WILEY | Hindawi

Research Article

Optimal Preference Satisfaction for Conflict-Free Joint Decisions

Hiroaki Shinkawa [iD],[1] **Nicolas Chauvet** [iD],[1] **Guillaume Bachelier** [iD],[2] **André Röhm** [iD],[1] **Ryoichi Horisaki** [iD],[1] **and Makoto Naruse** [iD][1]

[1]*Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
[2]*Université Grenoble Alpes, CNRS, Institut Néel, Grenoble 38000, France*

Correspondence should be addressed to Hiroaki Shinkawa; gokukyukyoku555@g.ecc.u-tokyo.ac.jp

When presented with multiple choices, we all have a preference; we may suffer loss because of conflicts arising from identical selections made by other people if we insist on satisfying only our preferences. Such a scenario is applicable when a choice cannot be divided into multiple pieces owing to the intrinsic nature of the resource. Earlier studies examined how to conduct fair joint decision-making while avoiding decision conflicts in terms of game theory when multiple players have their own deterministic preference profiles. However, probabilistic preferences appear naturally in relation to the stochastic decision-making of humans, and therefore, we theoretically derive conflict-free joint decision-making that satisfies the probabilistic preferences of all individual players. To this end, we mathematically prove conditions wherein the deviation of the resultant chance of obtaining each choice from the individual preference profile (loss) becomes zero; i.e., the satisfaction of all players is appreciated while avoiding conflicts. Further, even in scenarios where zero-loss conflict-free joint decision-making is unachievable, we present approaches to derive joint decision-making that can accomplish the theoretical minimum loss while ensuring conflict-free choices. Numerical demonstrations are presented with several benchmarks.

1. Introduction

Since the industrial revolution, we have witnessed important technological advances; however, we are currently faced with issues such as the finiteness of resources, which has led to the need for developing strategies of competition or resource sharing [1, 2]. We need to consider the choices of other people or any entities and those of ours. In this regard, decision conflicts are a critical issue that needs to be addressed. For example, assume that many people or devices attempt to connect to the same mobile network simultaneously. In this case, the wireless resources available per person or device will become significantly smaller, and it will result in poor communication bandwidth or even zero connectivity attributable to congestion [3]. This example considers a case in which an individual suffers from choice conflict with others. Here, the division of resources among entities is allowed, but in other cases, a separable allocation

to multiple entities is not permitted. As another example, we consider a draft in a professional football league. Each club has specific players to pick in the draft; however, in principle, only a single club can sign with one player. Thus, decision-making conflicts must be resolved. These examples highlight the importance of achieving individual satisfaction while avoiding conflict.

Sharpley and Scarf proposed the top trading cycle (TTC) allocation algorithm [4] to address this problem. A typical example is a house-allocation problem; each of the n students ranked the houses in which they wanted to live from a list of n houses. In this scenario, the TTC allocates one house to each student. The solution provided by the TTC is known to be game-theoretic core-stable; that is, arbitrary numbers of students can swap houses with each other and still not obtain a better allocation than the current one. There are many other mechanisms that originate in TTC, which include those that allow indifference in preference profiles

[5–7] and those where agents can be allocated fractions of houses [8].

Although TTC works well when players have deterministic rankings, humans or artificial machines can also have probabilistic preferences. That is, unlike the house-allocation problem, an agent with probabilistic preferences would be unsatisfied by always receiving its top preference. Instead, over repeated allocations, the distribution of outcomes and their probabilistic preferences should be similar. Such probabilistic preferences can appear naturally in relation to the stochastic decision-making of humans. The multiarmed bandit (MAB) problem [9, 10] encompasses probabilistic preferences.

In the MAB problem, one player repeatedly selects and draws one of the several slot machines based on a certain policy. The reward for each slot machine is generated stochastically based on a probability distribution that the player does not know a priori. The MAB problem aims to maximize the cumulative reward obtained by a player under such circumstances. First, the player draws machines, which include those that do not generate many rewards, to accurately estimate the probability distribution of the reward that each machine follows; this is referred to as exploration. However, the extensive drawing of the highest reward probability machine increases the cumulative reward after the player is confident in the estimation. This is known as exploitation. To successfully solve the MAB problem, it is necessary to strike a balance between exploration and exploitation, especially in a dynamically changing environment; this is called the exploration-exploitation dilemma [11].

An example of probabilistic preferences can be observed in the context of the MAB problem. The softmax algorithm, a well-known stochastic decision-making method, is an efficient method to solve this problem [12]. If the empirical reward for each machine i at time t is $\mu_i(t)$, the probability that a player selects machine i is

$$s_i(t) = \frac{e^{\mu_i(t)/\tau}}{\sum_k e^{\mu_k(t)/\tau}}. \qquad (1)$$

Here, τ represents the temperature, which controls the balance between exploration and exploitation. Thus, the player has a probabilistic preference at each time step about which machine to select and with what percentage.

Furthermore, we can extend the MAB problem to a case involving multiple players, which is called the competitive MAB problem [13, 14]. In a competitive MAB problem, when multiple players simultaneously select the same machine and the machine generates a reward, the reward is distributed among the players. The aim is to maximize the collective rewards of players while ensuring equality [15]. Decision conflicts inhibit players from maximizing the total rewards in such scenarios.

For example, wireless communications suffer from such difficulties, wherein the simultaneous usage of the same band by multiple devices results in the performance degradation of individual devices, as shown in Figure 1. Players make a probabilistic decision to explore and exploit
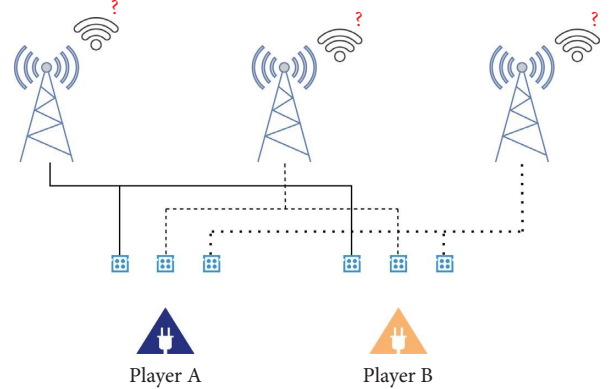

FIGURE 1: Internet connection through relays. Each player makes a probabilistic decision about which bandwidth to connect; however, when both select the same bandwidth, the line is shared, and there is a decrease in the transmission speed.

simultaneously because they do not know the speed of each bandwidth in advance (here, we assume that the speed changes probabilistically to reflect an uncertain environment). However, the line is shared if they choose the same bandwidth, and there is a decrease in the transmission speed.

For solving competitive MAB problems, Chauvet et al. theoretically and experimentally demonstrated the usefulness of quantum entanglement to avoid decision conflicts without direct communication [16, 17]. A powerful aspect of entanglement is that the optimization of a reward by one player leads to the optimization of the total rewards for the team; this is not easily achievable because if everyone attempts to draw the best machine, it can lead to decision conflicts that will diminish the total rewards in the competitive MAB problem. Further, Amakasu et al. proposed utilizing quantum interference such that the number of choices can be extended to an arbitrary number while perfectly avoiding decision conflicts [18]. A more general example is multiagent reinforcement learning [19, 20]. In this case, multiple agents can learn individually and make probabilistic choices while balancing exploration and exploitation at each time step. Successfully integrating these agents and making cooperative decisions without selection conflicts can help accelerate learning [21].

Given such motivations and backgrounds, it is important to clarify how the probabilistic preferences of individual players can be accommodated while avoiding choice conflicts. The question is whether it is possible to satisfy all player preferences while eliminating decision conflicts; if so, what are the conditions for realizing such requirements?

This study theoretically clarifies the condition under which joint decision-making probabilities exist such that the preference profiles of all players are satisfied perfectly. In other words, the condition that provides the loss (the deviation of the selection preference of each player from the resulting chance of obtaining each choice determined via the joint decision probabilities) becomes zero is clearly formulated. In addition, we derive joint decision-making probabilities that minimize the loss, even when such a condition is not satisfied.

A related problem of satisfying the needs of multiple players is a type of game theory called satisfaction games [22, 23]. Here, a minimum utility threshold is defined for each player, and a joint strategy is defined as a satisfaction equilibrium if all players obtain utility above the threshold. Like in other game theories, the payoff and reward or utility for each player can be defined by joint actions in the satisfaction games; however, in our study, there is no such external payoff. Instead, we define satisfaction as the correspondence between players' probabilistic preferences and the selection probability of each choice as a result of joint selection. In addition, players learn the optimal action by observing the outcomes of the environment and updating their actions in satisfaction games. However, in our study, we consider joint selection resulting from deliberation by players cooperating with each other rather than learning over multiple trials.

The consensus reaching process is another field that is closely related to our research [24–26]. In this area, consensus reaching is achieved by aggregating the preferences of multiple individual players using simple methods, such as weighted sums, to obtain a collective preference, and then, by calculating the distance between the collective preference and individual preferences and by adjusting the individual preferences through feedback [27]. The players continue this loop until they reach consensus. Meanwhile, our study focuses not on modifying individual preferences but on the aggregation of the preferences of multiple players in this process.

2. Theory

2.1. Problem Formulation. This section introduces the formulation of the study problem. To begin with, we examine cases where the number of players or agents is two: Player A and Player B. There are N choices for each player; these choices are called arms in the literature on the MAB problem, where N is a natural number greater than or equal to two.

Each player selects a choice probabilistically depending on his/her preference probability. Here, the preference of Player A is represented by

$$\mathbf{A} = \begin{pmatrix} A_1 & A_2 & \ldots & A_N \end{pmatrix}. \tag{2}$$

Similarly, the preference of player B is given by

$$\mathbf{B} = \begin{pmatrix} B_1 & B_2 & \ldots & B_N \end{pmatrix}. \tag{3}$$

These preferences have the typical properties of probabilities:

$$A_1 + A_2 + \cdots + A_N = B_1 + B_2 + \cdots + B_N = 1, \quad A_i \geq 0, \; B_i \geq 0 \; (i = 1, 2, \ldots, N). \tag{4}$$

The upper side of Figure 2 schematically illustrates a scenario where players A and B each have their own preferences in a four-arm case.

Now, we introduce the joint selection probability matrix, which is given as

$$\mathbf{P} = \begin{pmatrix} 0 & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & \cdots & \cdots & 0 \end{pmatrix}. \tag{5}$$

The nondiagonal element $p_{i,j}$, with $i \neq j$, denotes the probability that player A selects arm i and player B selects arm j. The diagonal terms are all zero because we consider nonconflict choices. Here, the summation of $p_{i,j}$ over all *is* and *js* is unity and $p_{i,j} \geq 0$.

$$\sum_{i,j} p_{i,j} = 1, \quad p_{i,j} \geq 0. \tag{6}$$

Our interest is to find $\mathbf{P}$ that meets the preferences of both players A and B. We formulate the degree of satisfaction of the players to incorporate this perspective as indicated in the following way. By summing up $p_{i,j}$s row-by-row in the joint selection probability matrix, a preference profile given by

$$\pi_A(i) = \sum_j p_{i,j}, \tag{7}$$

is obtained. $\pi_A(i)$ represents the probability that player A selects arm i as a result of the joint selection probability matrix $\mathbf{P}$. We call $\pi_A(i)$ "satisfied preference" (the link with preference A_i will be discussed hereafter).

$$\sum_i \pi_A(i) = 1, \tag{8}$$

holds based on the definition of the joint selection probability matrix. This structure is illustrated by the red-shaded elements in the second row of the joint selection probability matrix shown in Figure 2.

Similarly, satisfied preferences along the columns of $\mathbf{P}$

$$\pi_B(j) = \sum_i p_{i,j}, \tag{9}$$

represent the probability that player B selects arm j because of the joint selection probability matrix $\mathbf{P}$.

Our aim is to find the optimal $p_{i,j}$, wherein

$$\pi_A(i) \approx A_i, \quad \pi_B(j) \approx B_j, \tag{10}$$

holds for all $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, N$; that is, the players' preferences (A_i, B_j) are the same or close to the satisfied preferences $(\pi_A(i), \pi_B(j))$.

We define loss L akin to an L_2-norm to quantify the degree of satisfaction.

$$L = \sum_i \left(\pi_A(i) - A_i \right)^2 + \sum_j \left(\pi_B(j) - B_j \right)^2, \tag{11}$$

which comprises the sum of squares of the gap between the preferences and the satisfied preferences of the two players.

We consider the scenario shown in Figure 2 as an example. The preferences of players A and B are

$$
\begin{aligned}
A_1 &= 0.1, \\
A_2 &= 0.2, \\
A_3 &= 0.3, \\
A_4 &= 0.4, \\
B_1 &= 0.3, \\
B_2 &= 0.2, \\
B_3 &= 0.2, \\
B_4 &= 0.3.
\end{aligned}
\tag{12}
$$

Now, we consider the following joint selection probability matrix $\mathbf{P}$.

$$
\mathbf{P} = \begin{pmatrix}
0 & 0 & 0 & 0.1 \\
0 & 0 & 0.1 & 0.1 \\
0 & 0.2 & 0 & 0.1 \\
0.3 & 0 & 0.1 & 0
\end{pmatrix}.
\tag{13}
$$

By taking the sum of the first row, we can calculate the probability that player A chooses arm 1 as a result of this matrix.

$$\pi_A(1) = 0 + 0 + 0 + 0.1 = 0.1. \tag{14}$$

Similarly, if we calculate all satisfied preferences, we obtain

$$
\begin{aligned}
\pi_A(1) &= 0.1, \\
\pi_A(2) &= 0.2, \\
\pi_A(3) &= 0.3, \\
\pi_A(4) &= 0.4, \\
\pi_B(1) &= 0.3, \\
\pi_B(2) &= 0.2, \\
\pi_B(3) &= 0.2, \\
\pi_B(4) &= 0.3.
\end{aligned}
\tag{15}
$$

All preferences are equal to the original preferences of the players, which implies that loss L is zero. In other words, this joint selection probability matrix $\mathbf{P}$ completely satisfies the preferences of the player.

In the following sections, we prove that the loss defined previously can be zero; in other words, perfect satisfaction for the players can indeed be realized under certain conditions. In addition, even in cases where zero loss is not achievable, we can systematically derive a joint selection probability matrix that ensures a minimum loss.

Now, we define the popularity S_i for each arm and propose three theorems and one conjecture based on the values of S_i.

Definition 1. Popularity S_i is defined as the sum of the preferences of players A and B for arm i.

$$S_i := A_i + B_i \quad (i = 1, 2, \ldots, N). \tag{16}$$

Because preferences A_i and B_i are probabilities, it holds that

$$\sum_i S_i = \sum_i A_i + \sum_i B_i = 2. \tag{17}$$

Hereafter, the minimum loss is denoted by $L_{\min}$.

$$L_{\min} = \min_{\mathbf{P}} \{L\}. \tag{18}$$

Meanwhile, the loss function can be defined in other forms such as the Kullback–Leibler (KL) divergence, which is also known as relative entropy [28]. Extended discussions on such metrics should be considered in future studies; we consider that the fundamental structure of the problem under study will not change significantly.

Key notations adopted in this paper are summarized in Table 1.

2.2. Theorem 1

2.2.1. Statement

Theorem 1. *We assume that all popularities S_i are smaller than or equal to one. Then, it is possible to construct a joint selection probability matrix that makes the loss L equal to zero.*

$$\forall i; S_i \le 1 \Rightarrow L_{\min} = 0. \tag{19}$$

2.2.2. Auxiliary Lemma.
To prove Theorem 1, we first prove a lemma in which the problem settings are modified slightly. In the original problem, we treated the preference of each player as a probability. In this modified problem, however, their preferences do not have to be probabilities as long as they are nonnegative, and the sum of the preferences for each player is given by a constant T, which we refer to as the total preference. Their preferences are called $\widehat{A}$ and $\widehat{B}$. We will put a hat over each notation to avoid confusion about which definition we are referring to between the original and modified problems.
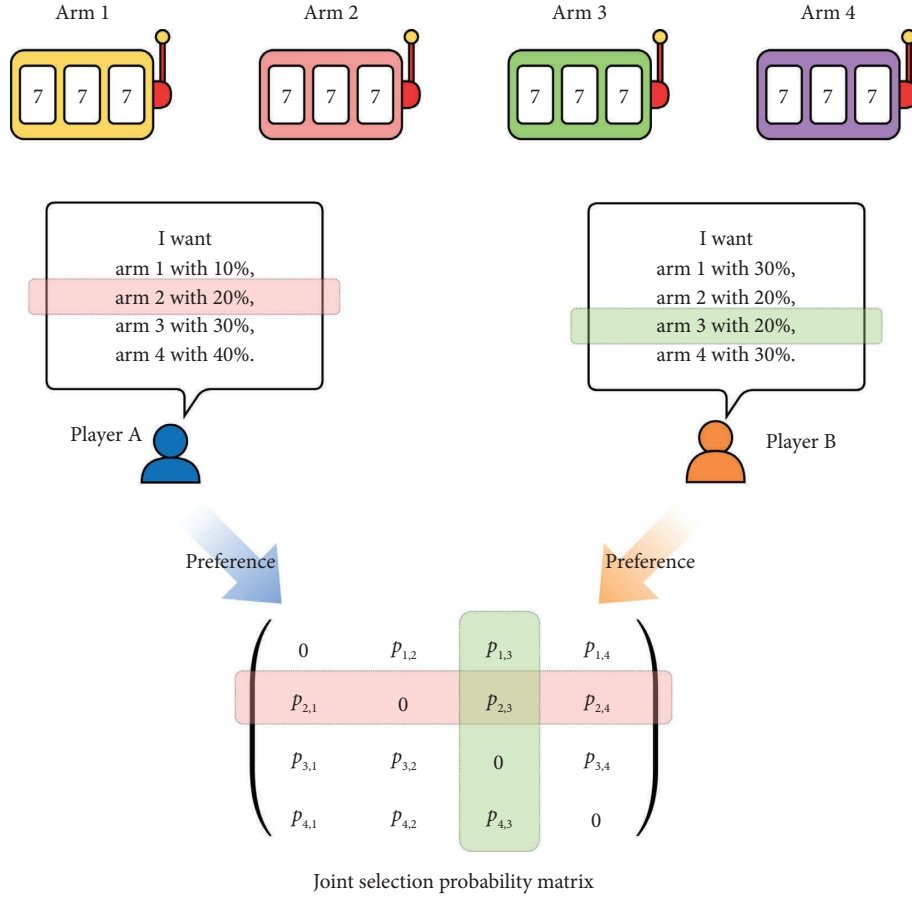
FIGURE 2: Problem settings. The sum of each row should be close to the corresponding preference of player A, and the sum of each column should be close to the corresponding preference of player B.

TABLE 1: Summary of key notations.

Notation	Description
N	Number of arms/choices
A_i, B_i	Preference of player on arm i
$\pi_A(i), \pi_B(i)$	Satisfied preference; probability of each player selecting arm i as a result of the joint selection probability matrix
L	Loss to measure the degree of satisfaction
S_i	Popularity; sum of the preferences of players A and B for arm i
T	Total preference used in Section 2.2; sum of the preference of each player
$\widehat{A}_i, \widehat{B}_i, \widehat{p}_{i,j}, \ldots$	Symbols used in the revised problem in Section 2.2
$A_i^{\dagger}, B_i^{\dagger}, p_{i,j}^{\dagger}, \ldots$	Symbols used in the general problem in Section 2.4

$$\widehat{\mathbf{A}} = \left(\widehat{A}_1 \widehat{A}_2 \ldots \widehat{A}_N \right),$$
$$\widehat{\mathbf{B}} = \left(\widehat{B}_1 \widehat{B}_2 \ldots \widehat{B}_N \right). \tag{20}$$

$$\widehat{A}_1 + \widehat{A}_2 + \cdots + \widehat{A}_N = \widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_N = T, \quad \widehat{A}_i \geq 0, \ \widehat{B}_i \geq 0 \ (i = 1, 2, \ldots, N). \tag{21}$$

We still refer to

$$\widehat{\mathbf{P}} = \begin{pmatrix} 0 & \widehat{p}_{1,2} & \cdots & \widehat{p}_{1,N} \\ \widehat{p}_{2,1} & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{p}_{N,1} & \cdots & \cdots & 0 \end{pmatrix}. \tag{22}$$

as a joint selection probability matrix, although technically, each $\widehat{p}_{i,j}$ is no longer a probability but a ratio of each joint selection occurring, and the sum of all entries is T, instead of 1. The other notations are defined in a manner similar to the original problem.

$$\text{Satisfied Preference: } \widehat{\pi}_A(i) = \sum_j \widehat{p}_{i,j},$$

$$\widehat{\pi}_B(j) = \sum_i \widehat{p}_{i,j},$$

$$\text{Loss: } \widehat{L} = \sum_i (\widehat{\pi}_A(i) - \widehat{A}_i)^2 + \sum_j (\widehat{\pi}_B(j) - \widehat{B}_j)^2, \tag{23}$$

$$\widehat{L}_{\min} = \min_{\widehat{\mathbf{P}}} \{\widehat{L}\},$$

$$\text{Popularity: } \widehat{S}_i = \widehat{A}_i + \widehat{B}_i.$$

Lemma 1. *We assume that all popularities $\widehat{S}_i$ are smaller than or equal to the total preference T. Then, it is possible to construct a joint selection probability matrix that makes the loss $\widehat{L}$ equal to 0.*

$$\forall i; \widehat{S}_i \le T \Rightarrow \widehat{L}_{\min} = 0. \tag{24}$$

Theorem 1 is a special case of Lemma 1 when $T = 1$.

2.2.3. Outline of the Proof. We prove Lemma 1 using mathematical induction. In Section 2.2.4, we show that Lemma 1 holds for $N = 2$ and $N = 3$. In Sections 2.2.5 and 2.2.6, we suppose that Lemma 1 has been proven for N arms, and then, we show that Lemma 1 also holds for $N + 1$ arms. In Sections 2.2.5, we assume the existence of $p_{i,j}$s that satisfy certain conditions and prove that perfect satisfaction for the players is achievable using these $p_{i,j}$s. Then, in Section 2.2.6, we verify their existence.

2.2.4. When $N = 2$ and When $N = 3$. When $N = 2$,

$$\widehat{S}_1 = \widehat{S}_2 = T, \tag{25}$$

is the only case in which the assumption of Lemma 1, i.e., $\forall i; \quad \widehat{S}_i \le T$ is fulfilled. Due to the constraint (21),

$$\widehat{A}_2 = T - \widehat{A}_1,$$
$$\widehat{B}_1 = T - \widehat{A}_1, \tag{26}$$
$$\widehat{B}_2 = \widehat{A}_1.$$

In this case,

$$\widehat{\mathbf{P}} = \begin{pmatrix} 0 & \widehat{A}_1 \\ \widehat{A}_2 & 0 \end{pmatrix}, \tag{27}$$

makes the loss equal to zero.

When $N = 3$, the loss becomes zero when we set the values in the joint selection probability matrix as

$$\widehat{\mathbf{P}} = \begin{pmatrix} 0 & \widehat{p}_{1,2} & -\widehat{p}_{1,2} + \widehat{A}_1 \\ T - \widehat{p}_{1,2} - \widehat{A}_3 - \widehat{B}_3 & 0 & \widehat{p}_{1,2} - \widehat{A}_1 + \widehat{B}_3 \\ \widehat{p}_{1,2} + \widehat{A}_3 - \widehat{B}_2 & -\widehat{p}_{1,2} + \widehat{B}_2 & 0 \end{pmatrix}. \tag{28}$$

The satisfied preferences for players A and B via $\widehat{\mathbf{P}}$ result in $(\widehat{A}_1, \widehat{A}_2, \widehat{A}_3)$ and $(\widehat{B}_1, \widehat{B}_2, \widehat{B}_3)$, which perfectly match the preferences of players A and B, respectively.

However, all elements in $\widehat{\mathbf{P}}$ must be nonnegative.

$$\widehat{p}_{1,2} \ge 0, \quad -\widehat{p}_{1,2} + \widehat{A}_1 \ge 0, \ T - \widehat{p}_{1,2} - \widehat{A}_3 - \widehat{B}_3 \ge 0,$$
$$\widehat{p}_{1,2} - \widehat{A}_1 + \widehat{B}_3 \ge 0, \ \widehat{p}_{1,2} + \widehat{A}_3 - \widehat{B}_2 \ge 0, \ -\widehat{p}_{1,2} + \widehat{B}_2 \ge 0. \tag{29}$$

To summarize these inequalities, the following inequality holds:

$$\max\{0, \widehat{A}_1 - \widehat{B}_3, -\widehat{A}_3 + \widehat{B}_2\} \le \widehat{p}_{1,2} \le \min\{\widehat{A}_1, \widehat{B}_2, T - \widehat{A}_3 - \widehat{B}_3\}. \tag{30}$$

If a nonnegative $\widehat{p}_{1,2}$ that satisfies (30) exists, $\widehat{\mathbf{P}}$ given in (28) is a valid joint selection probability matrix, which makes the loss zero. In other words, if we can prove that $\widehat{p}_{1,2}$, which is in the range of (30), exists, regardless of $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{B}}$, Lemma 1 holds for $N = 3$.

To prove that a nonnegative $\widehat{p}_{1,2}$ that satisfies (30) always exists, we consider the following three cases depending on the value of the left-hand term of (30). In each case, we show that the three candidates for the right-hand term of (30) are greater than the left-hand term.

[Case 1] When $\max\{0, \widehat{A}_1 - \widehat{B}_3, -\widehat{A}_3 + \widehat{B}_2\} = 0$.

$0 \le \widehat{A}_1$, $0 \le \widehat{B}_2$ are evident from the definitions in (21). In addition, $0 \le T - \widehat{A}_3 - \widehat{B}_3$ holds because of the assumption of lemma $\widehat{S}_3 = \widehat{A}_3 + \widehat{B}_3 \le T$. Thus,

$$0 \le \min\{\widehat{A}_1, \widehat{B}_2, T - \widehat{A}_3 - \widehat{B}_3\}. \tag{31}$$

[Case 2] When $\max\{0, \widehat{A}_1 - \widehat{B}_3, -\widehat{A}_3 + \widehat{B}_2\} = \widehat{A}_1 - \widehat{B}_3$.
First, the definition (21) guarantees the following:

$$\widehat{A}_1 - \widehat{B}_3 \le \widehat{A}_1. \tag{32}$$

Next, using the assumption $\widehat{S}_1 = \widehat{A}_1 + \widehat{B}_1 \le T$,

$$\widehat{A}_1 - \widehat{B}_3 \le (T - \widehat{B}_1) - \widehat{B}_3 = B_2. \tag{33}$$

Finally,

$$\widehat{A}_1 - \widehat{B}_3 = (T - \widehat{A}_2 - \widehat{A}_3) - \widehat{B}_3 \le T - \widehat{A}_3 - \widehat{B}_3. \tag{34}$$

Therefore,

$$\widehat{A}_1 - \widehat{B}_3 \le \min\{\widehat{A}_1, \widehat{B}_2, T - \widehat{A}_3 - \widehat{B}_3\}. \tag{35}$$

[Case 3] When $\max\{0, \widehat{A}_1 - \widehat{B}_3, -\widehat{A}_3 + \widehat{B}_2\} = -\widehat{A}_3 + \widehat{B}_2$.
$\widehat{S}_2 = \widehat{A}_2 + \widehat{B}_2 \le T$ implies the following equation:

$$-\widehat{A}_3 + \widehat{B}_2 \le -\widehat{A}_3 + (T - \widehat{A}_2) = A_1. \tag{36}$$

In addition, definition (21) guarantees $-\widehat{A}_3 + \widehat{B}_2 \le \widehat{B}_2$. Finally,

$$-\widehat{A}_3 + \widehat{B}_2 = -\widehat{A}_3 + (T - \widehat{B}_1 - \widehat{B}_3) \le T - \widehat{A}_3 - \widehat{B}_3, \tag{37}$$

holds because $\widehat{B}_1 \ge 0$. Hence,

$$-\widehat{A}_3 + \widehat{B}_2 \le \min\{\widehat{A}_1, \widehat{B}_2, T - \widehat{A}_3 - \widehat{B}_3\}. \tag{38}$$

From (31), (35), and (38) for any $\widehat{\mathbf{A}}, \widehat{\mathbf{B}}$,

$$\max\{0, \widehat{A}_1 - \widehat{B}_3, -\widehat{A}_3 + \widehat{B}_2\} \le \min\{\widehat{A}_1, \widehat{B}_2, T - \widehat{A}_3 - \widehat{B}_3\}. \tag{39}$$

Therefore, it is evident that $\widehat{p}_{1,2}$ that satisfies (30) exists. Hence, $\widehat{\mathbf{P}}$ given in (28) is a valid joint selection probability matrix that makes the loss $\widehat{L}$ zero.

First, we suppose that Assumption 1 is valid and show the way to construct a joint selection probability matrix that makes the loss $\widehat{L}$ zero, using these $\widehat{p}_{i,j}s$. Later, we will prove that Assumption 1 is indeed correct.

Here, we show that, using the $\widehat{p}_{i,j}s$ assumed to exist in Assumption 1, we can make the loss $\widehat{L}$ equal to zero. From conditions (41) and (42), $(\pi_A(K) - A_K)^2 + (\pi_B(K) - B_K)^2$, which are terms in the definition of $\widehat{L}$ regarding arm K, are zero.

2.2.5. Induction. In this and the following sections, we suppose that Lemma 1 has been proven to hold when there are N arms ($N \ge 3$). Here, we show that Lemma 1 also holds for $N + 1$ case. In the following argument, the arm with the lowest $\widehat{S}_i$ is denoted as arm K.

$$\min\{\widehat{S}_i\} = \widehat{S}_K. \tag{40}$$

Now, we make the following assumption, which focuses only on the values in the Kth row or Kth column.

Assumption 1. There exist $\widehat{p}_{K,1}, \widehat{p}_{K,2}, \ldots, \widehat{p}_{K,N+1}$, $\widehat{p}_{1,K}, \widehat{p}_{2,K}, \ldots, \widehat{p}_{N+1,K}$ that satisfy all of the following conditions (41)–(45).
The sum of the Kth row is equal to $\widehat{A}_K$.

$$\sum_{j=1}^{N+1} \widehat{p}_{K,j} = \widehat{A}_K. \tag{41}$$

The sum of the Kth column is equal to $\widehat{B}_K$.

$$\sum_{i=1}^{N+1} \widehat{p}_{i,K} = \widehat{B}_K. \tag{42}$$

The sum of the jth column without the Kth row is nonnegative.

$$\sum_{i \ne K} \widehat{p}_{i,j} \ge 0 \Leftrightarrow \widehat{p}_{K,j} \le \widehat{B}_j, \quad (j = 1, 2, \ldots, K-1, K+1, \ldots, N+1). \tag{43}$$

The sum of the ith row without the Kth column is nonnegative.

$$\sum_{j \ne K} \widehat{p}_{i,j} \ge 0 \Leftrightarrow \widehat{p}_{i,K} \le \widehat{A}_i, \quad (i = 1, 2, \ldots, K-1, K+1, \ldots, N+1). \tag{44}$$

In the gray-shaded area of the joint selection probability matrix in (48), all remaining popularities are less than or equal to the remaining total preference. We note that $\widehat{S}_K = \min\{\widehat{S}_i\}$.

$$\left(\widehat{A}_i - \widehat{p}_{i,K}\right) + \left(\widehat{B}_i - \widehat{p}_{K,i}\right) \le T - \widehat{S}_K \quad (i = 1, 2, \ldots, K-1, K+1, \ldots, N+1). \tag{45}$$

$$\left(\pi_A(K) - A_K\right)^2 + \left(\pi_B(K) - B_K\right)^2 = 0. \tag{46}$$

Next, we consider the loss for the remaining part of the joint selection probability matrix in the gray-shaded region described in (48), which is defined as

$$\widehat{L}_{\text{rem}} = \widehat{L} - \left\{\left(\widehat{\pi}_A(K) - \widehat{A}_K\right)^2 + \left(\widehat{\pi}_B(K) - \widehat{B}_K\right)^2\right\}. \tag{47}$$

$$\begin{pmatrix} 0 & * & \cdots & p_{1,K} & \cdots & * & * \\ * & 0 & \cdots & p_{2,K} & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots \\ p_{K,1} & p_{K,2} & \cdots & 0 & \cdots & p_{K,N} & p_{K,N+1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & p_{N,K} & \cdots & 0 & * \\ * & * & \cdots & p_{N+1,K} & \cdots & * & 0 \end{pmatrix}$$

$$(48)$$

$\widehat{L}_{\mathrm{rem}}$ represents the loss corresponding to a problem in which the preference setting is described in Table 2. $\widehat{A}_i^*$ and $\widehat{B}_i^*$ are defined as follows:

$$\widehat{A}_i^* = \widehat{A}_i - \widehat{p}_{i,K}, \quad \widehat{B}_j^* = \widehat{B}_j - \widehat{p}_{K,j}. \tag{49}$$

Now, we prove that the optimal $\widehat{L}_{\mathrm{rem}}$ is zero. A_i^* and B_i^* fulfill the requirements for the preference; that is, the preferences need to be nonnegative because $p_{i,K}s$ and $p_{K,j}s$ follow (43) and (44) and the total preferences for both players are the same. In this section, we suppose we have proven that Lemma 1 holds when there are N arms; furthermore, because of (45), all remaining popularities $\widehat{S}_i^* := \widehat{A}_i^* + \widehat{B}_i^* = (\widehat{A}_i - \widehat{p}_{i,K}) + (\widehat{B}_i - \widehat{p}_{K,i})$ are smaller than or equal to the remaining total preference $T - \widehat{S}_K$. Thus, Lemma 1 in the case of N arms verifies

$$\min\{\widehat{L}_{\mathrm{rem}}\} = 0. \tag{50}$$

From (46) and (50), with $\widehat{p}_{K,1}, \widehat{p}_{K,2}, \ldots, \widehat{p}_{K,N+1}, \widehat{p}_{1,K}, \widehat{p}_{2,K}, \ldots, \widehat{p}_{N+1,K}$ and the optimal $\widehat{p}_{i,j}s$ for the gray-shaded region,

$$\widehat{L}_{\min} = \min\{\widehat{L}_{\mathrm{rem}}\} + \left\{ \left(\widehat{\pi}_A(K) - \widehat{A}_K\right)^2 + \left(\widehat{\pi}_B(K) - \widehat{B}_K\right)^2 \right\} = 0. \tag{51}$$

Therefore, if Assumption 1 is correct, Lemma 1 holds for $N + 1$ arms.

2.2.6. Verification of Assumption 1. In this section, we prove that Assumption 1 is correct. We refer to the arm with the highest $\widehat{S}_i$ as arm V ($V \neq K$).

$$\max\{\widehat{S}_i\} = \widehat{S}_V. \tag{52}$$

First, by contradiction, we prove that there is at most one arm that violates

$$\widehat{S}_i \leq T - \widehat{S}_K, \tag{53}$$

and if there is such an arm, arm V is the one that breaks (53). Note that all the other arms, which satisfy (53), follow (45).

We assume that there are two arms that do not satisfy (53) (we call them arm V_1 and arm V_2), and we prove that this assumption leads to a contradiction.

$$\widehat{S}_{V_1} > T - \widehat{S}_K, \quad \widehat{S}_{V_2} > T - S_K. \tag{54}$$

If we add each side,

$$\widehat{S}_{V_1} + \widehat{S}_{V_2} > 2T - 2\widehat{S}_K \Leftrightarrow 2T < \widehat{S}_{V_1} + \widehat{S}_{V_2} + 2\widehat{S}_K. \tag{55}$$

As $\min\{\widehat{S}_i\} = \widehat{S}_K$ and $N \geq 3$,

$$\widehat{S}_{V_1} + \widehat{S}_{V_2} + 2\widehat{S}_K \leq \sum_{i=1}^{N+1} \widehat{S}_i = 2T. \tag{56}$$

Together with (55), we obtain the contradiction

$$2T < 2T. \tag{57}$$

Therefore, by contradiction, at most one arm violates (53). Such an arm follows

$$\widehat{S}_i > T - \widehat{S}_K, \tag{58}$$

and has the highest $\widehat{S}_i$ since other popularities are no bigger than $T - \widehat{S}_K$. We call it arm V so that $\max\{\widehat{S}_i\} = \widehat{S}_V$.

Now, we can systematically determine $\widehat{p}_{i,j}s$ in the Kth row or the Kth column to show that Assumption 1 is correct; that is, $\widehat{p}_{i,j}s$ fulfill the conditions (41)–(45). We cannot arbitrarily fill in the values in the Kth row or the Kth column such that their sum is $\widehat{A}_K$ and $\widehat{B}_K$, respectively; instead, we must consider the preferences of the other arms. (43)–(45) give us boundaries for $\widehat{p}_{K,j}s$ and $\widehat{p}_{i,K}s$ to exist. We must not only ensure that $\widehat{p}_{K,j}$ and $\widehat{p}_{i,K}$ do not exceed these boundaries, but we must also ensure that a sufficient probability is assigned to the most popular arm V; otherwise, the left-hand side of (45) can sometimes be greater than the right-hand side for arm V. We consider the following three cases depending on the size relation of $\widehat{A}_K, \widehat{B}_V, \widehat{B}_K, \widehat{A}_V$. For each case, it is possible to construct the optimal $\widehat{p}_{i,j}s$ that satisfy all conditions (41)–(45). Note that these three cases are collectively exhaustive.

　　Case 1. $\widehat{A}_K \leq \widehat{B}_V$ and $\widehat{B}_K \leq \widehat{A}_V$
　　Case 2. $\widehat{A}_K > \widehat{B}_V$
　　Case 3. $\widehat{B}_K > \widehat{A}_V$

As a reminder, arm K represents the arm with the lowest popularity and arm V represents the arm with the highest popularity. For detailed construction procedures, see Appendix A.

The previous three cases prove that $\widehat{p}_{K,1}, \widehat{p}_{K,2}, \ldots, \widehat{p}_{K,N+1}, \widehat{p}_{1,K}, \widehat{p}_{2,K}, \ldots, \widehat{p}_{N+1,K}$ which satisfy (41)–(45) always exist. In other words, Assumption 1 is indeed correct. Therefore, if we assume that Lemma 1 has been proven to hold for N arms, then Lemma 1 holds when there are $N + 1$ arms.

Hence, using mathematical induction, Lemma 1 holds for any number of arms.

2.3. Theorem 2

2.3.1. Statement. Here, we return to the original problem, where the preferences of players are probabilities. Popularity S_i can still be greater than one because it is the sum of the preferences for each arm.

TABLE 2: Preference setting followed by the remaining part of the joint selection probability matrix in the gray-shaded region. $\widehat{A}_i^* = \widehat{A}_i - \widehat{p}_{i,K}, \widehat{B}_j^* = \widehat{B}_j - \widehat{p}_{K,j}$.

Player	Arm 1	Arm 2	$\cdots$	Arm $(K-1)$	Arm $(K+1)$	$\cdots$	Arm $(N+1)$	Sum
A	$\widehat{A}_1^*$	$\widehat{A}_2^*$	$\cdots$	$\widehat{A}_{K-1}^*$	$\widehat{A}_{K+1}^*$	$\cdots$	$\widehat{A}_{N+1}^*$	$T - \widehat{S}_K$
B	$\widehat{B}_1^*$	$\widehat{B}_2^*$	$\cdots$	$\widehat{B}_{K-1}^*$	$\widehat{B}_{K+1}^*$	$\cdots$	$\widehat{B}_{N+1}^*$	$T - \widehat{S}_K$

Theorem 2. *If any value of S_i is greater than one, it is impossible to make the loss L equal to zero, and the minimum loss is given by*

$$L_{\min} = \frac{N}{2(N-1)} \cdot (S_V - 1)^2. \tag{59}$$

Note that $S_V = \max\{S_i\}$. In a case where the Nth arm is the most popular, i.e., $S_V = S_N > 1$, the following joint selection probability matrix is one of the matrices that minimize the loss:

$$\widetilde{\mathbf{P}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & A_1 + \epsilon \\ 0 & 0 & \cdots & 0 & A_2 + \epsilon \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & A_{N-1} + \epsilon \\ B_1 + \epsilon & B_2 + \epsilon & \cdots & B_{N-1} + \epsilon & 0 \end{pmatrix}, \quad \epsilon = \frac{S_N - 1}{2(N-1)}. \tag{60}$$

2.3.2. Outline of the Proof. We can prove that the loss function L is convex by showing that its Hessian is semi-definite. Appendix B provides the details of the proof. In Section 2.3.3, we rewrite the problem into an optimization problem with equality and inequality constraints and derive $\widetilde{p}_{i,j}$s that satisfy the Karush–Kuhn–Tucker conditions [29, 30] (hereinafter, KKT conditions). In a convex optimization problem, where the objective function and all constraints are convex, points that satisfy the KKT conditions provide global optima [31]. Therefore, the optimal joint selection probability matrix comprises the aforementioned $\widetilde{p}_{i,j}$s.

2.3.3. KKT Conditions. We derive the optimal joint selection probability matrix and calculate the minimum loss. The flattened vector of a joint selection probability matrix is defined as

$$\mathbf{p} = \begin{pmatrix} p_{1,2} & p_{1,3} & \cdots & p_{1,N} & p_{2,1} & p_{2,3} & \cdots & p_{2,N} & \cdots & p_{N,N-1} \end{pmatrix}^T. \tag{61}$$

Functions h and $g_{i,j}$ are defined as

$$h(\mathbf{p}) = \sum_{i,j} p_{i,j} - 1, g_{i,j}(\mathbf{p}) = -p_{i,j} \quad (i \neq j). \tag{62}$$

The problem of minimizing the loss while satisfying the constraints can be written as

$$\min_{\mathbf{p}} L(\mathbf{p})$$
$$\text{s.t. } h(\mathbf{p}) = 0, \quad g_{i,j}(\mathbf{p}) \leq 0. \tag{63}$$

Here, $L(\mathbf{p})$ is the loss corresponding to $\mathbf{p}$. Because the objective function and constraints are all convex, $\widetilde{\mathbf{p}}$ that satisfies the following KKT conditions yields the global minimum.

$$\nabla L(\widetilde{\mathbf{p}}) + \sum_{i,j} \lambda_{i,j} \nabla g_{i,j}(\widetilde{\mathbf{p}}) + \mu \nabla h(\widetilde{\mathbf{p}}) = 0,$$
$$\lambda_{i,j} g_{i,j}(\widetilde{\mathbf{p}}) = 0,$$
$$\lambda_{i,j} \geq 0,$$
$$g_{i,j}(\widetilde{\mathbf{p}}) \leq 0, \quad h(\widetilde{\mathbf{p}}) = 0. \tag{64}$$

The following parameters satisfy all conditions described in (64).

$$\epsilon = \frac{S_N - 1}{2(N-1)},$$

$$\mu = 2(N-2)\epsilon,$$

$$\lambda_{i,j} = \begin{cases} 2N\epsilon, & (i \neq N \wedge j \neq N), \\ 0, & \text{otherwise}, \end{cases}$$

$$\tilde{p}_{i,j} = \begin{cases} 0, & (i \neq N \wedge j \neq N), \\ A_i + \epsilon, & (j = N), \\ B_j + \epsilon, & (i = N). \end{cases} \tag{65}$$

See Appendix C for the verification of each condition.

Hence,

$$\tilde{\mathbf{P}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & A_1 + \epsilon \\ 0 & 0 & \cdots & 0 & A_2 + \epsilon \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & A_{N-1} + \epsilon \\ B_1 + \epsilon & B_2 + \epsilon & \cdots & B_{N-1} + \epsilon & 0 \end{pmatrix}, \quad \epsilon = \frac{S_N - 1}{2(N-1)}, \tag{66}$$

provides a global optimum. The minimum loss is

$$\begin{aligned} L_{\min} &= \sum_i \left(\pi_A(i) - A_i\right)^2 + \sum_j \left(\pi_B(j) - B_j\right)^2 \\ &= \sum_{i=1}^{N-1} \left(\pi_A(i) - A_i\right)^2 + \sum_{j=1}^{N-1} \left(\pi_B(j) - B_j\right)^2 + \left(\pi_A(N) - A_N\right)^2 + \left(\pi_B(j) - B_N\right) \\ &= \sum_{i=1}^{N-1} \epsilon^2 + \sum_{j=1}^{N-1} \epsilon^2 + \{-(N-1)\epsilon\}^2 + \{-(N-1)\epsilon\}^2 \\ &= (N-1)\epsilon^2 + (N-1)\epsilon^2 + (N-1)^2\epsilon^2 + (N-1)^2\epsilon^2 \\ &= \frac{N}{2(N-1)} \cdot (S_N - 1)^2. \end{aligned} \tag{67}$$

Note that $\tilde{\mathbf{P}}$ is an example of the global optima, and there could be other matrices that provide the same minimum loss.

Therefore, Theorem 2 was proven.

2.4. Theorem 3. Thus far, we presented two theorems concerning two players and N arms. In this section, we consider the case where M players exist and where M is greater than two. Let the players be called players A, B, C, etc. Moreover, the preference of player X for arm i is denoted by $X_i^\dagger$. As in the original problem,

$$\sum_i X_i^\dagger = 1, \quad X_i^\dagger \geq 0, \tag{68}$$

holds. The sum of the preferences for each arm is denoted as popularity:

$$S_i^\dagger = A_i^\dagger + B_i^\dagger + C_i^\dagger + \cdots. \tag{69}$$

We define d_x as the arm index selected by the xth player. Then, $p_{d_1,d_2,\ldots,d_M}^\dagger$ represents the joint selection probability of the xth player selecting arm d_x. The collection of joint selection probabilities is called the joint selection probability tensor. The satisfied preference or resultant selection

probability for player X comprises the sum of all probabilities of cases in which the xth player selects arm i:

$$\pi_x^\dagger(i) = \underbrace{\sum_{d_1 \notin \{i\}} \sum_{d_2 \notin \{i,d_1\}} \cdots \sum_{d_M \notin \{i,d_1,d_2,\ldots,d_{M-1}\}} p_{d_1,d_2,\ldots,d_x=i,\ldots,d_M}^\dagger}_{\text{Summation without the } x\text{th player}}. \tag{70}$$

The loss $L^\dagger$ is defined as the sum of the squares of the gap between the preferences and the satisfied preferences.

$$L^\dagger = \sum_x \sum_i \left(\pi_x^\dagger(i) - X_i^\dagger\right)^2. \tag{71}$$

Here, the xth player is called player X.

2.4.1. Statement

Theorem 3. *Suppose that there are M players and N arms. If any $S_i^\dagger$ is greater than one, it is not possible to make the loss $L^\dagger$ equal to zero.*

2.4.2. Proof by Contradiction. We suppose without loss of generality that $S_1^\dagger > 1$. We first assume that it is possible to make the loss $L^\dagger$ equal to zero and then prove that this leads to a contradiction. If the loss $L^\dagger$ becomes zero, the following are required:

$$A_1^\dagger = \sum_{d_2 \notin \{1\}} \sum_{d_3 \notin \{1,d_2\}} \cdots \sum_{d_M \notin \{1,d_2,d_3,\ldots,d_{M-1}\}} p_{1,d_2,d_3,\ldots,d_M}^\dagger, \tag{72}$$

$$B_1^\dagger = \sum_{d_1 \notin \{1\}} \sum_{d_3 \notin \{d_1,1\}} \cdots \sum_{d_M \notin \{d_1,1,d_3,\ldots,d_{M-1}\}} p_{d_1,1,d_3,\ldots,d_M}^\dagger, \tag{73}$$

$$C_1^\dagger = \sum_{d_1 \notin \{1\}} \sum_{d_2 \notin \{d_1,1\}} \cdots \sum_{d_M \notin \{d_1,d_2,1,\ldots,d_{M-1}\}} p_{d_1,d_2,1,\ldots,d_M}^\dagger. \tag{74}$$
$$\vdots$$

Terms that appear on the right-hand side of (72) are not identical to those on the right-hand side of (73) because $d_1 \neq 1$ in (73). Similarly, any term on the right-hand side of (72) is not identical to a variable on the right-hand side of all equations (73) and (74) .... Therefore, if we consider the sum of the right-hand sides of (72)–(74), ..., it should be smaller than or equal to the sum of all elements in the joint selection probability tensor, which is 1.

In contrast, if we take the sum of the left-hand sides of (72)–(74), ..., we get $A_1^\dagger + B_1^\dagger + C_1^\dagger + \cdots = S_1^\dagger$. Thus, by comparing both sides, we obtain $S_1^\dagger \leq 1$. This result contradicts the assumption that $S_1^\dagger > 1$.

Hence, with a proof by contradiction, Theorem 3 holds for any number of players and arms.

2.5. Conjecture. The construction method introduced in Theorem 1 is expected to be applicable to general M players. Here, we propose the following conjecture.

TABLE 3: Example of a preference setting.

Player	Arm 1	Arm 2	Arm 3
A	0.3	0.25	0.45
B	0.5	0.2	0.3

Conjecture 1. *Suppose that there are M players and N arms. If all the values of $S_i^\dagger$ are less than or equal to one, it is possible to make the loss $L^\dagger$ equal to zero.*

This conjecture seems true; however, thus far, the proof has not been completed in generality and is left for future studies.

3. Numerical Demonstrations

This section introduces several baseline models that output a joint probability selection matrix for the probabilistic preferences of two players. Then, we show the extent to which the loss can be improved by using the construction method of the optimal joint probability selection matrix introduced in Theorems 1 and 2 (henceforth, "the optimal satisfaction matrix"). The definitions of the notations, such as preference and loss, follow those in Section 2.1.

3.1. Baselines

3.1.1. Uniform Random. In what we call the "uniform random" method, the resulting joint selection probability matrix is such that all elements are equal except for the diagonals, which are filled with zeros. In other words, decision conflicts never occur; however, the selection is determined completely randomly by the two players. If we consider the preference settings shown in Table 3, the output of this method is

$$\begin{pmatrix} 0 & \dfrac{1}{6} & \dfrac{1}{6} \\[2mm] \dfrac{1}{6} & 0 & \dfrac{1}{6} \\[2mm] \dfrac{1}{6} & \dfrac{1}{6} & 0 \end{pmatrix}. \tag{75}$$

3.1.2. Simultaneous Renormalization. In this method, the product of the preferences of both players is considered first. Then, the diagonals where decision conflicts occur are modified to zero, and finally, the whole joint selection probability matrix is renormalized so that the sum is 1. The formula used is

$$r_{i,j} = \begin{cases} A_i \cdot B_j, & (i \neq j), \\ 0, & (i = j), \end{cases}$$
$$p_{i,j} = \frac{r_{i,j}}{\sum_{i,j} r_{i,j}}. \tag{76}$$

The joint selection probability matrix generated by this simultaneous renormalization method for the case in Table 3 is

$$
\begin{pmatrix} 0 & 0.06 & 0.09 \\ 0.125 & 0 & 0.075 \\ 0.225 & 0.09 & 0 \end{pmatrix} / 0.665 = \begin{pmatrix} 0 & 0.090 & 0.1353 \\ 0.1880 & 0 & 0.1128 \\ 0.3383 & 0.1353 & 0 \end{pmatrix}.
$$

(77)

3.1.3. Random Order. In what we call the "random order" method, the players first randomly determine in which order they will draw arms. This was inspired by the random priority mechanism proposed by Abdulkadiroğlu and Sönmez where preferences are deterministic [32]. Instead, we consider probabilistic preference profiles. Each player selects an arm based on a predetermined order; however, the arms already drawn by the previous players cannot be selected again. Thus, the selection probabilities for those arms are set to zero. Under the settings in Table 3, if we want to calculate $p_{1,2}$, two possible orders are considered. In the first case, player A draws first. In this case, the joint selection probability is

$$
0.3 \cdot \frac{0.2}{1 - 0.5} = 0.12.
$$

(78)

In the other case, player B draws first, and the probability is

$$
0.2 \cdot \frac{0.3}{1 - 0.25} = 0.08.
$$

(79)

Therefore, considering the average,

$$
p_{1,2} = \frac{0.12 + 0.08}{2} = 0.1.
$$

(80)

Similarly, we can calculate all joint selection probabilities, and the resulting matrix is

$$
\begin{pmatrix} 0 & 0.1 & 0.1718 \\ 0.1674 & 0 & 0.1151 \\ 0.3214 & 0.1243 & 0 \end{pmatrix}.
$$

(81)

3.2. Performance Comparison. We compare the loss L of uniform random, simultaneous renormalization, random order, and the optimal satisfaction matrix. Four preference settings are examined to evaluate the performance of each method. Namely, the degree of satisfaction with the players' preferences is investigated by comparing the loss L. c_i is used as a normalization term to ensure that the sum of the preference is one.

(1) Arithmetic progression and same preference.

$$
A_1 \colon A_2 \colon \cdots \colon A_N = B_1 \colon B_2 \colon \cdots \colon B_N = \frac{(1 \colon 2 \colon \cdots \colon N)}{c_1},
$$

$$
c_1 = \frac{(N+1)N}{2}.
$$

(82)

(2) Modified geometric progression with common ratio 2 and same preference.

$$
A_1 \colon A_2 \colon \cdots \colon A_N = B_1 \colon B_2 \colon \cdots \colon B_N = \frac{\left(1 \colon 1 \colon 2 \colon \cdots \colon 2^{N-2}\right)}{c_2},
$$

$$
c_2 = 2^{N-1}.
$$

(83)

(3) Modified geometric progression with common ratio 2 and reversed preference.

$$
A_1 \colon A_2 \colon \cdots \colon A_N = B_N \colon B_{N-1} \colon \cdots \colon B_1 = \frac{\left(1 \colon 1 \colon 2 \colon \cdots \colon 2^{N-2}\right)}{c_3},
$$

$$
c_3 = 2^{N-1}.
$$

(84)

(4) Geometric progression with common ratio 3 and same preference.

$$
A_1 \colon A_2 \colon \cdots \colon A_N = B_1 \colon B_2 \colon \cdots \colon B_N = \frac{\left(1 \colon 3 \colon \cdots \colon 3^{N-1}\right)}{c_4},
$$

$$
c_4 = \frac{3^N - 1}{2}.
$$

(85)

In cases (1)–(3), the optimal satisfaction matrix achieves $L = 0$ because $\forall i; \quad S_i \leq 1$, whereas in case (4), it is impossible to achieve $L = 0$ because

$$
S_N = 2 \cdot \frac{3^{N-1}}{\sum_{i=0}^{N-1} 3^i} = \frac{4 \cdot 3^{N-1}}{3^N - 1} > \frac{3^N}{3^N - 1} > 1.
$$

(86)

For N, the following numbers are used.

$$
N = 3, 4, 5, \ldots, 50.
$$

(87)

Figure 3 summarizes loss L as a function of the number of arms N accomplished by each method.

In all cases, the optimal satisfaction matrix performs the best, followed by random order, simultaneous renormalization, and uniform random. The result in case (1) shows that loss decreases with an increase in the number of arms for all methods. This trend is because our choice of loss is similar to the L_2-norm. The absolute value of each preference A_i, B_i decreases with an increase in the number of arms.

In the real world, the ratio of preference settings in case (2) is likely to appear more frequently than in case (1). For uniform random, simultaneous renormalization, and random order, the loss increases with the number of arms when the two players have the same preference. In contrast, the results show that the optimal satisfaction matrix consistently achieves 0-loss, and this underlines the importance of the construction method of the optimal joint selection probability matrix in the real world, where there are a vast number of choices.

The result in case (3) shows that simultaneous renormalization and random order exhibit good accuracy
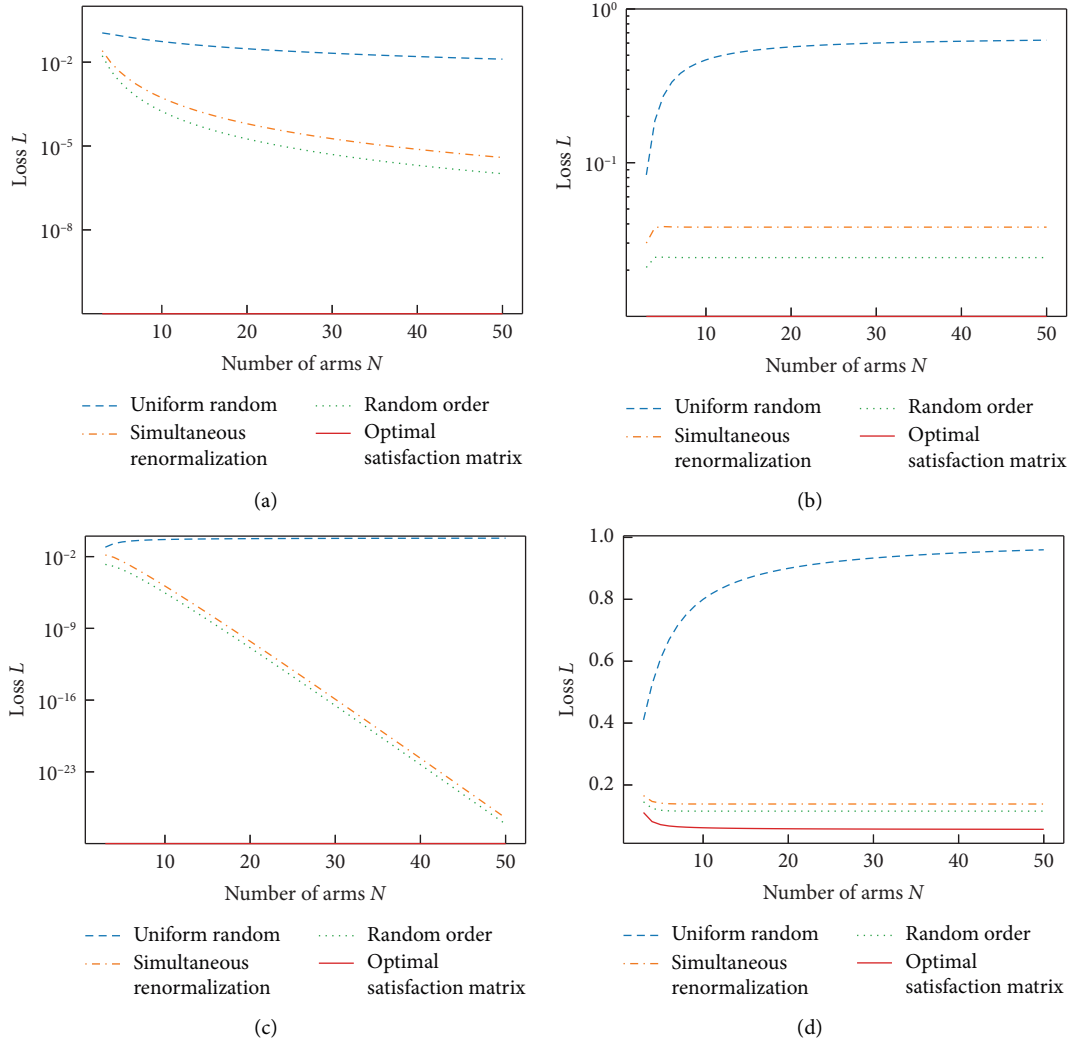
(a)

(b)

(c)

(d)

Figure 3: Loss comparison. Y axes are log scale for cases (1)–(3). Lines for the optimal satisfaction matrix overlap with X axes for these cases to show the loss is zero. (1) Arithmetic progression + same preference, (2) geometric progression with common ratio 2 + same preference, (3) geometric progression with common ratio 2 + reversed preference, and (4) geometric progression with common ratio 3 + same preference. For cases (1)–(3), the minimum loss is zero, while for case (4), it is greater than zero.

Table 4: Example of [Case 1].

Player	Arm 1	Arm 2	Arm 3	Arm 4	Total preference T
Player A	0.1	0.2	0.3	0.4	1.0
Player B	0.2	0.2	0.1	0.5	1.0
Popularity S	0.3	0.4	0.4	0.9	2.0

Table 5: Preference profile for the remaining part in [Case 1].

Player	Arm 2	Arm 3	Arm 4	Total preference T
Player A	0.2	0.3	0.2	0.7
Player B	0.2	0.1	0.4	0.7
Popularity S	0.4	0.4	0.6	1.4

together with the optimal satisfaction matrix. This result is due to the fact that the probability of decision conflicts happening is significantly smaller in this case (3), where the players have strong, reversed preferences. The diagonal terms are renormalized in simultaneous renormalization, which perturbs the other terms in the joint selection probability matrix. In case (3), these diagonal terms are smaller than in the other cases; therefore, perturbations to

TABLE 6: Example of [Case 2].

Player	Arm 1	Arm 2	Arm 3	Arm 4	Total preference T
Player A	0.25	0.1	0.15	0.5	1.0
Player B	0.1	0.35	0.35	0.2	1.0
Popularity S	0.35	0.45	0.5	0.7	2.0

TABLE 7: Preference profile for the remaining part in [Case 2].

Player	Arm 2	Arm 3	Arm 4	Total preference T
Player A	0.1	0.15	0.4	0.65
Player B	0.3	0.35	0.0	0.65
Popularity S	0.4	0.5	0.4	1.3

the other terms are reduced. In random order, the second player sets his/her preference for the arm drawn by the first player to zero; however, in the setup of case (3), this preference tends to be small because the first player is more likely to select an arm with a higher preference, which is an unfavored arm for the second player. This implies that the second player can select an arm based on a preference almost equal to his/her original preference.

In case (4), S_N is greater than 1, and thus, the optimal satisfaction matrix also does not achieve 0-loss. However, at approximately $N = 50$, the loss for random order is approximately 1.2 times smaller than the loss for simultaneous renormalization, whereas the optimal loss is almost twice as small as the loss for random order.

4. Conclusion

In this study, we theoretically examined how to maximize the satisfaction of each player by designing joint selection probabilities properly while avoiding decision conflicts when there are multiple players with probabilistic preferences. The present study demonstrated how to accomplish conflict-free stochastic decision-making among multiple players, wherein the preference of each player is highly appreciated. We clarified the condition when the optimal joint selection probabilities perfectly eliminate the deviation of the resulting choice selection probabilities and the probabilistic preference of the player in two player, N-choice situations, which leads to what we call zero-loss realizations. Furthermore, even under circumstances wherein zero loss is unachievable, we showed how to construct, what we call, the optimal satisfaction matrix, and whose joint selection probabilities minimize the loss. Furthermore, we generalized the theory to M-player situations ($M \geq 3$) and proved the conditions under which zero-loss joint selection was impossible. In addition, we numerically demonstrated the impact of the optimal satisfaction matrix by comparing several approaches that could provide conflict-free joint decision-making. The situation addressed in this study is very generic and can be widely applied to scenarios such as communication technology, multiagent reinforcement learning, and resource allocation, where multiple players make probabilistic choices, and choosing the same option is detrimental.

There are still many interesting future topics, including the mathematical proof of the conjecture shown in this study, which discusses the condition of zero-loss conflict-free joint stochastic selection involving an arbitrary number of players. The detailed analysis of situations when we replace loss with other metrics such as KL divergence also remains to be explored. Furthermore, extending the discussions to external environments and not just player satisfaction will be important for practical applications. This study paves the way for multiagent conflict-free stochastic decision-making.

Appendix

A. Construction Procedures for Assumption 1

In the induction step presented in Section 2.2, we assumed the existence of $p_{i,j}s$ that satisfy certain conditions, and we proved that perfect satisfaction is achievable using these $p_{i,j}s$. Here, we prove this assumption, namely, Assumption A.1 is true.

Assumption A.1. There exist $\widehat{p}_{K,1}, \widehat{p}_{K,2}, \ldots, \widehat{p}_{K,N+1}, \widehat{p}_{1,K}, \widehat{p}_{2,K}, \ldots, \widehat{p}_{N+1,K}$ that satisfy all of the following conditions (A.1)–(A.5).

The sum of the Kth row is equal to $\widehat{A}_K$.

$$\sum_{j=1}^{N+1} \widehat{p}_{K,j} = \widehat{A}_K. \tag{A.1}$$

The sum of the Kth column is equal to $\widehat{B}_K$.

$$\sum_{i=1}^{N+1} \widehat{p}_{i,K} = \widehat{B}_K. \tag{A.2}$$

The sum of the jth column without the Kth row is nonnegative.

$$\sum_{i \neq K} \widehat{p}_{i,j} \geq 0 \Leftrightarrow \widehat{p}_{K,j} \leq \widehat{B}_j \, (j \neq K). \tag{A.3}$$

The sum of the ith row without the Kth column is nonnegative.

$$\sum_{j \neq K} \widehat{p}_{i,j} \geq 0 \Leftrightarrow \widehat{p}_{i,K} \leq \widehat{A}_i \quad (i \neq K). \tag{A.4}$$

In the gray-shaded area of the joint selection probability matrix as follows, all remaining popularities are less than or

equal to the remaining total preference. Note that $\widehat{S}_K = \min\{\widehat{S}_i\}$.

$$(\hat{A}_i - \hat{p}_{i,K}) + (\hat{B}_i - \hat{p}_{K,i}) \leq T - \hat{S}_K \quad (i \neq K).$$

$$\begin{pmatrix} 0 & * & \cdots & p_{1,K} & \cdots & * & * \\ * & 0 & \cdots & p_{2,K} & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots \\ p_{K,1} & p_{K,2} & \cdots & 0 & \cdots & p_{K,N} & p_{K,N+1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & p_{N,K} & \cdots & 0 & * \\ * & * & \cdots & p_{N+1,K} & \cdots & * & 0 \end{pmatrix}$$ (A.5)

Now, we can systematically determine $\hat{p}_{i,j}$s in the Kth row or the Kth column of the joint selection probability matrix to show that Assumption A.1 is correct; that is, these $\hat{p}_{i,j}$s fulfill the conditions in (A.1)–(A.5). We cannot arbitrarily fill in the values in the Kth row or the Kth column such that their sum is $\hat{A}_K$ and $\hat{B}_K$, respectively; instead, we must consider the preferences of the other arms. (A.3)–(A.5) give us boundaries for $\hat{p}_{K,j}$s and $\hat{p}_{i,K}$s to exist. Not only must we ensure that $\hat{p}_{K,j}$ and $\hat{p}_{i,K}$ do not exceed these boundaries, but we must also ensure that a sufficient probability is assigned to the most popular arm V; otherwise, the left-hand side of (A.5) can sometimes be greater than the right-hand side for arm V. We consider the following three cases depending on the size relation of $\hat{A}_K, \hat{B}_V, \hat{B}_K, \hat{A}_V$. For each case, it is possible to construct the optimal $\hat{p}_{i,j}$s that satisfy all conditions (A.1)–(A.5). These three cases are collectively exhaustive.

[Case 1] $\hat{A}_K \leq \hat{B}_V$ and $\hat{B}_K \leq \hat{A}_V$
[Case 2] $\hat{A}_K > \hat{B}_V$
[Case 3] $\hat{B}_K > \hat{A}_V$

Arm K represents the arm with the lowest popularity, and arm V is the arm with the highest popularity. All arms except for arm V satisfy the following condition:

$$\widehat{S}_i \leq T - \widehat{S}_K. \tag{A.6}$$

[Case 1] $\hat{A}_K \leq \hat{B}_V$ and $\hat{B}_K \leq \hat{A}_V$

The following $\hat{p}_{i,j}$s on the Kth row or Kth column satisfy the conditions (A.1) and (A.2).

$$\begin{aligned} &\hat{p}_{K,V} \\ &\hat{p}_{V,K} = \hat{B}_K, \\ &\hat{p}_{K,j} = 0 \quad (j \neq V), \\ &\hat{p}_{i,K} = 0 \quad (i \neq V). \end{aligned} \tag{A.7}$$

In addition, (A.3) and (A.4) are satisfied because $\hat{A}_K \leq \hat{B}_V$ and $\hat{B}_K \leq \hat{A}_V$ in this particular case. Moreover, (A.5) is satisfied because for $i = V$,

$$\begin{aligned} &\hat{p}_{K,V} + \hat{p}_{V,K} \\ &= \widehat{S}_K \geq \widehat{S}_K - (T - \widehat{S}_V) = \widehat{S}_V + \widehat{S}_K - T, \end{aligned} \tag{A.8}$$

which leads to

$$(\hat{A}_V - \hat{p}_{V,K}) + (\hat{B}_V - \hat{p}_{K,V}) \leq T - \widehat{S}_K. \tag{A.9}$$

For the other is, (A.5) follows because they satisfy (A.6), and (A.6) is equivalent to (A.5) when $\hat{p}_{K,i} = \hat{p}_{i,K} = 0$.

Hence, $\hat{p}_{i,j}$s described in (A.7) fulfill all conditions (A.1)–(A.5) in [Case 1].

Table 4 illustrates an example of a case where $\hat{A}_K \leq \hat{B}_V$ and $\hat{B}_K \leq \hat{A}_V$. Here, the most popular arm K is arm 1 and the least popular arm V is arm 4. In this case, the first row and first column of the joint selection probability matrix should be filled as

$$\begin{pmatrix} 0 & 0 & 0 & 0.1 \\ 0 & * & * & * \\ 0 & * & * & * \\ 0.2 & * & * & * \end{pmatrix} \tag{A.10}$$

Then, the remaining gray-shaded region should satisfy the preference setting described in Table 5. Each popularity is less than or equal to the total preference in the remaining part, i.e., the assumption of Lemma 1 holds for this part.

[Case 2] $\hat{A}_K > \hat{B}_V$

In this case, as $\widehat{S}_K \leq \widehat{S}_V$, it follows that

$$\hat{B}_K < \hat{A}_V. \tag{A.11}$$

When $\hat{B}_V + \hat{B}_1 < \hat{A}_K$, let m be the arm index satisfying the inequality

$$\widehat{B}_V + \underbrace{\widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_{m-1}}_{\text{does not contain } \widehat{B}_V \text{ or } \widehat{B}_K} < \widehat{A}_K \le \widehat{B}_V + \underbrace{\widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_m}_{\text{does not contain } \widehat{B}_V \text{ or } \widehat{B}_K}. \tag{A.12}$$

When $\widehat{B}_V + \widehat{B}_1 \ge \widehat{A}_K$, we define $m = 1$. m always exists because

$$\widehat{B}_V + \underbrace{\widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_{N+1}}_{\text{does not contain } \widehat{B}_V \text{ or } \widehat{B}_K} = T - \widehat{B}_K \ge \widehat{A}_K. \tag{A.13}$$

Then,

$$\begin{cases} [\text{When } m = 1], \\ \widehat{p}_{K,V} = \widehat{B}_V, \quad \widehat{p}_{K,1} = \widehat{A}_K - \widehat{B}_V, \widehat{p}_{K,j} = 0 \ (j \notin \{1, V\}), \\ \widehat{p}_{V,K} = \widehat{B}_K, \quad \widehat{p}_{i,K} = 0 \ (i \ne V), \\ [\text{When } m \ge 2], \\ \widehat{p}_{K,K} = 0, \quad \widehat{p}_{K,V} = \widehat{B}_V, \widehat{p}_{K,j} = \widehat{B}_j \ (j < m \text{ and } j \notin \{K, V\}), \\ \widehat{p}_{K,m} = \widehat{A}_K - \underbrace{(\widehat{B}_V + \widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_{m-1})}_{\text{does not contain } \widehat{B}_K}, \quad \widehat{p}_{K,j} = 0 \ (m < j \text{ and } j \notin \{K, V\}), \\ \widehat{p}_{V,K} = \widehat{B}_K, \quad \widehat{p}_{i,K} = 0 \ (i \ne V), \end{cases} \tag{A.14}$$

satisfy conditions (A.1) and (A.2). Moreover, $\widehat{p}_{i,j}$s in (A.14) fulfill (A.3) because for $j = V, 1, 2, \ldots, m-1$,

$$\widehat{p}_{K,j} = \widehat{B}_j \le \widehat{B}_j. \tag{A.15}$$

For $j = m$, the definition of m satisfies the condition $\widehat{A}_K \le \widehat{B}_V + \underbrace{\widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_m}_{\text{does not contain } \widehat{B}_V \text{ or } \widehat{B}_K}$, which implies that

$$\widehat{A}_K - \underbrace{(\widehat{B}_V + \widehat{B}_1 + \widehat{B}_2 + \cdots + \widehat{B}_{m-1})}_{\text{does not contain } \widehat{B}_K} \le \widehat{B}_m. \tag{A.16}$$

For $j = m+1, m+2, \ldots, N+1$,

$$\widehat{p}_{K,j} = 0 \le \widehat{B}_j. \tag{A.17}$$

Furthermore, with $\widehat{p}_{i,j}$ described in (A.14), condition (A.4) is also satisfied because for $i = V$, (A.11) verifies

$$\widehat{p}_{V,K} = \widehat{B}_K \le \widehat{A}_K, \tag{A.18}$$

and for the other i,

$$\widehat{p}_{i,K} = 0 \le \widehat{A}_i. \tag{A.19}$$

Finally, (A.5) is satisfied because for $i = V$,

$$\begin{aligned} &\widehat{p}_{K,V} + \widehat{p}_{V,K} \\ &= \widehat{S}_V + \widehat{S}_K - T. \end{aligned} \tag{A.20}$$

$T - \widehat{A}_V - \widehat{A}_K$ is always nonnegative because $T \ge \widehat{A}_V + \widehat{A}_K$. (A.20) is equivalent to

$$\left(\widehat{A}_V - \widehat{p}_{V,K}\right) + \left(\widehat{B}_V - \widehat{p}_{K,V}\right) \le T - \widehat{S}_K. \tag{A.21}$$

The other is follow (A.6), and with the prerequisites $\widehat{p}_{i,K} \ge 0$, $\widehat{p}_{K,i} \ge 0$, it follows (A.5).

Therefore, $\widehat{p}_{i,j}$s given in (A.14) satisfy all conditions (A.1)–(A.5) in [Case 2]

Table 6 shows an example of [Case 2] Here, the most popular arm K is arm 1 and the least popular arm V is arm 4. In this case, $m = 2$ because

$$\widehat{B}_V (= 0.2) < \widehat{A}_K (= 0.25) \le \widehat{B}_V (= 0.2) + \widehat{B}_2 (= 0.35). \tag{A.22}$$

Therefore, the first row and first column of the joint selection probability matrix should be filled as

$$\begin{pmatrix} 0 & 0.05 & 0 & 0.2 \\ 0 & * & * & * \\ 0 & * & * & * \\ 0.1 & * & * & * \end{pmatrix} \tag{A.23}$$

Then, the remaining gray-shaded region should satisfy the preference setting described in Table 7. Each popularity is less than or equal to the total preference in

the remaining part, i.e., the assumption of Lemma 1 holds for this part.

[Case 3] $\widehat{B}_K > \widehat{A}_V$

This case is similar to [Case 2] If we swap $\widehat{A}_i$ and $\widehat{B}_i$ in the discussion of [Case 2], we obtain $\widehat{p}_{i,j}s$ that satisfy all conditions (A.1)–(A.5).

B. Convexity of loss L

When we derived the optimal joint selection probability matrix in Section 2.3 of the main text, we used the fact that the loss function L is convex. Here, we prove that the loss function L is convex.

$$L = \sum_i \left(\pi_A(i) - A_i\right)^2 + \sum_j \left(\pi_B(j) - B_j\right)^2. \tag{B.1}$$

We prove that each $\left(\pi_A(i) - A_i\right)^2$ and $\left(\pi_B(j) - B_j\right)^2$ are convex, and we know that L is convex because the sum of convex functions is also convex.

The first term of L is defined as

$$L_1 := \left(\pi_A(1) - A_1\right)^2 = \left(p_{1,2} + p_{1,3} + \cdots + p_{1,N} - A_1\right)^2. \tag{B.2}$$

The Hessian matrix for L_1 in terms of all $p_{i,j}s (i \neq j, \quad i = 1, 2, \ldots, N, j = 1, 2, \ldots, N)$ is

$$H_1 = \begin{pmatrix} D & O \\ O & D_O \end{pmatrix}, \tag{B.3}$$

where

$$D = \overbrace{\begin{pmatrix} 2 & 2 & \cdots & 2 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & \vdots & \vdots \\ 2 & 2 & \cdots & 2 \end{pmatrix}}^{N-1}, D_O = \overbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}^{(N-1)^2}. \tag{B.4}$$

As this is a block diagonal matrix, the eigenvalues of H_1 are the union of the eigenvalues of D and D_o. The eigenvalues of D_o are obviously $\overbrace{0, 0, \ldots, 0}^{(N-1)^2}$, and those of D are $2(N-1), \overbrace{0, 0, \ldots, 0}^{N-2}$.

Therefore, all eigenvalues of H_1 are nonnegative, which means that H_1 is positively semidefinite, and this implies that L_1 is convex. Similarly, each $\left(\pi_A(i) - A_i\right)^2$ and $\left(\pi_B(j) - B_j\right)^2$ is convex because of the symmetry of the loss function, given that $i = 1$ is not unique among all is. From the previous argument, we know that loss function L is convex.

C. Verification of the KKT Conditions

In Section 2.3.3 of the main text, where we derived the point which satisfies all the KKT conditions, we defined new

notations and functions. The flattened vector of a joint selection probability matrix is defined as

$$\mathbf{p} = \begin{pmatrix} p_{1,2} & p_{1,3} & \cdots & p_{1,N} & p_{2,1} & p_{2,3} & \cdots & p_{2,N} & \cdots & p_{N,N-1} \end{pmatrix}^T. \tag{C.1}$$

Functions h and $g_{i,j}$ are defined as

$$h(\mathbf{p}) = \sum_{i,j} p_{i,j} - 1, g_{i,j}(\mathbf{p}) = -p_{i,j} \quad (i \neq j). \tag{C.2}$$

The problem can be written as

$$\min_{\mathbf{p}}, L(\mathbf{p}), \\ \text{s.t. } h(\mathbf{p}) = 0, g_{i,j}(\mathbf{p}) \leq 0, \tag{C.3}$$

$L(\mathbf{p})$ represents the loss corresponding to $\mathbf{p}$. Because the objective function and constraints are all convex, $\widetilde{\mathbf{p}}$ that satisfies the KKT conditions as follows gives the global minimum.

$$\nabla L(\widetilde{\mathbf{p}}) + \sum_{i,j} \lambda_{i,j} \nabla g_{i,j}(\widetilde{\mathbf{p}}) + \mu \nabla h(\widetilde{\mathbf{p}}) = 0, \tag{C.4}$$

$$\lambda_{i,j} g_{i,j}(\widetilde{\mathbf{p}}) = 0, \tag{C.5}$$

$$\lambda_{i,j} \geq 0, \tag{C.6}$$

$$h(\widetilde{\mathbf{p}}) = 0. \tag{C.7}$$

Here, we show that

$$\epsilon = \frac{S_N - 1}{2(N-1)}, \tag{C.8}$$

$$\mu = 2(N-2)\epsilon, \tag{C.9}$$

$$\lambda_{i,j} = \begin{cases} 2N\epsilon, & (i \neq N \text{ and } j \neq N), \\ 0, & (\text{otherwise}), \end{cases} \tag{C.10}$$

$$\widetilde{p}_{i,j} = \begin{cases} 0, & (i \neq N \text{ and } j \neq N), \\ A_i + \epsilon, & (j = N), \\ B_j + \epsilon, & (i = N), \end{cases} \tag{C.11}$$

satisfy (C.4)–(C.7).

Condition C.1. Stationarity (C.4).

In the following argument, we call the element in each vector, which is in the same dimension as $p_{i,j}$ "the (i, j) th element" in a vector.

From the definition of L, the (i, j) th element of $\nabla L(\mathbf{p})$, denoted by $\nabla L(\mathbf{p})_{[i,j]}$, is

$$\nabla L(\mathbf{p})_{[i,j]} = 2\left(\pi_A(i) - A_i\right) + 2\left(\pi_B(j) - B_j\right). \tag{C.12}$$

With $\widetilde{p}_{i,j}$ described in (C.11),

$$(\nabla L(\widetilde{\mathbf{p}}))_{[i,j]} = \begin{cases} 2\epsilon + 2\epsilon = 4\epsilon, & (i \neq N \text{ and } j \neq N), \\ -2(N-1)\epsilon + 2\epsilon = -2(N-2)\epsilon, & (\text{otherwise}). \end{cases} \tag{C.13}$$

In addition, with $\widetilde{p}_{i,j}$ and $\lambda_{i,j}$ described in (C.10) and (C.11), the (i, j) th element of $\sum_{i,j} \lambda_{i,j} \nabla g_{i,j}(\widetilde{\mathbf{p}})$ is

$$\left(\sum_{i,j} \lambda_{i,j} \nabla g_{i,j}(\widetilde{\mathbf{p}}) \right)_{[i,j]} = \begin{cases} -2N\epsilon, & (i \neq N \text{ and } j \neq N), \\ 0, & (\text{otherwise}). \end{cases} \tag{C.14}$$

Moreover, the (i, j) th element of $\mu \nabla h(\widetilde{\mathbf{p}})$ is

$$(\mu \nabla h(\widetilde{\mathbf{p}}))_{[i,j]} = 2(N-2)\epsilon. \tag{C.15}$$

Therefore, stationarity is successfully achieved with $\widetilde{p}_{i,j}$, $\lambda_{i,j}$, and μ given in (C.8)–(C.11).

$$\nabla L(\widetilde{\mathbf{p}}) + \sum_{i,j} \lambda_{i,j} \nabla g_{i,j}(\widetilde{\mathbf{p}}) + \mu \nabla h(\widetilde{\mathbf{p}}) = 0. \tag{C.16}$$

Condition C.2. Complementary slackness (C.5).
From (C.10) and (C.11) in both cases where $i \neq N \wedge j \neq N$,

$$\lambda_{i,j} g_{i,j}(\widetilde{\mathbf{p}}) = -\lambda_{i,j} \widetilde{p}_{i,j} = 0 \geq 0. \tag{C.17}$$

Condition C.3. Dual feasibility (C.6).
Because $S_N - 1 > 0$, ϵ is positive and so is $2N\epsilon$. Thus,

$$\lambda_{i,j} \geq 0. \tag{C.18}$$

Condition C.4. Primal feasibility (C.7).
Because $\epsilon > 0$, it follows that $A_i + \epsilon$ and $B_j + \epsilon$ are nonnegative. Then,

$$g_{i,j}(\widetilde{\mathbf{p}}) \leq 0. \tag{C.19}$$

Furthermore,

$$\begin{aligned} h(\widetilde{\mathbf{p}}) &= \sum_{i,j} p_{i,j} - 1 \\ &= \sum_{i=1}^{N-1} (A_i + \epsilon) + \sum_{j=1}^{N-1} (B_j + \epsilon) - 1 \\ &= \sum_{i=1}^{N-1} (A_i + B_i) + 2(N-1)\epsilon - 1 \\ &= 2 - S_N + (S_N - 1) - 1 \\ &= 0. \end{aligned} \tag{C.20}$$

Therefore,

$$h(\widetilde{\mathbf{p}}) = 0, \tag{C.21}$$

holds.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Disclosure

A preprint was published in [33].

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] A. Hira, K. Reilly, and K. Reilly, "The emergence of the sharing economy: The Emergence of the Sharing Economy: Implications for Developmentmplications for development," *Journal of Developing Societies*, vol. 33, no. 2, pp. 175–190, 2017.

[2] G. George, S. J. Schillebeeckx, and T. L. Liak, "The management of natural resources: an overview and research agenda," *Managing Natural Resources*, 2018.

[3] L. Liu, P. Parag, J. Tang, W. Y. Chen, and J. F. Chamberland, "Resource allocation and quality of service evaluation for wireless communication systems using fluid models," *IEEE Transactions on Information Theory*, vol. 53, no. 5, pp. 1767–1777, 2007.

[4] L. Shapley and H. Scarf, "On cores and indivisibility," *Journal of Mathematical Economics*, vol. 1, pp. 23–37, 1974.

[5] J. Alcalde-Unzu and E. Molis, "Exchange of indivisible goods and indifferences: the top trading absorbing sets mechanisms," *Games and Economic Behavior*, vol. 73, pp. 1–16, 2011.

[6] H. Aziz and B. De Keijzer, "Housing markets with indifferences: a tale of two mechanisms," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, no. 1, pp. 1249–1255, 2021.

[7] D. Saban and J. Sethuraman, "House allocation with indifferences: a generalization and a unified view," in *Proceedings*

of the 14th ACM Conference on Electronic Commerce, pp. 803–820, Pennsylvania, PA, USA, June 2013.

[8] S. Athanassoglou and J. Sethuraman, "House allocation with fractional endowments," *International Journal of Game Theory*, vol. 40, no. 3, pp. 481–513, 2011.

[9] H. Robbins, "Some aspects of the sequential design of experiments," *Bulletin of the American Mathematical Society*, vol. 58, no. 5, pp. 527–535, 1952.

[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT press, Cambridge, MA, USA, 2018.

[11] J. G. March, "Exploration and exploitation in organizational learning," *Organization Science*, vol. 2, no. 1, pp. 71–87, 1991.

[12] J. Vermorel and M. Mohri, "Multi-armed bandit algorithms and empirical evaluation," in *European Conference on Machine Learning*, pp. 437–448, Springer, Berlin/Heidelberg, Germany, 2005.

[13] L. Lai, H. El Gamal, H. El Gamal, H. V. Poor, H. Jiang, and H. V. Poor, "Cognitive medium access: Cognitive Medium Access: Exploration, Exploitation, and Competitionxploration, exploitation, and competition," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 239–253, 2011.

[14] S. J. Kim, M. Naruse, and M. Aono, "Harnessing the computational power of fluids for optimization of collective decision making," *Philosophie*, vol. 1, no. 3, pp. 245–260, 2016.

[15] J. L. Deneubourg and S. Goss, "Collective patterns and decision-making," *Ethology Ecology & Evolution*, vol. 1, no. 4, pp. 295–311, 1989.

[16] N. Chauvet, D. Jegouso, B. Boulanger et al., "Entangled-photon decision maker," *Scientific Reports*, vol. 9, no. 1, Article ID 12229, 2019.

[17] N. Chauvet, G. Bachelier, S. Huant, H. Saigo, H. Hori, and M. Naruse, "Entangled N-photon states for fair and optimal social decision making," *Scientific Reports*, vol. 10, no. 1, Article ID 20420, 2020.

[18] T. Amakasu, N. Chauvet, G. Bachelier, S. Huant, R. Horisaki, and M. Naruse, "Conflict-free collective stochastic decision making by orbital angular momentum of photons through quantum interference," *Scientific Reports*, vol. 11, no. 1, Article ID 21117, 2021.

[19] M. Tan, "Multi-agent reinforcement learning: independent vs. cooperative agents," in *Proceedings of the 10th International Conference on Machine Learning*, pp. 330–337, San Francisco; CA, USA, June 1993.

[20] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156–172, 2008.

[21] V. Mnih, A. P. Badia, M. Mirza et al., "Asynchronous methods for deep reinforcement learning," in *International Conference on Machine Learning*, pp. 1928–1937, PMLR, New York City, NY, USA, 2016.

[22] S. Ross and B. Chaib-draa, "Satisfaction equilibrium: achieving cooperation in incomplete information games," in *Conference of the Canadian Society for Computational Studies of Intelligence*, pp. 61–72, Springer, Berlin/Heidelberg, Germany, 2006.

[23] M. Fasoulakis, E. E. Tsiropoulou, and S. Papavassiliou, "Satisfy instead of maximize: improving operation efficiency in wireless communication networks," *Computer Networks*, vol. 159, pp. 135–146, 2019.

[24] H. Zhang, S. Zhao, G. Kou, C. C. Li, Y. Dong, and F. Herrera, "An overview on feedback mechanisms with minimum adjustment or cost in consensus reaching in group decision making: research paradigms and challenges," *Information Fusion*, vol. 60, pp. 65–79, 2020.

[25] H. Zhang, C. C. Li, Y. Liu, and Y. Dong, "Modeling personalized individual semantics and consensus in comparative linguistic expression preference relations with self-confidence: an optimization-based approach," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 3, pp. 627–640, 2021.

[26] H. Zhang, F. Wang, Y. Dong, F. Chiclana, and E. Herrera-Viedma, "Social trust driven consensus reaching model with a minimum adjustment feedback mechanism considering assessments-modifications willingness," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 6, pp. 2019–2031, 2021.

[27] Y. Dong, Y. Xu, H. Li, and B. Feng, "The OWA-based consensus operator under linguistic representation models using position indexes," *European Journal of Operational Research*, vol. 203, no. 2, pp. 455–463, 2010.

[28] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[29] W. Karush, "Minima of Functions of Several Variables with Inequalities as Side Constraints," M. Sc. Dissertation, Dept. of Mathematics, Univ. of Chicago, Chicago, IL, USA, 1939.

[30] H. W. Kuhn and A. W. Tucker, *Nonlinear Programming Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Neyman pp. 481–492, 1951.

[31] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[32] A. Abdulkadiroğlu and T. Sönmez, "Random serial dictatorship and the core from random endowments in house allocation problems," *Econometrica*, vol. 66, no. 3, pp. 689–701, 1998.

[33] H. Shinkawa, N. Chauvet, G. Bachelier, A. Röhm, R. Horisaki, and M. Naruse, *Optimal preference satisfaction for conflict-free joint decisions*, https://arxiv.org/abs/2205.00799, 2022.