WILEY | Hindawi

Research Article

Memory Effect on Adaptive Decision Making with a Chaotic Semiconductor Laser

**Takatomo Mihana [ID],[1] Yuta Terashima,[1] Makoto Naruse [ID],[2]
Song-Ju Kim [ID],[3] and Atsushi Uchida [ID][1]**

[1]*Department of Information and Computer Sciences, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City,
Saitama 338-8570, Japan*
[2]*Network System Research Institute, National Institute of Information and Communications Technology, 4-2-1 Nukui-kita,
Koganei, Tokyo 184-8795, Japan*
[3]*Graduate School of Media and Governance, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-0882, Japan*

Correspondence should be addressed to Takatomo Mihana; t.mihana.229@ms.saitama-u.ac.jp

We investigate the effect of a memory parameter on the performance of adaptive decision making using a tug-of-war method with the chaotic oscillatory dynamics of a semiconductor laser. We experimentally generate chaotic temporal waveforms of the semiconductor laser with optical feedback and apply them for adaptive decision making in solving a multiarmed bandit problem that aims at maximizing the total reward from slot machines whose hit probabilities are dynamically switched. We examine the dependence of making correct decisions on different values of the memory parameter. The degree of adaptivity is found to be enhanced with a smaller memory parameter, whereas the degree of convergence to the correct decision is higher for a larger memory parameter. The relations among the adaptivity, environmental changes, and the difficulties of the problem are also discussed considering the requirement of past decisions. This examination of ultrafast adaptive decision making highlights the importance of memorizing past events and paves the way for future photonic intelligence.

1. Introduction

Artificial intelligence based on deep learning, as a type of supervised learning, has been rapidly deployed in society. Reinforcement learning is another branch of machine learning that involves trial-and-error processes to accommodate unknown agents in environments [1, 2]. Its areas of application include, but are not limited to, robotics [3] and computer gaming [4]. The multiarmed bandit (MAB) problem is a fundamental problem in reinforcement learning wherein the total reward (e.g., the number of coins) from multiple slot machines with unknown hit probabilities needs to be maximized [1, 2, 5]. In order to solve the MAB problem, it is important to estimate the slot machine that may exhibit the highest hit probability by playing a series of slot machines (called *exploration*) and to use the estimation to gain more rewards (called *exploitation*). A certain amount of exploration action is necessary to estimate the best slot machine, because

insufficient exploration results in the failure of the estimation of the best slot machine. Further, overexploration reduces the extent of exploitation and results in the loss of the total reward. Therefore, the trade-off of the *exploration-exploitation dilemma* [1, 2] has been known in the MAB problem.

A variety of techniques have been proposed to solve the MAB problem, such as ε-greedy [2], soft-max [2, 6], and upper confidence bound algorithms [7]. One of the promising approaches adopted to solve the MAB problem, the tug-of-war (TOW) method, was proposed by Kim et al. [8, 9]. The idea of the TOW method has been inspired by the behavior of the unicellular amoeba of the true slime mold [10]. The volume of the amoeba's body remains constant when the amoeba oscillates its branches to collect environmental information, and it realizes nonlocal correlation under fluctuation [9]. These characteristics can be used for exploration

and exploitation purposes to solve the MAB problem. It has been shown that the TOW method is superior to the soft-max algorithm for adaptive decision making when the hit probabilities of multiple slot machines are changed [8].

Physical implementation of the TOW method has been demonstrated in photonic systems with quantum dots at the nanoscale [11, 12] and single photons [13]. The nonlocality and fluctuation required for the TOW result from the physical attributes of the wave nature of an exciton-polariton and a single photon. However, the speed of the fluctuations is limited by the experimental measurement systems in the order of Hertz. Recently, ultrafast adaptive decision making based on the TOW method has been proposed using fast chaotic laser outputs over the gigahertz range [14, 15]. Decision making with zero-prior knowledge has been achieved at 1-GHz rate, and the sampling of temporal waveforms with negative correlation improves the performance of adaptive decision making [14].

For adaptive decision making where the hit probabilities of multiple slot machines are changed over time, it is important to determine how much of the knowledge accumulated via past exploration needs to be incorporated to make the present decision. For this purpose, the memory parameter (also known as the forgetting parameter [8]) has been introduced in the TOW method to include past exploration results. The tuning of the memory parameter is crucial for adaptive decision making because, for example, past exploration results may be useless after the hit probabilities are changed. However, no systematic investigation has been provided regarding the effect of the memory parameter for adaptive decision making using the TOW method.

In this study, we investigate the effect of the memory parameter on the performance of adaptive decision making using the TOW method with the chaotic output of a semiconductor laser. We experimentally generate chaotic temporal waveforms of the semiconductor laser with optical feedback, and we apply them for the TOW method for adaptive decision making. We investigate the performances of decision making for different values of the memory parameter.

2. Tug-of-War Method Using a Chaotic Semiconductor Laser

We consider an MAB problem under the following conditions as a simple case. We assume two slot machines A and B (referred to as S_A and S_B) with unknown hit probabilities (referred to as P_A and P_B). We consider that a player selects one of the slot machines each time, and the player either earns or loses the reward (e.g., coins) if the result of the selected slot machine is *"hit"* (or WIN) or *"miss"* (or LOSE), respectively. The player intends to maximize the reward by a good strategy of selecting one of the two slot machines. We assume that the total number of trials to play the slot machines is fixed. Besides, the sum of the hit probabilities of the two slot machines are supposed to be fixed at 1 ($P_A + P_B = 1$) as prior knowledge [14].

We use chaotic temporal waveforms generated experimentally from a semiconductor laser with optical feedback for decision making. Fast chaotic laser outputs have been
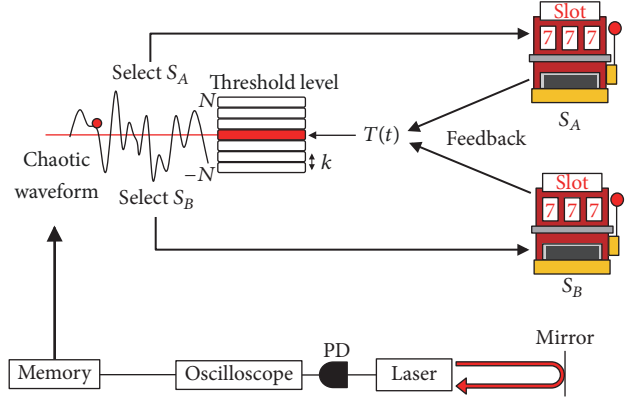

FIGURE 1: Decision-making system architecture based on the tug-of-war method using chaotic temporal waveforms of the semiconductor laser. PD, photodetector; S_A, slot machine A; S_B, slot machine B.

used for applications in fast physical random number generation [16–18], secure key distribution [19, 20], and reservoir computing [21, 22]. Figure 1 shows the decision-making scheme based on the TOW method using chaotic temporal waveforms of the semiconductor laser. First, we use a semiconductor laser with an external mirror to generate chaotic laser outputs. Chaotic temporal waveforms of the laser output are measured with a digital oscilloscope. We sample the chaotic temporal waveform by an analog-to-digital converter (ADC) with 8-bit vertical resolution. Meanwhile, we prepare a "threshold" value to be compared with the sampled data.

We introduce the following rule for selecting one of the two slot machines based on the sampled data of the chaotic temporal waveform. We decide to select the slot machine A (S_A) if the sampled data is *larger* than the threshold, and we decide to select the slot machine B (S_B) if the sampled data is *smaller* than the threshold.

We update the threshold value based on the betting results. For example, if we select S_A and the result is "hit," the threshold value *decreases* so that the probability of selecting S_A (i.e., the range above the threshold on the chaotic waveform) may increase in the next step. On the contrary, the threshold value *increases* if we select S_A and the result is "miss" so that the probability of selecting S_A may decrease. The threshold needs to be shifted to the opposite direction if we select S_B; that is, the threshold increases or decreases if the result is "hit" or "miss," respectively. After repeating this procedure by selecting one of the slot machines and changing the threshold value, the threshold value saturates to the bottom or top of the amplitude of the chaotic temporal waveform, which corresponds to an equilibrium of selecting S_A or S_B, respectively.

The threshold value, denoted by $T(t)$, is changed with respect to the chaotic temporal waveform as follows:

$$T(t) = \begin{cases} kN & ((\text{int})\,TA\,(t-1) > N) \\ k\,(\text{int})\,TA\,(t-1) & (|(\text{int})\,TA\,(t-1)| \le N) \quad (1) \\ -kN & ((\text{int})\,TA\,(t-1) < -N), \end{cases}$$

TABLE 1: Value of $X(t)$.

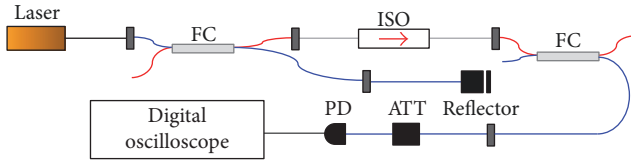| Selected slot machine	Hit	Miss
S_A	−1	+1
S_B	+1	−1

FIGURE 2: Experimental setup for the acquisition of chaotic temporal waveforms in the semiconductor laser with optical feedback. ATT, variable attenuator; FC, fiber coupler; ISO, optical isolator; PD, photodetector.

TABLE 2: Parameter values used for numerical simulation of decision making.

Parameters	Symbols	Values
Memory parameter	α	0.990 or 0.999
Width of threshold step	k	64
Number of threshold levels	N	2
Total number of threshold levels	$2N + 1$	5
Hit probability of S_A	P_A	0.4 or 0.6
Hit probability of S_B	P_B	0.6 or 0.4
Flip interval	FI	1000
Number of trials	m	5000
Number of cycles	n	1000

where k is the width of the threshold step and N is the number of threshold levels. The threshold step ranges from $-N$ to N (i.e., the total number of the threshold levels of $2N + 1$). These two parameters k and N limit the range of the threshold adjustment. In this study, we set $k = 64$ and $N = 2$ for 8-bit chaotic data ($2^8 = 256$ levels), and the total number of the threshold levels is $2N + 1 = 5$. TA(t) is the threshold adjuster variable that is determined by the result of the past decisions. The threshold adjuster variable TA(t) is defined as

$$\text{TA}(t) = X(t) + \alpha \text{TA}(t - 1), \qquad (2)$$

where $X(t)$ is the variable determined by the result of "hit" or "miss" from the selected slot machine, defined in Table 1. Here, α is the memory (or forgetting) parameter for weighting past threshold adjuster variables and ranges from 0 to 1. TA(t) is determined only by the present value of $X(t)$ for $\alpha = 0$, whereas TA(t) depends on the results of *all* of the past values of $X(t)$ if $\alpha = 1$. We investigate the influence of α on the performance of decision making in this study.

3. Experimental Setup for the Measurement of Chaotic Temporal Waveforms

We conducted a laser experiment to acquire chaotic temporal waveforms of the semiconductor laser output for decision making, as shown in Figure 2. We used a commercial semiconductor laser (NTT Electronics, KELD1C5GAAA, optical wavelength of ~1548 nm) for optical communication. The injection current of the semiconductor laser was set to 58.5 mA (5.0 I_{th}, where $I_{th} = 11.7$ mA is the lasing threshold). The wavelength of the laser was set to 1547.782 nm. The output light of the semiconductor laser was reflected by an external mirror (Reflector) and fed back to the laser to generate the chaotic fluctuation of the laser output. The feedback light power was set to 210 μW. The chaotic laser output was injected into a photodetector (PD, New Focus 1474A, 35 GHz bandwidth) and converted into an electric signal. The chaotic temporal waveforms were acquired by a high-speed digital oscilloscope (Tektronix, DPO73304D, 33 GHz bandwidth, 100 GigaSamples/s, 8-bit vertical resolution). The temporal

waveforms are stored in the memory of the oscilloscope and used for decision making.

Figure 3 shows a temporal waveform of the laser output and the histogram of the amplitude of the temporal waveform. From Figure 3(a), we observe that a fast chaotic irregular oscillation is obtained with the dominant oscillation period of 0.15 ns. The corresponding dominant frequency of chaotic oscillation is 6.6 GHz. In Figure 3(b), the histogram shows an asymmetric distribution with a long tail for the negative value of the amplitude with 8-bit resolution (from −128 to 128). The data of chaotic temporal waveforms are taken for 5 M points and used for decision making.

We set a sampling time of 10 ps for each decision, to demonstrate the adaptive decision making at the fastest rate. In our previous work [14], we found that the sampling time of 50 ps shows the best performance; however, this is only in the case of zero-prior knowledge (i.e., the sum of the two hit probabilities is unknown). Under the assumption of the presence of prior knowledge in this work (i.e., $P_A + P_B = 1$), the effect of sampling rate is not significant, at least within the problem investigated in this study.

4. Decision Making

4.1. Evaluation of Correct Decision Rate. We emulate the TOW method in numerical calculations with experimentally obtained chaotic temporal waveforms of the laser output (Figure 3). The parameter values used for the computation are summarized in Table 2. The number of trials for selecting machines S_A and S_B are set to $P_A = 0.4$ and $P_B = 0.6$, respectively. These two probabilities are switched every 1000 times to examine the adaptive characteristics of decision making (flip interval, FI = 1000). For example, $P_A = 0.4$ and $P_B = 0.6$ are used for the first 1000 trials, and $P_A = 0.6$ and $P_B = 0.4$ are used for the second 1000 trials, then $P_A = 0.4$ and $P_B = 0.6$ are used again for the third 1000 trials. We define 5000 trials of selecting the slot machines as one cycle, and we test 1000 cycles ($n = 1000$) to evaluate the average performance of decision making.

We evaluate the *correct decision rate* (CDR) defined as the ratio of the number of times of selecting the slot machine with the highest hit probability (best slot machine selection) to the
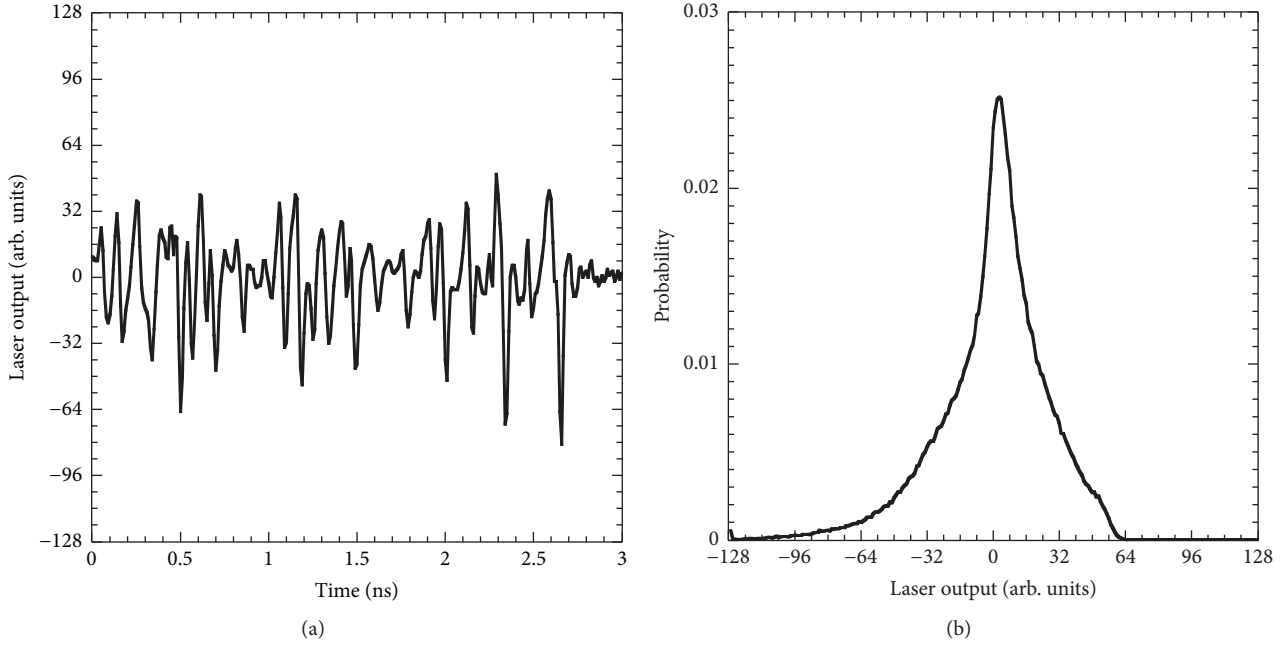
(a)

(b)

FIGURE 3: (a) Temporal waveform of the laser output and (b) the corresponding histogram of the amplitude of the temporal waveform.

(a)

(b)

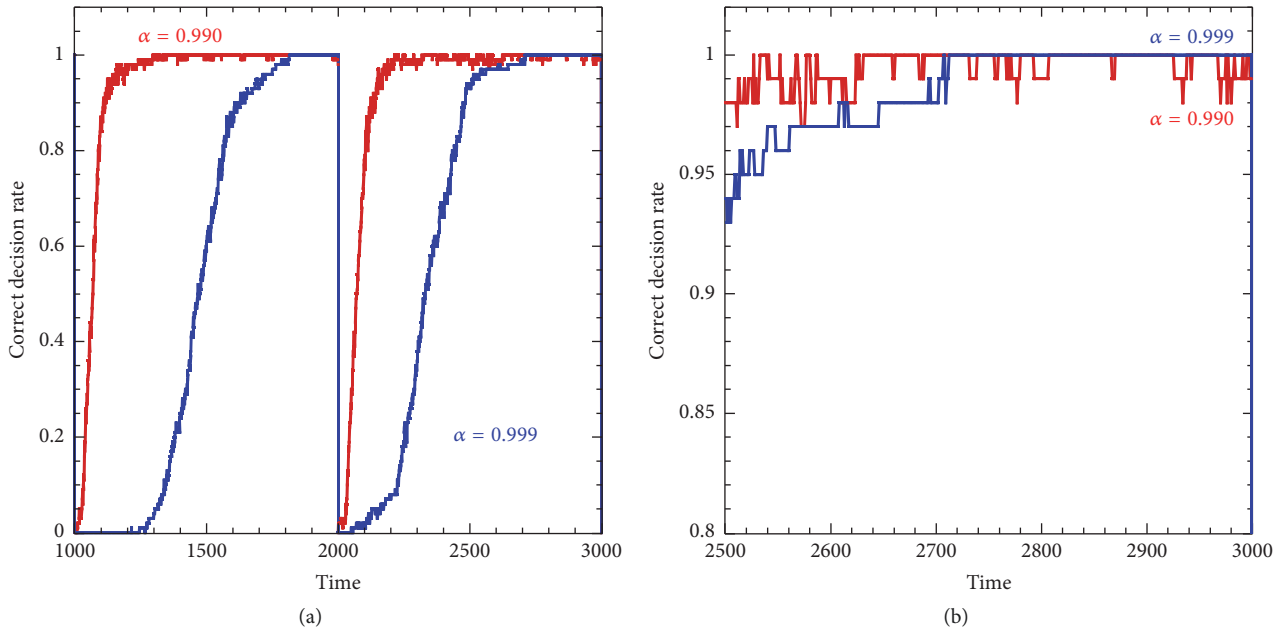FIGURE 4: (a) Temporal evolutions of the correct decision rate (CDR) for different values of the memory parameter [$\alpha = 0.990$ (red line) and 0.999 (blue line)] and (b) enlarged view of (a).

total number of times of selecting a slot machine and is given by

$$\text{CDR}(t) = \frac{1}{n} \sum_{i=1}^{n} C_i(t), \qquad (3)$$

where $C_i(t) = 1$ if the slot machine with the highest hit probability is selected; otherwise, $C_i(t) = 0$ at the tth trial and the ith cycle. n is the number of cycles ($n = 1000$).

Figure 4(a) shows the temporal evolutions of CDR with respect to two different memory parameters ($\alpha = 0.990$ and 0.999). With a smaller memory parameter $\alpha = 0.990$, the value of CDR quickly approaches 1.0 after the hit probabilities are switched (around $t = 1000$ and 2000); decisions are made adaptively to the swapping of the hit probabilities P_A and P_B (adaptation to environmental change). In contrast, with a larger memory parameter $\alpha = 0.999$, the recovery of the CDR
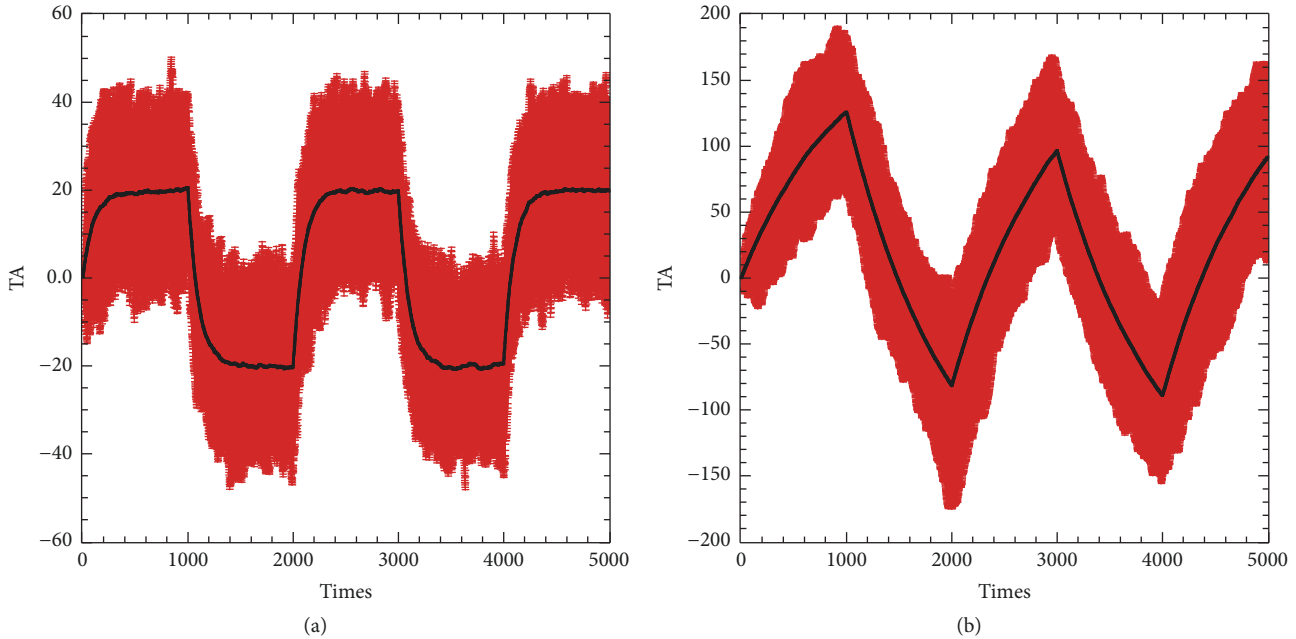
(a)

(b)

FIGURE 5: Time evolution of the threshold adjuster (TA) variable for (a) $\alpha = 0.990$ and (b) $\alpha = 0.999$. The red line indicates TA for 1000 cycles at the tth trial, and the black line indicates the average TA over 1000 cycles.

value is slow after the switching of the hit probabilities; that is, the adaptivity of decision making is inferior to the former case.

Figure 4(b) shows an enlarged view of Figure 4(a) to investigate the *convergence* of correct decision making. We found that the value of the CDR fluctuates around 1 in case of a small memory parameter ($\alpha = 0.990$) whereas the CDR converges to the value of 1 without fluctuations for a large memory parameter ($\alpha = 0.999$). This result indicates that correct decisions are made in a stable manner for a larger memory parameter after the transient of the change of the hit probabilities.

From Figures 4(a) and 4(b), we observe the trade-off between the adaptivity and the convergence regarding the CDR values. For a smaller α, a decision can be promptly made after environmental changes, but it suffers from a certain degree of instability. On the contrary, for a larger α, the adaptation to environmental changes is slow while the correct decision exhibits stable performances once the adaptation is made. Therefore, an adequate tuning of α is important to accomplish the demanded performances of decision making.

Figure 5 shows the time evolutions of the threshold adjuster (TA) variable (defined by (2)) for $\alpha = 0.990$ and $\alpha = 0.999$. The red line indicates TA for 1000 cycles at the tth trial, and the black line indicates the average TA over 1000 cycles. As can be seen in Figure 5(a) regarding the small α, the average TA (black line) saturates after the switching of hit probabilities where the TA value quickly reacted after the switching, leading to fast adaptation. However, the TA values for the 1000 cycles (red lines) show large deviations where some TA values exhibit the opposite sign (+ or −) below or above 0, corresponding to *wrong* selections of the slot machine. Such a deviation of TA values results in the

fluctuation of the CDR value after the change of the hit probabilities (Figure 4(b)). In contrast, for the large α in Figure 5(b), the average TA (black line) increases rapidly and monotonically before the occurrence of swapping of hit probabilities. As a result, it takes a certain duration to accomplish the change the sign of the TA value, leading to slow adaptation of decision making. At the same time, it is noteworthy that all the TA values (red lines) for 1000 cycles remain in either positive or negative, leading to the stable decision making observed in Figure 4(b). Therefore, we understand that the adaptivity and the convergence of decision making is determined by the evolution of TA values for different α.

To obtain further insights into the effect of α, we derive the following equation from (2):

$$\text{TA}\,(t) = \sum_{i=1}^{t} \alpha^{t-i} X\,(i)\,, \qquad (4)$$

which indicates that the TA value results from the past decision making with the exponential weight of the memory parameter α. Figure 6 shows the time evolution of the TA value described in (4) assuming $X(i) = 1$ for all i. For $\alpha = 0.990$, the TA value saturates before the 1000 trials, and the effect of the past results of the decision making decreases rapidly in the present TA value during 1000 trials. On the contrary, for $\alpha = 0.999$, the TA value increases monotonically over the 1000 trials, and most past results of the decisions are included in the present TA value. The memory parameter indicates how many of the past results are taken into account in updating the TA value; (4) (also Figure 6) provides a measure of how to determine an appropriate value of the memory parameter.
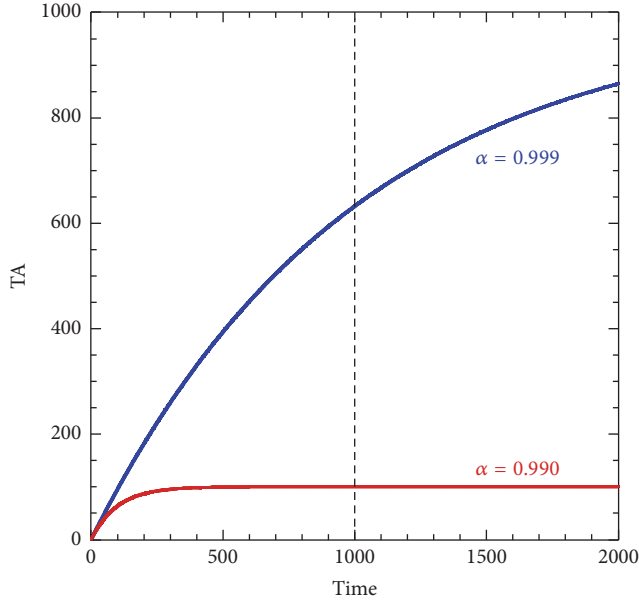
FIGURE 6: Time evolution of the threshold adjuster (TA) value described in (4) for $\alpha = 0.990$ (red line) and 0.999 (blue line) when $X(i) = 1$ is assumed for all i.
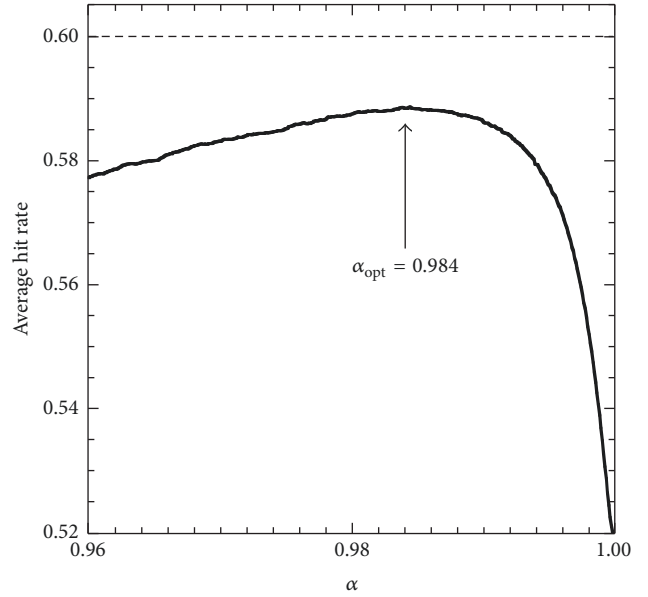

FIGURE 7: Average hit rate (AHR) as the memory parameter α is continuously changed when $P_A = 0.4$ and $P_B = 0.6$ are switched at every 1000 trials. The upper limit of AHR is 0.6 (dashed lines).

4.2. Evaluation of Average Hit Rate. Next, we investigate the memory parameter to maximize the total reward of decision making. We introduce the *average hit rate* (AHR), defined as follows:

$$\text{AHR} = \frac{1}{mn} \sum_{i=1}^{n} \sum_{t=1}^{m} H_i(t), \tag{5}$$

where $H_i(t) = 1$ if the selected slot machine is "hit" whereas $H_i(t) = 0$ if the selected slot machine is "miss" at the tth trial and the ith cycle. The AHR represents the total reward acquisition rate that is defined as the ratio of the number of "hits" (e.g., the number of acquired coins) to the total number of trials and cycles.

Figure 7 shows the AHR as the memory parameter α is continuously changed when $P_A = 0.4$ and $P_B = 0.6$ are switched every 1000 trials. The upper limit of the AHR is 0.6 because it is the maximum value of the hit probabilities. From Figure 7, we found that an *optimal* value of the memory parameter exists that maximizes the AHR. The maximum AHR is 0.589 when the memory parameter is set to $\alpha_{\text{opt}} = 0.984$. For larger α than α_{opt}, a longer transient time after the switching of P_A and P_B appears; hence AHR is reduced. For smaller α than α_{opt}, the transient time is shortened but incorrect decisions appear after the transient time due to the insufficient inclusion of past decision results, thereby reducing the AHR.

We investigate α_{opt} to maximize the AHR for different conditions of decision making. We first change the *flip interval* (FI) and investigate its impact to the maximum AHR and the corresponding optimal memory parameter α_{opt}. Figure 8 shows the maximum AHR and α_{opt} for different flip intervals. From Figure 8(a), the maximum AHR increases monotonically as the flip interval increases since the number of "hits"

increases during a larger flip interval. From Figure 8(b), the optimal memory parameter α_{opt} for maximizing the AHR also increases as the flip interval increases. For a larger flip interval, a longer memory is useful to determine the correct decision; the transient time after switching of the hit probabilities needs to be sufficient. Therefore, a larger α_{opt} is obtained when the flip interval increases.

Next, we change the hit probabilities of the two slot machines. Figure 9 shows the AHRs for three different combinations of the hit probabilities of the two slot machines: (i) $P_A = 0.8$ and $P_B = 0.2$, (ii) $P_A = 0.7$ and $P_B = 0.3$, and (iii) $P_A = 0.6$ and $P_B = 0.4$. That is, the *difficulty* of the given decision-making problem is configured. The maximum hit rates are, by definition, given as 0.8, 0.7, and 0.6, respectively. We found that an optimal memory parameter α_{opt} exists for all these cases; however, the values of α_{opt} are different. A larger α_{opt} is observed for case (iii) ($\alpha_{\text{opt}} = 0.988$) where the difference of the hit probabilities is small. This is a reasonable consequence that the increased difficulties of the given problem (i.e., smaller differences of the hit probabilities) require more memory of the past results to maximize the reward. In contrast, a smaller α_{opt} is obtained for case (i) ($\alpha_{\text{opt}} = 0.910$) when the problem is easier (large difference of the hit probabilities) because a smaller number of trials yield good decisions. When the hit probability of the slot machines is switched, it is important to estimate the correct slot machine quickly and change the decision rather than sticking to the old decision. In the case when the difference of the hit probabilities is large, the player often suffers from "miss" just after the switching; hence, the player can rather easily recognize the change of the slot machines, meaning that memorizing the past events is less important. In contrast, when the difference of the hit probabilities is small,
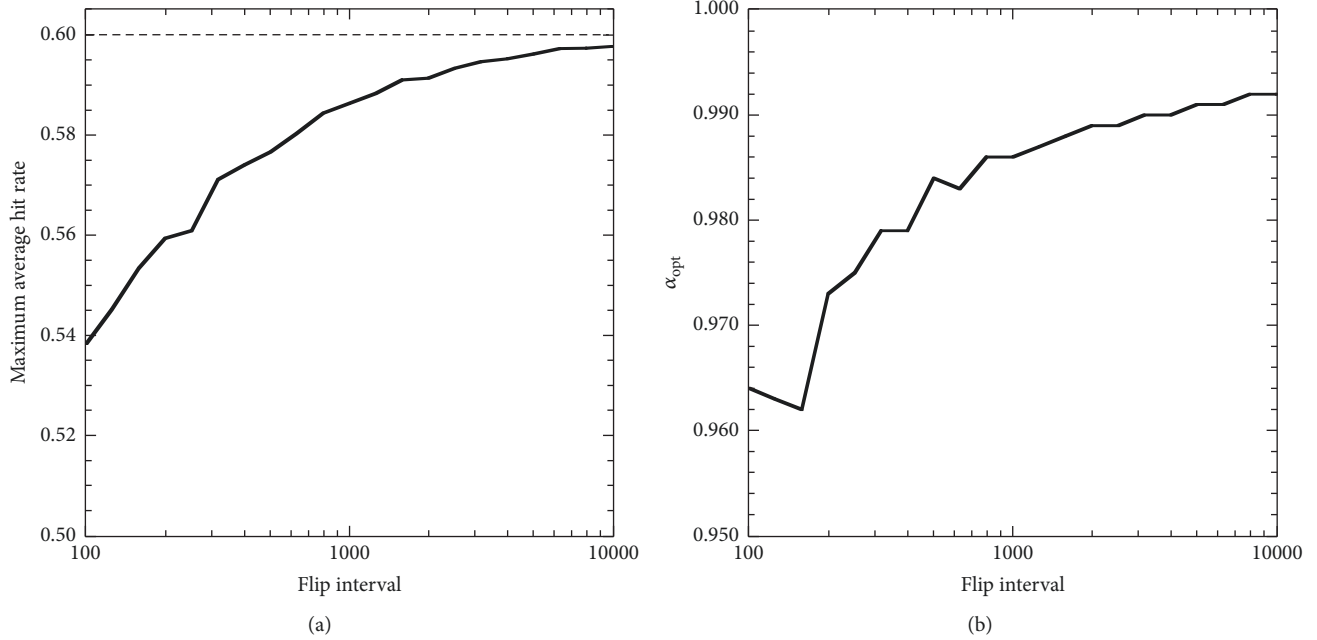
FIGURE 8: (a) Maximum average hit rate (AHR) and (b) the corresponding optimal memory parameter α_{opt} for different flip intervals. The upper limit of AHR is 0.6 (dashed lines).
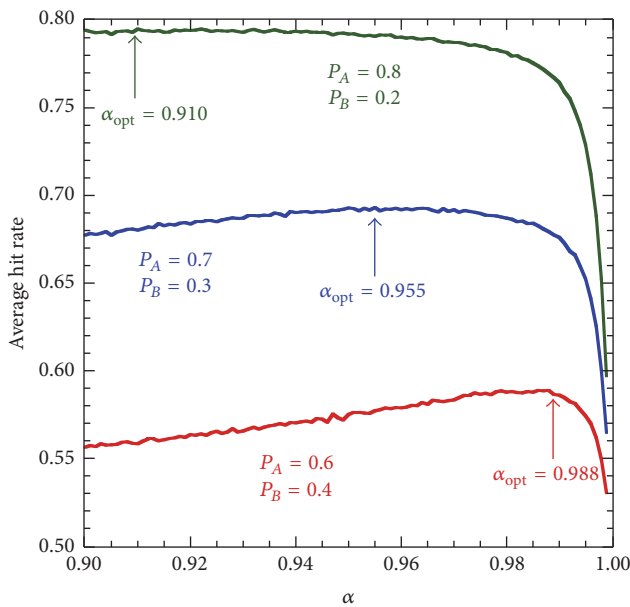

FIGURE 9: Average hit rates (AHR) for three different combinations of the hit probabilities of the two slot machines [(i) $P_A = 0.8$ and $P_B = 0.2$, (ii) $P_A = 0.7$ and $P_B = 0.3$, and (iii) $P_A = 0.6$ and $P_B = 0.4$] and their dependencies on the memory parameter. α_{opt} indicates the optimal memory parameter that maximizes AHR.

it is difficult to immediately recognize the change of the slot machines. A larger memory parameter is needed to figure out the correct decision.

These results show that using past results for decision making by using the memory parameter is crucial for improving the decision-making performance. The past decisions and their results are necessary for accurate decision making. However, we should note that they should not be taken into account too much concerning the environmental changes (e.g., switching of the hit probabilities). Taking account of the trade-off between the adaptation speed and the correctness of decision making, configuring the memory parameter is important to maximize the reward.

5. Conclusions

We investigated adaptive decision making based on the TOW method using the temporal waveforms of a chaotic semiconductor laser. We experimentally generated chaotic temporal waveforms of the semiconductor laser with optical feedback and applied them for adaptive decision making of the MAB problem aiming at maximizing the total reward from slot machines. We highlighted the requirements of memorizing the past in solving the MAB problem. We examined decision making in an uncertain environment, namely, the problem of choosing one of the two slot machines whose hit probabilities are dynamically switched and evaluated the effect of the memory parameter on the performance of adaptive decision making. Fast adaptation to the change in the hit probabilities can be obtained for a small memory parameter; however, the correct decision rate does not converge. In contrast, the correct decision rate converges perfectly for a large memory parameter, but the adaptation is slow. Thus, a trade-off exists between the adaptation speed and the convergence of the correct decision rate. An optimal memory parameter is found to maximize the average hit rate. We found that a larger memory parameter is needed for a larger flip interval (or slow environmental changes) and a smaller difference between the

hit probabilities of the slot machines (or difficulties of the decision-making problem).

Decision making using fast chaotic temporal waveforms generated from a semiconductor laser can be used for applications requiring an arbitration of resources at data centers [23] and resource allocation in wireless communications among others where decision making at the milliseconds level is required [24]. The use of chaotic lasers for decision making leads to a new research field of photonic intelligence.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

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

[2] R. S. Sutton and A. G. Barto, *Reinforcement Learning, An Introduction*, The MIT Press, Cambridge, London, UK, 1998.

[3] O. B. Kroemer, R. Detry, J. Piater, and J. Peters, "Combining active learning and reactive control for robot grasping," *Robotics and Autonomous Systems*, vol. 58, no. 9, pp. 1105–1116, 2010.

[4] D. Silver, A. Huang, C. J. Maddison et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[5] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[6] N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, and R. J. Dolan, "Cortical substrates for exploratory decisions in humans," *Nature*, vol. 441, no. 7095, pp. 876–879, 2006.

[7] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, no. 2-3, pp. 235–256, 2002.

[8] S.-J. Kim, M. Aono, and M. Hara, "Tug-of-war model for the two-bandit problem: nonlocally-correlated parallel exploration via resource conservation," *BioSystems*, vol. 101, no. 1, pp. 29–36, 2010.

[9] S.-J. Kim, M. Aono, and E. Nameda, "Efficient decision-making by volume-conserving physical object," *New Journal of Physics*, vol. 17, no. 8, p. 083023, 2015.

[10] T. Nakagaki, H. Yamada, and Á. Tóth, "Intelligence: maze-solving by an amoeboid organism," *Nature*, vol. 407, no. 6803, p. 470, 2000.

[11] S.-J. Kim, M. Naruse, M. Aono, M. Ohtsu, and M. Hara, "Decision maker based on nanoscale photo-excitation transfer," *Scientific Reports*, vol. 3, article no. 3297, 2013.

[12] M. Naruse, W. Nomura, M. Aono et al., "Decision making based on optical excitation transfer via near-field interactions between quantum dots," *Journal of Applied Physics*, vol. 116, no. 15, p. 154303, 2014.

[13] M. Naruse, M. Berthel, A. Drezet et al., "Single-photon decision maker," *Scientific Reports*, vol. 5, article no. 13253, 2015.

[14] M. Naruse, Y. Terashima, A. Uchida, and S.-J. Kim, "Ultrafast photonic reinforcement learning based on laser chaos," *Scientific Reports*, vol. 7, no. 1, article no. 8772, 2017.

[15] T. Mihana, Y. Terashima, M. Naruse, S.-J. Kim, and A. Uchida, "Adaptation of decision making with chaotic semiconductor laser," in *Proceedings of 2017 International Symposium on Nonlinear Theory and Its Applications*, vol. 1, pp. 275–278, 2017.

[16] A. Uchida, K. Amano, M. Inoue et al., "Fast physical random bit generation with chaotic semiconductor lasers," *Nature Photonics*, vol. 2, no. 12, pp. 728–732, 2008.

[17] R. Sakuraba, K. Iwakawa, K. Kanno, and A. Uchida, "Tb/s physical random bit generation with bandwidth-enhanced chaos in three-cascaded semiconductor lasers," *Optics Express*, vol. 23, no. 2, pp. 1470–1490, 2015.

[18] K. Ugajin, Y. Terashima, K. Iwakawa et al., "Real-time fast physical random number generator with a photonic integrated circuit," *Optics Express*, vol. 25, no. 6, pp. 6511–6523, 2017.

[19] K. Yoshimura, J. Muramatsu, P. Davis et al., "Secure key distribution using correlated randomness in lasers driven by common random light," *Physical Review Letters*, vol. 108, no. 7, p. 070602, 2012.

[20] H. Koizumi, S. Morikatsu, H. Aida et al., "Information-theoretic secure key distribution based on common random-signal induced synchronization in unidirectionally-coupled cascades of semiconductor lasers," *Optics Express*, vol. 21, no. 15, pp. 17869–17893, 2013.

[21] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature Communications*, vol. 4, article no. 1364, 2013.

[22] J. Nakayama, K. Kanno, and A. Uchida, "Laser dynamical reservoir computing with consistency: an approach of a chaos mask signal," *Optics Express*, vol. 24, no. 8, pp. 8679–8692, 2016.

[23] K. Bilal, S. U. R. Malik, S. U. Khan, and A. Y. Zomaya, "Trends and challenges in cloud datacenters," *IEEE Cloud Computing*, vol. 1, no. 1, pp. 10–20, 2014.

[24] K. Kuroda, H. Kato, S.-J. Kim, M. Naruse, and M. Hasegawa, "Improving throughput using multi-armed bandit algorithm for wireless LANs," *Nonlinear Theory and Its Applications, IEICE*, vol. 9, no. 1, pp. 74–81, 2018.