

# PHOTONICS Research

## Generation of multi-channel chaotic signals with time delay signature concealment and ultrafast photonic decision making based on a globally-coupled semiconductor laser network

YANAN HAN,<sup>1</sup> SHUIYING XIANG,<sup>1,2,\*</sup> YANG WANG,<sup>1</sup> YUANTING MA,<sup>1</sup> BO WANG,<sup>1</sup>  
AIJUN WEN,<sup>1</sup> AND YUE HAO<sup>2</sup>

<sup>1</sup>State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China

<sup>2</sup>State Key Discipline Laboratory of Wide Band Gap Semiconductor Technology, School of Microelectronics, Xidian University, Xi'an 710071, China

\*Corresponding author: jxxy@126.com

Received 21 July 2020; revised 3 September 2020; accepted 3 September 2020; posted 8 September 2020 (Doc. ID 403319); published 30 October 2020

We propose and demonstrate experimentally and numerically a network of three globally coupled semiconductor lasers (SLs) that generate triple-channel chaotic signals with time delayed signature (TDS) concealment. The effects of the coupling strength and bias current on the concealment of the TDS are investigated. The generated chaotic signals are further applied to reinforcement learning, and a parallel scheme is proposed to solve the multi-armed bandit (MAB) problem. The influences of mutual correlation between signals from different channels, the sampling interval of signals, and the TDS concealment on the performance of decision making are analyzed. Comparisons between the proposed scheme and two existing schemes show that, with a simplified algorithm, the proposed scheme can perform as well as the previous schemes or even better. Moreover, we also consider the robustness of decision making performance against a dynamically changing environment and verify the scalability for MAB problems with different sizes. This proposed globally coupled SL network for a multi-channel chaotic source is simple in structure and easy to implement. The attempt to solve the MAB problem in parallel can provide potential values in the realm of the application of ultrafast photonics intelligence. © 2020 Chinese Laser Press

<https://doi.org/10.1364/PRJ.403319>

### 1. INTRODUCTION

Since its advent, the laser has been applied in many fields due to the advantages of rapid response and rich dynamics [1]. For example, it is used in high-speed random bit generators [2,3], optical secure communication, and secret key distribution that requires synchronized chaotic signals [4–7]. Recently, photonic technologies have also been developed as efficient ways of solving some conventional problems in the area of artificial intelligence (AI) calculation such as reservoir computing [8,9], reinforcement learning [10–12], and brain-inspired photonic neuromorphic computing [13–16].

The security of information transmission has always been a focus of attention. In optical communication systems, chaotic signals can be generated by means of delayed optical feedback, optical injection, and other external disturbances [17–22]. However, a time delay signature (TDS) can be introduced (typically by external cavity feedback) and cause internal

periodicity of chaotic oscillations [23,24]. This feature can be analyzed by methods like permutation entropy (PE), delayed mutual information, autocorrelation functions (ACF), etc., and utilized for reconstruction of chaotic systems [25–29], which seriously threaten the security of communication. Many methods have been reported to complicate and suppress the TDS. For example, Lee *et al.* first proposed to complicate the TDS in a semiconductor laser (SL) subject to double optical feedback [30], and the result was experimentally demonstrated later by Wu *et al.* [31]. We also numerically achieved the suppression of TDS in a mutually coupled ring network with heterogeneous time delays [32]. Very recently, Jiang *et al.* proposed a new scheme for the generation of wideband laser chaos with excellent TDS suppression by using parallel-coupling ring resonators as reflector [33].

As one of the fundamental problems in reinforcement learning, adequate decision making in a dynamically changing environment is also required in frequency and channel assignments

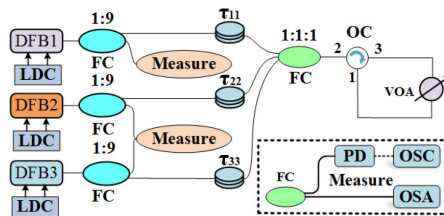
in communication networks [12,34,35]. The multiarmed bandit (MAB) problem is one of the most important issues in decision making. One remarkable method to solve the MAB problem was proposed by Kim *et al.*, called the tug-of-war (TOW) method, which was inspired by the unicellular amoeba of true slime mold [36,37]. In recent years, several works on ultrafast decision making have been reported based on the TOW method [38–41]. In our previous work, we have already proposed to solve a four-armed bandit problem in parallel by sampling dual-channel TDS-concealed chaotic signals simultaneously and found it works more efficiently [42]. However, the threshold value (TV) for each channel is set and adjusted dependently; therefore, the scheme is not completely parallel.

In this paper, we propose a scheme for the generation of laser chaos with TDS concealment and demonstrate its application in reinforcement learning. Our contribution includes three aspects. First, the new proposed scheme for the generation of complex laser chaos is simple in structure and easy to implement. Second, we propose a scheme to solve the MAB problem in parallel via using the generated laser chaos and verify its scalability and adaptability. Third, in order to solve the MAB problem in parallel, we propose a modified strategy and demonstrate its effectiveness.

## 2. SYSTEM MODEL AND RESULTS

### A. Experimental Setup

The experimental setup of three globally coupled SLs is presented in Fig. 1. Here, three distributed feedback (DFB) lasers are driven by laser diode controllers (LDCs) to control the current and temperature of the SLs. The wavelengths of free-running DFB lasers are precisely matched by adjusting the current and temperature. In this setup, the optical output from each DFB laser is divided into two parts through a 10:90 fiber coupler (FC). The smaller part is sent to the measure module, where the optical signal can be detected by a high-speed photodiode (PD, HP11982A, 15 GHz) and analyzed by a real-time oscilloscope (OSC) with 8-bit analog-to-digital converter (Keysight DSOV334A, 33 GHz, 80 GS/s), or directly sent to an optical spectrum analyzer (OSA, AndoAQ6317). The rest of the parts are combined into one with an FC through fiber jumpers with different lengths, then pass through a variable optical attenuator (VOA), and feed back to all the three DFB lasers via an optical circulator (OC). Thus, the coupling strength and feedback strength can be adjusted simultaneously by the VOA.



**Fig. 1.** Experimental setup of three globally coupled SLs. DFB1, DFB2, DFB3, three distributed feedback lasers; LDC, laser diode controller; FC, fiber coupler; OC, optical circulator; VOA, variable optical attenuator;  $\tau_{11}$ ,  $\tau_{22}$ ,  $\tau_{33}$ , feedback delay time; PD, photodiode; OSC, oscilloscope; OSA, optical spectrum analyzer.

For simplicity, they are referred to as coupling strength in the following.

### B. Experimental Results

The ACF is one of the effective methods for identifying the TDS of the measured chaotic signals [29,32], as defined in Eq. (1),

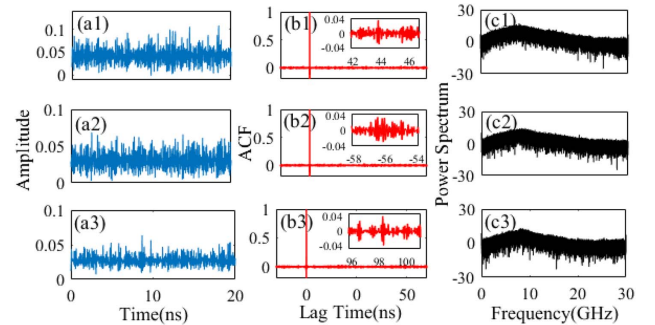
$$C_m(\Delta t) = \frac{\langle [I_m(t + \Delta t) - \langle I_m(t + \Delta t) \rangle][I_m(t) - \langle I_m(t) \rangle] \rangle}{\sqrt{\langle [I_m(t + \Delta t) - \langle I_m(t + \Delta t) \rangle]^2 \rangle \langle [I_m(t) - \langle I_m(t) \rangle]^2 \rangle}}, \quad (1)$$

where  $C_m(\Delta t)$  is the ACF value of the chaotic time series  $I_m(t)$  at time lag  $\Delta t$ , measured from DFB $m$  ( $m = 1, 2, 3$ ).  $\langle \rangle$  means time average. The TDS concealment can be reflected by the most pronounced residual peak, denoted as  $\rho_m$  ( $m = 1, 2, 3$ ), in the ACF. Better TDS concealment is indicated by a lower value of  $\rho_m$  [32].

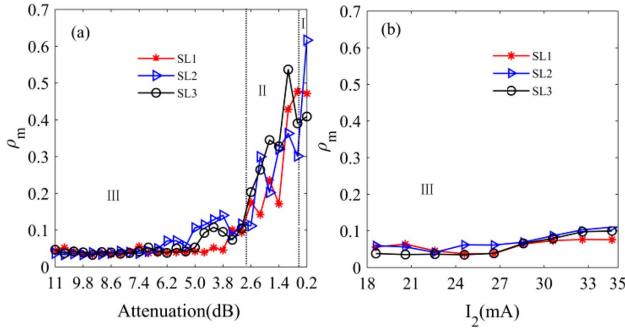
To identify the TDS of DFB1, we turn off DFB2 and DFB3, and calculate the ACF of the output intensity; the round-trip feedback time delay of DFB1 is indicated by the location of  $\rho_m$  in the ACF. By this method, the feedback time delays for DFB1, DFB2, and DFB3 are determined to be 97.4, 97.53, and 97.38 ns, respectively. Note that the time delay values are close, introduced by slightly different propagation paths, and need not be precisely adjusted by the variable optical delay line (VODL). The wavelengths of free-running DFB lasers are precisely set as 1552.250, 1552.265, and 1552.255 nm, respectively, by carefully adjusting the current and temperature.

Figure 2 shows the measured chaotic time series from the three DFB lasers, the calculated ACF as a function of  $\Delta t$ , as well as the power spectrum. The chaotic dynamics of the three SLs can be revealed by the time series shown in Figs. 2(a1)–2(a3) and the power spectrum in Figs. 2(c1)–2(c3). As can be seen in Figs. 2(b1)–2(b3), no pronounced peaks can be found in the ACFs except for that at time lag 0, which means the TDS is greatly concealed in all three channels.

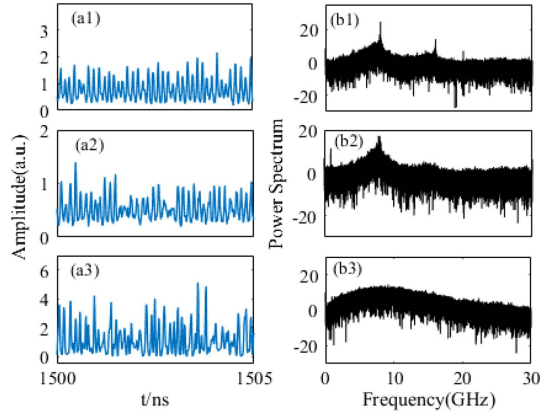
Then, in order to illustrate the effect of coupling strength on TDS concealment, the  $\rho_m$  as a function of attenuation is presented in Fig. 3(a). Region I (III) indicates that all three DFB lasers are in a quasi-periodic state (chaotic state). Region II represents the transition region where the states of the three DFB



**Fig. 2.** (a1)–(a3) The chaotic time series from the three DFB lasers; (b1)–(b3) the ACFs; (c1)–(c3) the power spectra. The attenuation is 9 dB,  $I_1, I_2, I_3 = 28.34, 24.5, 26.6$  mA,  $T_1, T_2, T_3 = 27.75, 15.5, 18^\circ\text{C}$ .



**Fig. 3.** (a)  $\rho_m$  as a function of attenuation; (b)  $\rho_m$  as a function of  $I_2$ .



**Fig. 4.** (a1)–(a3) Time series of signals at states I, II, and III, respectively; (b1)–(b3) the corresponding power spectrum.

lasers can be quasi-periodic, weakly chaotic, and chaotic, but not identical. Examples of the time series and the power spectrum of signals in each state are shown in Fig. 4. It can be seen that  $\rho_m$  is less than 0.1 when the attenuation is larger than 5.0 dB and increases with the decrease of attenuation, indicating that better TDS concealment can be achieved when the attenuation is large, namely, when the coupling strength is relatively small. The influence of bias currents on the TDS concealment is further investigated, as shown in Fig. 3(b). Here, the bias currents of the three DFBs are adjusted at the same time, and we simply present  $\rho_m$  as a function of  $I_2$  (which varies from 18.6 to 34.6 mA). It can be seen that  $\rho_m$  is less than 0.1 when  $I < 30.6$  mA, indicating that low TDS can be obtained in that region. However, when  $I > 30.6$  mA, the  $\rho_m$  values are larger than 0.1 and get larger with the increase of  $I_2$ , indicating reduced concealment of TDS for all three DFB lasers.

### C. Numerical Results

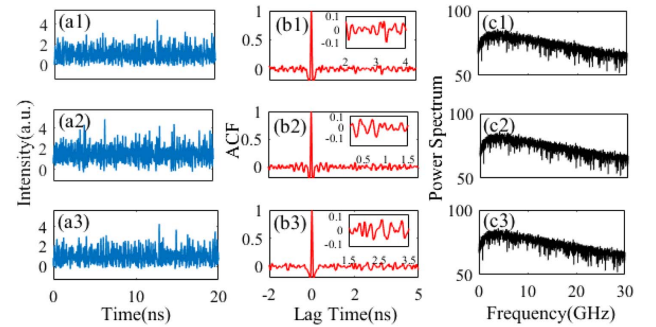
In addition, we also numerically verified the concealment of TDS in the proposed scheme. To model the dynamics of the three DFB lasers, the well-known Lang–Kobayashi equations are adopted, which describe the slowly varying complex electric-field  $E_m(t)$  and the carrier density  $N_m(t)$  in the active region [31,32]. The rate equations of our scheme can be written as

$$\begin{aligned} \frac{dE_m(t)}{dt} &= \frac{1 + i\alpha}{2} \left[ G_m(t) - \frac{1}{\tau_p} \right] E_m(t) \\ &\quad + k_{rn} \sum_{n=1}^3 E_n(t - \tau_{nm}) e^{-i(\omega_n \tau_i - \Delta\omega_{nm} t)} + \sqrt{2DN_m} \xi_m(t), \\ \frac{dN_m(t)}{dt} &= \frac{I_m}{q} - \frac{N_m(t)}{\tau_e} - G_m(t) |E_m(t)|^2, \\ G_m(t) &= \frac{g[N_m(t) - N_0]}{1 + s|E_m(t)|^2}, \end{aligned} \quad (2)$$

where  $m(n) = 1, 2, 3$  denotes the three SLs.  $\alpha = 5$  is the line-width enhancement factor [32,43],  $g = 1.5 \times 10^{-8} \text{ ps}^{-1}$  is the differential gain coefficient,  $s = 5 \times 10^{-7}$  is the nonlinear gain saturation coefficient, and  $\tau_p = 2 \text{ ps}$  and  $\tau_e = 2 \text{ ns}$  stand for the photon lifetime and the carrier lifetime, respectively.  $N_0 = 1.5 \times 10^8$  is the transparency carrier number. The variable  $I_m$  represents the bias current and  $k_{rn}$  describes the coupling strength. The coupling time delay  $\tau_{nm}$  ( $n \neq m$ ) from SL $n$  to SL $m$  can be calculated from the feedback time delay  $\tau_{nm}$  ( $n = m$ ) by  $\tau_{nm} = \tau_{mn} = (\tau_{nn} + \tau_{mm})/2$ .

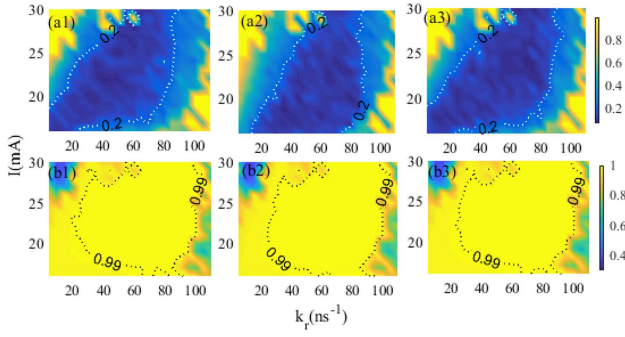
In Fig. 5, we present the time series, the ACF, and the power spectrum of the numerical results as in Fig. 2. The results show that the TDS can be concealed in such a scheme if the parameters are properly selected. Note that the mismatch of parameters is important to improve the concealment of the TDS. When the currents are the same for the three SLs, the region in which the TDS is concealed is quite narrow. To find a proper bias current, we can fix the currents of two SLs and change the other. In this way, we find that a current mismatch of 0.5–3.5 mA allows better TDS concealment in all three SLs. We choose a mismatch of 2.5 mA.

For a further exploration of the parameters' scope in which the TDS can be better suppressed, we show in Figs. 6(a1)–6(a3) the two-dimensional map of  $\rho_m$  for the three SLs as functions of the coupling strength and bias current of SL2 (for simplicity). The parameter region for  $\rho_m < 0.2$  is considered to have better TDS concealment and is marked by a white dotted line [32]. It can be seen that the evolution patterns for three  $\rho_m$  are similar, and the parameters for low TDS are mainly in the diagonal region, meaning that concealment is affected by both the current and the strength. The PE is also calculated as an indicator of the dynamical state of SLs [44] and is presented



**Fig. 5.** (a1)–(a3) The chaotic time series from the three SLs; (b1)–(b3) the ACFs; (c1)–(c3) the power spectra. The parameters are:  $I_m = 20, 22.5, 20 \text{ mA}$ ;  $k_{rm} = 11.7, 16.7, 11.7 \text{ ns}^{-1}$ ;  $\tau_{nm} = 2, 2.02, 2.04 \text{ ns}$ ;  $m = 1, 2, 3$ .

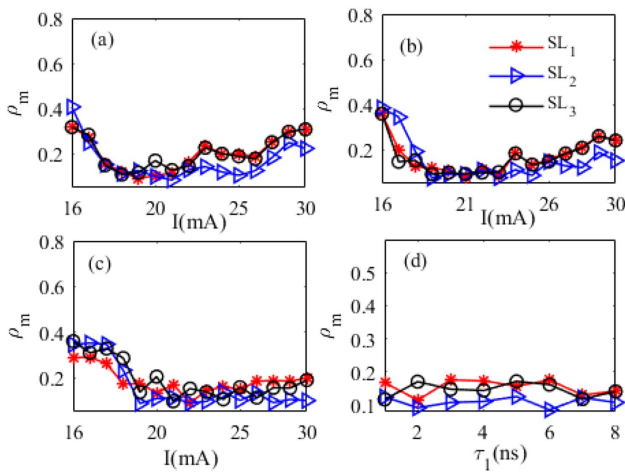




**Fig. 6.** (a1)–(a3) The two-dimensional map of  $\rho_m$  as functions of the coupling strength  $k_{r2}$  and bias current  $I_2$  of DFB1, DFB2, and DFB3, respectively; (b1)–(b3) the PE of DFB1, DFB2, and DFB3, respectively.  $I_1 = I_3, I_2 = I_1 + 2.5$  mA;  $k_{r1} = k_{r3}, k_{r2} = k_{r1} + 5$  ns<sup>-1</sup>;  $\tau_{mm} = 2, 2.02, 2.04$  ns.

in Figs. 6(b1)–6(b3). The dynamics of the SL is in chaotic oscillation when the PE value is larger than 0.99, marked by a black dotted line. As PE decreases, the dynamics goes through chaos to weak chaos and finally enters quasi-periodic oscillation.

Moreover, time delay is also an important factor that affects the dynamics of a system, and different time delays may cause different sensitivities to parameter mismatches. Hence, it is necessary to consider different coupling delays in the investigation of TDS concealment. Figures 7(a)–7(c) depict the  $\rho_m$  as a function of  $I$  with three different cases of time delay. We can see that in all three cases, the TDS can be concealed with properly selected parameters. Typically, we find that for a larger time delay, stronger coupling strength is required to achieve better TDS concealment. In Fig. 7(d), we further show the  $\rho_m$  as a function of  $\tau_{11}$  for all three SLs. It can be seen that for fixed current and coupling strength, the values of  $\rho_m$  remain



**Fig. 7.** TDS concealment with different time delays.  $I_1 = I_3$ . (a)  $I_2 = I_1 - 1$  mA,  $k_{rm} = 12, 10, 12$  ns<sup>-1</sup>,  $\tau_{mm} = 3, 3.02, 3.06$  ns,  $m = 1, 2, 3$ ; (b)  $I_2 = I_1 - 1$  mA,  $k_{rm} = 12, 11, 12$  ns<sup>-1</sup>,  $\tau_{mm} = 3, 3.1, 3.2$  ns; (c)  $I_2 = I_1 - 2$  mA,  $k_{rm} = 13.3, 12.3, 13.3$  ns<sup>-1</sup>;  $\tau_{mm} = 4, 4.07, 4.13$  ns; (d)  $\rho_m$  as a function of  $\tau_{11}$ .  $\tau_{22} = \tau_{11} + 0.3$  ns,  $\tau_{33} = \tau_{11} + 0.7$  ns,  $I_m = 21, 19, 21$  mA,  $k_{rm} = 12, 10, 12$  ns<sup>-1</sup>,  $m = 1, 2, 3$ .

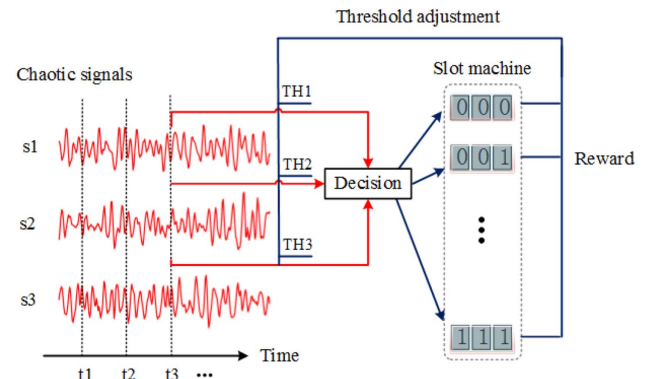
relatively small as  $\tau_{11}$  varies from 1 to 8 ns. The results indicate that in this scheme, the TDS concealment can be achieved with different time delays.

### 3. APPLICATION IN DECISION MAKING

In this section, we utilize the triple-channel chaotic signals generated from the above scheme to solve an eight-armed bandit problem in parallel. By choosing one of eight slot machines, there is a chance of getting a reward. The reward probabilities are different and unknown to users [40]. Users need to explore the slot machines to find the one that has the highest reward probability, which we call the target machine. Due to the trade-off known as the exploration-exploitation dilemma [40,41], the exploration needs to be effective so that the target machine can be found as quickly as possible and without the risk of missing it.

#### A. Scheme of Solving MAB Problem in Parallel

For an  $N$ -armed bandit problem, where  $N = 2^k$  with  $k$  being a natural number,  $k$ -bit binary number  $[D_1, D_2, \dots, D_k]$  can be used to distinguish the  $N$  slot machines [41]. When  $N = 8$  ( $k = 3$ ), the eight slot machines can be encoded by  $[D_1, D_2, D_3]$ . Figure 8 gives the schematic diagram for solving the eight-armed bandit problem in parallel. We propose a modified strategy for the implementation of the parallel scheme, in which the triple-channel chaotic signals  $s_1, s_2, s_3$  are simultaneously sampled and are, respectively, compared with the threshold values  $TH_1, TH_2, TH_3$  of each channel. Before sampling, the signals are standardized and normalized. A decision is made according to the comparison result, that is, if  $s_i(t) \leq TH_i$ ,  $D_i = 0$ , else  $D_i = 1$ . To be specific, suppose that the triple-channel chaotic signals sampled at  $t_1$  are  $s_1(t_1), s_2(t_1), s_3(t_1)$ ; then they are compared with the threshold values  $TH_1, TH_2, TH_3$ , respectively. If  $s_1(t_1) \leq TH_1$ , the most significant bit is determined as  $D_1 = 0$ ; if  $s_2(t_1) \leq TH_2$ , the second-most significant bit is  $D_2 = 0$ ; if  $s_3(t_1) \leq TH_3$ , the last-significant bit is  $D_3 = 0$ . Therefore, the slot machine 1, marked by  $D = [0, 0, 0]$ , is chosen. If a reward is given by choosing slot machine 1, then the threshold values are adjusted so that the same decision is more likely to be made in the next cycle. Otherwise, if no reward is yielded, the threshold values



**Fig. 8.** Architecture for the eight-armed bandit problem processed in parallel based on triple-channel chaos.

are adjusted to reduce the probability of making the same choice the next time.

### B. Threshold Value Adjustment

The threshold values of the three channels are independently updated according to  $TH_i = k[TV_i]$ ,  $i = 1, 2, 3$ , where  $[TV_i]$  is the threshold adjuster and takes the integer value from  $[-L, L]$ .  $L$  is a constant integer. Here we set  $L = 10$ .  $k$  is a constant factor to limit the range of  $TH_i$ . The threshold values are adjusted as follows.

If the selected slot machine yields a reward at  $t$ , the TV value is updated at  $t + 1$  by

$$\begin{cases} TV_i(t+1) = +\Delta + \alpha TV_i & \text{if } D_i = 0 \\ TV_i(t+1) = -\Delta + \alpha TV_i & \text{if } D_i = 1 \end{cases} \quad (3)$$

If the selected slot machine yields no reward at  $t$ , the TV value is updated at  $t + 1$  by

$$\begin{cases} TV_i(t+1) = -\Omega_i + \alpha TV_i & \text{if } D_i = 0 \\ TV_i(t+1) = +\Omega_i + \alpha TV_i & \text{if } D_i = 1 \end{cases} \quad (4)$$

where the increment parameter  $\Delta$  is fixed unity [41],  $\alpha = 0.99$  is a constant memory parameter,  $\Omega_i$  is determined based on the history of getting rewards, and is given by [42]

$$\Omega_i = \hat{P}_{D_i} = 0 + \hat{P}_{D_i} = 1, \quad (5)$$

$$\hat{P}_{D_i} = k = \frac{N_{D_i} = k, \text{hit}}{N_{D_i} = k, \text{total}}. \quad (6)$$

$N_{D_i} = k, \text{total}$  is the total number of times selecting  $D_i = k$  ( $i = 1, 2, 3; k = 0, 1$ ).  $N_{D_i} = k, \text{hit}$  is the number of times that one gets a reward by selecting  $D_i = k$ . The initial value of  $[TV_i]$  is set to 0. Note that for an  $N$ -armed bandit problem where  $N = 2^k$ , it only requires  $k$ -channel signals and  $k$ -threshold values, which greatly simplifies the implementation compared with the previous method that requires  $2^k - 1$  threshold values [41,42].

### C. Results and Discussion

To describe the decision-making performance, we define convergence cycle (CC) as the number of the first cycle that reaches a correct decision rate (CDR) of 0.9, where  $CDR = N_{\text{hit}}/N_{\text{total}}$  is the ratio of the times of getting a reward and the total number of selections. In practice, the average accuracy rate is often adopted to describe a short-time behavior, as the environment is always changing [45]. Here, the CDR is averaged over 400 repeated runs.

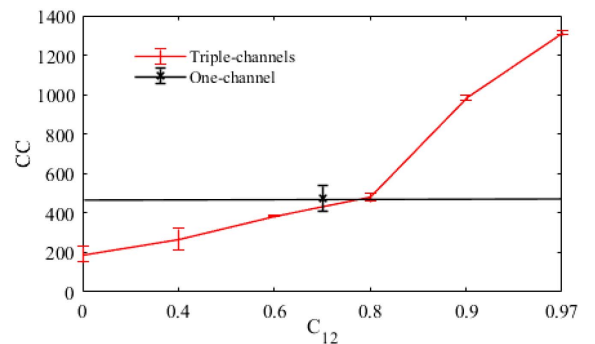
Due to the parallel structure of our scheme, the cross correlation among the triple-channel chaotic signals should be taken into account. The cross-correlation function is introduced as [5]

$$C_{mn}(\Delta t) = \frac{\langle [I_m(t + \Delta t) - \langle I_m(t + \Delta t) \rangle][I_n(t) - \langle I_n(t) \rangle] \rangle}{\sqrt{\langle [I_m(t + \Delta t) - \langle I_m(t + \Delta t) \rangle]^2 \rangle \langle [I_n(t) - \langle I_n(t) \rangle]^2 \rangle}}, \quad (7)$$

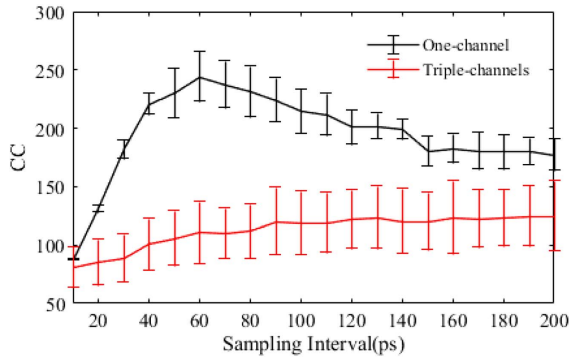
where  $C_{mn}(\Delta t)$  is the cross-correlation coefficient of signals from  $SLm$  and  $SLn$  at time lag  $\Delta t$ . The zero lag correlation ( $\Delta t = 0$ ) can be accurately controlled by shifting the signals in the time domain.

Three channels of zero-lag synchronized chaotic signals may cause an ultrafast convergence when the target is encoded as  $[0,0,0]$ , making it nearly impossible to recognize the target machine  $[0,1,0]$ . For simplicity, to investigate the impact of correlation on the performance of decision making, we only consider the effect of  $C_{12}(\Delta t)$ , and the values of  $C_{13}(\Delta t)$  and  $C_{23}(\Delta t)$  are kept close to 0. In Fig. 9, we show the CC as a function of  $C_{12}(\Delta t)$  for three sets of numerically generated signals with different correlations, where  $\Delta t = 0$ . Additionally, the result of the one-channel scheme is also calculated for a brief comparison. Here, the distribution of reward probability is  $P = [0.2, 0.2, 0.8, 0.2, 0.2, 0.2, 0.2, 0.2]$ . It can be seen that as the cross correlation decreases, the CC of the triple-channel scheme is smaller and becomes less than that of the one-channel scheme when  $C_{12}(\Delta t) < 0.8$ . This critical value may change with different distributions of reward probability and with different signals. The result shows, obviously, that the performance of the triple-channel scheme could outstrip the one-channel scheme when the correlation of the signals is quite low (which is easy to realize for chaos signals). Therefore, in order to reduce the impact of correlation among the triple-channel signals, we properly shift each set of signals in the time domain so that their cross-correlation coefficient at zero-time lag is around 0. Here, the time lags for the three signals to avoid the cross correlation are 0, 1, and 2 ns, respectively.

Next, we compare the decision-making performance of the one-channel scheme and the triple-channel scheme by calculating the CC with different sampling intervals. The results are illustrated in Fig. 10. It can be seen that for both schemes, it converges quickly when the sampling interval is as small as 10 ps, which requires the highest sampling rate that is currently available, but slows down with the increase of sampling interval. Hence, we choose a sampling rate of 10 ps in the following. Also note that the CC value of the triple-channel scheme is statistically lower and grows more slowly than that of the one-channel scheme, which means that in the proposed scheme, it can converge more quickly to the desired accuracy, and the performance is relatively stable against the variation of sampling interval. Note that in Fig. 10 and the following, the CC value of the one-channel scheme is the average of the results of three channel signals.

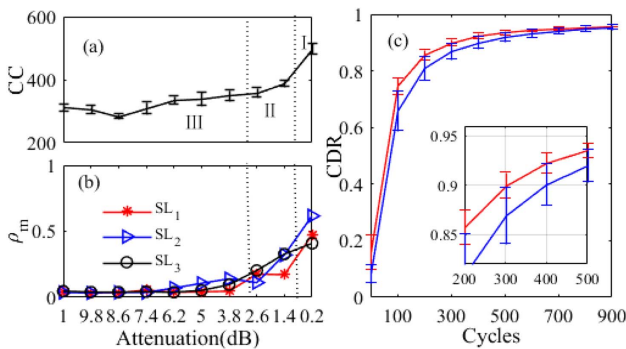


**Fig. 9.** Evolution of CDR for the triple-channel signals with different correlations and for the one-channel scheme. The vertical bars indicate the standard deviation around the mean value for three sets of simulated signals.  $P = [0.2, 0.2, 0.8, 0.2, 0.2, 0.2, 0.2, 0.2]$ .



**Fig. 10.** CC with different sampling intervals for the one-channel and triple-channel schemes, respectively. The vertical bars indicate the standard deviation around the mean value for eight sets of simulated signals.  $P = [0.8, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2]$ .

Then the experimentally generated signals with varying attenuation are utilized to investigate the influence of TDS on the decision-making performance. The CC as a function of attenuation is presented in Fig. 11(a), and in Fig. 11(b) we show the result of  $\rho_m$  for ease of comparison. The laser dynamics is clarified, as in Fig. 3. It is obvious that when  $\rho_m > 0.3$ , especially when it reaches about 0.6, the cycle to reach a CDR of 0.9 is quite large. When  $\rho_m < 0.3$ , the change of CC is not directly linked with  $\rho_m$ , but overall, a smaller CC appears with lower  $\rho_m$ . Note that the signals are normalized during preprocessing, so it is not the amplitude of the signals but the characteristics that affect the result. In addition, for a deeper understanding of the influence of TDS suppression on the decision-making performance, we statistically investigate the evolution of CDR using numerical signals with different TDS concealments, where the value of  $\rho_m$  is controlled by slightly changing the bias current, the coupling strength, or the coupling delay of the three SLs. In Fig. 11(c), we show the CDR as a function of the learning cycles using 11 sets of signals with  $\rho_m < 0.2$  and  $\rho_m > 0.3$ , respectively. It can be seen that there exist signals with larger  $\rho_m$  that still converge more quickly than those with lower  $\rho_m$ , showing that the decision-making performance does not entirely depend on the suppression of TDS. However, on the whole, it

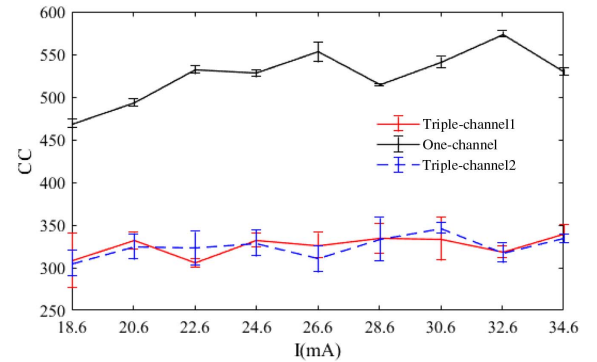


**Fig. 11.** (a), (b) CC and  $\rho_m$  as functions of attenuation; (c) CDR as a function of learning cycles. The vertical bars indicate the standard deviation around the mean value for 11 sets of signals with  $\rho_m < 0.2$  and  $\rho_m > 0.3$ , respectively. The sampling interval is 10 ps.  $P = [0.3, 0.2, 0.8, 0.1, 0.2, 0.3, 0.5, 0.4]$ .

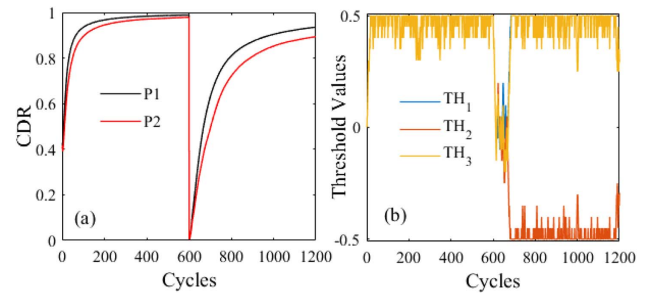
converges faster for signals with lower  $\rho_m$  in a decision-making problem, which indicates that the concealment of TDS can be helpful for better decision-making performance.

Next, we compare the decision-making performance of the one-channel scheme, the previously proposed parallel scheme [42], and the triple-channel scheme by calculating the CC, where experimentally generated signals with different bias currents are adopted. The results are illustrated in Fig. 12. Triple-channel1 and Triple-channel2 represent the new scheme and the previously proposed scheme, respectively. Three channels of signals are used to solve the eight-armed bandit problem. However, in the Triple-channel2 scheme, the adopted algorithm for threshold adjustment is the same as in the one-channel scheme. It can be seen that for both the triple-channel schemes, the CC is quite stable against the variation of bias current, and the performance is quite similar, whereas for the one-channel scheme, it takes more cycles to reach the desired CDR, and the CC value fluctuates more obviously with the change of bias current, indicating that the one-channel scheme may be more sensitive to the dynamics of signals.

In addition, it is necessary to make decisions accurately in a dynamically changing environment, where the slot machine with the highest reward probability may change with time. Figure 13(a) illustrates the evolution of the CDR in a changing



**Fig. 12.** CC as a function of bias current, for a comparison of the triple-channel scheme (red solid line), the previously investigated parallel scheme (blue dotted line), and the one-channel scheme (black solid line). The vertical bars indicate the standard deviation around the mean value for three runs.  $P = [0.3, 0.2, 0.8, 0.1, 0.2, 0.3, 0.5, 0.4]$ .



**Fig. 13.** (a) Evolution of the CDR for different distributions of reward probability in a changing environment.  $P_1 = [0.8, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2]$ ,  $P_2 = [0.7, 0.2, 0.3, 0.2, 0.2, 0.2, 0.2, 0.2]$ . (b) Threshold value adaption for  $P_2$ .



environment. We suppose that the target machine changes from slot machine 1 to 3 at the 600th cycle, and slot machines with different probability distributions are considered for comparison. It can be seen that after the sudden change of the target machine, the CDR drops to zero, and then increases rapidly. Meanwhile, one can see that it takes longer time to reach a CDR of 0.9 for  $P_2$  than that for  $P_1$ , because the former has less difference in the distribution of reward probability [12,46]. To further reveal the underlying process of the reinforcement learning, the adaption of the threshold values during the 1200 cycles is presented in Fig. 13(b). In the first 600 cycles where the target slot machine is encoded as [0,0,0], the threshold values  $TH_1$ ,  $TH_2$ , and  $TH_3$  all increase until they eventually fluctuate around a maximum value of 0.5. Hence, the chaotic signals  $s_1(t)$ ,  $s_2(t)$ ,  $s_3(t)$  are more likely to be lower than the threshold values  $TH_i$  ( $i = 1, 2, 3$ ), and the three significant bits  $[D_1, D_2, D_3]$  are more likely to be determined as [0,0,0]. When the target machine changes to [0,1,0], after temporary fluctuation around 0, the values of  $TH_1$  and  $TH_3$  return to about 0.5. The value of  $TH_2$  is reduced to about -0.5, which makes it more possible for  $s_2(t)$  to be larger than  $TH_2$ , and further results in an increase in the likelihood of choosing the slot machine [0,1,0].

Scalability is also very important for a decision-making scheme. Due to the chaotic dynamics of signals, it can be assumed that arbitrarily selected  $k$ -channel chaotic signals that are generated from the scheme as in Fig. 1 can be utilized to solve the  $N$ -armed bandit problem successfully. To demonstrate this, three channels of experimentally generated signals with varying bias current are randomly selected to solve the eight-armed bandit problem. The evolution of the CDR is presented in Fig. 14, denoted by a red solid line, and the vertical bars indicate the standard deviation around the mean value for 10 different selections. It can be seen that the average CDR is about 330, similar to the result in Fig. 12. Meanwhile, eight different selections of four-channel signals are successfully used to solve a 16-armed bandit problem. The evolution of the CDR is also shown in Fig. 14, represented by the dashed blue line. These results show that random combination of chaotic signals is

capable of solving the MAB problem efficiently, and the scalability of our scheme to larger decision problems is verified.

## 4. CONCLUSION

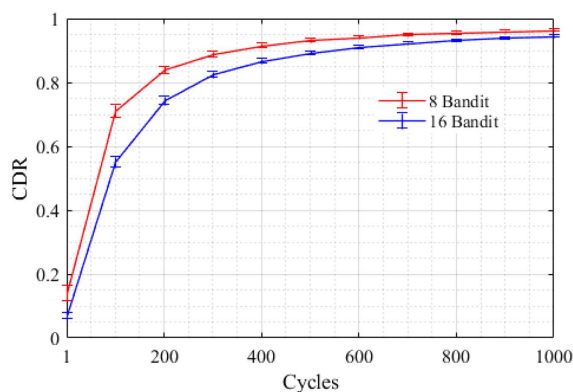
In conclusion, we propose a simple scheme of achieving triple-channel chaotic signals with TDS concealment and demonstrate it via experiment and numerical analysis. The parameters' range that contributes to better TDS concealment is explored by systematically changing the bias current and the coupling strength. Moreover, we utilize the generated triple-channel chaotic signals and a modified strategy for the realization of an eight-armed bandit problem in parallel; the influences of the signal correlation between each channel, the TDS concealment, and the sampling interval on the performance of decision making are investigated. In the proposed decision-making scheme, the simplified algorithm compared with the one-channel scheme and the previously studied parallel scheme makes it easier for implementation. However, it can perform even better given that the mutual-correlation is relatively low. Moreover, it has stabler performance for different sampling rates than the one-channel scheme. The proposed system is scalable to varying size of MAB problems and is adaptable in changing environments. This work may be helpful for potential applications in the ultrafast processing of AI.

**Funding.** National Natural Science Foundation of China (61974177, 61674119).

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

1. J. Ohtsubo, *Semiconductor Lasers: Stability, Instability and Chaos* (Springer, 2012).
2. P. Li, Y. Guo, Y. Q. Guo, Y. L. Fan, X. M. Guo, X. L. Liu, K. Y. Li, K. A. Shorel, Y. C. Wang, and A. B. Wang, "Ultrafast fully photonic random bit generator," *J. Lightwave Technol.* **36**, 2531–2540 (2018).
3. S. Y. Xiang, B. Wang, Y. Wang, Y. N. Han, A. J. Wen, and Y. Hao, "2.24-Tb/s physical random bit generation with minimal post-processing based on chaotic semiconductor lasers network," *J. Lightwave Technol.* **37**, 3987–3993 (2019).
4. G. D. Van Wiggeren and R. Roy, "Communication with chaotic lasers," *Science* **279**, 1198–1200 (1998).
5. C. Posadas-Castillo, R. M. López-Gutiérrez, and C. Cruz-Hernández, "Synchronization of chaotic solid-state Nd:YAG lasers: application to secure communication," *Commun. Nonlinear Sci. Numer. Simul.* **13**, 1655–1667 (2008).
6. N. Jiang, W. Pan, L. S. Yan, B. Luo, S. Y. Xiang, L. Yang, D. Zheng, and N. Q. Li, "Chaos synchronization and communication in multiple time-delayed coupling semiconductor lasers driven by a third laser," *IEEE J. Sel. Top. Quantum Electron.* **17**, 1220–1227 (2011).
7. C. Xue, N. Jiang, K. Qiu, and Y. Lv, "Key distribution based on synchronization in bandwidth-enhanced random bit generators with dynamic post-processing," *Opt. Express* **23**, 14510–14519 (2015).
8. J. Vatin, D. Rontani, and M. Sciamanna, "Experimental reservoir computing using VCSEL polarization dynamics," *Opt. Express* **27**, 18579–18584 (2019).
9. X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "Polarization multiplexing reservoir computing based on a VCSEL with polarized optical feedback," *IEEE J. Sel. Top. Quantum Electron.* **26**, 1700109 (2020).
10. M. Naruse, W. Nomura, M. Aono, M. Ohtsu, Y. Sonnefraud, A. Drezet, S. Huant, and S. J. Kim, "Decision making based on optical excitation



**Fig. 14.** Evolution of the averaged CDR with randomly selected signals for eight-armed and 16-armed bandit problems.  $P = [0.3, 0.2, 0.8, 0.1, 0.2, 0.3, 0.5, 0.4]$ , and  $P = [0.3, 0.2, 0.8, 0.1, 0.2, 0.5, 0.2, 0.2, 0.2, 0.3, 0.3, 0.4, 0.5, 0.1, 0.1, 0.2]$ , respectively.

- transfer via near-field interactions between quantum dots," *J. Appl. Phys.* **116**, 154303 (2014).
11. T. Mihana, Y. Mitsui, M. Takabayashi, K. Kazutaka, S. Sunada, M. Naruse, and A. Uchida, "Decision making for the multi-armed bandit problem using lag synchronization of chaos in mutually-coupled semiconductor lasers," *Opt. Express* **27**, 26989–27008 (2019).
  12. M. Naruse, N. Chauvet, A. Uchida, A. Drezet, G. Bachelier, S. Huant, and H. Hori, "Decision making photonics: solving bandit problems using photons," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700210 (2020).
  13. S. Y. Xiang, Y. Zhang, J. Gong, X. Guo, L. Lin, and Y. Hao, "STDP-based unsupervised spike pattern learning in a photonic spiking neural network with VCSELs and VCSOs," *IEEE J. Sel. Top. Quantum Electron.* **25**, 1700109 (2019).
  14. J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–214 (2019).
  15. S. Y. Xiang, Z. X. Ren, Y. H. Zhang, Z. W. Song, and Y. Hao, "All-optical neuromorphic XOR operation with inhibitory dynamics of a single photonic spiking neuron based on VCSEL-SA," *Opt. Lett.* **45**, 1104–1107 (2020).
  16. S. Y. Xiang, Z. X. Ren, Y. H. Zhang, X. X. Guo, G. Q. Han, and Y. Hao, "Computing primitive of fully-VCSELs-based all-optical spiking neural network for supervised learning and pattern classification," *IEEE Trans. Neural Netw. Learning Syst.*, 1–12 (2020).
  17. J. G. Wu, Z. M. Wu, X. Tang, X. D. Lin, T. Deng, G. Q. Xia, and G. Y. Feng, "Simultaneous generation of two sets of time delay signature eliminated chaotic signals by using mutually coupled semiconductor lasers," *IEEE Photon. Technol. Lett.* **23**, 759–761 (2011).
  18. A. B. Wang, Y. B. Yang, B. J. Wang, B. B. Zhang, L. Li, and Y. C. Wang, "Generation of wide band chaos with suppressed time-delay signature by delayed self-interference," *Opt. Express* **21**, 8701–8710 (2013).
  19. N. Q. Li, W. Pan, S. Y. Xiang, L. S. Yan, B. Luo, X. H. Zou, L. Y. Zhang, and P. H. Mu, "Photonic generation of wide band time-delay-signature-eliminated chaotic signals utilizing an optically injected semiconductor laser," *IEEE J. Sel. Top. Quantum Electron.* **48**, 1339–1345 (2012).
  20. T. Deng, Z. M. Wu, and G. Q. Xia, "Two-mode coexistence in 1550-nm VCSELs with optical feedback," *IEEE Photon. Technol. Lett.* **27**, 2075–2078 (2015).
  21. J. G. Wu, S. W. Huang, Y. J. Huang, H. Zhou, J. H. Yang, J. M. Liu, M. B. Yu, G. Q. Lo, D. L. Kwong, S. K. Duan, and C. W. Wong, "Mesoscopic chaos mediated by Drude electron-hole plasma in silicon optomechanical oscillators," *Nat. Commun.* **8**, 15570 (2017).
  22. N. Jiang, A. K. Zhao, S. Q. Liu, C. P. Xue, and K. Qiu, "Chaos synchronization and communication in closed-loop semiconductor lasers subject to common chaotic phase-modulated feedback," *Opt. Express* **26**, 32404–32416 (2018).
  23. M. J. Bünner, A. Kittel, J. Parisi, I. Fischer, and W. Elsässer, "Estimation of delay times from a delayed optical feedback laser experiment," *Europhys. Lett.* **42**, 353–358 (1998).
  24. S. S. Li and S. C. Chan, "Chaotic time-delay signature suppression in a semiconductor laser with frequency-detuned grating feedback," *IEEE J. Sel. Top. Quantum Electron.* **21**, 541–552 (2015).
  25. M. J. Bünner, M. Popp, T. Meyer, A. Kittel, and J. Parisi, "A tool to recover scalar time-delay systems from experimental time series," *Phys. Rev. E* **54**, 3082–3085 (1996).
  26. R. Hegger, M. J. Bünner, and H. Kantz, "Identifying and modeling delay feedback systems," *Phys. Rev. Lett.* **81**, 558–561 (1998).
  27. B. P. Bezruchko, A. S. Karavaev, V. I. Ponomarenko, and M. D. Prokhorov, "Reconstruction of time-delay systems from chaotic time series," *Phys. Rev. E* **64**, 056216 (2001).
  28. M. C. Soriano, L. Zunino, O. A. Rosso, I. Fischer, and C. R. Mirasso, "Timescales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis," *IEEE J. Quantum Electron.* **47**, 252–261 (2011).
  29. X. Porte, O. D'Huys, T. Jüngling, X. Porte, D. Brunner, M. C. Soriano, and I. Fischer, "Autocorrelation properties of chaotic delay dynamical systems: a study on semiconductor lasers," *Phys. Rev. E* **90**, 052911 (2014).
  30. M. W. Lee, P. Rees, K. A. Shore, S. Ortin, L. Pesquera, and A. Valle, "Dynamical characterisation of laser diode subject to double optical feedback for chaotic optical communications," *IEEE P-Optoelectron.* **152**, 97–102 (2005).
  31. J. G. Wu, G. Q. Xia, and Z. M. Wu, "Suppression of time delay signatures of chaotic output in a semiconductor laser with double optical feedback," *Opt. Express* **17**, 20124–20133 (2009).
  32. S. Y. Xiang, A. J. Wen, W. Pan, L. Lin, H. X. Zhang, H. Zhang, X. X. Guo, and J. F. Li, "Suppression of chaos time delay signature in a ring network consisting of three semiconductor lasers coupled with heterogeneous delays," *J. Lightwave Technol.* **34**, 4221–4227 (2016).
  33. N. Jiang, Y. J. Wang, A. Zhao, S. Q. Liu, Y. Q. Zhang, L. Chen, B. C. Li, and K. Qiu, "Simultaneous bandwidth-enhanced and time delay signature-suppressed chaos generation in semiconductor laser subject to feedback from parallel coupling ring resonators," *Opt. Express* **28**, 1999–2009 (2020).
  34. L. Lai, H. ElGamal, H. Jiang, and H. V. Poor, "Cognitive medium access: exploration, exploitation, and competition," *IEEE Trans. Mobile Comput.* **10**, 239–253 (2011).
  35. K. Kuroda, H. Kato, S.-J. Kim, M. Naruse, and M. Hasegawa, "Improving throughput using multi-armed bandit algorithm for wireless LANs," *Nonlinear Theory Its Applications IEICE J.* **9**, 74–81 (2018).
  36. K. Morihiro, N. Matsui, and H. Nishimura, "Chaotic exploration effects on reinforcement learning in shortcut maze task," *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **16**, 3015–3022 (2006).
  37. S. J. Kim, M. Aono, and E. Nameda, "Efficient decision-making by volume-conserving physical object," *New J. Phys.* **17**, 083023 (2015).
  38. S. J. Kim, M. Naruse, M. Aono, M. Ohtsu, and M. Hara, "Decision maker based on nanoscale photo-excitation transfer," *Sci. Rep.* **3**, 2370 (2013).
  39. M. Naruse, M. Berthel, A. Drezet, S. Huant, H. Hori, and S. J. Kim, "Single photon in hierarchical architecture for physical decision making: photon intelligence," *ACS Photon.* **3**, 2505–2514 (2016).
  40. T. Mihana, Y. Terashima, M. Naruse, S. J. Kim, and A. Uchida, "Memory effect on adaptive decision making with a chaotic semiconductor laser," *Complexity* **2018**, 4318127 (2018).
  41. M. Naruse, T. Mihana, H. Hori, H. Saigo, K. Okamura, M. Hasegawa, and A. Uchida, "Scalable photonic reinforcement learning by time-division multiplexing of laser chaos," *Sci. Rep.* **8**, 10890 (2018).
  42. Y. T. Ma, S. Y. Xiang, X. X. Guo, Z. W. Song, A. J. Wen, and Y. Hao, "Time-delay signature concealment of chaos and ultrafast decision making in mutually coupled semiconductor lasers with a phase-modulated Sagnac loop," *Opt. Express* **28**, 1665–1678 (2020).
  43. L. Zunino, O. A. Rosso, and M. C. Soriano, "Characterizing the hyper-chaotic dynamics of a semiconductor laser subject to optical feedback via permutation entropy," *IEEE J. Sel. Top. Quantum Electron.* **17**, 1250–1257 (2011).
  44. C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Phys. Rev. Lett.* **88**, 174102 (2002).
  45. 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* **101**, 29–36 (2010).
  46. M. Naruse, Y. Terashima, A. Uchida, and S. J. Kim, "Ultrafast photonic reinforcement learning based on laser chaos," *Sci. Rep.* **7**, 8772 (2017).