SCIENTIFIC REPORTS

OPEN

On-chip photonic decision maker using spontaneous mode switching in a ring laser

Ryutaro Homma[1], Satoshi Kochi[1], Tomoaki Niiyama[1,2], Takatomo Mihana[3], Yusuke Mitsui[3], Kazutaka Kanno[3], Atsushi Uchida [3], Makoto Naruse[4] & Satoshi Sunada[1,2]

Efficient and accurate decision making is gaining increased importance with the rapid expansion of information communication technologies including artificial intelligence. Here, we propose and experimentally demonstrate an on-chip, integrated photonic decision maker based on a ring laser. The ring laser exhibits spontaneous switching between clockwise and counter-clockwise oscillatory dynamics; we utilize such nature to solve a multi-armed bandit problem. The spontaneous switching dynamics provides efficient exploration to find the accurate decision. On-line decision making is experimentally demonstrated including autonomous adaptation to an uncertain environment. This study paves the way for directly utilizing the fluctuating physics inherent in ring lasers, or integrated photonics technologies in general, for achieving or accelerating intelligent functionality.

In the age of artificial intelligence, the requirements for efficient and intelligent processing of massive amount of data are continuously increasing. Present technologies to accommodate these demands rely on digital electronics; however, hardware scaling in electronics, as foreseen by Moore's law, is predicted to be unsustainable[1,2]. As a consequence, studies on novel computing principles and architectures beyond the present Turing–von-Neumann computing paradigm[2,3] are gaining importance. These include neuromorphic computing[4–6], photonic deep learning[7,8], reservoir computing[9,10], molecular computing[11], and quantum and coherent ising machines[12,13], where the underlying complexity and fluctuations in natural systems have been used for cognitive processing, prediction, and solving large-scale combinatorial optimization problems.

From the view of information processing, most of the above-mentioned studies have focused on supervised learning or optimization processing. Reinforcement learning is another emergent and important branch of machine learning[14], where utilization of physical processes can enhance or accelerate their performance[15–18].

As a foundation of reinforcement learning, decision making plays a key role in engineering applications such as cognitive wireless communication[19,20], online advertisements[21], and Monte-Carlo searches[22]. Herein, the decision-making problems under study are called multi-armed bandit (MAB) problems[23]; the goal is to maximize the total reward from multiple slot machines whose reward probabilities are unknown. A key point of the MAB problems is to resolve the *exploration-exploitation dilemma* inherent in decision making under uncertainty: sufficient exploratory actions may inform the best slot machine, but it may be accompanied by a significant amount of losses. In contrast, insufficient exploration may result in missing the best machine.

Recently, we have experimentally revealed that optical fluctuation dynamics can be used for exploring and making an optimal decision in the MAB problems[24–27]. Particularly, it has been found that complex temporal waveforms generated from a chaotic laser are useful for making decisions at a fast rate in the gigahertz regimes[26]. Previous studies suggest the potential of using complex laser dynamics with ultra-wide bandwidth for fast decision making. However, important issues remain open ranging from novel fundamental principles, system architectures, to device implementations for photonic decision making. For instance, the former studies using laser chaos[26,27] *only* exploit chaotic waveforms as correlated random numbers for a decision making software

[1]Graduate School of Natural Science and Technology, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa, 920-1192, Japan. [2]Faculty of Mechanical Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa, 920-1192, Japan. [3]Department of Information and Computer Sciences, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama, 338-8570, Japan. [4]Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan. Correspondence and requests for materials should be addressed to S.S. (email: sunada@se.kanazawa-u.ac.jp)
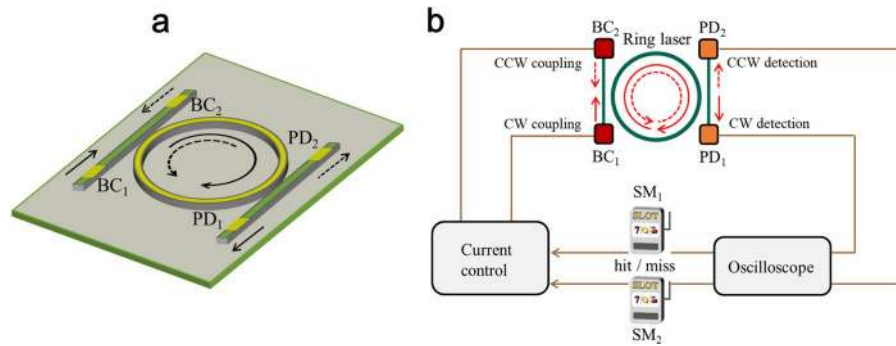
Figure 1. Ring-laser-based decision-making. (**a**) Schematic of the ring laser device coupled to waveguides with contact electrodes BC_i and PD_i ($i = 1, 2$). The ring radius is 1 mm, and the waveguide width is approximately $2\,\mu m$. $PD_{1(2)}$ was used as the photodetectors to monitor the CW(CCW) mode intensity in the ring laser, whereas $BC_{1(2)}$ was used for introducing an asymmetry and changing the mode-dynamics. (**b**) Setup for the proof-of-concept experiment on ring-laser-based decision-making. The two PD signals are sent to a digital oscilloscope, and the current is applied to either of BC_1 or BC_2 according to the results of the slot machine playing. In the experiment, the slot machines are numerically simulated in the embedded signal processing unit in the oscilloscope.

algorithm; the physical dynamic itself was *not* directly engineered, even though a variety of dynamical features are inherent in laser systems[28]. Furthermore, the previous studies have used a long fiber optic delay line to generate chaotic waveforms[26,27]; such use could lead to impractically large systems, inhibit stable operation, and may prevent practical deployments.

We here propose a compact ($<5\,mm^2$ area), on-chip photonic decision maker based on a ring laser structure. Unlike previous studies[26,27], the laser structure can generate fast, complex, but controllable dynamics at a *chip scale*, without a long delay line. The origin of the dynamics is a spontaneous switching phenomenon, i.e., noise-induced mode-hopping[29,30]; the phenomenon is used for exploring an optimal solution under uncertainty. We demonstrate that optimal decision-making is efficiently achieved by opto-electronically controlling the spontaneous-switching dynamics.

Principle of Ring-Laser-Based Decision-Making

Ring laser dynamics and device structure. The device structure used for decision-maker is shown in Fig. 1a. A ring laser is coupled to adjacent waveguides that are integrated on the same chip as a GaAs/AlGaAs single quantum well structure. The resonator of the ring laser supports clockwise (CW) and counter-clockwise (CCW) propagating waves, and can exhibit various operating regimes, such as bidirectional operation and bistability, depending on the pump current[31,32]. Spontaneous switching between the CW and CCW modes is an interesting dynamic that appears in the transition from the stable bidirectional regime to the bistable regime. Spontaneous switching has been regarded as an *obstacle* for deterministic optical switching applications[33–35]. Conversely, in this work, it is preferably utilized for decision making with feedback control of the CW and CCW modes, as discussed later.

The two waveguides with contact electrodes (denoted by PD_i and BC_i, $i = 1, 2$, as shown in Fig. 1a, are used for independent input/output control of the two modes in the ring laser: PD_1 and PD_2 are used as the photodetectors to monitor the intensities of the CW and CCW modes, whereas BC_1 and BC_2 with current injections are used for introducing an asymmetry and changing the dynamics of the CW and CCW modes, respectively[30]. (See *Methods* section for details.) We note that a similar optoelectronic control method has been used for deterministic optical switching[31] and random number generation[36]. However, unlike the previous studies, we use this method for changing statistical characteristics of spontaneous switching dynamics, as demonstrated later in detail.

Principle of decision-making. Here, we consider a two-armed bandit (TAB) problem, i.e., the issue is to select the machine with the higher reward probability among two machines, denoted by SM_1 and SM_2 (Fig. 1b). We examine a TAB problem, the simplest MAB problem, so that we can validate the principle of the first ring-laser-based decision making. Meanwhile, the scalability of photonic decision making has been studied in the literature[25,27], which would be applied to ring-laser-based device architectures. Our decision-making method is based on the tug-of-war (TOW) model, exhibiting highly efficient decision making compared to conventional algorithms[15,16]. Based on the model principle, we solve the TAB problem by repeating the following four steps:

(i) **Signal detection**: The intensity level of CW and CCW outputs, denoted respectively by I_{CW} and I_{CCW}, are detected by photodetector PD_1 and PD_2, respectively.

(ii) **Decision of the machine selection**: If I_{CW} is larger than I_{CCW}, the decision is to select SM_1. Otherwise, the decision is to choose SM_2.

(iii) **Playing the selected machine**.

(iv) **Learning and feedback**: If a reward is provided by playing SM_1 or if a reward is not provided by playing SM_2, the current (or voltage) applied to BC_1 is increased to facilitate the lasing in the CW mode. Consequently, the probability of selecting SM_1 slightly increases in the next decision making. On the other

hand, if a reward is provided by playing SM_2 or if a reward is not provided by playing SM_1, the current (or voltage) applied to BC_2 is increased so that the CCW lasing is facilitated, leading to a slight increase of the probability of choosing SM_2 in the next step.

Repeating steps (i)–(iv), we can finally choose the best slot machine.

As described in step (iv) above, an important point for the decision making is how to change currents J_1 and J_2 to activate controllers BC_1 and BC_2, respectively. In this study, we control J_1 and J_2 by the following rules in which a dimensionless, time-dependent control parameter $C(t)$ is introduced:

$$\begin{aligned} J_1 = KC(t), \quad J_2 = 0, \qquad & \text{if } C(t) \geq 0, \\ J_1 = 0, \qquad J_2 = -KC(t), \quad & \text{if } C(t) \leq 0, \end{aligned} \tag{1}$$

where K is a gain parameter. If $C \geq 0$ at the t-th play, the current $J_1 = KC$ (mA) is injected to controller BC_1 whereas $J_2 = -KC$ is injected to BC_2 if $C < 0$. The amount of $C(t)$ is updated in accordance with the results of slot machine playing as follows:

$$C(t + 1) = \alpha C(t) + \Delta C, \tag{2}$$

and

$$\Delta C = \begin{cases} +\Delta & \text{if } SM_1 \text{ wins} \\ -\Delta & \text{if } SM_2 \text{ wins} \\ -\Omega\Delta & \text{if } SM_1 \text{ fails} \\ +\Omega\Delta & \text{if } SM_2 \text{ fails}, \end{cases} \tag{3}$$

where $\alpha \in [0, 1]$ is the memory parameter (typically, ≈ 0.99–0.999)[37], and Δ is an incremental parameter ($\Delta = 1$ in this study). Ω in Eq. (3) is determined based on the estimated reward probability $\hat{P}_i$ for SM_i ($i = 1, 2$) from the history of the betting results. $\hat{P}_i$ is given by L_i/S_i, where S_i is the total number of times of playing SM_i and L_i is the number of wins in selecting SM_i. Ω is then given as,

$$\Omega = \frac{\hat{P}_1 + \hat{P}_2}{2 - (\hat{P}_1 + \hat{P}_2)}. \tag{4}$$

The details of the derivation of Eq. (4) are shown in[15].

Results

Optoelectronic control of spontaneous switching dynamics.
In our ring laser, a spontaneous switching phenomenon used for the above decision-making method appears when the pump current J_p exceeded ~1.3 times of the laser threshold current J_{th}. Figure 2a shows the examples of the switching dynamics, where the CW and CCW intensities stochastically change due to internal laser noise. For convenience, we hereafter refer to the state of $I_{CW(CCW)} > I_{CCW(CW)}$ as the CW (CCW) mode. A statistical analysis reveals that the mode switching is characterized by a characteristic time $\tau_c \approx 43$ ns; in a timescale longer than τ_c, the switching process is treated as a Poisson (random) process, and the duration time in the CW (CCW) mode obeys an exponential distribution (see Supplementary Fig. A1 for details). We refer to τ_c as the correlation time of the switching process. When current J_1 to BC_1 increases with $J_2 = 0$, the duration time in the CW mode increases [Fig. 2a(i,ii)]. In particular, we found that for $J_1 > 20$ mA, the duration time diverges, and a stable CW mode operation is achieved [Fig. 2a(iii)]. Otherwise, increasing J_2 can lead to an increase of the duration time in the CCW mode [Fig. 2a(iv,v)], and a stable CCW mode operation is achieved for $J_2 > 25$ mA [Fig. 2a(vi)].

On-chip decision making: proof-of-concept demonstration.
We conducted decision-making experiments based on the controllable dynamics in the ring laser by repeating the processes (i-iv) described in the previous section. In the experimental setup shown in Fig. 1b, the two machines SM_1 and SM_2 were emulated in a computer with the reward probabilities of $(P_1, P_2) = (0.7, 0.3)$. The gain K, step Δ, and memory parameter α were set to be 1, 1, and 0.99, respectively. A machine is selected and played once, and the reward dispensed from SM_1 and SM_2 is assumed to be both 1. The goal of the experiment is to confirm whether the ring-laser-based decision maker selects SM_1 (rather than SM_2) since SM_1 has a higher reward probability ($P_1 > P_2$). We assume the situation of zero prior knowledge, where the sum of the two hit probabilities is unknown, unlike ref.[26].

The experimental results on the decision-making process are displayed in Fig. 3. At first, I_{CW} and I_{CCW} are randomly switched when the number of plays $t < 100$ [Fig. 3a(i)], suggesting the exploration to choose the best machine. The accumulated knowledge is used for estimating the reward probabilities and setting the Ω-value, and then the C-value is appropriately updated [Fig. 3a(ii)]. The updated C-value affects the dynamics, and the dynamical state change from the switching mode to the CW mode. Consequently, the best machine (SM_1 in this case) is selected. We repeated the decision-making experiment $n_T = 200$ times and evaluated the correct decision rate (CDR), which is defined as the ratio of the number of selecting the slot machine with higher reward probability at the t-th play in n_T trials[24]. As shown in Fig. 3b, the CDR monotonically increases and approaches 1, suggesting the achievement of correct decision making.

We also conducted similar decision-making experiments with respects to different reward probabilities and parameters; we found that with appropriately tuned parameters (K and Δ), the decision-making performances
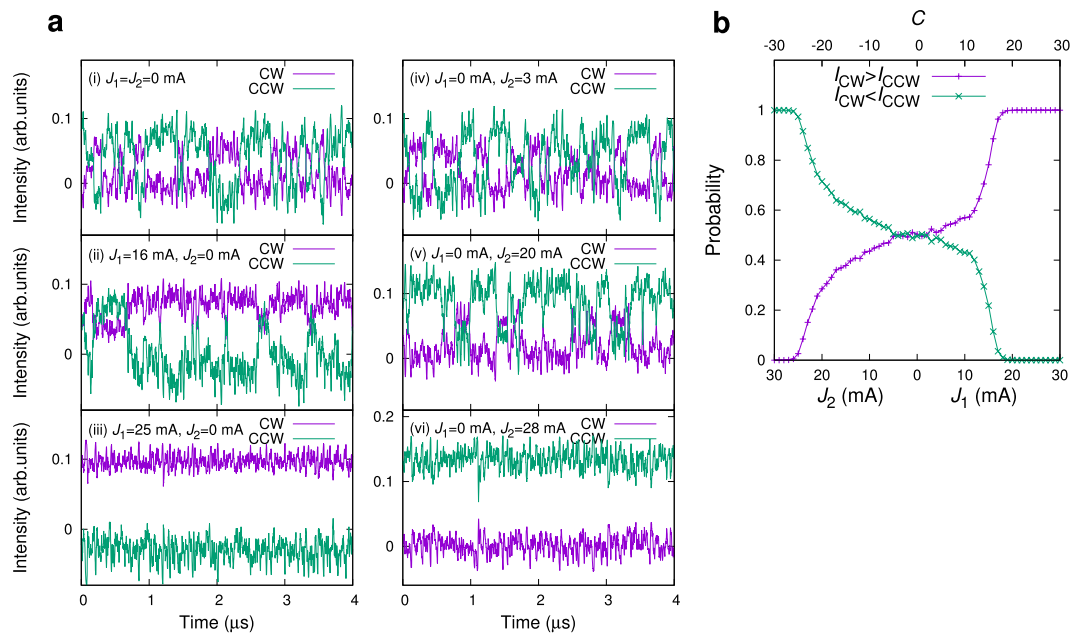
Figure 2. Spontaneous switching dynamics. (**a**) Temporal waveforms of the CW intensity I_{CW} and CCW intensity I_{CCW} for various values of currents J_1 and J_2. (**b**) Probability P_{CW} of $I_{CW} > I_{CCW}$ (purple color) as a function of currents J_1 and J_2 and the corresponding control parameter C with $K = 1$. For reference, the probability of $I_{CCW} > I_{CW}, 1 - P_{CW}$, is indicated by the green color.
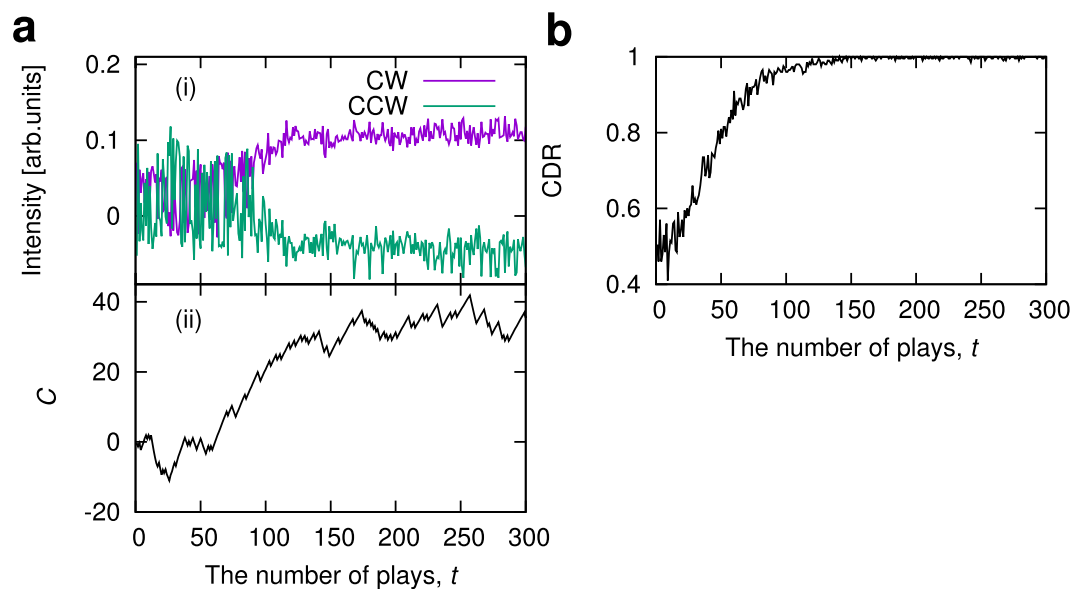

Figure 3. Experimental demonstration of ring-laser-based decision-making. (**a**) Evolution of (i) $I_{CW(CCW)}$ and (ii) control parameter C for a single trial. In this experiment, the reward probabilities of SM_1 and SM_2 are set as $(P_1, P_2) = (0.7, 0.3)$. $K = 1$, $\alpha = 0.99$, and $\Delta = 1$. (**b**) Evolution of CDR. The CDR was evaluated with $n_T = 200$ trials.

could be comparable to existing decision-making algorithms such as a modified softmax[16] and upper confidence bound 1-tuned (UCB1-tuned)[38,39]. As shown in Fig. 4, the CDR of the ring laser-based method can exceed those of the other methods in some cases.

Discussion

Decision-making strategy and its control. In our decision-making method, the strategy for making good decisions is characterized by the probability function of inducing CW mode configured by the control parameter $C(t)$, denoted by $P_{CW}(C)$. As observed in Fig. 2b, $P_{CW}(C)$ of the ring laser has a plateau region in the range of around $-21 \leq C \leq 12$, where $P_{CW}(C)$ moderately changes when C-value is changed. The plateau region
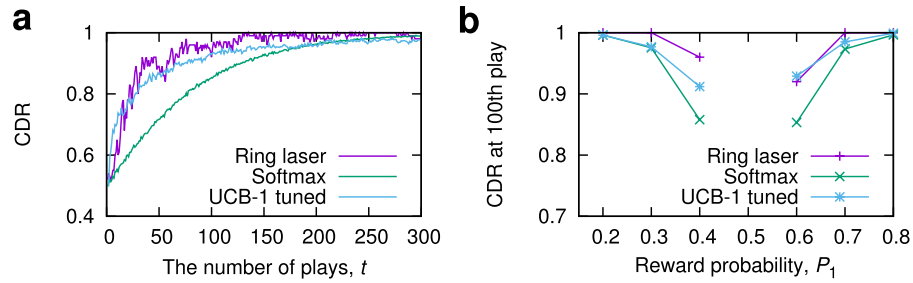
Figure 4. Performance comparison. (**a**) Comparison of CDRs with the modified Softmax and UCB1-tuned. The reward probabilities are set as $(P_1, P_2) = (0.6, 0.4)$. The CDR of the UCB1-tuned is better than the ring laser-based method for the first few ten plays, but the CDR of the ring laser-based method more quickly approaches closely to 1 even before the 100th play. In this experiment, $K = 4$, $\alpha = 0.99$, and $\Delta = 1$ were used. The parameters used in the modified softmax are similar to those in ref.[16]. The UCB1-tuned is a non-parameter algorithm[38,39]. (**b**) The CDRs at the 100th play. Here, P_2 was set to be $1 - P_1$.

plays a role in explorations to estimate the reward probability (and hence an appropriate Ω-value), and can lead to a correct decision after many slot plays, as demonstrated in Fig. 4; however, it may also lead to a slow convergence of CDR. A better alternative strategy (i.e., the design of $P_{CW}(C)$) satisfying both fast adaptation speed and decision accuracy can theoretically be estimated in the case when we can obtain prior knowledge on the sum of the reward probabilities, $P_1 + P_2$, such as when either of two events inevitably occurs with the probabilities P_1 and $P_2 = 1 - P_1$.

Let us here assume that the value of $P_1 + P_2$ is *a priori* known and Ω in Eq. (4) is a constant value. For simplicity, we consider $\alpha = 1$ and assume that the mode switching is random. Under these assumptions, we can treat the time evolution of C as a random walk. The random walk model gives an analytical expression of CDR and suggests that fast and correct decision is made when the probability distribution $P_{CW}(C)$ is close to 1 for $C > 0$ and 0 for $C < 0$, and steeply vary from 0 to 1 near $C = 0$. (See Sec. 2 of Supplementary Information).

In an actual experiment, such a $P_{CW}(C)$ is effectively realized by modifying the relationship between the control parameter C and $J_{1(2)}$ (Eq. 1) as follows:

$$
\begin{aligned}
J_1 &= 0, & J_2 &= 0, & \text{if } C(t) = 0, \\
J_1 &= KC(t) + K_1, & J_2 &= 0, & \text{if } C(t) > 0, \\
J_1 &= 0, & J_2 &= -KC(t) + K_2, & \text{if } C(t) < 0,
\end{aligned}
\tag{5}
$$

where K_1 and K_2 are chosen such that the plateau region of $P_{CW}(C)$ shown in Fig. 2b is reduced and the desirable $P_{CW}(C)$ results. Figure 5a shows $P_{CW}(C)$ with $(K_1, K_2) = (0, 0)$, (5, 9) and (13, 17), depicted by Types I, II, and III, respectively. As predicted by the random walk model, CDR in Type III most quickly increases and the convergence value is higher than the other types, regardless of the reward probabilities P_1 and P_2 (Fig. 5b,c). Thus, we conclude that the decision-making performance can be enhanced by changing the intrinsic characteristics ($P_{CW}(C)$) of the physical devices with an appropriate mode-control.

Decision-making rate. The rate of decision-making, i.e., the number of decision-making per unit time, in principle, depends on the sampling rate of the CW- and CCW-signal detections. Thus, fast decision making is possible by increasing the sampling rate; however, sampling too rapidly may degrade the accuracy of the decision making because nearly identical signal levels will be observed due to the limitation of the ring laser dynamics. It is important to know how rapidly decision making can be made without degrading the performance. In order to address this question and obtain an insight into the effect of the switching dynamics on the decision-making performance, we numerically examine decision-making processes by standard ring laser model equations[32]. See *Methods* section for details of the modeling.

Figure 6a shows the evolution of the CDR for various values of the sampling rate $1/\tau_{sam}$ when $(P_1, P_2) = (0.7, 0.3)$, where is the sampling time interval of the signal detections. The CDRs at the 30th-play are shown as a function of τ_{sam} in Fig. 6b. These numerical results clearly show that the decision-making performance (accuracy and adaptation) degrades when τ_{sam} is much shorter than the correlation time τ_c of the ring laser. Actually, the auto-correlation of the switching signals sampled at $\tau_{sam} \ll \tau_c$ exhibits a positive value [See Supplementary Fig. A1(d)]. In the decision-making, the positive correlation may result in repetition of the same choice even when the choice is wrong. In contrast, when $\tau_{sam} \gg \tau_c$, the correlation becomes close to zero, which enables an exploration without repeating wrong choices. Accordingly, the sampling time interval (i.e., inverse of the decision-making rate) can be shorter up to the correlation time without degrading the performance. The correlation time can be shorter in principle, allowing faster decision making by increasing the noise strength and activating mode-hopping phenomenon. In an actual experiment, this can be achieved by coupling the laser to an external amplified spontaneous emission noise source; the experimental verification will be an interesting future study.
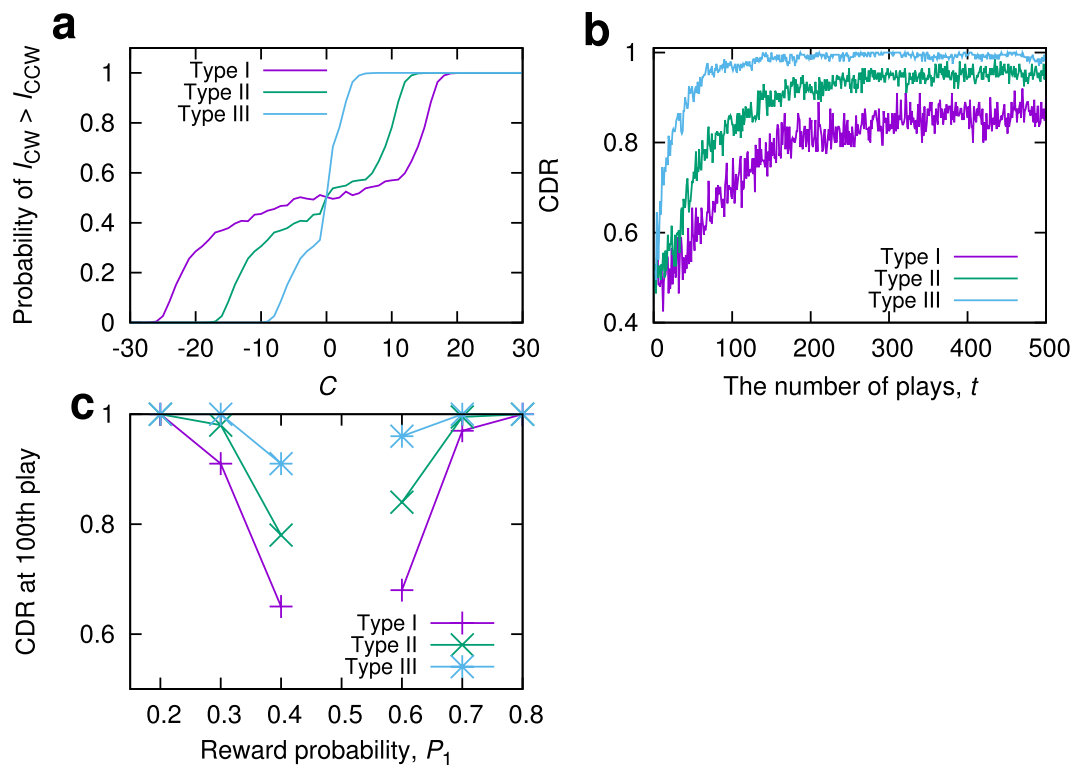
Figure 5. Decision-making strategy. (**a**) C-dependence of the occurrence probability of the state $I_{CW} > I_{CCW}$. Types I, II, and III represent $P_{CW}(C)$ obtained for $(K_1, K_2) = (0, 0)$, $(5, 9)$, and $(13, 17)$, respectively. (**b**) Time evolution of CDR for each type. In this experiment, the reward probabilities were set as $(P_1, P_2) = (0.6, 0.4)$, and prior knowledge of $P_1 + P_2 = 1$ was assumed. As predicted by the random walk model, the CDR for Type III is superior than the other types. $K = 1$, $\alpha = 0.99$, and $\Delta = 1$. (**c**) The CDRs at the 100th play were compared as a function of the given reward probability, where Type III outperforms other cases.
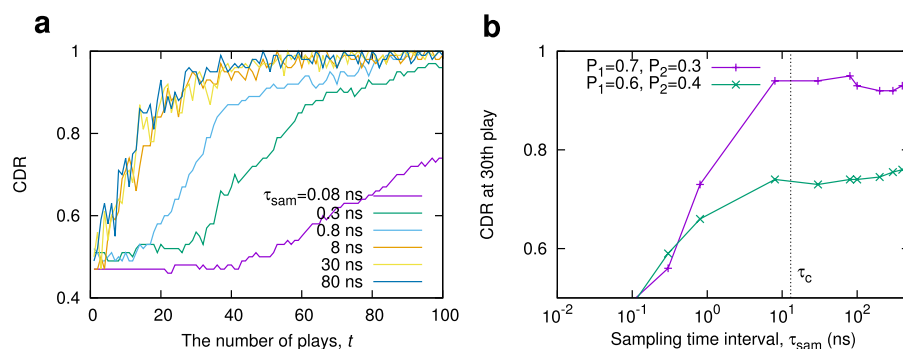

Figure 6. Ultimate operation rate analysis of ring-laser-based decision making. (**a**) Time evolution of CDRs with respects to different sampling time interval τ_{sam}. Too small τ_{sam} or too fast operation degrades CDRs. The reward probabilities were set as $(P_1, P_2) = (0.7, 0.3)$. The CDRs were evaluated from the results of $n_T = 100$ trials. (**b**) Comparison of CDR at the 30th play as a function of τ_{sam}. The reward probabilities were set as $(P_1, P_2) = (0.7, 0.3)$ and $(0.6, 0.4)$. In this simulation, the correlation time τ_c was ≈ 13 ns, which is indicated by dotted line. We can clearly observe that the CDRs are degraded in the regime where τ_{sam} is smaller than τ_c.

Summary

In this study, we proposed and experimentally demonstrated on-chip photonic decision making by an integrated ring laser. Ring lasers generate statistical characteristics regarding the CW and CCW lasing, which are optoelectronically controllable; we directly utilize such inherent spontaneous dynamics of ring lasers for decision-making functionalities. Correct decision making was successfully demonstrated with appropriate optoelectronic control of the dynamics, and it is found that the performance can be enhanced by changing the decision-making strategy with the statistical characteristics ($P_{CW}(C)$). These results would open novel research perspectives of controlling complex dynamics based on environmental changes.

One interesting and important future study is to increase the decision-making rate by using faster and more complex switching dynamics. In addition to the above-mentioned method on increasing the noise strength, the use of the delayed feedback structure will be useful. Interestingly, semiconductor ring lasers can exhibit chaotic switching in the GHz regimes by delayed feedback even with *a short time delay*[40,41]. Combination of noise-induced switching with delayed feedback instability indicates a promising research direction.

As for the ring laser structure, we emphasize that in addition to miniaturization, it would be beneficial for all optical realization of decision-making devices because all photonic components required for decision making can be monolithically integrated on a chip. Instead of the optoelectronic control methods employed in the present study, it would be interesting to use an optical injection method because ring lasers subjected to optical injection enable low power and ultrafast switching at picosecond time scales[33–35].

Another interesting and important future study is to tackle larger-scale MAB problems. MAB problems can be solved based on a hierarchical TOW principle[25,27]. The decision-making based on the hierarchical principle can be achieved by using a number of independent two-choice decision-makers (for two-armed bandit problems) or using a time-division multiplexing scheme[27]. Compact ring lasers could offer a good experimental platform for implementing the hierarchical principle and addressing the MAB problems.

We believe that the combination of photonic integration technologies and competitive fluctuating dynamics, as demonstrated by the proposed ring laser, will shed light on a way toward novel photonic intelligent computing paradigms.

Methods

Device structure and operating regime.
The ring laser device used in this study was fabricated in a graded-index separate-confinement-heterostructure (GRIN-SCH) single-quantum well GaAs/Al$_x$Ga$_{1-x}$As structure, the emission wavelength of which is designed to be 850 nm. The fabricated laser device was thermally controlled by a heat-sink with an accuracy of 0.01 °C. The ring radius is 1 mm, and the waveguide width is 2 μm. In an actual device, multiple waveguides with independent electrical contacts are coupled to the ring with an angle to the cleaved facet. We used the waveguides with contacts, PD$_i$ and BC$_i$ (i = 1, 2), as shown in Fig. 1a. The CW and CCW intensity signals are detected with PD$_1$ and PD$_2$ in the waveguide, respectively, and sent to a digital oscilloscope (Tektronix TDS7154B, bandwidth 1.5 GHz, 20 GSample/s) via the bonding wires attached to PD$_1$ and PD$_2$. Bias contacts BC$_1$ and BC$_2$ were used for the mode-control inside the ring laser. Sending current to BC$_1$ and BC$_2$ reduces the absorption loss of the waveguide. Thus, the light coupled from the CCW(CW) mode in the ring to the waveguide is back-reflected at the BC$_{1(2)}$-side end of the waveguide and re-coupled to the ring in the CW(CCW) direction. In addition, BC$_{1(2)}$ can enhance the spontaneous emission noise coupled to the CW(CCW) mode, and consequently, facilitates the laser operation in the CW(CCW) mode[31,36].

When $J_1 = J_2 = 0$ mA, the threshold current J_{th} of the ring laser used in the experiment was approximately estimated to be 210 mA at 25 °C. The large threshold may partly be attributed to non-optimal etching depth of the ring waveguide[32]. For $J/J_{th} < 1.3$, the ring laser operated on a bidirectional state of the CW and CCW modes. For larger J-value, a transition to spontaneous switching regime occurred.

Intensity adjustment.
In the experiment, the PD couplings to the CW and CCW modes are not essentially equal to each other due to an imperfect device fabrication. In order to reduce the effect of the asymmetry of the PD-couplings and appropriately evaluate the decision-making performance, the CW and CCW intensities, I_{CW} and I_{CCW}, were adjusted by adding constant biases so that the occurrence probability is calibrated being around 0.5 when $J_1 = J_2 = 0$ mA. This way would realize easy tuning of both intensities, while we should also note that there is another simpler way, which is to measure either of I_{CW} or I_{CCW} only and adjust the switching probability to 0.5 by bias currents J_1 and J_2 without the intensity biases.

Decision-making experiment.
First, the BC$_1$ and BC$_2$ were connected to a standard current source. Discrete-valued electrical currents were applied to BC$_1$ or BC$_2$. Then, the CW and CCW optical intensity signals for different values of J_1 and J_2 were recorded by a digital oscilloscope. In the decision-making experiment, the slot machines were numerically simulated in the embedded signal processing unit in the oscilloscope using pseudorandom numbers. The decision is immediately made based on the sampling. The controllers BC$_1$ and BC$_2$ were also connected to a two-channel function generator (Tektronix AFG3152C), which reconfigures the oscillation dynamics of the ring laser in an on-line or real-time manner.

Rate-equation model for semiconductor ring laser.
The numerical simulation was conducted by using a set of dimensionless semiclassical equations for the two slowly varying complex amplitudes of CW and CCW waves, E_1 and E_2[32].

$$\frac{dE_{1,2}}{dt} = (1 + i\widetilde{\alpha})[\xi_{1,2}N - 1]E_{1,2} - k_{1,2}E_{2,1} + \eta_{1,2}(t),$$
(6)

$$\xi_{1,2} = 1 - s|E_1|^2 - c|E_2|^2,$$
(7)

where $\widetilde{\alpha}$ accounts for phase-amplitude coupling, s and c are the self- and cross-saturation coefficients, and $k_{1,2}$ represents the complex backscattering coefficients. We model internal optical noises as complex Gaussian noise satisfying $\langle\eta_i\rangle = 0$ and $\langle\eta_i(t)\eta_j^*(t')\rangle = 2D\delta_{ij}\delta(t - t')$ (i = 1, 2). $\langle\,\cdot\,\rangle$ represents the ensemble average, and D represents the noise strength. Carrier density N obeys the following equation:

$$\frac{dN}{dt} = 2\gamma \left[\mu - N \left(1 - \xi_1 |E_1|^2 - \xi_2 |E_2|^2 \right) \right], \tag{8}$$

where μ is the dimensionless pumping power ($\mu = 1$ at the laser threshold). In the above equations, t is dimensionless time rescaled by photon lifetime τ_p. γ is the ration of τ_p to carrier lifetime τ_s.

In Eq. (6), the asymmetric coupling caused by activating BC_1 and BC_2 is simply modeled as an asymmetric backreflection effect such that $k_1 = \beta_1 k_b$ and $k_2 = \beta_2 k_b$, where k_b denotes the backreflection coefficient when $C = 0$, and $\beta_{1,2}$ denotes a dimensionless asymmetry parameter, depending on C as follows:

$$\begin{aligned} \beta_1 = 1 + kC(t), \quad \beta_2 = 1, &\qquad \text{if } C(t) \geq 0, \\ \beta_1 = 1, \qquad\qquad \beta_2 = 1 - kC(t), &\quad \text{if } C(t) \leq 0, \end{aligned} \tag{9}$$

where $C(t)$ is updated by Eq. (2). This is the simplest model of the asymmetric backscattering, although a real asymmetry may be introduced in a more complex way in the actual experiment. We confirmed that regardless of the details of the asymmetry model, the control of spontaneous switching can be achieved. The detailed investigation using more realistic model will be a future work.

In this study, we set some of the parameters as follows: $\tilde{\alpha} = 3.5$, $2s = c = 0.006$, $k_b = 0.004 + i0.001$, $k = 0.025$, $D = 5 \times 10^{-5}$, $\mu = 2.0$, $\tau_p = 10$ ps, $\gamma = 0.01$. With these parameter values, we obtained stochastic switching dynamics with the correlation time $\tau_c \approx 13$ ns when $C = 0$. In the decision-making simulation, we assume that the slot machines provide a reward without any time delay and use Eqs (2–4) and (6–9).

References

1. Vetter, J. S., DeBenedictis, E. P. & Conte, T. M. Architectures for the Post-Moore Era. *IEEE Micro.* **37**(4), 6–8 (2017).
2. Nahmias, M. A., Shastri, B. J., Tait, A. N., Ferreira de Lima, T. & Prucnal, P. R. Neuromorphic Photonics. *Optics & Photonics News* **29**(1), 34–41 (2018).
3. Peper, F. The End of Moore's Law: Opportunities for Natural Computing? *New Gener. Comput.* **35**, 253 (2017).
4. Merolla, P. A. *et al.* A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**(6197), 668–673 (2014).
5. Esser, S. K. *et al.* Convolutional networks for fast, energy efficient neuromorphic computing. *Proc. Natl Acad. Sci. USA* **113**, 11441–11446 (2016).
6. Tait, A. N. *et al.* Neuromorphic photonic networks using silicon photonic weight banks. *Sci. Rep.* **7**, 7430 (2017).
7. Shen, Y. *et al.* Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).
8. Lin, X. *et al.* All-optical machine learning using diffractive deep neural networks. *Science* **361**(6406), 1004–1008 (2018).
9. Appeltant, L. *et al.* Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, 468 (2011).
10. Van der Sande, G. *et al.* Advances in photonic reservoir computing. *Nanophotonics* **6**, 561 (2017).
11. Bandyopadhyay, A., Pati, R., Sahu, S., Peper, F. & Fujita, D. Massively parallel computing on an organic molecular layer. *Nat. Phys.* **6**(5), 369–375 (2010).
12. Berkley, A. J. *et al.* Quantum annealing with manufactured spins. *Nature* **473**, 194–198 (2011).
13. Inagaki, T. *et al.* A coherent Ising machine for 2000-node optimization problems. *Science* **354**(6312), 603–606 (2016).
14. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An introduction.* (The MIT Press, Massachusetts, 1998).
15. Kim, S.-J. M., Aono & Nameda, E. Efficient decision-making by volume-conserving physical object. *New. J. Phys.* **17**, 083023 (2015).
16. Kim., S.-J., Aono, M. & Hara, M. Tug-of-war model for the two-bandit problem: Nonlocally-correlated parallel exploration via resource conservation. *Biosystems* **101**(1), 29–36 (2010).
17. Hu, W., Wu, K., Shum, P. P., Zheludev, N. I. & Soci, C. All-optical implementation of the ant colony optimization algorithm. *Sci. Rep.* **6**, 26283 (2016).
18. Alonzo, M. *et al.* All-Optical Reinforcement Learning In Solitonic X-Junctions. *Sci. Rep.* **8**, 5716 (2018).
19. Jouini, W., Ernst, D., Moy, C. & Palicot, J. Multi-armed bandit based policies for cognitive radio's decision making issues. *2009 3rd International Conference on Signals, Circuits and Systems (SCS), Medenine*, pp. 1–6, https://doi.org/10.1109/ICSCS.2009.5412697 (2009).
20. Lai, L., El Gamal, H., Jiang, H. & Poor, H. V. Cognitive Medium Access: Exploration, Exploitation, and Competition. *IEEE Trans. on Mob. Comput.* **10**(2), 239–253 (2011).
21. Agarwal, D., Chen, B. & Elango, P. Explore/Exploit Schemes for Web Content Optimization. *2009 Ninth IEEE International Conference on Data Mining*, pp. 1–10, https://doi.org/10.1109/ICDM.2009.52 (2009).
22. Kocsis, L. & Szepesvári, C. Bandit Based Monte-Carlo Planning. In: Furnkranz, J., Scheffer, T. & Spiliopoulou, M. (eds) Machine Learning: ECML 2006. ECML 2006. Lecture Notes in Computer Science, vol 4212. Springer, Berlin, Heidelberg, https://doi.org/10.1007/11871842_29.
23. Mahajan, A. & Teneketzis, D. Multi-Armed Bandit Problems. In: Hero, A. O., Castanon, D. A., Cochran, D. & Kastella, K. (eds) Foundations and Applications of Sensor Management. Springer, Boston, MA.
24. Naruse, M. *et al.* Single-photon decision maker. *Sci. Rep.* **5**, 13253 (2015).
25. Naruse, M. *et al.* Single Photon in Hierarchical Architecture for Physical Decision Making: Photon Intelligence. *ACS Photonics* **3**(12), 2505–2514 (2016).
26. Naruse, M., Terashima, Y., Uchida, A. & Kim, S.-J. Ultrafast photonic reinforcement learning based on laser chaos. *Sci. Rep.* **7**, 8772 (2017).
27. Naruse, M. *et al.* Scalable photonic reinforcement learning by time-division multiplexing of laser chaos. *Sci. Rep.* **8**, 10890 (2018).
28. Uchida, A. *Optical Communication with Chaotic Lasers: Applications of Nonlinear Dynamics and Synchronization* (Wiley-VCH, 2012).
29. Beri, S. *et al.* Topological Insight into the Non-Arrhenius Mode Hopping of Semiconductor Ring Lasers. *Phys. Rev. Lett.* **101**, 093903 (2008).
30. Gelens, L. *et al.* Exploring Multistability in Semiconductor Ring Lasers: Theory and Experiment. *Phys. Rev. Lett.* **102**, 193904 (2009).
31. Sorel, M., Laybourn, P. J. R., Giuliani, G. & Donati, S. Unidirectional bistability in semiconductor waveguide ring lasers. *Appl. Phys. Lett.* **80**, 3051–3053 (2002).
32. Sorel, M. *et al.* Operating regimes of GaAs- AlGaAs semiconductor ring lasers. *IEEE J. Quantum Electron.* **39**, 1187–1195 (2003).
33. Pérez, T., Scirè, A., Van der Sande, G., Colet, P. & Mirasso, C. R. Bistability and all-optical switching in semiconductor ring lasers. *Opt. Express* **15**, 12941–12948 (2007).
34. Hill, M. T. *et al.* A fast low-power optical memory based on coupled micro-ring lasers. *Nature* **432**, 206–209 (2004).
35. Liu, L. *et al.* An ultra-small, low-power, all-optical flip-flop memory on a silicon chip. *Nat. Photonics* **4**, 182–187 (2010).

36. Sunada, S. *et al*. Random optical pulse generation with bistable semiconductor ring lasers. *Opt. Exp.* **19**, 7439 (2011).
37. Mihana, T., Terashima, Y., Naruse, M., Kim, S.-J. & Uchida, A. Memory Effect on Adaptive Decision Making with a Chaotic Semiconductor Laser. *Complexity* **2018**, 4318127 (2018).
38. Auer, P., Cesa-Bianchi, N. & Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* **47**, 235 (2002).
39. Kuleshov, V. & Precup, D. Algorithms for multi-armed bandit problems. *Journal of Machine Learning Research.* **1**, 1–48, arXiv:1402.6028 (2014).
40. Ermakov, I., Van der Sande, G. & Danckaert, J. Semiconductor ring laser subject to delayed optical feedback: Bifurcations and stability. *Commun. Nonlinear Sci. Numer. Simul.* **17**, 4767 (2012).
41. Sunada, S. *et al*. Chaos laser chips with delayed optical feedback using a passive ring waveguide. *Opt Express* **19**(7), 5713 (2011).

Acknowledgements

Author Contributions

S.S., A.U. and M.N. directed the project and designed the decision-making principle. R.H., S.K., T.M., Y.M., K.K., A.U. and S.S. designed and performed ring laser experiments. R.H., S.K., T.N. and S.S. analyzed the data. S.S. conducted numerical evaluation of the performance. R.H., M.N., A.U. and S.S. wrote the paper, and all of the authors contributed to the preparation of the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-45754-3.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.