Quantum Decision Making with Small Sample for Network Monitoring and Control

Tatsuya Otoshi[¶], Masayuki Murata[†]

***Abstract*—With the development and diversification of applications on the Internet, applications that require high responsiveness, such as video streaming, are becoming mainstream. Application responsiveness is not only a matter of communication delay but also a matter of time required to grasp changes in network conditions. The tradeoff between accuracy and measurement time is a challenge in network control. We people make countless decisions all the time, and our decisions seem to resolve tradeoffs between time and accuracy. When making decisions, people are known to make appropriate choices based on relatively small samples. Although there have been various studies on models of human decision-making, a model that integrates various cognitive biases, called "quantum decision making," has recently attracted much attention. However, the modeling of small samples has not been examined much so far. In this paper, we extend the model of quantum decision making to model decision making with small sample. In the proposed model, the state is updated by value-based probability amplitude amplification. By analytically obtaining a lower bound on the number of samples required for decision making, we show that decision making with a small number of samples is feasible.**

***Index Terms*—Quantum Decision Making, Small Sample, MPEG-DASH**

I. Introduction

With the development and diversification of applications on the Internet, applications that require high responsiveness, such as video streaming and tactile internet [1], are becoming mainstream. Such requirements have led to the development of technologies to reduce communication latency, such as edge computing [2] and Ultra-Reliable Low-Latency Communication (URLLC) for 5G [3]. With these technologies, it is possible to maintain low communication latency by changing the settings appropriately according to the network conditions. On the other hand, the responsiveness of an application depends not only on the communication delay but also on the time required to grasp the changes in the network status. In general, an accurate grasp of the situation requires a long time to be measured, depending on the accuracy. Therefore, network control to improve responsiveness needs to handle the trade-off between accuracy and measurement time appropriately.

The most familiar example that deals with the trade-off between the amount of information to be acquired and the understanding of the situation is human cognition. People are known to make appropriate choices based on a relatively small sample of information when making decisions, and research has been done on modeling this [4], [5]. The literature [4] discusses sample-based decision making under conditions of constrained cognitive resources, using the time required to obtain a sample as a cost relative to the time to make a decision. In the literature [5], the effect of a small sample is that the difference in expected gain is amplified. Learning from this mechanism of utilizing a small sample of people may be a shortcut to solving the trade-off between accuracy and measurement time.

It is known that various information processing biases exist in human decision-making, not only in a small sample. As a model that integrates various cognitive biases, a model called quantum decision-making has recently attracted much attention [6]–[8]. Quantum mechanics and quantum decision-making are not directly related, but rather model human decision-making by analogy through mathematical formulas that represent quantum behavior in quantum mechanics. Our research group uses quantum decision-making to model the QoE of users while watching streaming video and applies it to bitrate selection [9]. In this method, the agent uses the user's model to guide the user to an appropriate choice, thus combining the user's own choice with the system's guidance. Because of its versatility, quantum decision-making is also the first choice for modeling decision-making with small samples.

However, the modeling of small samples by quantum decision-making has not been studied much so far. In the literature [10], decision making is modeled as an iterative process of amplifying the probability amplitude, increasing the probability of taking a particular option, and making a final decision. When there is a bias in the probability of the alternatives in the initial state, the probability of taking a particular alternative increases with fewer iterations. The availability heuristic is expressed by viewing the bias in the initial state as the ease of recalling the alternatives and the number of iterations of amplification of the probability amplitude as the time required to make a decision. In this model, the initial state only affects the bias of the choice, and the nature of the sample is not taken into account.

In this paper, we model decision-making with a small sample using quantum decision-making. For changes in the cognitive state with samples, amplitude amplification [10] is extended to reflect sample values. We analytically find the number of samples required to make the correct choice in the model and show that a small number of samples is appropriate, as in the literature [4].

The organization of the intentions of this paper is as follows: Section II provides an overview of decision making based on a small sample of people and maps network control to the problem of decision making based on a small sample; Section III models decision making based on a small sample of people using quantum decision making; Section IV shows the analytical result. In Section V, we summarize the paper and discuss future work.

¶ Graduate School of Economics, Osaka University, 1-7 Machikaneyama-Cho, Toyonaka, Osaka 565-0043, Japan

†Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

II. Decision Making with Small Sample

A. Human Decision

In conditioning by reward, people are known to use the phenomenon of probability matching as an example where the animal's choice does not necessarily result in optimal behavior [11]. Probability matching is the phenomenon that when repeated choices are made, each choice is made with a probability corresponding to the ratio of the rewards. Since the optimal action is to decisively choose the option with the highest expected reward, the reward cannot be maximized when probability matching is occurring. Such probability matching has also been reported to occur in perceptual tasks [12]. In the literature [12], when subjects were asked to estimate the location of a sound source based on visual stimuli (pictures) and auditory stimuli (sounds), probability matching was found to select the location of the sound source probabilistically according to the intensity of the stimuli.

On the other hand, it is also known that people make decisions based on a small number of samples, and it is believed that a small number of samples can have the effect of making probability matching choices more like deterministic choices [5]. When the reward obtained for each choice and its probability are unknown, a small number of outcomes may be sampled to make a decision. Even if the rewards and probabilities are known, it is also possible to simulate a small number of samples in your mind to make decisions as a shortcut to thinking in calculating expected values [4]. In such a case, the difference in reward for each option in the sample tends to be larger than the difference in actual expected value. As the difference in rewards becomes larger, the ratio of rewards will also become more biased, so that the choice made by probability matching will also be biased toward a particular choice. This means that the bias caused by probability matching will be corrected to some extent in decision-making using a small sample.

However, in decision-making with a small sample, the choice that is more likely to be selected as a result of an expanded reward difference is not necessarily the choice with the highest expected value. By integrating such small sample decision-making into quantum decision-making, we aim to build a model that includes bias in the small sample and correct it by user agents.

1) Advantage of Small Sample: The literature [4] discusses sample-based decision making under conditions of constrained cognitive resources, using the time required to obtain a sample as a cost relative to the time to make a decision. In a situation where cognitive resources are constrained, deciding in a short time, although not optimal, based on a small sample is a rational decision. The decision-making in the literature [4] deals with the problem of choosing one of several alternatives with a certain probability of being a winner, and in this case, the accuracy is greatly improved by the first sample. Therefore, when the cost of cognitive resources is included, a single sample may be appropriate.

In the literature [5], the effect of a small sample is that the difference in expected gain is amplified. When the sample is small, the absolute difference in gain estimated from a small number of samples may be larger than the difference in the expected gain. In the conditioning of actions and gains, there is a known phenomenon called probability matching, where the probability of choosing an action is proportional to the average of the gains obtained. In the situation of probability matching, when the difference in expected gain is small, the choice of action approaches random, and when the difference in expected gain is large, the choice of a particular action becomes more likely. Thus, the difference in expected gain with a small sample may facilitate the exclusive selection. However, in the literature [5], they only focus on the expansion of the absolute difference in expected gain, which does not necessarily amplify the correct choice.

2) Disadvantage of Small Sample: When the number of samples is small, the probability distribution of the gain cannot be completely obtained, and thus there is a possibility of making a wrong decision. In cases such as guessing the flip side of a coin, it is considered advantageous to believe the sample as it is because what is observed reflects the choice that is likely to come up. The literature [4] is just such a setting, and it was possible to obtain sufficient information even with a very small sample of one.

The decision-making problem dealt with in the literature [5]is a bit more complex: choosing one lottery among several lotteries that reward x with a certain probability p and y otherwise. In this case, one sample is taken from each lottery.

For the two lotteries $i = 1, 2$, if the probabilities are p_1, p_2 and the rewards are $x_1, x_2 (x_1 \geq x_2 \geq 0)$ and $y_1 = y_2 = 0$, the gain X_1, X_2 and the difference between them for one sample from each lottery are shown in Table I.

TABLE I
Difference of revenues from 1 sample

probability	X_1	X_2	$X_1 - X2$
$p_1 p_2$	x_1	x_2	$x_1 - x_2$
$p_1(p_1 - p_2)$	x_1	0	x_1
$(1 - p_1)p_2$	0	x_2	$-x_2$
$(1 - p_1)(1 - p_2)$	0	0	0

If we exclude the case where $X_1 - X_2 = 0$ as not choosing either lottery, then the probability of choosing lottery 1 is p_1 and the probability of choosing lottery 2 is $(1 - p_1)p_2$.

The condition under which lottery 1 is more likely to be chosen than lottery 2 based on one sample is $p_1 > (1 - p_1)p_2$, which can be organized as follows.

$$p_1 > \frac{1}{\frac{1}{p_2} + 1} \quad (1)$$

In other words, if p_1 has a probability greater than or equal to the right-hand side, lottery 1 is likely to be selected. This condition on p_1 is the strictest when $p_2 = 1$, and lottery 1 is more likely to be selected if $p_1 > 0.5$.

For example, in the case of getting $x_1 = 1000$ with $p_1 = 0.1$ and getting $x_2 = 10$ for sure ($p_2 = 1$), lottery 2 is more likely to be chosen because $p_1 > 0.5$ is not satisfied. This makes it easier to make an irrational choice as the expected value $p_1 x_1 = 100, p_2 x_2 = 10$.

Although the size of x_1, x_2 affects the decision of which lottery is more likely to be chosen, the difference between the two does not affect the decision. In other words, when choosing with one sample, there will be a bias to only look at the large and small relationship of the gains and ignore the difference in the gains.

The same lottery problem is often addressed in quantum decision making, but often the true expected value of the lottery is considered, i.e., it is discussed in situations where the sample is sufficiently large. The situation of limited thinking time in quantum decision making is expressed in the literature [10] by placing a limit on the number of times the amplification of the probability amplitude. In the literature [10], there is an assumption that positive information is always obtained, but in reality, as mentioned above, information can be obtained that encourages false choices through probability. Therefore, in the lottery problem, as described above, we examine the quantum decision-making under less information in the literature [10] to study the quantum decision-making interpretation of small samples.

B. Small Sample on Network

1) Bitrate Selection on Streaming: In bitrate selection, if the picture quality is selected and the QoE corresponding to the selected picture quality is the reward, it corresponds to the lottery described above. At this time, the QoE corresponding to the selected image quality will change probabilistically depending on the streaming environment. For example, if playback stops in the middle of streaming, the QoE is expected to drop significantly, and the QoE will change depending on the stochastic phenomenon of whether it stops or not.

Given the choice between high quality and low quality, the following options are given.

- Select high quality: QoE: x_1 without stopping with probability p_1, QoE: y with stopping with probability $1-p_1$.
- Select low quality: QoE: x_2 without stopping with probability p_2, QoE: y with stopping with probability $1-p_2$.

where p_1, p_2 is the probability that streaming can be played back without stopping when the respective image quality is selected, and $p_1 < p_2$. x_1, x_2 is the QoE of the respective image quality when not stopped. y is the QoE when stopped, and the QoE when stopped is assumed to be constant regardless of the image quality.

At this point, we can consider the same gain difference as in Table I, since the generality is not lost by $y = 0$. If p_1 is close to 1, choosing high quality will result in a higher expected QoE. In this case, high image quality is more likely to be chosen because it satisfies equation 1, and the choice based on one sample induces the correct choice.

In the case of $p_1 = 0.51, p_2 = 1$, which is close to the boundary condition of equation 1, it is easy to select a high quality even though the high quality is more likely to stop. Therefore, since the loss of QoE in the case of stopping is generally considered to be large, in this case, the user will be induced to choose a high quality even though the expected value of QoE is higher for low quality.

By integrating the decision-making of such a small sample into quantum decision-making, we can build a model that includes bias in the small sample and can be corrected by user agents.

2) Bitrate Selection with Optimal Selection: In the literature [4], the problem was like guessing the roll of a k-sided die, where there is only one choice that is the correct answer, and the rest are outliers. In the case of bitrate selection, we can correspond to the problem of choosing the optimal bitrate based on the observed situation.

Let p_i be the probability that i is the optimal bitrate when the ith bitrate is selected from among k types of bitrates. The optimal bitrate is defined as the maximum bitrate at which the video does not stop when played at that bitrate. This optimal bitrate is the i^*th bitrate. In this case, i^* is the roll of a k-sided die, which corresponds to the problem of guessing the roll.

What we measure as a sample is the throughput R_t at time t. The i-th bitrate is R_i, and the R_i that is the largest among the bitrates under R_t is the optimal bitrate R_{i^*}, so i^* is determined from the throughput R_t measured as follows.

$$i^*(r_t) = \arg\max_i \{R_i | R_i < r_t\} \quad (2)$$

The more samples, the more accurate the throughput can be measured and the more accurate the selection can be, but the more time is required for measurement. Therefore, it is desirable to make the appropriate selection with a small number of samples.

The difficulty in discriminating the appropriate choice depends on the magnitude of the bias in p_i, but in the present case, the temporal stability of the measured throughput r_t determines the bias in p_i. In other words, when the throughput is stable, the probability p_{i^*} of obtaining the observation corresponding to the optimal choice is high, making it easy to discriminate it from other choices. On the other hand, when the throughput is unstable, the probability of obtaining observations corresponding to truly suboptimal choices also increases to some extent, making it difficult to discriminate among those choices. This means an increase in the number of samples required for discrimination, but if there is little difference between the alternatives, the benefit of sampling will be small, so a small number of samples is desirable when considering the balance with sampling cost.

3) Network Monitoring: An example of decision making by a small sample of people in their daily lives is when deciding between a route with a bridge or a tunnel to a destination, choosing the less crowded route based on the experience of a small number of people, considering whether the bridge or the tunnel is more crowded. The above bitrate selection is based on this. The bitrate selection above corresponds to such an example, where a small sample of successful and unsuccessful experiences of whether the chosen bitrate was optimal is used to select the more successful choice.

Another example of how people use a small sample is to learn a word that refers to an object, such as [4], or to recognize a specific state. This corresponds to monitoring the network and estimating its state from a small amount of monitoring data.

Adaptive streaming, such as MPEG-DASH, changes the bitrate adaptively by measuring the network throughput. The throughput is usually measured passively when downloading a video segment from a server [13]. That is, the throughput is estimated by dividing the file size of the downloaded video segment by the time taken to download it. However, the number of video segments downloaded at a time depends on the capacity of the buffer on the player, and in extreme cases, there are problems such as no throughput measurement when the buffer is full. For this reason, active throughput measurement using probe packets has been proposed [14]. However, although sending a large number of probe packets enables accurate measurement, the traffic for measurement strains the network bandwidth, so here again, it is desirable to estimate the bandwidth from a small number of probe packet samples.

If we divide the throughput into k discrete levels and estimate which level it is from the samples by probe packets, we can treat it as equivalent to the problem of guessing the roll on a k-sided die. In the case of adaptive streaming, the k-step partitioning of the throughput is obtained by Voronoi partitioning (partitioning bounded by points equidistant from the reference point) concerning the bitrate that can be selected.

III. Quantum Decision Making with Small Sample

In the literature [10], quantum decision-making under insufficient information is represented using Grover's algorithm and state update with probability amplitude amplification. Therefore, we model sample-based decision-making as quantum decision-making by mapping the acquisition of information about alternatives by samples to state updating.

A. Grover's Algorithm

Grover's algorithm is a search algorithm in quantum computers that uses quantum superposition to speed up the search of a database. In this case, the search refers to finding one target element from N items, and the search is performed by querying the function $f(x)$ whether an element x is the target of the search or not. In the classical algorithm, $O(k)$ queries are required, but in Grover's algorithm, the search can be done with $O(\sqrt{k})$ queries.

The search with Grover's algorithm is performed by iterating the diffusion U_D and querying the database U_ω for the k-dimensional state vector $|x\rangle$.

$$\begin{aligned} |x_{t+1}\rangle &= U_D U_\omega |x_t\rangle && (3)\\ U_D &= 2|D\rangle\langle D| - I && (4)\\ U_\omega &= I - 2|\omega\rangle\langle\omega| && (5)\end{aligned}$$

Here, I is the unit matrix, $|\omega\rangle$ is the target state of the search, and $|D\rangle = \frac{1}{\sqrt{k}}\sum_x |x\rangle$ is the uniform superposition state.

At each step, the probability amplitude of $|\omega\rangle$ is amplified, so that after a sufficient number of iterations, $|x_t\rangle \simeq |\omega\rangle$ is obtained. By making observations in this final state, we can obtain the target ω of the search with high probability.

B. Amplitude Amplification

The literature [10] uses probability amplitude amplification, a generalization of Grover's algorithm, to express early and late decisions by making the time required for a decision correspond to the number of iterations of probability amplification. Since the initial state $|x_0\rangle$ influences the decision in the case of early decision, the probability amplitude in the initial state is regarded as the availability of the alternatives to express the availability bias.

In the stochastic amplitude amplification, instead of $|D\rangle$ in Grover's algorithm, we use $|D(\theta)\rangle$, which is generalized by introducing the parameter θ.

$$|D(\theta)\rangle = \sin\theta\,|\omega\rangle + \cos\theta\,|\omega^\perp\rangle \tag{6}$$

Here, $\omega^\perp$ is the orthogonal state to $|\omega\rangle$. If the initial state is $|x\rangle = \sin\theta\,|\omega\rangle + \cos\theta\,|\omega^\perp\rangle$, then after n amplifications, the probability amplitude of $|\omega\rangle$ is amplified to $\sin(2n+1)\theta$. In the case of Grover's algorithm, $\theta = \pi/4N$.

In Grover's algorithm and the literature [10], it is assumed that the correct state $|\omega\rangle$ is always amplified at each step. In the case of sample-based decision making, since the information obtained is probabilistic, it is natural that the amplified state $|\omega\rangle$ changes according to the information obtained.

For example, in the lottery example above, let $|\omega_1\rangle, |\omega_2\rangle$ be the state that selects each lottery, and amplify the state ω_i that selects the lottery with the largest gain $i(X_i > X_j)$ obtained from the samples. At this time, if the initial state is $|x_0\rangle = |D\rangle$ and the probability amplitude is uniform, the bias of the selection will reflect the bias due to a small number of samples.

Thus, in amplifying the probability amplitude, by determining the target to be amplified from the sample, it is possible to map sample-based decision making to quantum decision making. When using probability amplitude amplification, there is a degree of freedom in determining θ, but since the magnitude of the reward affects the choice probability according to probability matching, it is considered appropriate that θ also reflects the magnitude of the reward. In other words, we set θ so that the ratio of rewards corresponds to the ratio of amplified choice probabilities.

$$\sin^2((2n+1)\theta) : \cos^2((2n+1)\theta) = \frac{R_\omega}{n} : \frac{R_{\omega^\perp}}{n} \tag{7}$$

Here, $R_\omega, R_{\omega^\perp}$ represents the reward for choosing ω and the reward for choosing otherwise. Thus, the state update of quantum decision-making can naturally incorporate the state update equivalent of probability matching, which gives a kind of optimal selection policy in the multi-armed bandit problem.

IV. Analysis

A. Choice Problem Setting

In the literature [4], the problem of sample-based selection in multiple choice is addressed. In the multiple-choice problem, choosing option i out of k options yields a reward $x > 0$ with probability p_i and a reward $y = 0$ with probability $1 - p_i$ where $\sum_i p_i = 1$. However, when choice i gives a reward x, the other choices $j \neq i$ give a reward $y = 0$. This corresponds to the choice of guessing the roll of the k-sided die. It also does not lose generality as $p_1 > p_2 > \cdots > p_k$. In this case, the correct answer is selecting the choice $i = 1$ with maximum p_1 from the obtained samples.

If there are k alternatives, let n_i be the number of samples in which choosing alternative i rewards x. Then $n_1, \cdots, n_k$ follows a multinomial distribution with the total number of samples $n = \sum_i n_i$ and the probability $p_1, \cdots, p_k$ as a parameter . Therefore, the expected value, variance, and covariance of each follow.

$$\begin{aligned} E[n_i] &= np_i && (8)\\ V[n_i] &= np_i(1-p_i) && (9)\\ Cov[n_i, n_j] &= -np_ip_j. && (10)\end{aligned}$$

B. Quantum Decision Making Setting

We build a model with quantum decision-making by amplifying the probability amplitude of each alternative according to the sample of events. In other words, when an event (e_H) occurs with probability p, the probability of choosing the option (ω_e) is amplified.

In database search, Grover's algorithm required $O(\sqrt{k})$ times amplification of the probability amplitude when k items are present. On the other hand, in the literature [4], the number of samples required is proportional to $\log k$ for k number of choices. This difference arises mainly from the difference in the percentage of correct answers we are trying to achieve.

The number of amplifications in Grover's algorithm is calculated based on the assumption of a sufficiently high probability of getting the correct answer. On the other hand, in the literature [4], the correct response rate decreases as k increases, with a correct response rate of about 50% for $k = 4$ and a reward to sampling cost ratio of $\gamma = 1000$. In the literature [4], sampling is stopped when a certain percentage of correct answers is achieved because the change in the percentage of correct answers due to obtaining samples is commensurate with the sampling cost.

C. Condition of Correct Choice

Suppose that we choose the option with the largest n_i as a sample-based decision, we make the choice with the largest expected value of reward when n_1 is the largest. In other words, making the right choice is equivalent to the following event S.

$$S = \{n_1 > n_i | i = 2, \cdots, k\} \tag{11}$$

Therefore, the goal is to find the required number of samples as n such that the probability $P(S)$ of S occurring is above a certain level.

Although $P(S)$ depends on the probability $p_1, \cdots, p_k$ of each choice, depending on the value of $p_1, \cdots, p_k$, the sampling cost may exceed the gain of distinguishing the choices. For this reason, it makes sense to analyze in the parameter region where the gain from distinguishing the alternatives exceeds the sampling cost.

If the cost per sample is a, then for each option i, the condition that the gain is greater than or equal to the sampling cost is

$$an < x(p_1 - p_i) \; i = 2, \cdots, k. \tag{12}$$

but $p_2 > p_i (i = 3, \cdots, k)$, so we end up with

$$an < x(p_1 - p_2). \tag{13}$$

Thus, when there is gap between the probability of the most probable option and that of the second option, it makes sense to obtain a sample.

D. Approximation of Optimal Number of Samples

To discuss the factors that determine the number of samples required, we derive an approximate formula for the number of samples.

Expanding $P(S)$ by conditional probability, we get

$$\begin{aligned} P(S) &= P(n_1 > n_2, \cdots, n_1 > n_k) & (14)\\ &= P(n_1 > n_2 | n_1 > n_3, \cdots) \cdots P(n_1 > n_k) & (15) \end{aligned}$$

Now, if we approximate $n_1 > n_i$ as being independent of each other, then

$$P(S) = P(n_1 > n_2) \cdots P(n_1 > n_k). \tag{16}$$

Therefore, the condition that $P(S) > \theta$ is obtained by taking the logarithm of both sides

$$\sum_i \log P(n_1 > n_i) > \log \theta. \tag{17}$$

The expected value and variance of $n_1 - n_i$ is, due to its linearity

$$\begin{aligned} E[n_1 - n_i] &= n(p_1 - p_i) & (18)\\ V[n_1 - n_i] &= n\{p_1(1 - p_1) + p_i(1 - p_i) + 2p_1 p_i\} & (19) \end{aligned}$$

From the one-sided Chebyshev's inequality, we get

$$\begin{aligned} P(n_1 > n_i) &= 1 - P(n_1 \le n_i) & (20)\\ &= 1 - P(n_i - n_1 \ge 0) & (21)\\ &\ge \frac{E[n_1 - n_i]^2}{V[n_1 - n_i] + E[n_1 - n_i]^2}. & (22) \end{aligned}$$

Since the expected value and variance are proportional to n, we can divide n and the proportionality constant (depending on p_1, p_i) as $E[n_1 n_i] = nC_E(p_i), V[n_1 - n_i] = nC_V(p_i)$. Thus, we can rewrite the inequality as

$$P(n_1 > n_i) \ge \frac{1}{\frac{1}{n}\frac{C_V(p_i)}{C_E(p_i)} + 1}. \tag{23}$$

E. Optimal Number of Samples under Multiple Choices

In order to examine the effect of increasing the number of choices, we will analyze the situation where the effect of the number of choices is large, i.e., when there is a certain amount of probability that the reward will be obtained with the second or later choice. The case with fewer choices is included as a case where the probability is biased toward a particular choice, as in $p_i = 0 (i > k')$. Therefore, the effect of having more choices appears largest when the bias of p_i is small. Assuming that $p_2 = \cdots = p_k = p$, in which the effect of the number of choices are maximized, then

$$\begin{aligned} \log P(S) &\ge \sum_i \log \frac{1}{\frac{1}{n}\frac{C_V(p_i)}{C_E(p_i)} + 1} & (24)\\ &= (k-1) \log \frac{1}{\frac{1}{n}\frac{C_V(p)}{C_E(p)} + 1} > \log \theta. & (25) \end{aligned}$$

Therefore, we obtain the following conditional expression for n.

$$n \ge \frac{C_V(p)}{C_E(p)} \frac{1}{\theta^{\frac{1}{1-k}} - 1} \tag{26}$$

The $C_E(p) = p_1 - p$ represents the difference in the probability of getting a reward for choosing the correct option, and when this is small, n becomes infinitely large, but the sampling cost requires that $p_1 - p$ be somewhat large within meaningful parameters.

F. Numerical Simulation and Analysis Result

To check the validity of the analysis results, we compare the required number of samples determined by the analysis (Equation 26) with the number of samples calculated by the numerical simulation.

In the numerical simulation, the operation of selecting the option with the largest number of samples when n samples are obtained from the k-nomial distribution is taken as one trial, and sample-based selection of $N = 1000$ trials was performed for each n, k combination. For each n, k, the number of times the choice with the highest probability of winning is selected is $N_s(n, k)$, and the required number of samples n^* is n such that the proportion of correct choices $P(\bar{S}) = \frac{N_s(k,n)}{N}$ exceeds a certain probability θ.

$$n^* = min\{n|P(\bar{S}) = \frac{N_s(k, n)}{N} > \theta\} \quad (27)$$

The required number of samples n^* obtained by this numerical simulation is compared to the lower bound of n in Equation 26. The solid line represents the result of the numerical simulation, and the dashed line represents the lower bound of equation 26. The results for different values of θ are also shown in different colors. The probability was set to $\frac{p_1}{p} = 1.75$ $(p = p2 = \cdots = p_k)$.

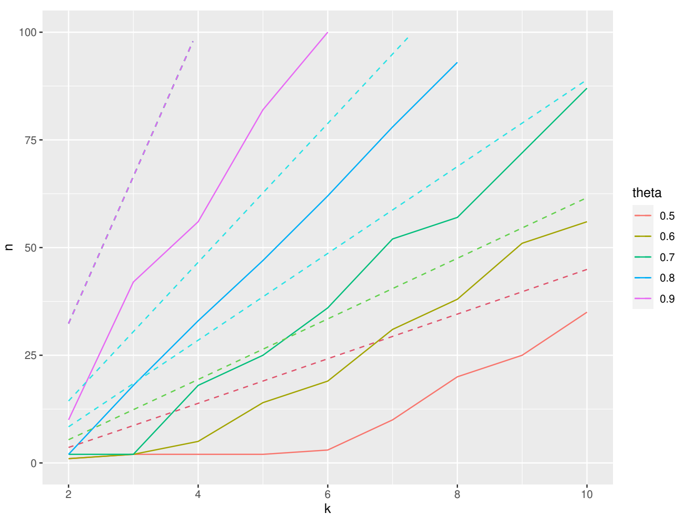

Fig. 1. Numerical Results and Approximation of Optimal Number of Samples

From the figure, we can see that the lower bound on the number of samples in Equation 26 is larger than the actual number of samples required. This is because, in the process of analysis, the lower limit of the probability is estimated by Chebyshev's inequality for the difference $n_1 - n_i$ between the number of times the best option and the other options are hit. Since Chebyshev's inequality holds for any probability distribution, it tends to give a conservative lower bound for a specific probability distribution. Therefore, a large number of samples is obtained, which is estimated on the safe side than the actual sample required.

Although there is a difference in the values, there is a similar trend in the way the required number of samples increases when the number of choices increases. In other words, the results of both numerical simulation and analysis show that the required number of samples increases almost linearly with the number of choices.

The number of samples required to achieve $\theta = 0.5$, that is, the probability of selecting the correct answer is more than 50%, even when there are 10 choices, the number of samples does not exceed 50 in both the numerical simulation and analysis results, and only a relatively small number of samples are required.

This indicates that the lower limit of the number of samples in Equation 26 is appropriate to be used as a safe estimate of the required number of samples, and that the number of samples is small even if it is estimated on the safe side.

V. Conclusion

In this paper, we model small sample-based decision-making by quantum decision-making for network control with a small sample. In the proposed model, the cognitive state specifies the probability of choosing each option, and the change in the sample-based choice is modeled by updating the cognitive state based on the sample. Specifically, the state is updated by a process of quantum amplification of probability amplitudes for the choices associated with the sampled values. We also analyzed the behavior of the proposed model. The results of the analysis show that the number of samples required to make a correct choice is sufficiently small, as in the literature [4].

In future work, we will work on the application of the proposed model for decision making with a small number of samples in actual network control, such as bitrate selection in streaming and network measurement.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP20K14735.

References

[1] G. P. Fettweis, "The tactile internet: Applications and challenges," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, 2014.

[2] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, and A. Ahmed, "Edge computing: A survey," *Future Generation Computer Systems*, vol. 97, pp. 219–235, 2019.

[3] Z. Li, M. A. Uusitalo, H. Shariatmadari, and B. Singh, "5G URLLC: Design challenges and system concepts," in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, 2018, pp. 1–6.

[4] E. Vul, N. Goodman, T. L. Griffiths, and J. B. Tenenbaum, "One and done? optimal decisions from very few samples," *Cognitive science*, vol. 38, no. 4, pp. 599–637, 2014.

[5] R. Hertwig and T. J. Pleskac, "Decisions from experience: Why small samples?" *Cognition*, vol. 115, no. 2, pp. 225–237, 2010.

[6] V. Yukalov, E. Yukalova, and D. Sornette, "Information processing by networks of quantum decision makers," *Physica A: Statistical Mechanics and its Applications*, vol. 492, pp. 747–766, 2018.

[7] V. I. Yukalov and D. Sornette, "Quantitative predictions in quantum decision theory," *IEEE Transactions on systems, man, and cybernetics: systems*, 2016.

[8] J. M. Yearsley and J. R. Busemeyer, "Quantum cognition and decision theories: A tutorial," *Journal of Mathematical Psychology*, vol. 74, pp. 99 – 116, 2016, foundations of Probability Theory in Psychology and Beyond.

[9] T. Otoshi and M. Murata, "Implementation of quantum decision-making based recommendation method for adaptive bitrate streaming," in *2019 22nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*. IEEE, 2019, pp. 25–30.

[10] R. Franco, "Quantum amplitude amplification algorithm: an explanation of availability bias," in *International Symposium on Quantum Interaction*. Springer, 2009, pp. 84–96.

[11] N. Vulkan, "An economist ' s perspective on probability matching," *Journal of economic surveys*, vol. 14, no. 1, pp. 101–118, 2000.

[12] D. R. Wozny, U. R. Beierholm, and L. Shams, "Probability matching as a computational strategy used in perception," *PLoS computational biology*, vol. 6, no. 8, p. e1000871, 2010.

[13] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, "A control-theoretic approach for dynamic adaptive video streaming over http," in *Proceedings of SIGCOMM*, 2015, pp. 325–338.

[14] D. Yun, K. Chung, and J. Hong, "Efficient bandwidth estimation for http adaptive streaming," in *The International Conference on Information Networking 2014 (ICOIN2014)*. IEEE, 2014, pp. 464–468.