the elementary Indian Buffet Game.

Algorithm 1 BR_EIBG($\mathbf{p}, n_i, i$)

```
if Customer i == N then
  //******For customer N******//
  if U_N = Σ_Θ Σ_Q u_N(q, n_N + 1) f(q|θ) p(θ) > 0 then
    d_N ← 1
  else
    d_N ← 0
  end if
  m_N ← 0
else
  //******For customer 1, 2, ..., N − 1******//
  //***Predicting***//
  (d_{i+1}, m_{i+1}) ← BR_EIBG(p, n_i + 1, i + 1)
  m_i ← m_{i+1} + d_{i+1}
  //***Making decision***//
  if U_i = Σ_Θ Σ_Q u_i(q, n_i + m_i + 1) f(q|θ) p(θ) > 0 then
    d_i ← 1
  else
    (d_{i+1}, m_{i+1}) ← BR_EIBG(p, n_i, i + 1)
    m_i ← m_{i+1} + d_{i+1}
    d_i ← 0
  end if
end if
return (d_i, m_i)
```

B. Subgame Perfect NE

In this section, we will show that Algorithm 1 leads to a subgame perfect NE for the elementary Indian Buffet Game. In the following, we first give the formal definitions of NE, subgame, and subgame perfect NE.

Definition 1: Given the belief $\mathbf{p} = \{p(\theta), \theta \in \Theta\}$, the action profile $\mathbf{d}^* = \{d_1^*, d_2^*, \ldots, d_N^*\}$ is a NE of the N-customer elementary Indian Buffet Game if and only if $\forall\, i \in \{1, 2, \ldots, N\}$, $d_i^* = \mathrm{BR}_i(\mathbf{p}, \sum_{k \neq i} d_k^*)$ as given in (12).

Definition 2: A subgame of the N-customer elementary Indian Buffet Game consists of the following three elements: 1) it starts from customer i with $i = 1, 2, \ldots, N$; 2) it has the belief $\mathbf{p}$ at the current time slot; and 3) it has the current observation, n_i, which is the decisions of the previous customers.

Definition 3: A NE is a subgame perfect NE if and only if it is a NE for every subgame.

According to the above definitions, we show in the following theorem that the action profile derived by Algorithm 1 is a subgame perfect NE of the elementary Indian Buffet Game.

Theorem 1: Given the belief $\mathbf{p} = \{p(\theta), \theta \in \Theta\}$, the action profile $\mathbf{d}^* = \{d_1^*, d_2^*, \ldots, d_N^*\}$, with d_i^* being determined by BR_EIBG($\mathbf{p}, n_i, i$) and $n_i = \sum_{k=1}^{i-1} d_k^*$, is a subgame perfect NE for the elementary Indian Buffet Game.

Proof: We first show that d_k^* is the best response of customer k in the subgame starting from customer $i, \forall\ 1 \le i \le k \le N$.

If $k = N$, we can see that BR_EIBG($\mathbf{p}, n_N, N$) assigns the value of d_N^* directly as

$$d_N^* = \begin{cases} 1, & \text{if } U_N = \sum_\Theta \sum_Q u_N(q, n_N + 1) f(q|\theta) p(\theta) > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Since $n_N = n_{-N}$, we have $d_k^* = \mathrm{BR}_k(\mathbf{p}, n_{-k})$ in the case of $k = N$ according to (12), i.e., d_k^* is the best response of customer k.

If $k < N$, suppose d_k^* is the best response of customer k derived by BR_EIBG($\mathbf{p}, n_k, k$). If $d_k^* = 0$, denoting $d_k' = 1$ as the contradiction, we can see from BR_EIBG($\mathbf{p}, n_k, k$) that

$$U_k|_{d_k'=1} = \sum_\Theta \sum_Q u_k(q, n_k + m_k + 1) f(q|\theta) p(\theta) < 0 = U_k|_{d_k^*=0} \quad (19)$$

which means customer k has no incentive to deviate from $d_k^* = 1$ given the prediction of other customers' decisions. If $d_k^* = 1$, denoting $d_k' = 0$ as the contradiction, we can see from BR_EIBG($\mathbf{p}, n_k, k$) that

$$U_k|_{d_k'=0} = 0 < U_k|_{d_k^*=1} = \sum_\Theta \sum_Q u_k(q, n_k + m_k + 1) f(q|\theta) p(\theta) \quad (20)$$

which means that customer k has no incentive to deviate from $d_k^* = 0$ given the prediction of other customers' decisions. Therefore, $d_k^* = $ BR_EIBG($\mathbf{p}, n_k, k$) is the best response of customer k in the subgame of the elementary Indian Buffet Game starting with customer i. Moreover, since the statement is true for $\forall\ k$ satisfying $i \le k \le N$, we know that $\{d_i^*, d_{i+1}^*, \ldots, d_N^*\}$ is the NE for the subgame starting from customer i. Therefore, according to the definition of subgame perfect NE, we can conclude that Theorem 1 is true. ■

C. Homogeneous Case

From the previous section, we know that a recursive procedure is required to determine the best responses of the elementary Indian Buffet Game. This is due to the fact that one customer needs to predict the decisions of all subsequent customers to determine the best response of a certain customer. In this section, we simplify the game with a homogeneous setting to derive a more concise best response.

In the homogeneous case, we assume that all customers have the same form of utility function, i.e., $u_i(q, n) = u(q, n)$, for all i, q, n. Under such a setting, the equilibrium can be characterized in a much simpler way as shown in the following lemma.

Lemma 1: In the N-customer elementary Indian Buffet Game under a homogeneous setting, if $\mathbf{d}^* = \{d_1^*, d_2^*, \ldots, d_N^*\}$ is the NE action profile specified by BR_EIBG(), then we have $d_i^* = 1$ if and only if $0 \le i \le n^*$, where $n^* = \sum_{k=1}^N d_k^*$.

Proof: Suppose the best response of customer i, $d_i^* = 0$. Then, according to Algorithm 1, we have

$$U_i = \sum_\Theta \sum_Q u(q, n_i + m_i|_{d_i=1} + 1) f(q|\theta) p(\theta) \le 0. \quad (21)$$

The prediction of m_i under the condition of $d_i = 1$ relies on the recursive estimations of all subsequent customers' decisions. In particular, we have $m_i|_{d_i=1} = d_{i+1}|_{d_i=1} + m_{i+1}|_{d_i=1}$, where the value of $d_{i+1}|_{d_i=1}$ can be computed as follows:

$$d_{i+1}|_{d_i=1} = \begin{cases} 1, & \text{if } U_{i+1}|_{d_i=1} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

with

$$U_{i+1}|_{d_i=1} = \sum_\Theta \sum_Q u(q, n_i + 1 + m_{i+1}|_{d_i=1} + 1) f(q|\theta) p(\theta). \quad (23)$$

Since $n_i + 1 + m_{i+1}|_{d_i=1} + 1 \ge n_i + m_i|_{d_i=1} + 1$ and $u(q, n)$ is a decreasing function in terms of n, we have $d_{i+1}|_{d_i=1} = 0$ according to (21) and (23). Following the same argument, we can show that $d_k|_{d_i=1} = 0$ for all $k = i + 1, i + 2, \ldots, N$. Therefore, we have

$$m_i|_{d_i=1} = \sum_{k=i+1}^N d_k|_{d_i=1} = 0. \quad (24)$$

Then, let us consider the best response of customer $i + 1$, which can be calculated by

$$d_{i+1}^* = \begin{cases} 1, & \text{if } U_{i+1} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

where

$$U_{i+1} = \sum_\Theta \sum_Q u(q, n_{i+1} + m_{i+1}|_{d_{i+1}=1} + 1) f(q|\theta) p(\theta). \quad (26)$$

Since $n_{i+1} = n_i + d_i$, $m_i|_{d_i=1} = 0$ and $m_{i+1}|_{d_{i+1}=1} \ge 0$, we have $n_{i+1} + m_{i+1}|_{d_{i+1}=1} + 1 \ge n_i + m_i|_{d_i=1} + 1$. According to (21) and (26), and the decreasing property of utility function in terms of the number of customers sharing the same dish, we have $d_{i+1}^* = 0$.

Following the same argument, we can show that if $d_i^* = 0$, then $d_k^* = 0$ for all $k \in \{i+1, i+2, \ldots, N\}$. Since all decisions can take values of either 0 or 1, we have $d_i^* = 1$ if and only if $0 \le i \le \sum_{k=1}^N d_k^*$. This completes the proof. ■

From Lemma 1, we can see that there is a threshold structure in the NE of the elementary India Buffet Game with the homogeneous setting. The threshold structure is embodied in the fact that if $d_i^* = 0$, then $d_k^* = 0, \forall\ k \in \{i+1, i+2, \ldots, N\}$, and if $d_i^* = 1$, then $d_k^* = 1, \forall\ k \in \{1, 2, \ldots, i-1\}$. The result can be easily extended to the Indian Buffet Game without budget constraint under the homogeneous setting as shown in the following theorem.

Theorem 2: In the M-dish and N-customer Indian Buffet Game without budget constraint, if all the customers have the same utility functions, there is a threshold structure in the NE matrix $\mathbf{D}^*$ denoted by (2), i.e., for any row $j \in \{1, 2, \ldots, M\}$ of $\mathbf{D}^*$, there is a $T_j \in \{1, 2, \ldots, N\}$ satisfying that

$$d^*_{i,j} = \begin{cases} 1, \ \forall\ i < T_j \\ 0, \ \forall\ i \geq T_j. \end{cases} \tag{27}$$

Proof: This theorem directly follows by extending Lemma 1 into M independent dishes case. ■

IV. Indian Buffet Game With Budget Constraint

In this section, we study the Indian Buffet Game with budget constraint, which is corresponding to the case with $L < M$ in (4). Unlike the previous case, when there is a budget constraint for each customer, the selection among different dishes is no longer independent but coupled. In the following, we will first discuss a recursive algorithm that can characterize the subgame perfect NE of the Indian Buffet Game with budget constraint. Then, we discuss a simplified case with a homogeneous setting to gain more insights.

A. Recursive Best Response Algorithm

In the budget constraint case, we assume that each customer can at most request L dishes at each time slot with $L < M$. In such a case, the best response of customer i can be found by the following optimization problem:

$$\mathbf{d}^*_i = \text{BR}_i(\mathbf{P}, \mathbf{n}_{-i}) = \arg\max_{\mathbf{d}_i \in \{0,1\}^M} \sum_{j=1}^{M} d_{i,j} \cdot U_{i,j}$$
$$\text{s.t.} \ \sum_{j=1}^{M} d_{i,j} \leq L < M \tag{28}$$

where

$$U_{i,j} = \sum_{\Theta}\sum_{Q} u_{i,j}(q_j, n_{-i,j} + d_{i,j}) f_j(q_j|\theta_j) p_j(\theta_j). \tag{29}$$

From (28), we can see that customer is decision on dish r_j is coupled with all other dishes, and thus (28) cannot be decomposed into M subproblems. Nevertheless, we can still find the best response of each customer by comparing all possible combinations of L dishes. Let $\Phi = \{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_H\}$ denote the set of all combinations of l ($1 \leq l \leq L$) dishes out of M dishes, where $H = \sum_{l=1}^{L} C^l_M = \sum_{l=1}^{L} M!/l!(M-l)!$ and $\boldsymbol{\phi}_h = (\phi_{h,1}, \phi_{h,2}, \ldots, \phi_{h,M})'$ is one possible combination with $\phi_{h,j}$ representing whether dish r_j is requested, that is

$$\boldsymbol{\phi}_h = (\underbrace{1, 1, \ldots, 1}_{l}, \underbrace{0, 0, \ldots, 0}_{M-l})' \tag{30}$$

means the customer requests dishes $r_1, r_2, \ldots, r_l$ ($1 \leq l \leq L$). In other words, Φ is the candidate strategy set of each customer with constraint L.

Let us define customer is observation of previous customers' decisions as

$$\mathbf{n}_i = \{n_{i,1}, n_{i,2}, \ldots, n_{i,M}\} \tag{31}$$

where $n_{i,j} = \sum_{k=1}^{i-1} d_{k,j}$ is the number of customers choosing dish r_j before customer i. Let $\mathbf{m}_i$ denote the subsequent customers' decisions after customer i, we have its recursive form as

$$\mathbf{m}_i = \mathbf{m}_{i+1} + \mathbf{d}_{i+1}. \tag{32}$$

Then, let

$$\mathbf{m}_i|_{\mathbf{d}_i=\boldsymbol{\phi}_h} = \left\{m_{i,1}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}, m_{i,2}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}, \ldots, m_{i,M}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}\right\} \tag{33}$$

with $m_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}$ be the predicted number of subsequent customers who will request dish r_j under the condition that $\mathbf{d}_i = \boldsymbol{\phi}_h$, where $\mathbf{d}_i = (d_{i,1}, d_{i,1}, \ldots, d_{i,M})'$ and $\boldsymbol{\phi}_h \in \Phi$. In such a case, the predicted number of customers choosing different dishes excluding customer i is

$$\hat{\mathbf{n}}_{-i}|_{\mathbf{d}_i=\boldsymbol{\phi}_h} = \mathbf{n}_i + \mathbf{m}_i|_{\mathbf{d}_i=\boldsymbol{\phi}_h}. \tag{34}$$

According to the definitions above, we can get customer is expected utility by obtaining dish r_j when $\mathbf{d}_i = \boldsymbol{\phi}_h$ as

$$U_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h} = \phi_{h,j} \sum_{\Theta}\sum_{Q} u_{i,j}(q_j, n_{i,j} + m_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h} + \phi_{h,j}) \times f_j\left(q_j|\theta_j\right) p_j\left(\theta_j\right). \tag{35}$$

Then, the total expected utility customer i can obtain with $\mathbf{d}_i = \boldsymbol{\phi}_h$ is the sum of $U_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}$ over all M dishes, that is

$$U_i|_{\mathbf{d}_i=\boldsymbol{\phi}_h} = \sum_{j=1}^{M} U_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}. \tag{36}$$

In such a case, we can find the optimal $\boldsymbol{\phi}^*_h$ which maximizes customer is expected utility $U_i|_{\mathbf{d}_i=\boldsymbol{\phi}_h}$ as follows:

$$\boldsymbol{\phi}^*_h = \arg\max_{\boldsymbol{\phi}_h \in \Phi}\{U_i|_{\mathbf{d}_i=\boldsymbol{\phi}_h}\} \tag{37}$$

which is the best response of customer i.

To obtain the best response in (37), each customer needs to calculate the expected utilities defined in (35), which in turn requires to predict $m_{i,j}|_{\mathbf{d}_i=\boldsymbol{\phi}_h}$, i.e., the number of customers who choose dish r_j after customer i. When it comes to customer N who has already known all the previous customers' decisions, no prediction is required. Therefore, similar to Algorithm 1, given belief $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$ at current time slot and the current observation $\mathbf{n}_i = \{n_{i,1}, n_{i,2}, \ldots, n_{i,M}\}$, we can design another recursive best response algorithm BR_IBG($\mathbf{p}, \mathbf{n}_i, i$) for solving the Indian Buffet Game with budget constraint in Algorithm 2. As one can see, customer N only needs to compare the expected utilities of requesting all M dishes and choose L or less than L dishes with highest positive expected utilities. Note that $\max^L$ means finding the highest L values. For other customers, each one of them needs to first recursively predict the following customers' decisions, and then makes his/her own decision based on the prediction and the current observations.

B. Subgame Perfect NE

Similar to the elementary Indian Buffet Game, we first give the formal definitions of NE and subgame of the Indian Buffet Game with budget constraint.

Definition 4: Given the belief $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$, the action profile $\mathbf{D}^* = \{\mathbf{d}^*_1, \mathbf{d}^*_2, \ldots, \mathbf{d}^*_N\}$ is a NE of the M-dish and N-customer Indian Buffet Game with budget constraint L, if and only if $\mathbf{d}^*_i = \text{BR}_i(\mathbf{P}, \sum_{k \neq i} \mathbf{d}^*_k)$ as defined in (28) for all i.

Definition 5: A subgame of the M-dish and N-customer Indian Buffet Game with budget constraint L consists of the following three elements: 1) it starts from customer i with

Algorithm 2 BR_IBG($\mathbf{P}, \mathbf{n}_i, i$)

```
if Customer i == N then
  //******For customer N******//
  for j = 1 to M do
    U_{i,j} = Σ_Θ Σ_Q u_{N,j}(q_j, n_{N,j} + 1) f_j(q_j|θ_j) p_j(θ_j)
  end for
  j = {j_1, j_2, ..., j_L} ← arg max^L_{j∈{1,2,...,M}} {U_{i,j}}
  for j = 1 to M do
    if (U_{i,j} > 0)&&(j ∈ j) then
      d_{N,j} ← 1
    else
      d_{N,j} ← 0
    end if
  end for
  m_N = 0
else
  //******For customer 1, 2, ..., N − 1******//
  //***Predicting***//
  for φ_h = φ_1 to φ_H do
    (d_{i+1}, m_{i+1}) ← BR_IBG(P, n_i + φ_h, i + 1)
    m_i ← m_{i+1} + d_{i+1}
    U_i(φ_h) = Σ_M φ_{h,j} Σ_Θ Σ_Q u_{i,j}(q_j, n_{i,j} + m_{i,j} + φ_{h,j})
              ·f_j(q_j|θ_j) p_j(θ_j)
  end for
  //***Making decision***//
  φ*_h ← arg max_{φ_h∈Φ} {U_i(φ_h)}
  (d_{i+1}, m_{i+1}) ← BR_IBG(P, n_i + φ*_h, i + 1)
  d_i ← φ*_h
  m_i ← m_{i+1} + d_{i+1}
end if
return (d_i, m_i)
```

$i = 1, 2, \ldots, N$; 2) it has the belief $\mathbf{P}$ at current time slot; and 3) it has the current observation, $\mathbf{n}_i$, which comes from the decisions of the previous customers.

Based on Definitions 3–5, we show in the following theorem that the action profile obtained by Algorithm 2 is a subgame perfect NE of the Indian Buffet Game with budget constraint.

Theorem 3: Given the belief $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$, the action profile $\mathbf{D}^* = \{\mathbf{d}_1^*, \mathbf{d}_2^*, \ldots, \mathbf{d}_N^*\}$, where $\mathbf{d}_i^*$ determined by BR_IBG($\mathbf{P}, \mathbf{n}_i, i$) and $\mathbf{n}_i = \sum_{k=1}^{i-1} \mathbf{d}_k^*$, is a subgame perfect NE for the Indian Buffet Game.

Proof: The proof of this theorem is similar to that of Theorem 1, the details of which are omitted due to the page limitation. The proof outline is that first to show $\forall\ i, k$ such that $1 \le i \le N$ and $i \le k \le N$, $\mathbf{d}_k^*$ is the best response of customer k in the subgame starting from customer i by analyzing two cases: $k = N$ and $k < N$. Then, we can know that $\{\mathbf{d}_i^*, \mathbf{d}_{i+1}^*, \ldots, \mathbf{d}_N^*\}$ is the NE for the subgame starting from customer i. Finally, according to the definition of subgame perfect NE, we can conclude that Theorem 3 is true. ■

C. Homogenous Case

In the homogenous case, we assume that all customers' utility functions are the same, i.e., $u_{i,j}(q, n) = u(q, n)$; and all dishes are in the same state, i.e., the dish state $\boldsymbol{\theta} = \{\theta, \theta, \ldots, \theta\}$. Under such circumstances, we can find some special properties in the NE action profile $\mathbf{D}^*$ of the Indian Buffet Game with budget constraint. First, let us define a parameter n_T which satisfies

$$\begin{cases} \sum_{\Theta} \sum_{Q} u(q, n) f(q|\theta) p(\theta) > 0, & \text{if } n \le n_T \\ \sum_{\Theta} \sum_{Q} u(q, n) f(q|\theta) p(\theta) \le 0, & \text{if } n > n_T. \end{cases} \tag{38}$$

From (38), we can see that n_T is the critical value such that the utility of n_T customers sharing a certain dish is positive but becomes nonpositive with one extra customer, i.e., each dish can be requested by at most n_T customers. In the following theorem, we will show that, under the homogeneous setting, all dishes will be requested by nearly equal number of customers, i.e., the equal-sharing is achieved.

Theorem 4: In the M-dish and N-customer Indian Buffet Game with budget constraint L, if all M dishes are in the same states and all N customers have the same utility function, the NE matrix $\mathbf{D}^*$ denoted by (2) satisfies that, for all dishes $\{r_j, j = 1, 2, \ldots, M\}$

$$\sum_{i=1}^{N} d_{i,j}^* = \begin{cases} n_T, & \text{if } n_T \le \left\lfloor \frac{NL}{M} \right\rfloor \\ \left\lfloor \frac{NL}{M} \right\rfloor \text{ or } \left\lceil \frac{NL}{M} \right\rceil, & \text{if } n_T \ge \left\lceil \frac{NL}{M} \right\rceil. \end{cases} \tag{39}$$

Proof: We prove this theorem by contradiction as follows.

Case 1: $n_T \le \lfloor NL/M \rfloor$.

Suppose that there exists a NE $\mathbf{D}^*$ that contradicts (39). That is, there is a dish $r_{j'}$ such that $\sum_{i=1}^{N} d_{i,j'}^* > n_T$ or $\sum_{i=1}^{N} d_{i,j'}^* < n_T$. From (38), we know that each dish can be requested by at most n_T customers, which means that only $\sum_{i=1}^{N} d_{i,j'}^* < n_T$ may hold. If $\sum_{i=1}^{N} d_{i,j'}^* < n_T \le \lfloor NL/M \rfloor$, we have $\sum_{j=1}^{M} \sum_{i=1}^{N} d_{i,j}^* < NL$, which means that there exists at least one customer i' who requests less than L dishes, i.e., $\sum_{j=1}^{M} d_{i'j}^* < L$. However, according to (38), we have $\sum_{\Theta} \sum_{Q} u(q, \sum_{i=1}^{N} d_{i,j'}^* + 1) f(q|\theta) p(\theta) > 0$, which means that the utility of customer i' can increase if he/she requests dish $r_{j'}$, i.e., his/her utility is not maximized unless $\mathbf{D}^*$ is not a NE. This contradicts our assumption. Therefore, we have $\sum_{i=1}^{N} d_{i,j}^* = n_T$ for all dishes when $n_T \le \lfloor NL/M \rfloor$.

Case 2: $n_T \ge \lceil NL/M \rceil$.

Similar to the arguments in Case 1, we cannot have $\sum_{j=1}^{M} \sum_{i=1}^{N} d_{i,j}^* < NL$, which means that $\sum_{j=1}^{M} \sum_{i=1}^{N} d_{i,j}^* = NL$. Let us assume that there exists a NE $\mathbf{D}^*$ that contradicts (39). Since $\sum_{j=1}^{M} \sum_{i=1}^{N} d_{i,j}^* = NL$, there is a dish r_{j_1} with $\sum_{i=1}^{N} d_{i,j_1}^* < \left\lfloor \frac{NL}{M} \right\rfloor$ and a dish r_{j_2} with $\sum_{i=1}^{N} d_{i,j_2}^* > \lceil NL/M \rceil$. In such a case, we have $\sum_{i=1}^{N} d_{i,j_2}^* > \sum_{i=1}^{N} d_{i,j_1}^* + 1$, which leads to

$$\sum_{\Theta} \sum_{Q} u\left(q, \sum_{i=1}^{N} d_{i,j_1}^* + 1\right) f(q|\theta) p(\theta) > \sum_{\Theta} \sum_{Q} u\left(q, \sum_{i=1}^{N} d_{i,j_2}^*\right) f(q|\theta) p(\theta). \tag{40}$$

From (40), we can see that the customer who has requested dish r_{j_2} can obtain higher utility by unilaterally deviating his/her decision by requesting dish r_{j_1}. Therefore, $\mathbf{D}^*$ is not a NE of the Indian Buffet Game with budget constraint L, and thus we have $\sum_{i=1}^{N} d_{i,j}^* = \lfloor NL/M \rfloor$ or $\lceil NL/M \rceil$, when $n_T \ge \lceil NL/M \rceil$. This completes the proof of the theorem. ■

V. Non-Bayesian Social Learning

In the previous two sections, we have analyzed the proposed Indian Buffet Game and characterized the corresponding equilibrium. From the analysis, we can see that the equilibrium highly depends on customers' belief $\mathbf{P} = \{\mathbf{p}_j, j = 1, 2, \ldots, M\}$, i.e., the estimated distribution of the dish state $\boldsymbol{\theta} = \{\theta_j, j = 1, 2, \ldots, M\}$. The more accurate the belief is, the better best response customers can make and thus the better utility customers can obtain. Therefore, it is very important for customers to improve their belief by exploiting their inferred signals. In this section, we will discuss the learning process in the proposed Indian Buffet Game. Specifically, we propose an effective non-Bayesian social learning algorithm that can guarantee customers to learn the true system state. In social learning, the users' rationality and the way they process signal information is the essential difference between non-Bayesian learning and Bayesian learning, while the implementation difference between them is that customers exchange their signal information in Bayesian learning rule instead of exchanging the intermediate belief information in our proposed non-Bayesian learning rule. The motivation of designing the non-Bayesian learning rule is that customers can first distributedly process their own signals and then cooperatively estimate the belief regarding the dish state, which can greatly decrease the computational cost of each customer. Note that since the learning process of one dish state θ_j are independent of others, in the rest of this section, we omit the dish index j for notation simplification.

Suppose the true dish state is θ^*, given customers' belief at time slot t, $\mathbf{p}^{(t)} = \{p^{(t)}(\theta), \forall\, \theta \in \Theta\}$, their belief at time slot $t+1$, $\mathbf{p}^{(t+1)} = \{p^{(t+1)}(\theta), \forall\, \theta \in \Theta\}$, can be updated by

$$p^{(t+1)}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left[d_i^{(t+1)}\mu_i^{(t+1)}(\theta) + \left(1 - d_i^{(t+1)}\right)p^{(t)}(\theta)\right] \tag{41}$$

where $d_i^{(t+1)} = 1$ or 0 is customer is decision, and $\mu_i^{(t+1)}(\theta)$ is the intermediate belief updated by the Bayesian learning rule for customers who have requested the dish and inferred some signal $s_i^{(t+1)} \sim f(\cdot|\theta^*)$, that is

$$\mu_i^{(t+1)}(\theta) = \frac{f(s_i^{(t+1)}|\theta)p^{(t)}(\theta)}{\sum_{\Theta} f(s_i^{(t+1)}|\theta)p^{(t)}(\theta)}, \quad \forall\, \theta \in \Theta. \tag{42}$$

Definition 6: A learning rule has the strong convergence property if and only if the learning rule can learn the true state in probability such that

$$\begin{cases} p^{(t)}(\theta^*) \to 1, \\ p^{(t)}(\forall\, \theta \neq \theta^*) \to 0, \end{cases} \quad \text{as } t \to \infty. \tag{43}$$

By reorganizing some terms, we can rewrite the non-Bayesian learning rule in (41) as

$$p^{(t+1)}(\theta) = p^{(t)}(\theta) + \frac{1}{N}\sum_{i=1}^{N} d_i^{(t+1)}\left(\frac{f\left(s_i^{(t+1)}|\theta\right)}{\lambda\left(s_i^{(t+1)}\right)} - 1\right)p^{(t)}(\theta) \tag{44}$$

with

$$\lambda\left(s_i^{(t+1)}\right) = \sum_{\Theta} f\left(s_i^{(t+1)}|\theta\right)p^{(t)}(\theta). \tag{45}$$

From (45), we can see that $\lambda(s_i^{(t+1)})$ is the estimation of the probability distribution of the signal $s_i^{(t+1)}$ at the next time slot. With $\lambda(s_i^{(t+1)})$, we can define a weak convergence, compared with the strong convergence in (43), as follows.

Definition 7: A learning rule has the weak convergence property if and only if the learning rule can learn the true state in probability such that

$$\lambda(s) = \sum_{\Theta} f(s|\theta)p^{(t)}(\theta) \to f(s|\theta^*), \forall\, s \in Q, \text{ as } t \to \infty. \tag{46}$$

Notice that the weak convergence is sufficient for the proposed Indian Buffet Game since the objective of learning is to find an accurate estimate of the expected utilities of customers and thus derive the true best response. According to (9), we can see that the signal distribution $\sum_{\Theta} f_j(q_j|\theta_j)p_j(\theta_j)$ is a sufficient statistic of the expected utility function. Therefore, if it can be shown that the proposed social learning algorithm has the weak convergence property, then we are able to derive the true best response for customers in the proposed Indian Buffet Game. In the following theorem, we will show and prove that the proposed learning algorithm in (41) indeed has the weak convergence property. We will also show with simulation that the proposed learning algorithm in (41) has the strong convergence property.

Theorem 5: In the Indian Buffet Game, suppose that the true dish state is θ^*, all customers update their belief $\mathbf{p}$ using (41) and their prior belief $\mathbf{p}^{(0)}$ satisfies $p^{(0)}(\theta^*) > 0$, then, the belief sequence $\{p^{(t)}(\theta)\}$ ensures a weak convergence, i.e., for $\forall\, s \in Q$

$$\lambda(s) = \sum_{\Theta} f(s|\theta)p^{(t)}(\theta) \to f(s|\theta^*), \quad \text{as } t \to \infty. \tag{47}$$

Proof: See the Appendix. ■

VI. Simulation Results

In this section, we conduct simulations to verify the performance of the proposed non-Bayesian social learning rule and recursive best response algorithms. We simulate an Indian Buffet restaurant with five dishes $\{r_1, r_2, r_3, r_4, r_5\}$ and five possible dish states $\theta_j \in \{1, 2, 3, 4, 5\}$. Each dish is randomly assigned with a state. After requesting a specific dish r_j, customer i can infer the quality of the dish and a signal $s_{i,j} \in \{1, 2, 3, 4, 5\}$ obeying the conditional distribution that

$$f_j(s_{i,j}|\theta_j) = \begin{cases} w, & \text{if } s_{i,j} = \theta_j \\ (1-w)/4, & \text{if } s_{i,j} \neq \theta_j. \end{cases} \tag{48}$$

The parameter w can be interpreted as the quality of the signal or customers' detection probability. When the signal quality w is close to 1, the customers' inferred signal is more likely to reflect the true dish state. Note that w must satisfy $w \geq 1/5$; otherwise, the true state can never be learned correctly. With the signals, customers can update their belief $\mathbf{P}$ cooperatively at the next time slot and then make their decisions sequentially. Once the ith customer makes the dish selection, he/she reveals his/her decisions to other customers. After all customers make their decisions, they begin to share the corresponding dishes they have requested. The customer is utility of requesting dish r_j is given by

$$u_{i,j} = \gamma_i \frac{s_{i,j}R}{N_j} - c_j \tag{49}$$

TABLE I
NE MATRIX $\mathbf{D}^*$

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
r_1	1	1	0	0	0	0	0	0	0	0
r_2	1	1	1	1	0	0	0	0	0	0
r_3	1	1	1	1	1	0	0	0	0	0
r_4	1	1	1	1	1	1	0	0	0	0
r_5	1	1	1	1	1	1	1	1	0	0

where γ_i is the utility coefficient for customer i since different customers may have different utilities regarding same reward, $s_{i,j}$ is a realization of dish quality, as well as the signal inferred by customer i, R is the basis award of requesting each dish as $R = 10$, N_j is the overall number of customers requesting dish r_j and c_j is the cost of requesting dish r_j as $\{c_j = 1, \forall j\}$. From (48) and (49), we can see that by requesting dish with higher level of state, e.g., $\theta_j = 5$, customers can obtain higher utilities. However, the dish state is unknown to customers and they have to estimate it through social learning. On the other hand, we can also see that the more customers requesting a same dish, the less utility each customer can obtain, which manifests the negative network externality.

A. Indian Buffet Game Without Budget Constraint

In this section, we evaluate the performance of the proposed best response algorithm for Indian Buffet Game without budget constraint. We first simulate the homogenous case to verify the threshold property of the NE matrix, i.e., Theorem 2, and the impact of different decision making orders on customers' utilities, i.e., making decisions earlier may have advantage. Then, we compare the performance of the proposed best response algorithm, i.e., Algorithm 1, with the performance of other algorithms in heterogeneous settings.

For the homogenous case, we set all customers' utility coefficients as $\gamma_i = 1$. The customers' prior belief regarding the dish state starts with a uniform distribution, i.e., $\{p_j^{(0)}(\theta) = 0.2, \forall j, \theta\}$. The dish state is set as $\Theta = [1, 2, 3, 4, 5]$, i.e., $\theta_j = j$, in order to verify different threshold structures for different dish states as illustrated in Theorem 2. At each time slot, we let customers sequentially make decisions according to Algorithm 1 and then update their belief according to the non-Bayesian learning rule. The game is played time slot by time slot until customers' belief $\mathbf{P}^{(t)}$ converges. In the first simulation, we set the number of customers as $N = 10$ to specifically verify the threshold structure of NE matrix. Table I shows the NE matrix $\mathbf{D}^*$ derived by Algorithm 1 after customers' belief $\mathbf{P}^{(t)}$ converges, where each column contains one customer's decisions $\{d_{i,j}, \forall j\}$ and each row contains all customers' decisions on one specific dish r_j, i.e., $\{d_{i,j}, \forall i\}$. From the table, we can see that once a customer does not request some specific dish, all the subsequent customers will not request that dish, which is consistent with the conclusion in Theorem 2. Moreover, since requesting the dish with higher level of state, e.g., $\theta_5 = 5$, can obtain more utilities, we can see that most customers decided to request dish r_5. From Table I, we can see that customers who make decisions early have advantage, e.g., customer 1 can request all dishes while customer 8 can only request one dish. Therefore, in the second simulation of the homogenous case, we dynamically adjust the order of decision making to ensure the fairness. In this simulation, we assume that there are five customers with a common utility coefficient $\gamma_i = 0.4$. In Fig. 3, we show all customers' utilities along with the simulation, where the order of decision making changes every 100 time slots. In the first 100 time slots, where the order of decision making is $1 \to 2 \to 3 \to 4 \to 5$, we can see that customer 1 obtains the highest utility and customer 4 and 5 receive 0 utility since they have not requested any dish. In the second 100 time slots, we reverse the decision making order to $5 \to 4 \to 3 \to 2 \to 1$, which results customer 1 and 2 receiving 0 utility. Therefore, by periodically changing the order of decision making, we can infer that the expected utilities of all customers will be the same after a period of time.

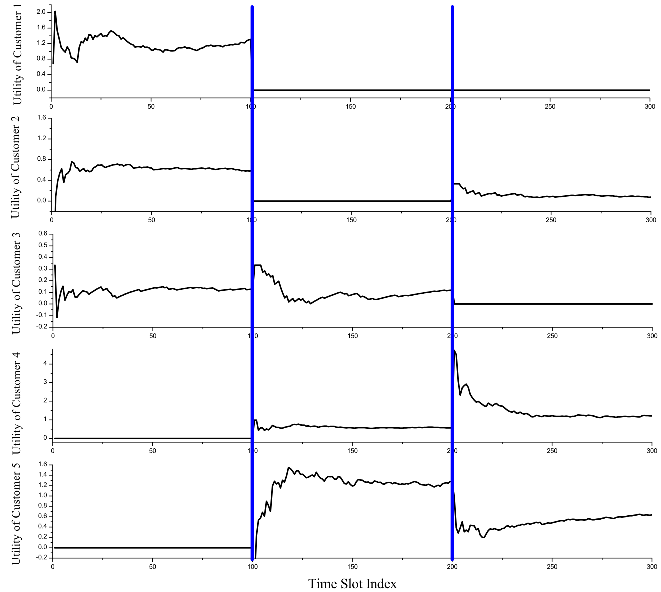


Fig. 3. Each customer's utility in homogenous case without budget constraint.

For the heterogeneous case, we randomize each customer's utility coefficient γ_i between 0 and 1 and set their prior belief as $\{p_j^{(0)}(\theta) = 0.2, \forall j, \theta\}$. In this simulation, we compare the performance in terms of customers' social welfare, which is defined as the total utilities of all customers, among different kinds of algorithms listed as follows.

1) *Best Response:* The proposed recursive best response algorithm in Algorithm 1 with non-Bayesian learning.
2) *Myopic:* At each time slot, customer i requests dishes according to current observation $\mathbf{n}_i = \{n_{i,j}, \forall j\}$ without social learning.
3) *Learning:* At each time slot, each customer requests dishes purely based on the updated belief $\mathbf{P}^{(t)}$ using non-Bayesian learning rule without considering the negative network externality.
4) *Random:* Each customer randomly requests dishes.

For the myopic and learning strategies, customer is expected utility of requesting dish r_j can be calculated by

$$U_{i,j}^{\mathrm{m}} = \sum_{\Theta} \sum_{Q} u_{i,j}(q_j, n_{i,j} + d_{i,j}) f_j(q_j|\theta_j) p_j^{(0)}(\theta_j) \quad (50)$$

$$U_{i,j}^{\mathrm{l}} = \sum_{\Theta} \sum_{Q} u_{i,j}(q_j, d_{i,j}) f_j(q_j|\theta_j) p_j^{(t)}(\theta_j). \quad (51)$$

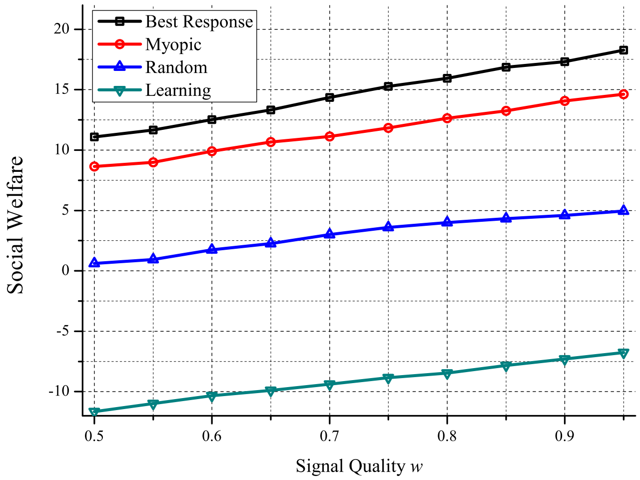


Fig. 4. Social welfare comparison without budget constraint.

With these expected utilities, both myopic and learning algorithms can be derived by (8). We can see that the myopic strategy does not take into account social learning while the learning strategy does not involve negative network externality. In the simulation, we average these four algorithms over hundreds of realizations. Fig. 4 shows the performance comparison result, where the x-axis is the signal quality w varying from 0.5 to 0.95 and y-axis is the social welfare averaged over hundreds of time slots. From the figure, we can see with the increase of signal quality, the social welfare keeps increasing for all algorithms. Moreover, we can also see that our best response algorithm performs the best while the learning algorithm performs the worst. This is because, with the learning algorithm, customers can gradually learn the true dish states and then request the dish without considering other customers' decisions. In such a case, too many customers may request the same dishes and each customer's utility is dramatically decreased due to the negative network externality. For the myopic algorithm, although customers can not learn the true dish states, by considering other customers' decisions, each customer can avoid requesting dishes which have been over-requested. Therefore, we can conclude that our proposed best response algorithm achieves the best performance through considering the negative network externality and using social learning to estimate the dish state.

B. Indian Buffet Game With Budget Constraint

In this section, we evaluate the performance of the proposed best response algorithm for Indian Buffet Game with budget constraint $L = 3$. Similar to the previous section, we start from the homogenous case, where all customers' utility coefficients are set as $\gamma_i = 1$. In the first simulation, we set all dish states as $\theta_j = 5$ to verify the property of the NE matrix illustrated in Theorem 4. Table II shows the NE matrix $\mathbf{D}^*$ derived by Algorithm 2. We can see that each dish has been requested by $N * L/M = 10 * 3/5 = 6$ customers, which is consistent with the conclusion in Theorem 4. In the second simulation, we dynamically change the order of customers' sequential decision making and illustrate each customer's utility along with simulation in Fig. 5, from which we can see

TABLE II
NE MATRIX $\mathbf{D}^*$

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
r_1	1	1	1	1	1	1	0	0	0	0
r_2	1	1	1	1	0	0	1	1	0	0
r_3	1	1	1	1	0	0	0	0	1	1
r_4	0	0	0	0	1	1	1	1	1	1
r_5	0	0	0	0	1	1	1	1	1	1

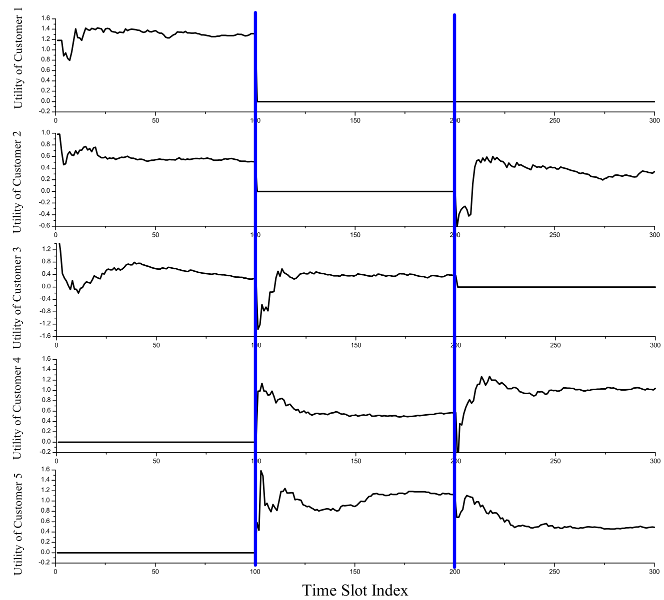


Fig. 5. Each customer's utility in homogenous case with budget constraint.

similar phenomenon to the Indian Buffet Game without budget constraint.

For the heterogeneous case, we randomize each customer's utility coefficient γ_i within [0, 1] and compare the performance of our proposed best response algorithm, i.e., Algorithm 2, with myopic, learning and random algorithms in terms of customers' social welfare. For the myopic, learning and random algorithms, same budget constraint is adopted, i.e., each customer can at most request 3 dishes. Fig. 6 shows the performance comparison result, from which we can see that our best response algorithm performs the best while the learning algorithm performs the worst.

C. Non-Bayesian Social Learning Performance

In this section, we evaluate the performance of the proposed non-Bayesian social learning rule. At the beginning of the simulation, we randomize the states of five dishes and assign customers' prior belief regarding each dish state with uniform distribution, i.e., $\{p_j(\theta = 0.2), \forall j, \theta\}$. After requesting the chosen dishes, each customer can infer signals following the conditional distribution defined in (48) with signal quality $w = 0.6$. Fig. 7 shows the learning curve of the Indian Buffet Game with and without budget constraint, respectively. The y-axis is the difference between customers' belief at each time slot $\mathbf{P}^{(t)}$ and the true belief $\mathbf{P}^o = \{\mathbf{p}_j = (p_j(\theta_j^*) = 1, p_j(\theta_j \neq \theta_j^*) = 0), \forall j\}$, which can be calculated by $||\mathbf{P}^{(t)} - \mathbf{P}^o||_2$. From the figure, we can see that

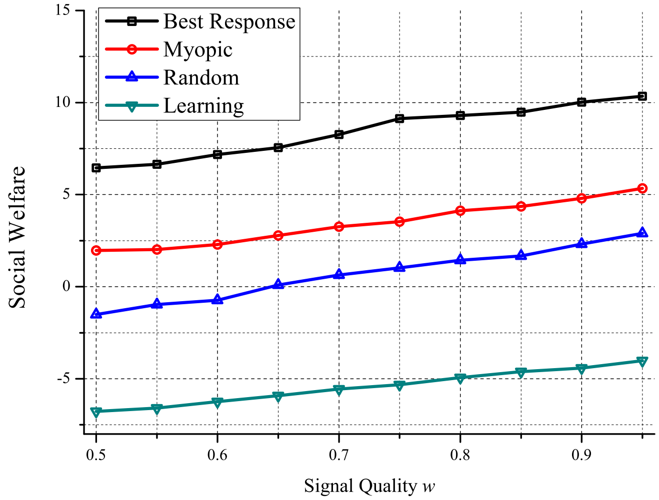


Fig. 6. Social welfare comparison with budget constraint.

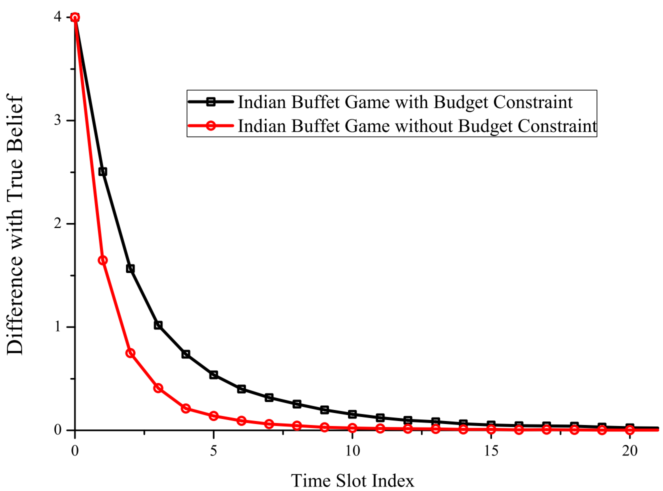


Fig. 7. Performance of the non-Bayesian social learning rule.

customers can learn the true dish states within 15 time slots. Moreover, the convergence rate of the case without budget constraint is faster than that of the case with budget constraint. This is because, due to the budget constraint, each customer requests fewer dishes at each time slot and thus can infer fewer signals regarding the dish state, which will inevitably slow down the customers' learning speed.

D. Application in Relay Selection of Cooperative Communication

In this section, we discuss an application of the Indian Buffet Game in the relay selection of cooperative communications. In the application, we consider a wireless network with N source nodes or users, which aim at sending their messages to the destination nodes. There are M potential relay nodes with different relay capabilities, given the different channel conditions, transmission power, and processing speed constraints. Each source node can select at most L relays in each time slot that help them relay their messages to the destination node. All the source nodes are assumed to be rational, i.e., each of them selects the relays that can maximize its own expected data rate. First, the more source nodes select the same relay in one time slot, the less

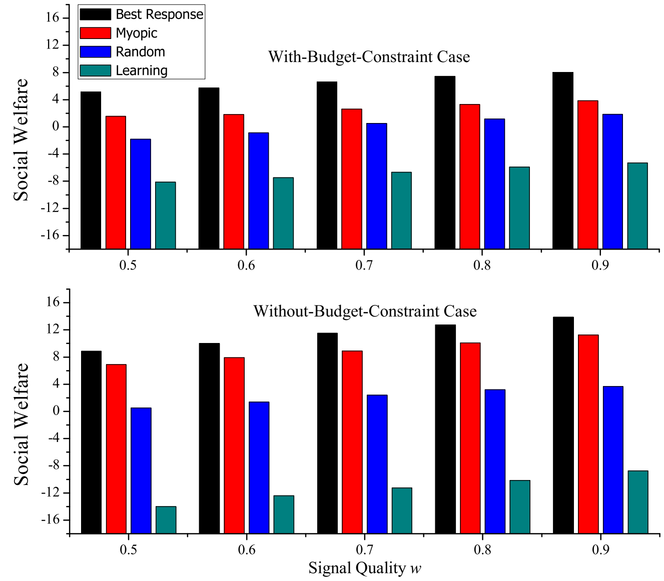


Fig. 8. Social welfare comparison in the application of relay selection.

throughput can be obtained by each source node, i.e., the negative externality. Second, since the source nodes may not exactly know the capacity of each relay node, they need to estimate the relay state by learning from the history and the current signals that reflect the relay properties. Third, the source nodes are not necessarily synchronized, which means that they may make the relay selection in a sequential manner. Considering these three properties, we can see that the proposed Indian Buffet Game is an ideal tool to formulate the relay selection problem, where the source nodes are the players and the relay nodes are corresponding to the dishes in the restaurant.

In the simulation, we set five source nodes and five relay nodes with five possible relay state $\theta_j \in \{1, 2, 3, 4, 5\}$. Each source node can receive a signal on the capacity of the selected relay nodes after data transmission, which obeys the distribution $f_j(s_{i,j}|\theta_j)$ as in (48). Assuming that the source nodes share the same relay in a time-division manner, we can define the utility of the ith source node selecting relay node j as

$$u_{i,j} = \frac{s_{i,j}\bar{g}_{i,j}}{N_j} - c_j \quad (52)$$

where $\bar{g}_{i,j}$ is the gain of the ith source node by selecting relay j which depends on the channel gain, N_j is the total number of source nodes sharing relay j and c_j is the cost of selecting relay j which can be considered as the price of relay service. With the utility definition, we can conduct simulation to evaluate the performance of relay selection with Indian Buffet Game and compare it with that of myopic strategy, random strategy and learning strategy. Fig. 8 shows the social welfare comparison results, including two cases: budget-constraint case where each source node can at most select $L = 3$ relays and without-constraint-case where each source node can at most select $L = 5$ relays. From the figure, we can see that, similar to the previous comparison results, our best response algorithm has the best performance in both cases. Therefore, our proposed Indian Buffet Game can be well applied in the relay selection of cooperative communications.

VII. Conclusion

In this paper, we proposed a general framework, called Indian Buffet Game, to study how users make multiple concurrent decisions under uncertain system state. We studied the Indian Buffet Game under two different scenarios: customers request multiple dishes without budget constraint and with budget constraint, respectively. We designed best response algorithms for both cases to find the subgame perfect NE, and discussed the simplified homogeneous cases to better understand the proposed Indian Buffet Game. We also designed a non-Bayesian social learning rule for customers to learn about the dish state and theoretically prove its convergence. Simulation results show that our proposed algorithms achieve much better performance than myopic, learning and random algorithms. Moreover, the proposed non-Bayesian learning algorithm can help customers learn the true system state with faster convergence speed.

We would like to emphasize that our focus in this paper is to study the interactions and decision making behaviors of agents in an uncertain negative-externality environment with the concept of social learning, which is a fundamental problem existing in the signal processing, wireless communication networks, and social networks. Meanwhile, our work in this paper also has limitations due to some assumptions in the analysis. One is that all the customers are assumed to share their belief information with each other, given that sharing the belief information can enhance the learning performance based on social learning. We assume that the rational customers have the incentive to do so. In addition, one of our ongoing works is to study a more general scenario where each user does not reveal his/her belief information and only action information can be observed. The other main assumption is that each customer knows all others' utility function forms in Algorithms 1 and 2. Note that Algorithms 1 and 2 are just designed to find the NE of the Indian Buffet Game. When each user have no knowledge about other users' utility function forms, he/she can take an expectation over the types of users according to some empirical user-type distribution, which is also one of our ongoing works.

The proposed Indian Buffet Game can be applied in various fields. One important application is the relay selection problem in cooperative communications, as shown in Section VI-D, Indian Buffet Game can well model and solve the problem of such relay selection in cooperative communications. Moreover, the multichannel sensing and access problem in cognitive radio networks can also be modeled by Indian Buffet Game, where each secondary user needs to learn the primary channel state and tries to access the channel with least secondary users, i.e., considering the negative externality [24]. For those problems, existing algorithms or models either only consider the negative externality or only consider the social learning, without integrating the learning and decision making with negative externality together. In such a case, the performance of those algorithms perform worse than the proposed Indian Buffet Game model, which can be seen from the comparison results in Figs. 4, 6, and 8. Therefore, the research value of Indian Buffet Game model lie in that it presents a general tool for agents to make optimal sequential decisions when confronted with uncertain system state and negative externality characteristic.

Appendix
Proof of Theorem 5

Proof of Lemma

Let us first define a probability triple $(\Omega, \mathcal{F}, \mathbb{P}^{\theta})$ for some specific dish state $\theta \in \Theta$, where Ω is the space containing sequences of realizations of the signals $s_i^{(t)} \in Q$, $\mathcal{F}$ is the σ-field generated by Ω, i.e., a set of subsets of Ω, and $\mathbb{P}^{\theta}$ is the probability measure induced over sample paths in Ω, i.e., $\mathbb{P}^{\theta} = \bigotimes_{t=1}^{\infty} f(\cdot|\theta)$. Moreover, we use $\mathbb{E}^{\theta}[\,\cdot\,]$ to denote the expectation operator associated with measure $\mathbb{P}^{\theta}$, and define $\mathcal{F}_t$ as the smallest σ-field generated by the past history of all customers' observations up to time slot t. To prove the weak convergence in (46), we start by showing that the belief sequence $\{p^{(t)}(\theta^*)\}$ converges to a positive number as $t \to \infty$ by the following lemmas.

Lemma 2: Suppose the true dish state is θ^*, all customers update their belief $\mathbf{p}$ according to the non-Bayesian learning rule in (41) and their prior belief $\mathbf{p}^{(0)}$ satisfies $p^{(0)}(\theta^*) > 0$, then, the belief sequence $\{p^{(t)}(\theta^*)\}$ converges to a positive number as $t \to \infty$.

Proof: From (41) and (42), we can see that if $p^{(t)}(\theta) > 0$, then $p^{(t+1)}(\theta) > 0$. Since the prior belief satisfies $p^{(0)}(\theta^*) > 0$, according to the method of induction, we have the belief sequence $\{p^{(t)}(\theta^*)\} > 0$.

According to (44), for the true dish state θ^*, we have

$$p^{(t+1)}(\theta^*) = p^{(t)}(\theta^*) + \frac{1}{N}\sum_{i=1}^{N} d_i^{(t+1)}\left(\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)} - 1\right) \times p^{(t)}(\theta^*). \quad (53)$$

By taking expectation over $\mathcal{F}_t$ on both sides of (53), we have

$$\mathbb{E}^{\theta^*}\left[p^{(t+1)}(\theta^*)|\mathcal{F}_t\right] = p^{(t)}(\theta^*) + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}^{\theta^*} \times \left[d_i^{(t+1)}\left(\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)} - 1\right)\Bigg|\mathcal{F}_t\right] p^{(t)}(\theta^*). \quad (54)$$

According to the time slot structure shown in Fig. 1, each customer's decision at time slot $t+1$ is made according to the belief updated at the end of last time slot t. In such a case, when given all the history information up to time slot t, $\mathcal{F}_t$, the decision $d_i^{(t+1)}$ is independent from the signal $s_i^{(t+1)}$ which is received at the end of time slot $t+1$. In such a case, we can separate the expectation in the second term of (54) as

$$\mathbb{E}^{\theta^*}\left[d_i^{(t+1)} \cdot \left(\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)} - 1\right)\Bigg|\mathcal{F}_t\right] = \mathbb{E}^{\theta^*}\left[d_i^{(t+1)}\Big|\mathcal{F}_t\right] \cdot \mathbb{E}^{\theta^*}\left[\left(\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)} - 1\right)\Bigg|\mathcal{F}_t\right]. \quad (55)$$

In (55), $\mathbb{E}^{\theta}\left[d_i^{(t+1)}\Big|\mathcal{F}_t\right] \geq 0$ since $d_i^{(t+1)}$ can only be 1 or 0. Moreover, since $g(x) = 1/x$ is a convex function, according to Jensen's inequality, we have

$$\mathbb{E}^{\theta^*}\left[\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}\middle|\mathcal{F}_t\right] \geq \left(\mathbb{E}^{\theta^*}\left[\frac{\lambda\left(s_i^{(t+1)}\right)}{f\left(s_i^{(t+1)}|\theta^*\right)}\middle|\mathcal{F}_t\right]\right)^{-1}$$
$$= \left(\sum_{Q}\frac{\lambda\left(s_i^{(t+1)}\right)}{f\left(s_i^{(t+1)}|\theta^*\right)}f\left(s_i^{(t+1)}|\theta^*\right)\right)^{-1} = 1. \quad (56)$$

In such a case, the equation in (55) is nonnegative, which means that in (54)

$$\mathbb{E}^{\theta^*}\left[p^{(t+1)}(\theta^*)|\mathcal{F}_t\right] \geq p^{(t)}(\theta^*). \quad (57)$$

Since customers' belief $p^{(t)}(\theta^*)$ is bounded within interval [0, 1], according to the martingale convergence theorem [25], we can conclude that the belief sequence $\{p^{(t)}(\theta^*)\}$ converges to a positive number as $t \to \infty$. ∎

Proof of Theorem 5

Proof: Let $\mathcal{N}^{(t+1)}$ denote the set of customers who request the dish at time slot $t+1$. In such a case, we can rewrite (53) as

$$p^{(t+1)}(\theta^*) = \frac{1}{|\mathcal{N}^{(t+1)}|}\sum_{i\in\mathcal{N}^{(t+1)}}\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}p^{(t)}(\theta^*) \quad (58)$$

where $|\cdot|$ means the cardinality. By taking logarithmic operation on both sides of (58) and utilizing the concavity of the logarithm function, we have

$$\log p^{(t+1)}(\theta^*) \geq \log p^{(t)}(\theta^*) + \frac{1}{|\mathcal{N}^{(t+1)}|}\sum_{i\in\mathcal{N}^{(t+1)}}\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}. \quad (59)$$

Then, by taking expectation over $\mathcal{F}_t$ on both sides of (59), we have

$$\mathbb{E}^{\theta^*}\left[\log p^{(t+1)}(\theta^*)|\mathcal{F}_t\right] - \log p^{(t)}(\theta^*) \geq \frac{1}{|\mathcal{N}^{(t+1)}|}\sum_{i\in\mathcal{N}^{(t+1)}}\mathbb{E}^{\theta^*}\left[\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}\middle|\mathcal{F}_t\right]. \quad (60)$$

As to the left hand of (60), according to Lemma 2, we know that $p^{(t)}(\theta^*)$ will converge as $t \to \infty$, and thus

$$\mathbb{E}^{\theta^*}\left[\log p^{(t+1)}(\theta^*)|\mathcal{F}_t\right] - \log p^{(t)}(\theta^*) \to 0. \quad (61)$$

As to the right hand of (60), it follows:

$$\mathbb{E}^{\theta^*}\left[\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}\middle|\mathcal{F}_t\right] = -\mathbb{E}^{\theta^*}\left[\log\frac{\lambda\left(s_i^{(t+1)}\right)}{f\left(s_i^{(t+1)}|\theta^*\right)}\middle|\mathcal{F}_t\right]$$
$$\geq -\log\mathbb{E}^{\theta^*}\left[\frac{\lambda\left(s_i^{(t+1)}\right)}{f\left(s_i^{(t+1)}|\theta^*\right)}\middle|\mathcal{F}_t\right]$$
$$= 0. \quad (62)$$

In such a case, combining (61) and (62), as $t \to \infty$, we have

$$0 \geq \frac{1}{|\mathcal{N}^{(t+1)}|}\sum_{i\in\mathcal{N}^{(t+1)}}\mathbb{E}^{\theta^*}\left[\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}\middle|\mathcal{F}_t\right] \geq 0. \quad (63)$$

By squeeze theorem [26], we have for $\forall\ i \in \mathcal{N}^{(t+1)}$, as $t \to \infty$

$$\mathbb{E}^{\theta^*}\left[\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)}\middle|\mathcal{F}_t\right] = \sum_{\Theta}f\left(s_i^{(t+1)}|\theta^*\right)\log\frac{f\left(s_i^{(t+1)}|\theta^*\right)}{\lambda\left(s_i^{(t+1)}\right)} \to 0. \quad (64)$$

According to Gibbs' inequality [27], the (64) converges to 0 if and only if as $t \to \infty$

$$\lambda\left(s_i^{(t+1)}\right) \to f\left(s_i^{(t+1)}|\theta^*\right), \quad \forall\ s_i^{(t+1)} \in Q. \quad (65)$$

This completes the proof of the theorem. ∎

References

[1] M. L. Katz and C. Shapiro, "Technology adoption in the presence of network externalities," *J. Polit. Econ*, vol. 94, no. 4, pp. 822–841, 1986.

[2] W. H. Sandholm, "Negative externalities and evolutionary implementation," *Rev. Econ. Stud.*, vol. 72, no. 3, pp. 885–915, 2005.

[3] Y. Yang, Y. Chen, C. Jiang, C. Wang, and K. J. R. Liu, "Wireless access network selection game with negative network externality," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5048–5060, Oct. 2013.

[4] H. V. Zhao, W. S. Lin, and K. J. R. Liu, *Behavior Dynamics in Media-Sharing Social Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[5] Y. Chen and K. J. R. Liu, "Understanding microeconomic behaviors in social networking: An engineering view," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 53–64, Mar. 2012.

[6] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *Rev. Econ. Stud.*, vol. 78, no. 4, pp. 1201–1236, 2011.

[7] V. Krishnamurthy and H. V. Poor, "Social learning and Bayesian games in multiagent signal processing: How do local and global decision makers interact?" *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 43–57, May 2013.

[8] V. Bala and S. Goyal, "Learning from neighbours," *Rev. Econ. Stud.*, vol. 65, no. 3, pp. 595–621, 1998.

[9] D. Gale and S. Kariv, "Bayesian learning in social networks," *Games Econ. Behav.*, vol. 45, no. 11, pp. 329–346, 2003.

[10] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dyn. Games Appl.*, vol. 1, no. 1, pp. 3–49, 2008.

[11] L. G. Epstein, J. Noor, and A. Sandroni, "Non-Bayesian learning," *B.E. J. Theor. Econ.*, vol. 10, no. 1, pp. 1–16, 2010.

[12] B. Golub and M. O. Jackson, "Naive learning in social networks and the wisdom of crowds," *Amer. Econ. J. Microecon.*, vol. 2, no. 1, pp. 112–149, 2010.

[13] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games Econ. Behav.*, vol. 76, no. 1, pp. 210–225, 2012.

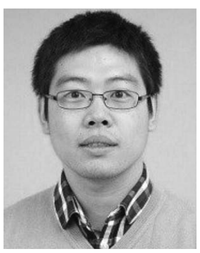
[14] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese restaurant game," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 898–901, Dec. 2012.

[15] B. Zhang, Y. Chen, C.-Y. Wang, and K. J. R. Liu, "Learning and decision making with negative externality for opportunistic spectrum access," in *Proc. IEEE Glob. Commun. Conf. (Globecom)*, Anaheim, CA, USA, 2012, pp. 1–6.

[16] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Sequential Chinese restaurant game," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.

[17] D. Aldous, I. Ibragimov, J. Jacod, and D. Aldous, "Exchangeability and related topics," *Lect. Notes Math.*, vol. 1117, no. 12, pp. 1–198, 1985.

[18] C. Jiang, Y. Chen, Y.-H. Yang, C.-Y. Wang, and K. J. R. Liu, "Dynamic Chinese restaurant game: Theory and application in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1960–1973, Apr. 2014.
[19] T. L. Griffiths and Z. Ghahramani, "The Indian buffet process: An introduction and review," *J. Mach. Learn. Res.*, vol. 12, pp. 1185–1224, Feb. 2011.
[20] A. Aradillas-Lopez, "Semiparametric estimation of a simultaneous game with incomplete information," *J. Econ.*, vol. 157, no. 2, pp. 409–431, 2010.
[21] A. Savikhin and R. M. Sheremeta, "Simultaneous decision-making in competitive and cooperative environments," *Econ. Inquiry*, vol. 51, no. 2, pp. 1311–1323, 2013.
[22] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1987.
[23] R. Swinburne, *Bayes's Theorem*. Oxford, U.K.: Oxford Univ. Press, 2002.
[24] C. Jiang, Y. Chen, and K. J. R. Liu, "Multi-channel sensing and access game: Bayesian social learning with negative network externality," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2176–2188, Apr. 2014.
[25] R. Isaac, "Shorter notes: A proof of the martingale convergence theorem," *Proc. Amer. Math. Soc.*, vol. 16, no. 4, pp. 842–844, 1965.
[26] J. Stewart, "Chapter 15.2 limits and continuity," in *Multivariable Calculus*. Emeryville, CA, USA: Key Curriculum Press, 2008.
[27] L. M. Surhone, M. T. Tennoe, and S. F. Henssonow, *Gibbs' Inequality*. Saarbrücken, Germany: Betascript, 2010.



Chunxiao Jiang (S'09–M'13) received the B.S. (Hons.) degree in information engineering from the Beijing University of Aeronautics and Astronautics (Beihang University), Beijing, China, and the Ph.D. (Hons.) degree from Tsinghua University (THU), Beijing, in 2008 and 2013, respectively.

From 2011 to 2013, he visited the Signals and Information Group, Department of Electrical and Computer Engineering, University of Maryland, College Park, College Park, MD, USA, with Prof. K. J. R. Liu. He is currently a Postdoctoral Fellow with the Electrical Engineering Department, THU, with Prof. Y. Ren. His current research interests include the applications of game theory and queuing theory in wireless communication and networking and social networks.

Dr. Jiang was the recipient of the Best Paper Award from the IEEE GLOBECOM in 2013, the Beijing Distinguished Graduated Student Award, the Chinese National Fellowship, and the Tsinghua Outstanding Distinguished Doctoral Dissertation in 2013.

Yan Chen (S'06–M'11–SM'14) received the bachelor's degree from the University of Science and Technology of China, Hefei, China, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, and the Ph.D. degree from the University of Maryland, College Park, College Park, MD, USA, in 2004, 2007, and 2011, respectively.

His current research interests include data science, network science, game theory, social learning and networking, as well as signal processing, and wireless communications.

Dr. Chen was the recipient of multiple honors and awards, including the Best Paper Award from the IEEE GLOBECOM in 2013, Future Faculty Fellowship, and the Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering, in 2010 and 2011, respectively, a finalist of Dean's Doctoral Research Award from the A. James Clark School of Engineering, University of Maryland, in 2011, and the Chinese Government Award for Outstanding Students Abroad in 2011.

Yang Gao (S'12) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2009. He is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Maryland, College Park, College Park, MD, USA.

His current research interests include strategic behavior analysis and incentive mechanism design for crowdsourcing and social computing.

Mr. Gao was the recipient of the Silver Medal of the 21st National Chinese Physics Olympiad, the Honor of the Excellent Graduate of Tsinghua University in 2009, the University of Maryland Future Faculty Fellowship in 2012, and the IEEE Globecom 2013 Best Paper Award.

K. J. Ray Liu (F'03) was as a Distinguished Scholar-Teacher at the University of Maryland, College Park, College Park, MD, USA, in 2007, where he is the Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative and cognitive communications, social learning and network science, information forensics and security, and green information and communications technology. He is an ISI highly cited author.

Dr. Liu was the recipient of the IEEE Signal Processing Society 2009 Technical Achievement Award and various best paper awards. He was also the recipient of various teaching and research recognitions from the University of Maryland including the University-Level Invention of the Year Award, the Poole and Kent Senior Faculty Teaching Award, the Outstanding Faculty Research Award, and the Outstanding Faculty Service Award, all from the A. James Clark School of Engineering. He is a Distinguished Lecturer of the IEEE Signal Processing Society. He is the Past President of the IEEE Signal Processing Society, where he has served as a President, Vice President—Publications, and as the Board of Governor. He was the Editor-in-Chief of the IEEE SIGNAL PROCESSING MAGAZINE and a Founding Editor-in-Chief of the *EURASIP Journal on Advances in Signal Processing*. He is a Fellow of the American Association for the Advancement of Science.