Title

Unveiling the Synergistic Effects: A Unified Autonomous Synaptic Development Mechanism for Reservoir Computing

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Zuo, Zhihao
Gan, Zhongxue

Publication Date

2024

Unveiling the Synergistic Effects: A Unified Autonomous Synaptic Development Mechanism for Reservoir Computing

Zhihao Zuo (19110860013@fudan.edu.cn)
Institute of AI and Robotics, Academy for Engineering and Technology, Fudan University
Shanghai 200433, China

Zhongxue Gan * **(ganzhongxue@fudan.edu.cn)**
Institute of AI and Robotics, Academy for Engineering and Technology, Fudan University
Shanghai 200433, China

Abstract

Reservoir computing (RC) offers distinct advantages in extracting spatiotemporal information with low training costs by separating recurrent neural networks into a fixed network with recurrent connections. The quality of the fixed network, known as the reservoir, plays a pivotal role in the performance of the RC system. Our work aims to provide a unified synaptic development framework for RC, constructing a more biologically plausible reservoir to model and understand the neural networks development within the human brain. In this paper, we propose an Autonomous Synaptic Development Reservoir Computing model (ASD-RC) based on an adaptive network of phase oscillators. The reservoir autonomously adjusts the distribution of connection weights in response to external stimuli, forming a task-specific structure. Through experiments and theoretical analyses, we demonstrate that ASD-RC can emulate various synaptic development rules of biological neural networks in *vivo*, including the Hebbian rule and STDP. Furthermore, experiments reveal that combining different development rules can enhance performance on prediction tasks compared to using a single development rule, showcasing the emergence and effects of synergistic development that improve information processing capacity.

Keywords: reservoir computing; autonomous synaptic development; intelligent emergence; synergistic effects

Introduction

In recent decades, researchers have strived to understand the mechanisms behind intelligence emergence in the human brain, contributing to artificial intelligence (AI) development (Feulner et al., 2021; Fusi, Miller, & Rigotti, 2016). Since the introduction of the Hopfield network in 1982, the artificial neural networks (ANN) community has made significant progress in emulating human brain functionalities (Hopfield, 1982). Recurrent neural networks (RNNs), mimicking the feedback signals between neurons, have become a popular architecture for processing temporal data like time-series prediction and speech recognition (Mandic & Chambers, 2001). However, recurrent connections in RNNs present challenges such as high computational cost, vanishing gradient, or exploding grading (Morales, Mirasso, & Soriano, 2021).

Reservoir computing (RC), a novel RNN approach, offers fast and computational efficiency by training only at the linear output layer (Jaeger & Haas, 2004; Lukoeviius & Jaeger, 2009). The recurrent networks in RC, called ”reservoirs,” are randomly initialized and fixed during the training process. The reservoirs provides advantages in extracting spatiotemporal information across various applications (Kong, Weng, Glaz, Haile, & Lai, 2023; Chen et al., 2023). Conventional reservoirs typically employ fixed random connections, a key advantage of RC (Schrauwen, Verstraeten, & Campenhout, 2007a). However, recent evidence suggests reservoirs operate best ”at the edge of chaos” (Lukoeviius, Jaeger, & Schrauwen, 2012; Ferreira, Ludermir, & de Aquino, 2013). Randomly achieving optimal performance in the RC error landscape is practically zero, making the specification of a reservoir a challenging task requiring numerous trials and even luck (Schäfer & Zimmermann, 2006; Moon, Wu, & Lu, 2021).

While machine learning (ML) algorithms enhance reservoir performance for specific applications, they introduce inherent drawbacks of ML algorithm, such as being black box models, high computation costs and slow convergence (Lukoeviius et al., 2012). Biologically inspired methods, like the Hebbian rule and Intrinsic Plasticity, have been proposed for RC (Chatzidimitriou & Mitkas, 2013; Triesch, 2005). Nevertheless, the application of Hebbian rules to RC does not lead to a significant improvement in performance (Yao & Wang, 2019). In practice, integrating these biological methods into RC is laborious due to differences in network structure and neuron models.

How to construct a more biologically plausible RC for modeling and understanding neural networks development within human brain remains an open question. The human brain, housing approximately 86 billion neurons, exhibits a specific structure crucial for information processing (Song, Miller, & Abbott, 2000). At the microscopic level, brain synapses possess activity-dependent plasticity, leading to significant changes in neuronal connectivity patterns (Choi & Kim, 2019). On a macroscopic scale, structural changes in the brain follow nonlinear trajectories as neurons develop throughout life (Bethlehem et al., 2021). The evolution of network topology is heavily influenced by the states of elements in many biological networks, and vice versa. Adaptive networks, incorporating synaptic plasticity, play a fundamental role in learning, memory, and neural network development (Seoane, 2018). In these co-evolving dynamic systems, coupling weights between nodes and the states of active elements interact and evolve together (Aoki & Aoyagi, 2009).

Motivated by these biological and physical findings, we propose a unified synaptic development mechanism for RC called Autonomous Synaptic Development Reservoir Com-

---

*Corresponding author

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

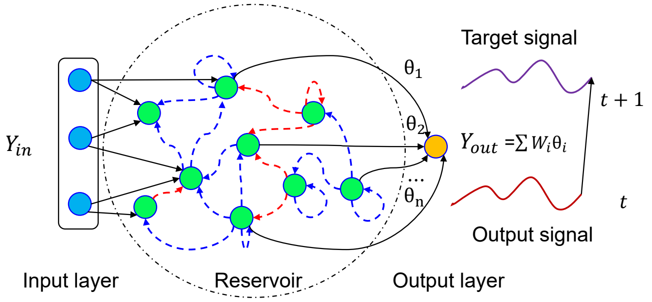


Figure 1: Configuration of ASD-RC: The dotted lines within the reservoir signify plastic connections, with red and blue colors denoting excitatory (positive) and inhibitory (negative) synapses, respectively.

puting (ASD-RC) to incorporate structural and connectivity plasticity. During the development stage, the dynamics of neurons and synapses in the reservoir continuously evolve in response to external inputs. ASD mechanism enables the reservoir to adapt its internal dynamics to specific tasks, resulting in the development of a customized structure that combines physics and information processing (Sung, Kim, Shin, Im, & Lee, 2022). To the best of our knowledge, most current studies focus on the evolution of networks from generation to generation, while the development of reservoir structures with external stimuli has not received sufficient attention.

In this paper, we detail ASD-RC construction and the algorithm in the Methods section. Then, we test ASD-RC performance in chaotic time series forecasting, demonstrating compatibility with biological neural networks. Furthermore, we explore the combination of different development rules, revealing the emergence and effects of multi-rule synergistic development. We shift focus from the finest-performance parameter, concentrating instead on understanding how ASD fosters the reservoir to reshape more biologically plausible synaptic connections and boost its overall performance.

Methods

We investigate an adaptive network comprising N oscillators influenced by external inputs, where the coupling weights among nodes and the active states of the elements at the nodes coevolve as Fig. 1 shows. When data is introduced into the reservoir, the reservoir nodes extract and transform information from a low-dimensional input space to a high-dimensional feature space. Concurrently, the connections autonomously adjust their coupling strength through the autonomous development mechanism. The core concept of the ASD-RC lies in the assumption that, given sufficient information for accurate computation, the reservoir state trajectories should converge under the impact of input data. Moreover, the internal connection structure of the reservoir can evolve into a task-specific network guided by the ASD mechanism.

Overall Architecture

We begin by considering a network of N Kuramoto-like phase oscillators. The dynamics of each oscillator can be described by a phase $\theta_i(t) \in [0, 2\pi)$ given by the equation:

$$\frac{d\theta_i}{dt} = \omega_i + \lambda \sum_{j=1}^{N} k_{ij} \sin(\theta_j - \theta_i + y(t)), \tag{1}$$

here, ω_i represents the natural frequency of oscillator i, $\lambda > 0$ is the global coupling strength, k_{ij} is the entry of the adjacency matrix K of the network, and $y(t)$ is the input data. We employ this extended Kuramoto model to simulate neuronal behavior in the reservoir.

Next, we propose a dynamical adaptive model for the coupling weights k_{ij}. Since the dynamics of the weights only depend on the relative timing of the oscillators, we introduce the following reasonable development function:

$$\frac{dk_{ij}}{dt} = -\varepsilon \sin(\theta_j - \theta_i + \beta), \quad |k_{ij}| \leq 1. \tag{2}$$

The parameter ε represents the time scale of this dynamics, which is much longer than that of the oscillator dynamics ($i.e.$, $\varepsilon \ll 1$). The constraint condition $|k_{ij}| \leq 1$ ensures that if k_{ij} exceeds the interval $[-1, 1]$, it is immediately set to the appropriate limiting value (-1 or 1). This limitation is reasonable as coupling weights cannot grow indefinitely.The development of coupling weights is governed by the function dictating how changes in weights correlate with phase differences among oscillators. We designate β as the character parameter, as it regulates the attributes of the development function—a topic we delve into with further scrutiny later.

The ASD-RC Algorithm

In conventional training methods, the reservoir undergoes random initialization, and the initial 100 to 200 steps of neuronal states are typically discarded to reach a stable state, known as the washout phase. However, In many scenarios, data is too expensive to get more, discarding data entails inefficiency. To optimize data utilization, we propose that the neural network can adapt and modify the distribution of synaptic weights with external stimulation during the washout phase, ultimately shaping a task-specific structure. For conciseness, we denote the lengths of the autonomous development, training, and test sequences as L_{adev}, L_{train}, and L_{test}, respectively. In this context, $\hat{y}(t)$ signifies the output at time t, and W_{out} denotes the reservoir-to-output weight matrix.

Algorithm 1 outlines the procedure of the ASD mechanism employed by the RC. Initially, we initialize the ASD-RC network with a random adjacency matrix. Subsequently, L_{train} values are input into the reservoir. The first L_{adev} input data is provided to the reservoir to aid in the development of a specific structure. Neuron synapses autonomously adjust their synaptic weights distribution according to the ASD mechanism. Then, the remaining values are input into the reservoir

Algorithm 1: The ASD-RC Algorithm

```
Data: Time series $y(t), t = (1, 2, \dots, T)$, initial
      adjacency matrix $K$, initial neuronal states $\Theta$,
      spectral radius $\rho$
Result: Prediction time series $\hat{y}(t)$
while $t \leq L_{train}$ do
    Update the neuronal states $\Theta$ according to $Eq.1$:
      $\Theta \leftarrow \Omega + \lambda K sin(\Delta\Theta + y(t))$;
    if $t < L_{adev}$ then
        Update adjacency matrix $K$ according to $Eq.2$:
          $K \leftarrow K - \varepsilon sin(\Delta\Theta + \beta)$;
    else
        Collect the neuronal states: $\Sigma(t - L_{adev}) \leftarrow \Theta$;
    end
end
Compute the output weight $W_{out}$ by ridge regression:
  $W_{out} \leftarrow Y \setminus \Sigma$;
while $t \leq L_{test}$ do
    Collect prediction value according to $W_{out}$:
      $\hat{y}(t) \leftarrow W_{out}\Theta$;
    Update the neuronal states $\Theta$ according to $Eq.1$:
      $\Theta \leftarrow \Omega + \lambda K sin(\Delta\Theta + \hat{y}(t))$;
end
```

to train the output weights for the prediction task. To assess the performance of the developed reservoir, we initially feed the ASD-RC with the last input of the training set $y(t)$ and then run the network for a number of L_{test} steps, using the predicted output at time t as the next input at time $t+1$ (i.e., $y(t+1) = \hat{y}(t)$). This way, the testing phase is carried out in generative mode with output feedback.

Experiments

In this study, we evaluate the performance of ASD-RC in chaotic time series forecasting using the Mackey-Glass system (MG). Through experiments and theoretical analyses, we demonstrate the model's compatibility with biological neural networks in *vivo*. Additionally, We explore the influence of integrating multiple development rules. By analyzing the information processing capacity of developed RC, we unveil emergence and effects of multi-rule synergistic development.

Experimental Settings

We commence with a single-layer reservoir comprising a fixed number of neuronal nodes, denoted as $m = 100$, and an initialized spectral radius[1] for each model set to $\rho = 0.95$. Our initial evaluation focuses on assessing ASD-RC performance in time series prediction tasks under the following configurations. We choose $\varepsilon \in= 0.1$ and initialize $\Theta = [\theta_1, \theta_2, \dots, \theta_m] = [\vec{\mathbf{0}}]$. The matrix $\Omega = [\omega_1, \omega_2, \cdots, \omega_m]$ is sampled from a standard normal distribution. The density s of the adjacency matrix K is set to 0.05, indicating that only 5% of the entries in K are non-zero, while the rest are set to zero. In this study, we consider weighted connections, where $k_{ij} \in [-1, 1]$, rather than a binary $0-1$ network. The non-zero entries k_{ij} in K are uniformly initialized from the interval $[-1, 1]$. Equations 1 and 2 are solved using the Euler method with a time step of $\Delta t = 0.1ms$.

Compatibility with Biological Neural Network

The character parameter β is thoroughly investigated to comprehend its impact on the ASD-RC's performance. The reservoir connections are initially randomized and then evolve into a task-specific structure through the application of the ASD mechanism. Fig. 2 visually represents the structures of the developed reservoir networks under different values of the character parameter β. Notably, the reservoir structures themselves exhibit asymmetry concerning β. The initial reservoir shown in Fig. 2a undergoes continuous development, guided by different character parameters ($-\pi/2$, 0, and $\pi/2$), along with constant external input stimulation. The corresponding developed networks are depicted in Fig. 2b, c, and d. Simulations reveal substantial adjustments in synaptic weights within the first 3-5 time steps. Despite a 100-step development process, the distribution of synaptic weights stabilizes after the initial 10 steps, with subsequent adjustments becoming imperceptible. This observation confirms the robust convergence property of ASD mechanism.

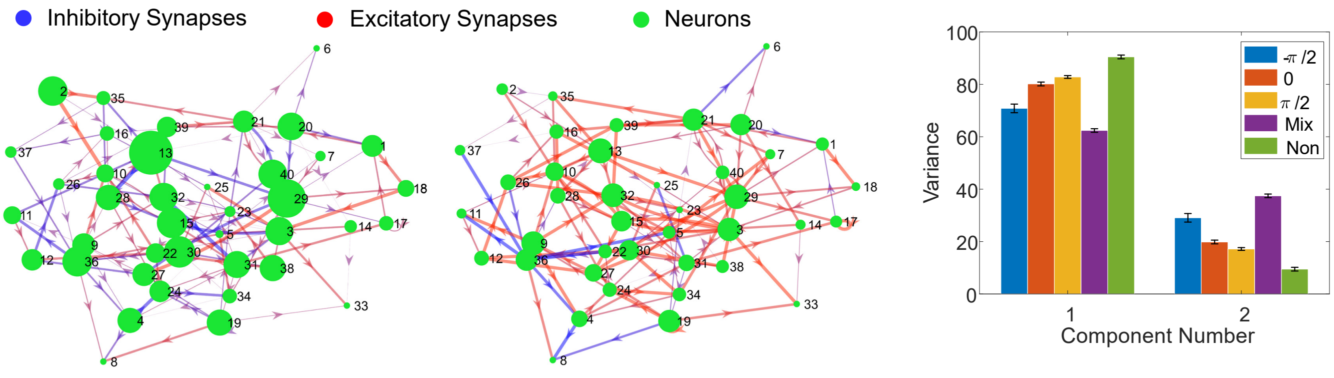
In the proximity of $\beta = -\pi/2$, Equation 2 approximates $cos(\Delta\theta)$, where $\Delta\theta = \theta_j - \theta_i$. This leads to an increase (decrease) in the coupling weight when the phase difference between two oscillators is small (large). Essentially, the synaptic development in the reservoir networks adheres to a like-and-like (different-and-different) rule, reminiscent of Hebbian learning (Mikkelsen, Imparato, & Torcini, 2014). Consequently, synaptic weights converge to a stable two-cluster state, with positive weights tending to align in the positive cluster when positive global coupling weights are present, and negative weights aligning in the negative cluster. Conversely, for $\beta = \pi/2$, synaptic weights tend to reside more in the negative cluster. At $\beta = 0$, Equation 2 approximates $-sin(\Delta\theta)$, leading to opposite changes in the two coupling weights for $\Delta\theta = \pm\Delta\theta^*$. This behavior reflects similarities to spike-timing-dependent plasticity (STDP) rule (Zhang, Boccaletti, Guan, & Liu, 2015). Due to the sinusoidal nature of the function in Equation 2, synaptic weights acquire opposite signs with the same strength, resulting in the convergence of weights to a stable symmetric two-cluster state.

Comparing Synaptic Weights Distribution to *in vivo*

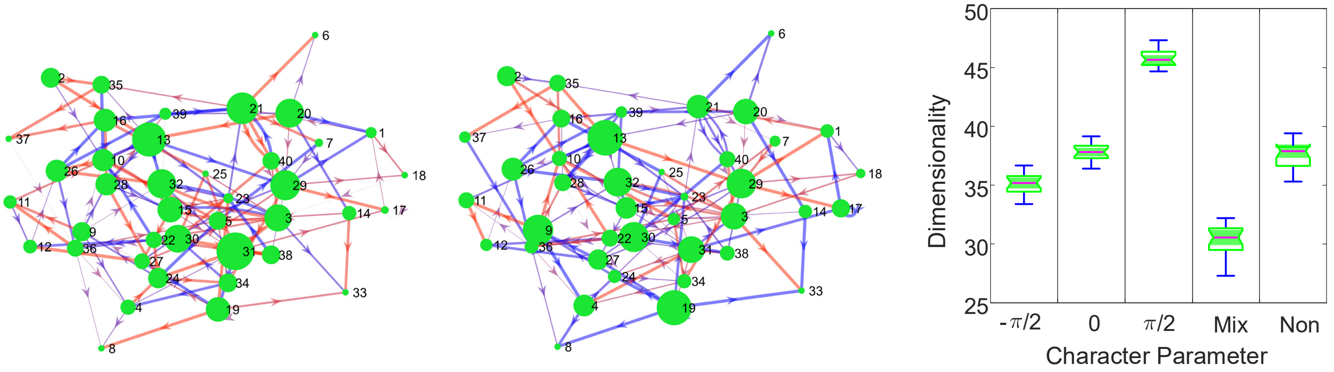
Theoretical analysis confirms that the ASD-RC can effectively simulate various development rules observed in biological neural networks. Furthermore, we conduct a comparative analysis of the synaptic weight distributions between the developed reservoir networks and those observed in *vivo*. The histograms illustrating the synaptic weights in reservoir networks and *vivo* are presented in Fig. 3. We initialize the reservoir networks with extreme synaptic weight distributions

---

[1] The maximum absolute eigenvalue of the reservoir matrix.



(a) Initial Network (b) β = -π/2 (e) PCA of Network Dynamics



(c) β = 0 (d) β = π/2 (f) Dimensionality of Network Structure

Figure 2: The reservoir structure is influenced by the character parameter. (a) Initial network; Developed network with (b) $\beta = -\pi/2$, (c) $\beta = 0$, and (d) $\beta = \pi/2$. Excitatory (positive) synapses are represented in red, while inhibitory (negative) synapses are in blue. The width of synapses corresponds to the absolute value of connection weights. Neuron size is proportional to their degree. To assess the impact of the autonomous development mechanism, (e) PCA is performed on the neuronal states set Θ, and (f) the dimensionality of the adjacency matrix K is analyzed with different character parameter settings. Averages were computed over 20 independent realizations for each character parameter.

(see Fig. 3a(1)-a(6)). The initial synaptic weights are drawn from a *Beta* distribution $B(a,b)$ with different combinations of parameters a and b. Notably, regardless of the initial distributions, the synaptic weights of reservoir networks consistently converge to a highly similar specific distribution for a given character parameter as shown in Fig. 3b(1)-b(3), validating the robust convergence property of the model. Moreover, the distributions in Fig. 3c(1) and c(3) are consistent with Hebbian rule, and c(2) is consistent with the STDP rule (Mikkelsen et al., 2014). The distribution of synaptic weights in the developed reservoir networks exhibits a striking resemblance to synaptic weights found in *vivo*, providing further evidence that our model effectively simulates the developmental processes of biological neural networks in *vivo*.

Performance in Prediction Tasks

Both theoretical analysis and experimental simulations demonstrate that the ASD-RC model is highly biologically plausible. To assess the information processing capability of the developed reservoir networks, we applied the model to time series prediction tasks. We initiated by comparing the performance of the reservoir with different development rules, specifically by initializing β with different values while keeping the other hyperparameters constant. Subsequently, we analyzed the effects of the ASD mechanism on the properties of the RC system.

Mackey-Glass System The MG time series serves as a prevalent benchmark in the literature to evaluate models' performance in dynamical system identification(Schrauwen, Verstraeten, & Campenhout, 2007b). It originates from a time-delay differential system represented by the following equation:

$$\frac{dy(t)}{dt} = \frac{ay(t-\tau)}{1+y^n(t-\tau)} + by(t). \tag{3}$$

In this equation, $y(t)$ denotes the output at time step t, and τ signifies the time delay. The parameters of the equation are specified as $n = 10$ and $b = -0.1$. Chaotic behavior in the system emerges when the value of τ surpasses 16.8. For this study, we generate a training set with $\tau = 17$ (MG-17), frequently employed as an illustration of chaotic time series. The fourth-order Runge-Kutta method is applied for solving

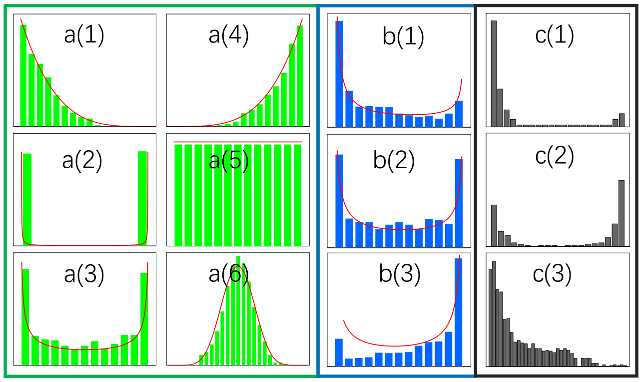


Figure 3: Synaptic weights distribution in reservoir networks. a(1)-a(6): Initial synaptic distributions for reservoir networks generated from the Beta distribution. b(1)-b(3): Synaptic weight distributions in developed reservoir networks controlled by character parameter b(1) $\beta = -\pi/2$, b(2) $\beta = 0$, and b(3) $\beta = \pi/2$, respectively. These distributions are fitted by the Beta distribution (shown in red curve). c(1)-c(3): Synaptic weights of biological neural networks in *vivo*.

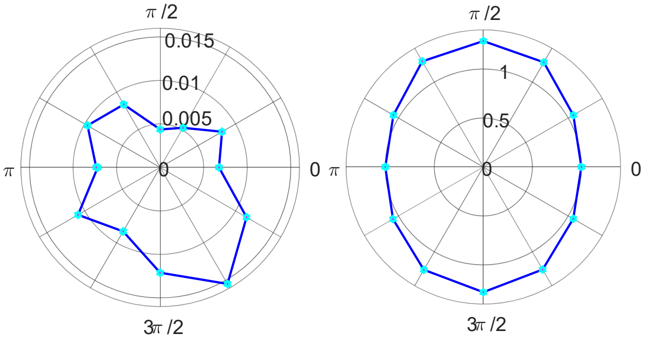
the equation with random initialization. applied for solving the equation with random initialization.

The root mean square error (RMSE) is employed to evaluate the performance of the ASD-RC:

$$RMSE = \sqrt{\frac{1}{L_{test}} \sum_{i=1}^{L_{test}} \|y(t_i) - \hat{y}(t_i)\|^2}, \tag{4}$$

where $\hat{y}(t_i)$ represents the readout output (generation), $y(t_i)$ represents the desired output (target), and $\|\bullet\|$ denotes the Euclidean distance (or norm).

Influence of Development Duration Based on above analysis, we emulate different developmental rules observed in biological neural networks in *vivo* by setting the character parameter β to $-\pi/2$, 0, and $\pi/2$ to correspond to Hebbian rule, STDP, and anti-Hebbian rule, respectively. Additionally, considering that different developmental rules can be easily applied to the model, unlike setting an identical value for the character parameter β in the entire reservoir as in the previous case, we set a unique value for β_i of each neuron i ($i = 1, 2, \ldots, m$) in the reservoir. β_i is sampled from a uniform distribution in the interval $[-\pi, \pi]$. We term the latter as 'mixing mode' since various synaptic development rules are applied simultaneously to the reservoir. In Fig. 5a, it is evident that the generation signal closely aligns with the target signal, indicating the ASD-RC's effectiveness in handling time series dynamics. Over time, the error between the generation and target signals gradually increases. This phenomenon is reasonable, given the absence of real signal input and the reservoir's limited memory capacity, causing the time-series dynamics to be gradually forgotten by the reservoir.



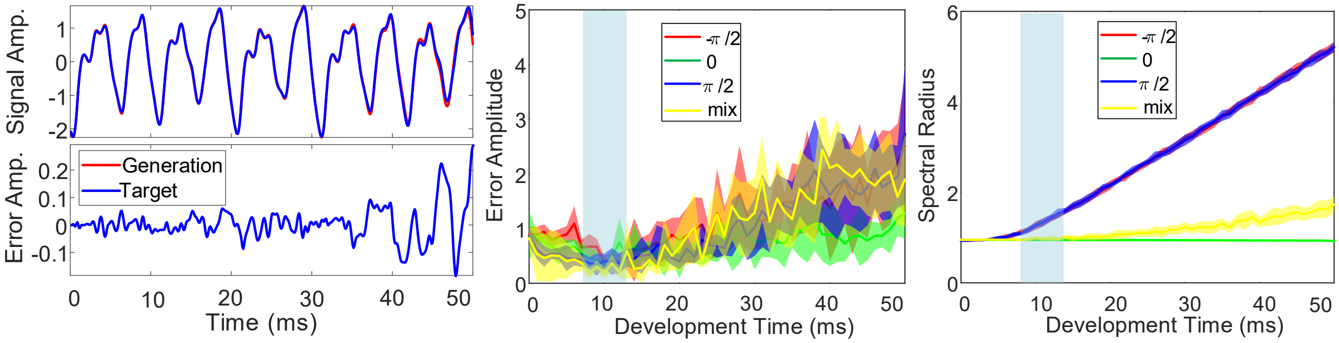
(a) Prediction Performance (b) Spectral Radius

Figure 4: The influence of the character parameter on the reservoir performance. The polar axis represents (a) the prediction performance RMSE of ASD-RC and (b) the corresponding spectral radius of the reservoir.

We applied the ASD-RC with four character parameter settings to the prediction task to compare performance differences. In Fig. 5b, we observe that prediction errors initially decrease, then dramatically increase over the development duration. All four error curves reach their lowest points in the vicinity of 10ms, indicating the optimal development duration for various autonomous synaptic development rules to exhibit the best performance. Correspondingly, as seen in Fig. 5c, the spectral radius of the reservoir with character parameters $-\pi/2$ and $\pi/2$ initially remains stable, then increases sharply. In contrast, the curve of the 'mixing mode' constantly grows mildly, and the curve of $\beta = 0$ remains relatively stable. The spectral radius less than one has often been regarded as the source of the 'echo state property' in RC. Nevertheless, subsequent studies have shown that the echo state property can actually be maintained over a unity spectral radius, which could explain why we find optimal performance slightly above $\rho = 1$ for $\beta = -\pi/2$ and $\beta = \pi/2$ in the vicinity of a 10ms development duration.

Character Parameter To investigate the impact of the character parameter, we varied it from 0 to 2π as a polar angle in polar coordinates, as shown in Fig. 4. We initialized the spectral radius $\rho = 1.0$, and the development duration $L_{adv} = 10ms$ for the ASD-RC. We observe that although the curves of the prediction error and spectral radius are symmetrical about the character parameter in the polar coordinates, they have different symmetry axes. This indicates that the spectral radius is not the sole parameter affecting the performance of the ASD-RC. This finding coincides with the results in Fig. 5, where the model with different character parameter settings shows optimal performance at different spectral radii.

Exploring the Emergence and Effects of Multi-rule Synergistic Development in Reservoir While ASD mechanism can enhance reservoir performance with an appropriate development duration, distinct development rules exhibit vary-



(a) Time Series Prediction (b) Prediction Performance (c) Spectral Radius

Figure 5: The performance of the ASD-RC in time series prediction is influenced by the development duration. (a) Upper: The generated signal (red) by ASD-RC closely aligns with target signal (blue); Lower: Error between generated and target signals. (b) Four ASD-RC models with character parameters $-\pi/2$, 0, $\pi/2$, and ‘mixing mode’ were set. Averages were computed at each development duration over 20 independent realizations of each model. (c) Spectral radius corresponding to (b).

ing effects. Notably, the ‘mixing mode’ outperforms others (see Fig. 5b), based on multi-rule synergistic development. We aimed to understand the impact of the ASD mechanism on the reservoir’s information processing capacity. Specifically, we explored how different development rules, based on changes in connection weights within the reservoir, alter network activity compared to activity without development.

In Fig. 2f, we illustrate the dimensionality of the connection matrix in the developed network with different character parameter settings and the non-development condition (‘Non’). The ‘mixing mode’ exhibits significantly lower dimensionality compared to the non-development, whereas character parameter $\beta = -\pi/2$ shows higher dimensionality. The $\beta = 0$ setting shows a smaller variation with an equivalent median value. Additionally, we estimate the dimensionality of network activity space using techniques related to Principal Component Analysis (PCA). In Fig. 2e, the first principal component shows a similar trend to Fig. 2f. The variance primarily lies in the first principal component for the non-development, indicating that network activity mainly resides in a low-dimensional dynamic space. In contrast, the dynamic space of the ‘mixing mode’ enjoys higher dimensionality.

High-dimensional neural representations are crucial for the reservoir, as they relate to richer behavior and better working memory (Fusi et al., 2016). However, high-dimensional representations are not always desirable, as a higher-dimensional space is more likely to amplify the influence of noise, resulting in unstable performance. For example, the brain needs to reduce dimensionality to eliminate irrelevant factors and recast the remaining factors into a high-dimensional space for processing, generating complex behavior. Therefore, the emergence and effects of multi-rule synergistic development may be partially attributed to the ASD mechanism placing the network structure in a lower-dimensional space to filter noise, while network dynamics evolve in a higher-dimensional space, allowing for more complex behavior.

Conclusion

We introduce an Autonomous Synaptic Development mechanism for Reservoir Computing using an adaptive network of phase oscillators. At the micro level, synapses autonomously adjust their weights distribution. At the mesoscopic level, the neural networks within the reservoir continuously evolve with external stimuli. Guided by the ASD mechanism, the reservoir forms a task-specific neural network. ASD-RC offers a novel approach to construct a more biologically plausible reservoir, aiming to model and understand the development of neural networks in the human brain.

Establishing a Unified Synaptic Development Framework for RC ASD-RC showcases remarkable compatibility with biological neural networks in *vivo*, establishing a unified synaptic development framework for RC. The reservoir autonomously evolves and adapts the distribution of synaptic weights under external stimulation. By manipulating the character parameter, diverse development rules can be effortlessly applied to RC.

Shedding Light on Multi-Rule Synergistic Development Furthermore, experiment results reveal that the reservoir’s information processing capability is significantly enhanced through multi-rule synergistic development, as opposed to a single development rule. Our work unveils the emergence and effects of synergistic development, showcasing the benefits of a collaborative application of various development rules.

Bridging the Gap Between Nonlinear Dynamics and Neuroscience Communities Built upon the Kuramoto-like oscillator system, ASD-RC integrates the strengths of neuromorphic computing and nonlinear dynamics, effectively emulating the characteristics and functionalities of the brain. The ASD mechanism facilitates the easy application of various development rules from the neuroscience community to RC. Additionally, it provides an avenue for exploring undiscovered neural network development rules in *vivo*.

Acknowledgments

This work was supported in part by Shanghai Municipal Science and Technology Major Project No.2021SHZDZX0103, in part by the Shanghai Engineering Research Center of AI & Robotics, Fudan University, China, and in part by the Engineering Research Center of AI & Robotics, Ministry of Education, China.

References

Aoki, T., & Aoyagi, T. (2009). Co-evolution of phases and connection strengths in a network of phase oscillators. *Physical review letters*, *102 3*, 034101.

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C. L., ... Alexander-Bloch, A. F. (2021). Brain charts for the human lifespan. *Nature*, *604*, 525 - 533.

Chatzidimitriou, K. C., & Mitkas, P. A. (2013). Adaptive reservoir computing through evolution and learning. *Neurocomputing*, *103*, 198-209.

Chen, Z., Li, W., Fan, Z., Dong, S., Chen, Y., Qin, M., ... Liu, J.-M. (2023). All-ferroelectric implementation of reservoir computing. *Nature Communications*, *14*.

Choi, J., & Kim, P. (2019, apr). Critical neuromorphic computing based on explosive synchronization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(4), 043110.

Ferreira, A. A., Ludermir, T. B., & de Aquino, R. R. (2013). An approach to reservoir computing design and training. *Expert Systems with Applications*, *40*(10), 4172-4182.

Feulner, B., Perich, M. G., Chowdhury, R. H., Miller, L. E., Gallego, J. A., & Clopath, C. (2021). Small, correlated changes in synaptic connectivity may facilitate rapid motor learning. *Nature Communications*, *13*.

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66-74.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79 8*, 2554-8.

Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, *304*, 78 - 80.

Kong, L.-W., Weng, Y., Glaz, B., Haile, M. A., & Lai, Y.-C. (2023). Reservoir computing as digital twins for nonlinear dynamical systems. *Chaos*, *33 3*, 033111.

Lukoeviius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.*, *3*, 127-149.

Lukoeviius, M., Jaeger, H., & Schrauwen, B. (2012). Reservoir computing trends. *KI - Künstliche Intelligenz*, *26*, 365-371.

Mandic, D. P., & Chambers, J. A. (2001). Recurrent neural networks for prediction: Learning algorithms, architectures and stability. In *Recurrent neural networks for prediction* (p. 69-89). Hillsdale, NJ: John Wiley & Sons, Ltd.

Mikkelsen, K., Imparato, A., & Torcini, A. (2014, Jun). Sisyphus effect in pulse-coupled excitatory neural networks with spike-timing-dependent plasticity. *Phys. Rev. E*, *89*, 062701.

Moon, J., Wu, Y., & Lu, W. D. (2021). Hierarchical architectures in reservoir computing systems. *Neuromorphic Computing and Engineering*, *1*.

Morales, G. B., Mirasso, C. R., & Soriano, M. C. (2021). Unveiling the role of plasticity rules in reservoir computing. *Neurocomputing*, *461*, 705-715.

Schäfer, A. M., & Zimmermann, H. G. (2006). Recurrent neural networks are universal approximators. In *Artificial neural networks – icann 2006* (pp. 632–640). Berlin, Heidelberg: Springer Berlin Heidelberg.

Schrauwen, B., Verstraeten, D., & Campenhout, J. (2007b, 01). An overview of reservoir computing: Theory, applications and implementations. In (p. 471-482).

Schrauwen, B., Verstraeten, D., & Campenhout, J. M. V. (2007a, 01). An overview of reservoir computing: theory, applications and implementations. In *The european symposium on artificial neural networks* (p. 471-482).

Seoane, L. F. (2018). Evolutionary aspects of reservoir computing. *Philosophical Transactions of the Royal Society B*, *374*.

Song, S., Miller, K., & Abbott, L. (2000, 10). Competitive hebbian learning through spike timing-dependent plasticity. *Nature neuroscience*, *3*, 919-26.

Sung, S. H., Kim, T. J., Shin, H., Im, T. H., & Lee, K. J. (2022). Simultaneous emulation of synaptic and intrinsic plasticity using a memristive synapse. *Nature Communications*, *13*.

Triesch, J. (2005). A gradient rule for the plasticity of a neuron's intrinsic excitability. In *International conference on artificial neural networks* (Vol. 3696). Berlin, Heidelberg: Springer.

Yao, X., & Wang, Z. (2019). An intelligent interconnected network with multiple reservoir computing. *Applied Soft Computing*, *78*, 286-295.

Zhang, X., Boccaletti, S., Guan, S., & Liu, Z. (2015, Jan). Explosive synchronization in adaptive and multilayer networks. *Phys. Rev. Lett.*, *114*, 038701.