






Adaptive time-delayed photonic reservoir computing based on Kalman-filter training

JIAOYANG JIN,¹ NING JIANG,^{1,*}  YIQUN ZHANG,^{1,2} WEIZHOU FENG,¹ ANKE ZHAO,¹  SHIQIN LIU,¹ JIAFA PENG,¹ KUN QIU,¹ AND QIANWU ZHANG³ 

¹*School of Information and Communication Engineering, University of Electronic Science and Technology of China, 2006 Xiyuan Avenue, West High-Tech District, Chengdu 611731, China*

²*James Watt School of Engineering, University of Glasgow, Glasgow, UK*

³*Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, China*

*uestc_nj@uestc.edu.cn

Abstract: We propose an adaptive time-delayed photonic reservoir computing (RC) structure by utilizing the Kalman filter (KF) algorithm as training approach. Two benchmark tasks, namely the Santa Fe time-series prediction and the nonlinear channel equalization, are adopted to evaluate the performance of the proposed RC structure. The simulation results indicate that with the contribution of adaptive KF training, the prediction and equalization performance for the benchmark tasks can be significantly enhanced, with respect to the conventional RC using a training approach based on the least-squares (LS). Moreover, by introducing a complex mask derived from a bandwidth and complexity enhanced chaotic signal into the proposed RC, the performance of prediction and equalization can be further improved. In addition, it is demonstrated that the proposed RC system can provide a better equalization performance for the parameter-variant wireless channel equalization task, compared with the conventional RC based on LS training. The work presents a potential way to realize adaptive photonic computing.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Reservoir computing (RC) which is originated from the recurrent neural networks (RNN), is a novel biologically inspired computing method for processing time-dependent information [1]. While different from the conventional neural networks, the connection weights between the input and the internal network in RC are fixed at a set of random values, and only the readout weights deserve training, which greatly simplifies the training process and reduces the power consumption [2–4]. The RC has been extensively applied in many areas, such as financial time-series prediction, nonlinear channel equalization [3,4], human action recognition [5], and reconstruction of complex dynamic system [6], etc. However, due to the energy loss and the noise interference among different devices, it is difficult to construct a fully-connected RC system with a large number (e.g. 10^2 - 10^3) of physical nodes.

In recent years, a new RC architecture based on a time-delayed feedback system has been proposed, in which a time-delayed nonlinear component is used to construct a virtual network, replacing the real network used in the original RC [7]. Under such a scenario, fully-connected RC can be easily implemented without the requirements of complex physical equipment. For this reason, the time-delayed RC has been a hotspot which attracted plenty of attention. Up to present, several classic time-delayed RC structures on the basis of the Mackey-Glass electronic [7], optoelectronic [8–10], and all-optical [11–20] delay feedback systems have been proposed. Brunner and coworkers realized a time-delayed RC scheme by utilizing the transient response of a semiconductor laser (SL) with optical feedback [12]. Romain et al. demonstrated a time-delayed RC structure supporting parallel processing, in virtue of two modes of semiconductor ring laser

(SRL) [13]. S. Xiang and colleagues proposed to use a semiconductor nanolaser (SNL) with double phase conjugate feedbacks (PCF) to construct a time-delayed RC system [14]. G. Q. Xia et al. investigated a time-delayed RC system under electrical information injection and further optimized its performance [15].

In general, performance of time-delayed RC greatly relies on the training technique that is adopted to obtain readout weights. In the previously-reported literatures, two typical training methods based on the least squares (LS) and the ridge regression, are commonly adopted, to obtain the readout weights. With these two training methods, the RC structure can achieve satisfactory performance without complex calculations [21–26]. Nevertheless, since the readout weights remain invariable once the training using these two methods is completed, these RC structures are usually adopted to process the parameter-invariant tasks. While for the parameter-variant tasks, even though a satisfactory initial performance can be achieved by the RC structure with a proper training, the long-term performance would deteriorate if the RC cannot update the readout weights according to the variation of task states. Therefore, it is valuable to explore novel RC structures that are capable of processing parameter-variant tasks.

In this paper, we propose an adaptive time-delayed photonic RC structure in which Kalman filter (KF) algorithm is adopted as the training approach. A chaotic mask signal with enhanced complexity and bandwidth is used as the mask signal of input signal. The performance of the proposed RC structure is evaluated by two benchmark tasks, namely the Santa Fe time-series prediction task and the nonlinear channel equalization task. Moreover, the feasibility of the proposed adaptive RC structure to process the parameter-variant wireless channel equalization is demonstrated.

2. Principles and system model

Figure 1 shows the schematic of the proposed adaptive time-delay RC structure. Similar to typical SL-based RC structures, the reservoir consists of a semiconductor laser (referred to as response laser) subject to a time-delayed optical feedback and an external optical injection from a driver laser (DL). With the transient response of the response laser (RL), the reservoir can map the input information into a high-dimensional state space, and simultaneously, the fading-memory feature can be implemented in the feedback loop.

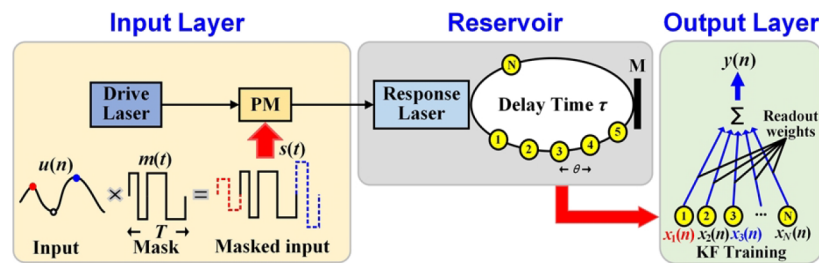


Fig. 1. Schematic of the time-delayed RC based on a response laser (RL) under optical feedback and optical injection from a drive laser (DL). PM: phase modulator.

The proposed adaptive RC structure is composed of three parts: the input layer, the reservoir, and the output layer. In the input layer, the input signal is firstly sampled and each sample is expanded for time duration T , the expanded samples are denoted as $u(n)$, where n is the discrete time index. Then, a temporal masking signal $m(t)$ with the length of T is multiplied with $u(n)$, to obtain a masked input signal $s(t)$ (i.e., $s(t) = u(n) \times m(t) \times \gamma$, here γ is a scaling factor). After that, the masked input signal is sent into the reservoir by modulating the output of DL with an opt-electronic phase modulator (PM), as those in [16,17,22–24]. This process can be

mathematically described as:

$$E_d(t) = \sqrt{I_d} \exp(i\pi s(t)). \quad (1)$$

where I_d represents the output optical intensity of DL and $E_d(t)$ is the optical field injected into the reservoir. To fully make use of the dynamics of the phase of optical field of response laser, here the masked input signal is not normalized. In the reservoir, the transient response of RL with an interval θ is taken as a state of a virtual node. Within a period of feedback time τ , N virtual nodes $x_i(n)$ ($i = 1, 2, 3, \dots, N$) for the n -th input data can be implemented, on the basis of a de-synchronization scheme, in which the feedback time τ of RL is chosen as $\tau = T + \theta$, and $T = N \times \theta$ is satisfied [8, 11, 16, 18]. In the output layer, the reservoir output $y(n)$ for the n -th input data is calculated by a linear combination of virtual nodes $x_i(n)$ with readout weights w_i for each time duration T as follows:

$$y(n) = \sum_{i=1}^N w_i x_i(n). \quad (2)$$

It is the read-out estimator, and the readout weights can be optimized by minimizing the mean square error between the target value $\bar{y}(n)$ that is defined in Eq. (6) and the RC output $y(n)$. The nonlinear dynamics of RL in the reservoir are simultaneously affected by the time-delayed optical feedback and the phased-modulated masked input signal, which can be illustrated by the following Lang-Kobayashi rate equations [27, 28]:

$$\begin{aligned} \frac{dE(t)}{dt} = & \frac{1}{2}(1 + i\alpha) \left[\frac{g(N(t) - N_0)}{1 + \varepsilon|E(t)|^2} - \frac{1}{\tau_p} \right] E(t) + k_{inj} E_d(t) \exp(i2\pi \Delta v t) \\ & + k_f E(t - \tau) \exp(-i2\pi \nu \tau) + \sqrt{2\beta N(t)} \chi(t) \end{aligned} \quad (3)$$

$$\frac{dN(t)}{dt} = J - \frac{N(t)}{\tau_s} - \frac{g(N(t) - N_0)}{1 + \varepsilon|E(t)|^2} |E(t)|^2 \quad (4)$$

where $E(t)$ is the complex electric field amplitude, $N(t)$ is the corresponding carrier density, $E(t - \tau)$ is the feedback signal with feedback time τ . α is the linewidth enhancement factor, g is the differential gain coefficient, N_0 represents the transparent carrier density, ε denotes the gain saturation coefficient, τ_p is the photon lifetime, τ_s is the carrier lifetime, k_f is the feedback strength of RL, k_{inj} is the injection strength from DL to RL, J means the injection current, ν is the frequency of RL, and Δv is the frequency detuning between DL and RL. A white Gaussian noise χ with zero mean and unity variance is used to model the spontaneous emission noise, and β is the spontaneous emission rate.

In a conventional RC structure using the least-squares algorithm (LS-RC) for training, the readout weights w_1, w_2, \dots, w_N are fixed, once the training process is completed, as such this type of RC structure cannot afford to process the parameter-variant tasks that require adaptive readout-weight updating. To enable a reservoir to estimate the readout weights adaptively, a KF training method is introduced in the proposed RC structure. Differing from the traditional LS algorithm, the KF algorithm is a kind of typical method that adopts the state-space model to solve the error minimization problem, and it can provide adaptive recursive estimation without large data storage requirement [29]. In the proposed RC structure, the nonlinear part of the task is processed by the reservoir, as such the residual output weights training is a linear regression task, and then a simple classic KF algorithm is able to update the output weights adaptively. The states of virtual nodes $x(n)$ (i.e., $x(n) = [x_1(n), x_2(n), \dots, x_N(n)]^T$, here T means the transpose operation of matrix) are directly sent into the Kalman filter, and the fading memory of the reservoir provides

memory capacity for the filter. The state-space of KF is described by the following equations:

$$w(n) = w(n-1) + \xi(n-1) \quad (5)$$

$$\bar{y}(n) = w^T(n)x(n) + \eta(n). \quad (6)$$

where $w(n)$ (i.e., $w(n) = [w_1(n), w_2(n), \dots, w_N(n)]^T$) is the readout weights, $\bar{y}(n)$ is the target value (teacher output), $\xi(n)$ and $\eta(n)$ are two independent Gaussian noise sources with zero mean, and their covariance matrices are denoted as $Q(n)$ and $R(n)$, respectively. In the proposed RC structure, the readout weights are recursively obtained by the following equations [29,30]:

$$w(n|n-1) = w(n-1) \quad (7)$$

$$P(n|n-1) = P(n-1) + Q(n-1) \quad (8)$$

$$K(n) = P(n|n-1)x(n)[R(n) + x^T(n)P(n|n-1)x(n)]^{-1} \quad (9)$$

$$w(n) = w(n|n-1) + K(n)[\bar{y}(n) - w(n|n-1)x(n)] \quad (10)$$

$$P(n) = [I - K(n)x(n)]P(n|n-1) \quad (11)$$

where $P(n)$ is the covariance matrix of the estimation error of output weights, $K(n)$ is Kalman gain matrix, $w(n|n-1)$ and $P(n|n-1)$ are the estimations of $w(n)$ and $P(n)$ based on the previous measurements from the time index 1 to $n-1$. Through this process, the readout weights can be adaptively updated with the variation of the task states, and the RC output can be obtained by Eq. (2). The proposed RC structure is abbreviated as KF-RC.

3. Results and discussion

To thoroughly discuss the properties and demonstrate the advantages of the proposed structure of KF-RC, other two relevant structures, namely KF and LS-RC, are also considered for comparison. The fourth-order Runge-Kutta algorithm with a step of 1 ps is adopted to solve the rate equations of RL shown in Eqs. (3)-(4). The laser parameters used in our simulations are as follows: $\alpha = 3.0$, $g = 8.40 \times 10^{-13} \text{ m}^3 \text{ s}^{-1}$, $N_0 = 1.4 \times 10^{24} \text{ m}^{-3}$, $\varepsilon = 2 \times 10^{-23}$, $\tau_p = 1.927 \text{ ps}$, $\tau_s = 2.04 \text{ ns}$, $k_{inj} = 12.43 \text{ ns}^{-1}$, $k_f = 15.53 \text{ ns}^{-1}$, $J = 1.037 \times 10^{33} \text{ m}^{-3} \text{ s}^{-1}$, $I_d = 6.56 \times 10^{20}$, $\nu = 1.96 \times 10^{14} \text{ s}^{-1}$, $\Delta\nu = -4.0 \text{ GHz}$, and $\beta = 1.5 \times 10^{-6}$. Here, these parameters correspond to an optimum RC operation condition, which is selected by repeating simulations with the parameters reported in [16]. The value of T is set as 1ns, which corresponds to a processing rate of 1GSamples/s.

3.1. Conventional and complexity-enhanced chaotic mask signals

Two chaotic masks are introduced to investigate the impact of the complexity of chaotic mask signals on the performance of RC. One is the conventional chaotic mask that is derived from an external-cavity semiconductor laser (ECSL) and referred to as the C-chaotic mask. The other one is the complexity-enhanced chaotic mask that is referred to as the E-chaotic mask and obtained by a self-phase-modulated feedback ECSL cascaded with dispersive component. The generation of the E-chaotic signal has been thoroughly investigated in our recent work [31].

Figure 2 presents the temporal waveforms, the power spectra, and the autocorrelation function (ACF) of the analog C-chaotic mask and E-chaotic mask signals generated by using identical ECSL (the laser and dispersive component parameters can be found in [31], and the feedback delay is set as 8 ns here). For the C-chaotic mask signal generated by conventional ECSL (first row), the power is mainly concentrated nearby the relaxation oscillation frequency, as such the effective bandwidth is relatively low, which is only 7.45 GHz. Moreover, an obvious time-delay signature (TDS) appears at the position of feedback delay [32–34]. While for the E-chaotic mask signal (second row), the spectrum is much flatter, and the effective bandwidth is enhanced to 36.9

GHz. Besides, it shows Delta-like autocorrelation, as that shown in Fig. 2(b3). Therefore, with respect to the C-chaotic mask, both of the bandwidth and complexity of the E-chaotic mask signal have been significantly enhanced. In the following discussions, as that in [16], the amplitude of both chaotic masks is rescaled so that the mean value of the chaos masks is set to 0, and its standard deviation is set to 1.

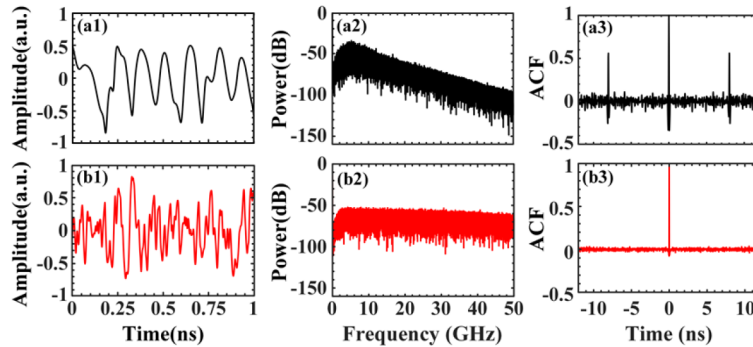


Fig. 2. Temporal waveforms (first column), RF spectra (second column), and ACF traces (third column) of (a) C-chaotic mask signal and (b) E-chaotic mask signal

Figure 3 presents the temporal waveforms of the masked input signals, as well as the corresponding transient responses of RL and virtual node states, in the cases with C-chaotic mask signal and E-chaotic mask signal. The temporal waveforms of the masked input signals are similar to those of the mask signals for the linear relationship between them [$s(t)=u(n)\times m(t)\times\gamma$, in each period of $T=1\text{ns}$, $u(n)$ is a constant, and $s(t)$ is proportional to $m(t)$]. Regarding the C-chaotic mask case, both of the masked input signal and the output of RL fluctuate smoothly, which induces that some virtual nodes do not display transient dynamics. While for the case with E-chaotic mask signal, since the bandwidth and complexity of mask signal are significantly enhanced, it is more efficient for the masked input signal to activate the high-dimensional mapping capability. Consequently, with respect to that in the C-chaotic mask case, the output of RL fluctuates more dramatically, and richer virtual node states can be obtained from the more complex and frequent dynamical response signal. Based on this, with the E-chaotic mask, better performance can be achieved, and this will be confirmed in the following discussions.

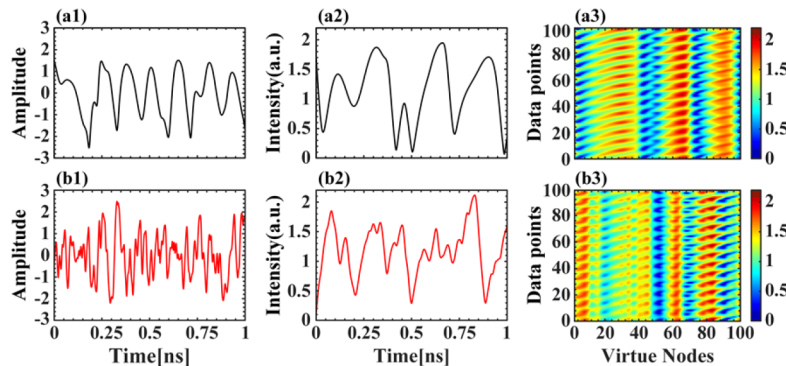


Fig. 3. Temporal waveforms of masked input signal (first column), transient response of RL (second column), and virtual node states (third column), in the cases with (a) C-chaotic mask signal and (b) E-chaotic mask signal.

3.2. Performance of Santa Fe chaotic time-series prediction

To evaluate the prediction performance of the proposed RC structure, the Santa Fe time-series prediction task which is a typical benchmark task that has been extensively used in the machine learning area is employed. The target for this task is to perform a one-step forward prediction of the Santa Fe chaotic time series that is experimentally derived from a far-infrared laser [35]. The prediction performance of RC is quantitatively evaluated by using the normalized mean square error (NMSE) defined as follows:

$$NMSE = \frac{1}{L} \frac{\sum_{n=1}^L (\bar{y}(n) - y(n))^2}{\text{var}(\bar{y})} \quad (12)$$

where L is the total number of the input data, n is the index of input data, $y(n)$ is the prediction result of RC, $\bar{y}(n)$ denotes the target value, and var represents the variance. Generally, when $NMSE \leq 0.1$, the prediction performance of RC can be considered good [14,15,22–24]. In the following investigations on the performance of the Santa Fe chaotic time-series prediction, 3000 steps in the Santa Fe data set are used for training, 1000 steps are adopted for testing, the number of virtual nodes is set as $N = 100$, the virtual node interval is fixed at $\theta = 0.01$ ns, and the scaling factor is chosen as $\gamma = 0.6$.

Figure 4 presents the results of the Santa Fe time-series prediction using direct Kalman filtering, as well those using LS-RC and KF-RC with C-chaotic mask and E-chaotic mask. As shown in Fig. 4(b), for the direct Kalman filtering, the prediction output fluctuates in a relatively small range and a relatively large prediction error is observed ($NMSE = 0.1478$). That is, using a single Kalman filter cannot achieve accurate time-series prediction. For the prediction with LS-RC structure, as those shown in Figs. 4(c) and 4(d), the prediction error is significantly improved, accurate prediction with small prediction error can be observed in a large time duration except that nearby the sharp jump point in the envelope of the Santa Fe time series. The NMSEs are respectively reduced to 0.0166 and 0.0084 for the cases with C-chaotic mask and E-chaotic mask. For the prediction with KF-RC, as those shown in Figs. 4(e) and 4(f), the NMSEs are further reduced to 0.0080 and 0.0028, which means the prediction performance is further improved. The performance improvement of KF-RC (with respect to LS-RC) is because the KF algorithm uses the state-space model to recursively update the readout weights in each training step, as such the effects of multicollinearity that would reduce the accuracy of the readout weight estimation in the conventional LS method can be avoided [30]. On the other hand, it is indicated that in both of the LS-RC and KF-RC structures, using E-chaotic mask can obtain better prediction performance than that of the cases with C-chaotic mask. Overall, the proposed RC structure using Kalman filter algorithm for training and adopting E-chaotic mask with enhanced bandwidth and complexity shows better prediction performance.

In addition to the linear readout layer defined in Eq. (2), a recurrent readout layer can also be introduced into the proposed RC structure. Figure 5 presents the NMSE values of the Santa Fe time-series prediction task using the KF-RC + E-chaotic mask RC structures with the linear readout layer in Eq. (2) and a recurrent readout layer defined as $y(n) = w^T(n)[1; x(n); x(n-1)]$. Here, the sampling period of input data T is fixed at 1 ns. It is observed that for both of the RC structures, the more the virtual nodes, the smaller the NMSE, and the RC system with recurrent readout layer behaves better prediction performance than that with linear readout layer. This is because the RC structure with recurrent readout layer can utilize the states of virtual nodes in multiple time indexes, as such the fade memory feature of RC is enhanced.

3.3. Performance of nonlinear channel equalization

In this subsection, the nonlinear channel equalization that is a practical task of great significance in the field of wireless communication [3], is studied to evaluate the equalization performance of

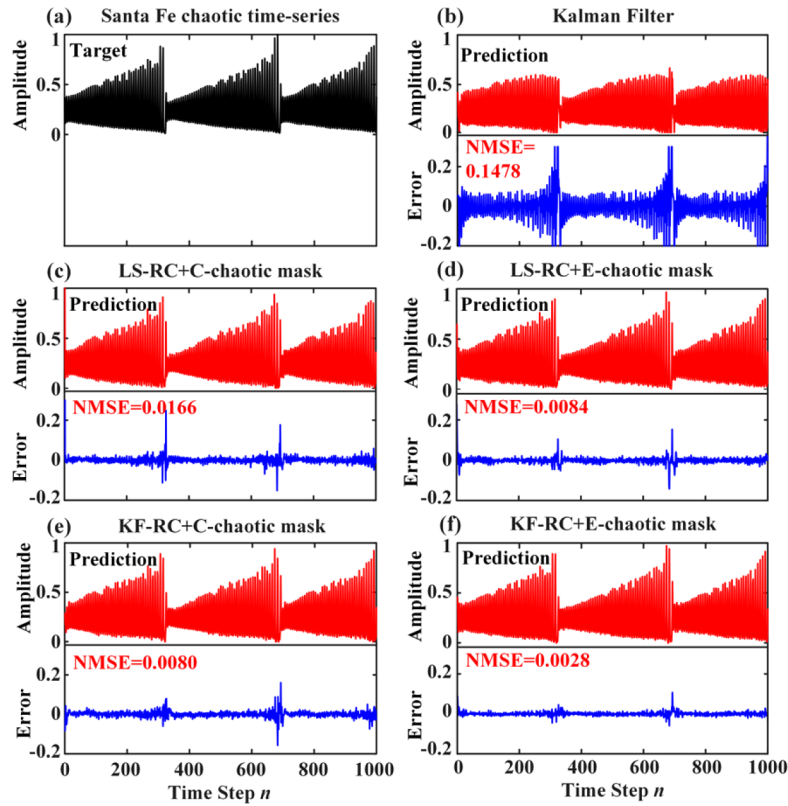


Fig. 4. (a) Santa Fe chaotic time series; Results of the Santa Fe time-series prediction task using the structure of (b) KF, (c) LS-RC + C-chaotic mask, (d) LS-RC + E-chaotic mask, (e) KF-RC + C-chaotic mask, and (f) KF-RC + E-chaotic mask. The blue curve denotes the error signal, representing the deviation between the target value and the prediction value.

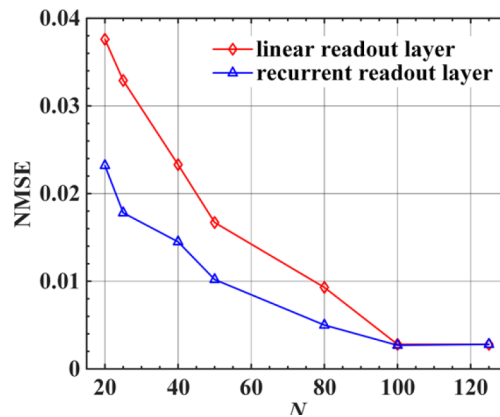


Fig. 5. NMSEs as a function of the number of virtual nodes N by utilizing the KF-RC + E-chaotic mask structure with linear readout layer (red curve) and recurrent readout layer (blue curve).

the proposed KF-RC structure. The original signal sequence $d(n)$ is randomly chosen from four values $\{-3, -1, 1, 3\}$ and propagated through a standardized multipath radio frequency channel with nonlinear distortions, defined as follows [36]:

$$\begin{aligned} q(n) = & 0.08d(n+2) - 0.12d(n+1) + d(n) + 0.18d(n-1) \\ & - 0.1d(n-2) + 0.091d(n-3) - 0.05d(n-4) \\ & + 0.04d(n-5) + 0.03d(n-6) + 0.01d(n-7) \end{aligned} \quad (13)$$

$$m(n) = q(n) + 0.036q^2(n) - 0.011q^3(n) + \zeta(n) \quad (14)$$

where $\zeta(n)$ is a Gaussian noise with zero mean, adjusted to yield signal-to-noise ratios (SNRs) ranging from 12 and 32 dB. Our goal is to recover the original symbol signal $d(n)$ from the disturbed signal $m(n)$. For this task, 10^5 symbols are generated, wherein 3×10^3 symbols are used for training and 9.7×10^4 symbols are applied to test the performance of RC. The number of virtual nodes is set as $N = 50$, which corresponds to a virtual-nodes interval of $\theta = 20$ ps, and the scaling factor is chosen as $\gamma = 0.3$. The equalization performance of RC is quantitatively evaluated in terms of the symbol error rate (SER), which is the fraction of error symbols.

Figure 6 presents the SER performance of the nonlinear channel equalization task under the five scenarios identical to those in Fig. 4. Here the SER values are obtained by averaging the results of 10 repeating testing procedures. The results indicate that for all the five equalization schemes, as the increase of SNR, the SER decreases accordingly, and the performances of RC structures are always better than that of individual Kalman filtering. When the SNR is smaller than 20dB, the equalization performance differences between different RC structures are not significant. However, when the SNR further increases, the equalization performance gaps raise up gradually. When the SNR is larger than 28dB, a SER floor can be observed in the two cases of LS-RC structures (red curves), which is in line with the experimental results reported in [36]. While for the KF-RC structures (blue curves), the SERs are further decreased, and better performance is achieved. Comparing to the LS-RC structures, the KF-RC structures with same chaotic mask show better equalization performances. On the other hand, the comparisons of the performances of RC structures with C-chaotic mask and E-chaotic mask show that the latter RC structures with identical training method can achieve better equalization performance. These phenomena are similar to those shown in Fig. 4, therefore it can be concluded that the proposed KF-RC structure with E-chaotic mask can support better prediction performance and equalization performance, with respect to the conventional LS-RC structure.

3.4. Performance of adaptive parameter-variant channel equalization

The model presented in Eqs. (13) and (14) is a static wireless communication channel model. Under such a scenario, the channel state remains invariable during the information transmission. However, in practice, the wireless channel is a typical parameter-variant channel. Either the change of environment or the relative offset of emitters and receivers would induce channel state variation. Here, in our simulations, the channel parameter variation is introduced by randomly tuning the coefficients in the channel model shown in Eqs. (13) and (14), which is mathematically expressed as

$$\begin{aligned} q'(n) = & 0.08(1 + \lambda_1)d(n+2) - 0.12(1 + \lambda_2)d(n+1) + (1 + \lambda_3)d(n) \\ & + 0.18(1 + \lambda_4)d(n-1) - 0.1(1 + \lambda_5)d(n-2) + 0.091(1 + \lambda_6)d(n-3) \\ & - 0.05(1 + \lambda_7)d(n-4) + 0.04(1 + \lambda_8)d(n-5) + 0.03(1 + \lambda_9)d(n-6) \\ & + 0.01(1 + \lambda_{10})d(n-4) \end{aligned} \quad (15)$$

$$m'(n) = (1 + \lambda_{11})q'(n) + 0.036(1 + \lambda_{12})q'^2(n) - 0.011(1 + \lambda_{13})q'^3(n) + \zeta(n) \quad (16)$$

where λ_i ($i = 1, 2, 3, \dots, 13$) is the variation factor that follows the Gaussian distribution with zero mean, and the probability distribution for different variation factors are independent. To

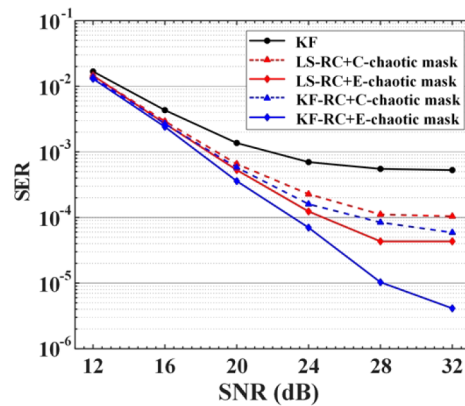


Fig. 6. Performance for the nonlinear channel equalization task versus the SNR by utilizing individual KF (black curve), LS-RC + C-chaotic mask (red dotted line), LS-RC + E-chaotic mask (red curve), KF-RC + C-chaotic mask (blue dotted line), and KF-RC + E-chaotic mask (blue curve).

thoroughly investigate the equalization performances under different variation-strength scenarios, four different values for the variance of variation factor that is referred to as σ^2 , namely $\sigma^2 = 0.1, 0.2, 0.5,$ and 1 , are simultaneously considered. In the following discussions, for the sake of simplicity, only the RC structures with E-chaotic mask are considered for its advantages of prediction and statistic channel equalization.

Figure 7 presents the SER performance of the parameter-variant channel equalization task utilizing the LS-RC and KF-RC structures with E-chaotic mask, for the case of SNR = 24 dB. Here, six variation periods including one static period and five subsequent varying periods are discussed, and the variation period of λ_i ($i = 1, 2, 3, \dots, 13$) is set as 0.2ms (which is equivalent to that the channel state varies five times, and in each period, there are 2×10^5 symbols being transmitted in a static channel). In the initial static period T_1 , the equalization performance of KF-RC is better than that of LS-RC, this is in line with the results shown in Fig. 6. Furthermore, in the periods from T_2 to T_6 , as the channel states are varied periodically, the SERs for both of LS-RC and KF-RC vary accordingly. For the LS-RC structure, due to the lack of adaptability, the SER maintains at a higher level with respect to that in the initial period. While for the KF-RC structure, even though there is a short-term high-SER peak occurring at the sudden channel state switching moment (in which the readout weights have not been updated accordingly), the SER decreases quickly after that. This phenomenon is because after a short transient process, the readout weights in the KF-RC structure are adaptively updated according to the new channel state, in virtue of the high-speed adaptive learning ability of the KF algorithm. It is worth mentioning that in some cases, such as the periods T_3 in Fig. 7(a), T_4 in Fig. 7(b), and T_3 in Fig. 7(d), the SER can be improved to a level better than that of the initial period T_1 , which is due to that in these cases the coefficients for the nonlinear terms in Eqs. (15) and (16) are decreased, as such the channel state is more ideal than that in the initial period. In general, for the parameter-variant channel equalization tasks, the KF-RC structure shows significantly better performance with respect to the LS-RC structure, for its ability of adaptive readout-weights updating.

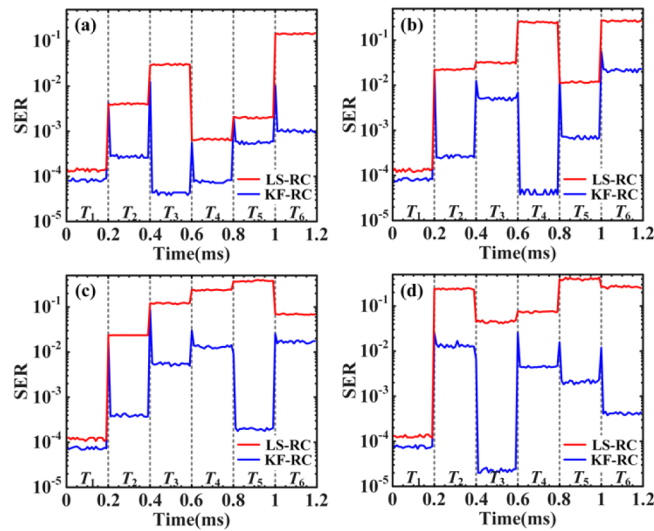


Fig. 7. SERs of the parameter-variant channel equalization task versus SNR by utilizing LS-RC (red curve) and KF-RC (blue curve). The variances of scaling factors are selected as (a) $\sigma^2 = 0.1$, (b) $\sigma^2 = 0.2$, (c) $\sigma^2 = 0.5$, and (d) $\sigma^2 = 1$, respectively.

4. Conclusion

In conclusion, we propose and demonstrate an adaptive time-delayed RC structure by applying a KF algorithm as the training approach. The performance of the proposed KF-RC structure is evaluated by the classic Santa Fe time-series prediction task and nonlinear channel equalization task. It is numerically demonstrated that the proposed KF-RC structure can provide better prediction and equalization performance than the conventional LS-RC system, and moreover, utilizing a bandwidth and complexity enhanced chaotic signal as the mask signal, the performances for prediction and channel equalization can be further improved, since it activates a faster transient response and richer dynamics for the response laser, and then improves the high-dimensional mapping capability of the reservoir. Furthermore, as the readout weights can be updated adaptively according to the variation of channel states, the proposed KF-RC structure is efficient to process the parameter-variant tasks.

Funding. National Natural Science Foundation of China (61671119, 62171087); Sichuan Province Science and Technology Support Program (2021JDJQ0023); Ministry of Education of the People's Republic of China (ZYGX2021K010); Fundamental Research Funds for the Central Universities (ZYGX2019J003); Science and Technology Commission of Shanghai Municipality (SKLSFO2020-05).

Disclosures. The authors declare no conflicts of interest.

Data availability. The data used to support the findings of this study are available from the corresponding author upon request.

References

1. G. Van der Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," *Nanophotonics* **6**(3), 561–576 (2017).
2. D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural Networks* **20**(3), 391–403 (2007).
3. H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science* **304**(5667), 78–80 (2004).
4. W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Computation* **14**(11), 2531–2560 (2002).
5. P. Antonik, N. Marsal, D. Brunner, and D. Rontani, "Human action recognition with a large-scale brain-inspired photonic computer," *Nat.* *Nat Mach Intell* **1**(11), 530–537 (2019).

6. J. Pathak, B. Hunt, M. Girvan, Z. X. Lu, and E. Ott, "Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach," *Phys. Rev. Lett.* **120**(2), 024102 (2018).
7. L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nat. Commun.* **2**(1), 468 (2011).
8. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *Sci. Rep.* **2**(1), 287 (2012).
9. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," *Opt. Express* **20**(3), 3241–3249 (2012).
10. R. Martinenghi, S. Rybalko, M. Jacquot, Y. K. Chembo, and L. Larger, "Photonic nonlinear transient computing with multiple-delay wavelength dynamics," *Phys. Rev. Lett.* **108**(24), 244101 (2012).
11. F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Opt. Express* **20**(20), 22783–22795 (2012).
12. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nat. Commun.* **4**(1), 1364 (2013).
13. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Simultaneous computation of two independent tasks using reservoir computing based on a single photonic nonlinear node with optical feedback," *IEEE Trans. Neural Netw. Learning Syst.* **26**(12), 3301–3307 (2015).
14. X. X. Guo, S. Y. Xiang, Y. Qu, Y. N. Han, A. J. Wen, and Y. Hao, "Enhanced prediction performance of a neuromorphic reservoir computing system using a semiconductor nanolaser with double phase conjugate feedbacks," *J. Lightwave Technol.* **39**(1), 129–135 (2021).
15. D. Z. Yue, Z. M. Wu, Y. S. Hou, B. Cui, Y. H. Jin, M. Dai, and G. Q. Xia, "Performance optimization research of reservoir computing system based on an optical feedback semiconductor laser under electrical information injection," *Opt. Express* **27**(14), 19931–19939 (2019).
16. J. Nakayama, K. Kanno, and A. Uchida, "Laser dynamical reservoir computing with consistency: an approach of a chaotic mask signal," *Opt. Express* **24**(8), 8679–8692 (2016).
17. Y. Kuriki, J. Nakayama, K. Takano, and A. Uchida, "Impact of input mask signals on delay-based photonic reservoir computing with semiconductor lasers," *Opt. Express* **26**(5), 5777–5788 (2018).
18. A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J. L. Oudar, M. Haelterman, and S. Massar, "All-optical reservoir computer based on saturation of absorption," *Opt. Express* **22**(9), 10868–10881 (2014).
19. S. Ortín and L. Pesquera, "Delay-based reservoir computing: tackling performance degradation due to system response time," *Opt. Lett.* **45**(4), 905–908 (2020).
20. C. Sugano, K. Kanno, and A. Uchida, "Reservoir computing using multiple lasers with feedback on a photonic integrated circuit," *IEEE J. Sel. Top. Quantum Electron.* **26**(1), 1–9 (2020).
21. B. Penkovsky, X. Porte, M. Jacquot, L. Larger, and D. Brunner, "Coupled nonlinear delay systems as deep convolutional neural networks," *Phys. Rev. Lett.* **123**(5), 054101 (2019).
22. Y. S. Hou, G. Q. Xia, W. Y. Yang, D. Wang, E. Jayaprasath, Z. F. Jiang, C. X. Hu, and Z. M. Wu, "Prediction performance of reservoir computing system based on a semiconductor laser subject to double optical feedback and optical injection," *Opt. Express* **26**(8), 10211–10219 (2018).
23. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Fast photonic information processing using semiconductor lasers with delayed optical feedback: Role of phase dynamics," *Opt. Express* **22**(7), 8672–8686 (2014).
24. X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "Four-channels reservoir computing based on polarization dynamics in mutually coupled VCSELs system," *Opt. Express* **27**(16), 23293–23306 (2019).
25. S. Y. Xiang, Z. X. Ren, Z. W. Song, 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.* **32**(6), 2494–2505 (2021).
26. Q. Cai, Y. Guo, P. Li, A. Bogris, K. A. Shore, Y. M. Zhang, and Y. C. Wang, "Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing," *Photonics Res.* **9**(1), B1–B8 (2021).
27. R. Lang and K. Kobayashi, "External optical feedback effects on semiconductor injection laser properties," *IEEE J. Quantum Electron.* **16**(3), 347–355 (1980).
28. N. Li, W. Pan, L. Yan, B. Luo, X. Zou, and S. Xiang, "Enhanced two-channel optical chaotic communication using isochronous synchronization," *IEEE J. Sel. Top. Quantum Electron.* **19**(4), 1–9 (2013).
29. R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.* **82**(1), 35–45 (1960).
30. J. P. Nobrega and A. Oliveira, "Kalman filter-based method for online sequential extreme learning machine for regression problems," *Engineering Applications of Artificial Intelligence* **44**(1), 101–110 (2015).
31. A. K. Zhao, N. Jiang, S. Q. Liu, C. P. Xue, and K. Qiu, "Wideband time delay signature-suppressed chaos generation using self-phase-modulated feedback semiconductor laser cascaded with dispersive component," *J. Lightwave Technol.* **37**(19), 5132–5139 (2019).

32. M. F. Cheng, X. J. Gao, L. Deng, L. F. Liu, Y. S. Deng, S. N. Fu, M. M. Zhang, and D. M. Liu, "Time-delay concealment in a three-dimensional electro-optic chaos system," *IEEE Photonics Technol. Lett.* **27**(9), 1030–1033 (2015).
33. X. J. Gao, W. Zhu, Q. Yang, D. Zeng, L. Deng, Q. Chen, and M. F. Cheng, "Time delay estimation from the time series for optical chaos systems using deep learning," *Opt. Express* **29**(5), 7904–7915 (2021).
34. A. K. Zhao, N. Jiang, S. Q. Liu, Y. Q. Zhang, and K. Qiu, "Physical layer encryption for WDM optical communication systems using private chaotic phase scrambling," *J. Lightwave Technol.* **39**(8), 2288–2295 (2021).
35. Y. Huang, P. Zhou, Y. G. Yang, and N. Q. Li, "High-speed photonic reservoir computer based on a delayed Fano laser under electrical modulation," *Opt. Lett.* **46**(24), 6035–6038 (2021).
36. Q. Vinckier, F. Duport, A. Smerieri, K. Vandoorne, P. Bienstman, M. Haelterman, and S. Massar, "High-performance photonic reservoir computer based on a coherently driven passive cavity," *Optica* **2**(5), 438–446 (2015).