Reinforcement Learning for Synchronization Analysis of Biological Rhythms

Ojashwini Dubey
AI Research Centre, Woxsen University
ojashwini.dubey_2028@woxsen.edu.in

Palak Dwivedi
AI Research Centre, Woxsen University
palak.dwivedi_2028@woxsen.edu.in

Tiasha Roy
AI Research Centre, Woxsen University
tiasha.roy_2028@woxsen.edu.in

Manjeet Rege
Department of Software Engineering and Data Science, University of St. Thomas, USA
rege@stthomas.edu

Shyam Krishan Joshi
AI Research Centre, Woxsen University
shyam.krishan@woxsen.edu.in

Hemachandran K
AI Research Centre, Woxsen University
hemachandran.k@woxsen.edu.in

Raul Villamarin Rodriguez
AI Research Centre, Woxsen University
raul.rodriguez@woxsen.edu.in

***Abstract*—Traditional synchronization of nonlinear biological oscillators often depends on fixed or manually tuned coupling parameters and this tends to be less adaptive when the physiological environment keeps changing over time. In contrast, existing reinforcement learning (RL) methods like Q-learning usually work with discrete action spaces, so they create kind of stepwise changes which can feel biologically unrealistic. Here we suggest a continuous-action Deep Deterministic Policy Gradient (DDPG) scheme, meant for adaptive identification and also control of a time-varying coupling strength in mixed or heterogeneous oscillator networks. In the setup , each oscillator is treated as a nonlinear dynamical system, while the states are linked through a single scalar, time dependent term $k(t)$. That coefficient is updated in real time by the DDPG agent, so the coupling can track the changes instead of staying frozen. Our aim is to optimize two goals at the same time: minimize Mean Squared Error (MSE) and maximize the Phase-Locking Value (PLV) which together help lock the phases and keep a stable synchronization regime. We test the approach using five well known biological oscillator models: Circadian (Goodwin), Glycolytic (Sel'kov), Intracellular Calcium, NF-κB, and p53–Mdm2. The simulations show good phase entrainment online, plus data-driven discovery of the coupling term across all five systems. In terms of use cases, it could support cardiac rhythm regulation, neural coordination, and metabolic pathway organization.**

***Index Terms*—biological oscillators, synchronization, reinforcement learning, deep deterministic policy gradient, adaptive coupling, phase-locking value, physiological rhythms, biological stability**

I. Introduction

Biological systems demonstrate synchronized oscillatory patterns which extend their rhythmical movements to both daily time periods and various body parts [1]–[3]. The physiological rhythms of humans develop from the combined actions of multiple nonlinear oscillators whose synchronized movements maintain essential biological processes. Glass demonstrated that biological systems use physiological oscillator synchronization as a core mechanism which helps different subsystems work together to maintain balance throughout various organizational levels of the body [4]. The inability to maintain phase synchronization results in medical conditions which include cardiac arrhythmias and epilepsy and immune system problems and circadian rhythm disturbances according to research findings [5], [6]. The study of synchronization in systems with interconnected oscillators relies on the Kuramoto model as one of its most common mathematical frameworks [7]. The system describes how oscillators reach phase alignment through their interactions which now includes stochastic effects and adaptive coupling and complex network structures for more accurate modeling of actual systems. However, classical synchronization models typically assume fixed coupling parameters and simplified phase interactions which limits their ability to represent real physiological systems where coupling strengths vary over time due to biochemical signaling, environmental disturbances, and internal feedback mechanisms [8], [9]. While these approaches show promising results, many existing methods rely on discrete action spaces that introduce discontinuities when applied to continuously evolving oscillator dynamics, making them unsuitable for modeling biologically realistic adaptive coupling. Therefore, there remains a need for a learning based framework that can model continuous, time-varying coupling between nonlinear oscillators. In this work, we propose a continuous reinforcement learning approach for adaptive synchronization in networks of heterogeneous oscillators, enabling smooth control of coupling dynamics and providing a more realistic representation of synchronization in biological systems.

A. Limitations of Existing Approaches

Despite recent RL progress [10]–[12], discrete action frameworks such as Q-learning introduce key limitations in biological oscillator control:

(i) Stepwise control discontinuities: Discrete action spaces produce abrupt parameter updates that interrupt the smooth evolution of oscillator dynamics.

(ii) Biologically unrealistic behavior: Physiological coupling strengths vary continuously, whereas discretized policies restrict control to fixed parameter values.

(iii) Approximation error: Discretization introduces quantization errors in the regulation of continuous-valued coupling parameters.

These limitations create a mismatch between discrete RL frameworks and the continuous dynamics governing biological oscillators.

B. Motivation

Existing synchronization control methods do not provide a unified framework capable of simultaneously:

i Adaptively estimating coupling strength in real time.

ii Maintaining smooth and stable synchronization in nonlinear oscillator networks.

iii Generalizing across heterogeneous oscillators with distinct dynamical structures.

This is critical as maladaptive synchronization underlies cardiac arrhythmias, epileptic seizures, Parkinsonian tremors, and immune dysregulation [5], [6].

C. Novel Contributions

The primary contributions of this work are summarized as follows:

i) Continuous-action DDPG framework: This work presents a continuous reinforcement learning framework which uses Deep Deterministic Policy Gradient (DDPG) to achieve synchronization between different biological oscillators.

ii) Autonomous coupling identification: The learning agent automatically discovers optimal coupling strengths without the need for manual parameter adjustments.

iii) Star-topology oscillator architecture: The system uses a calcium oscillator as its core synchronization element within a biologically inspired network design.

iv) Dual-objective reward formulation: The reward function

$$r_t = \exp(-\text{MSE}_t) + PV_t \tag{1}$$

jointly optimizes amplitude accuracy and phase coherence.

v) *Empirical convergence characterization:* Simulation results demonstrating observed exponential decay of synchronization error $e_i(t)$ across all five oscillator models under the learned DDPG policy, consistent with the error dynamics.

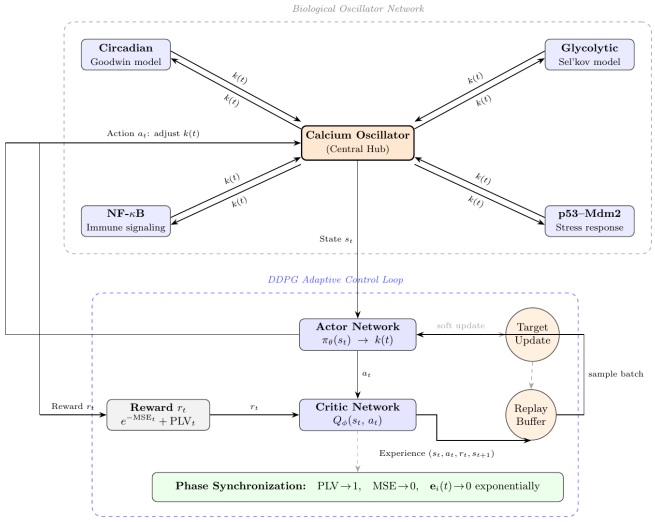

Fig. 1. Workflow of the DDPG-based adaptive synchronization controller

II. Problem Statement

Biological oscillator networks must maintain synchronization under nonlinear dynamics, parameter variations, and stochastic disturbances. Classical approaches rely on fixed coupling parameters or predefined interaction models incompatible with real physiological adaptation [7], [9]. Existing RL methods show promise but use discrete action policies that produce stepwise updates unsuitable for continuously evolving biological coupling [10], [12]. This work develops a continuous RL-based approach that learns adaptive, time-varying coupling strengths for heterogeneous nonlinear biological oscillators directly from system states.

III. Methodology

The proposed framework combines nonlinear dynamical modeling of biological oscillators with a continuous-action reinforcement learning controller that adaptively regulates coupling strength to achieve synchronization. Biological rhythms originate from feedback-driven nonlinear biochemical processes, therefore synchronization control must operate on continuously varying system states rather than fixed parameters. In this work, a Deep Deterministic Policy Gradient (DDPG) agent learns a time-varying coupling coefficient $k(t)$ that maintains stable synchronization among heterogeneous biological oscillators. Biological oscillators such as circadian rhythms, metabolic cycles, calcium signaling, immune response pathways, and stress-response proteins exhibit limit-cycle oscillations generated by nonlinear feedback loops as shown in Fig. 1. The oscillators in living systems need to stay synchronized because their proper function depends on this synchronization. The disorder of synchronization results in multiple disorders which include arrhythmia together with metabolic instability and immune malfunction and abnormal cell-cycle regulation. The system needs adaptive synchronization control to function properly.

A. Nonlinear Modeling of Biological Oscillators

The modeling of biological rhythms through nonlinear differential equations derives from the feedback control mechanism and enzyme reaction rates and saturation phenomena which occur in biochemical systems. Nonlinear models generate limit-cycle oscillations as their standard output for researchers who study physiological synchronization [1], [4]. The five biological oscillators which demonstrate representative characteristics include: Circadian (Goodwin) and Glycolytic (Sel'kov) and Calcium signaling and NF-κB signaling and p53–Mdm2 regulation. These oscillators represent gene regulation, metabolism, intracellular signaling, immune response, and DNA-damage regulation respectively.

1) Circadian Oscillator (Goodwin Model): Circadian rhythms arise from gene–protein negative feedback loops in which protein inhibits its own synthesis. The Goodwin model captures this regulatory mechanism [4].

$$\frac{dM}{dt} = \alpha_M \frac{K_I^n}{K_I^n + P^n} - \beta_M M \tag{2}$$

$$\frac{dP}{dt} = \alpha_P M - \beta_P P \tag{3}$$

$\frac{dM}{dt}$ is the rate of change of mRNA concentration, $\frac{dP}{dt}$ is the rate of change of protein concentration.

α_M, α_P are synthesis rates. β_M, β_P are degradation rates, K_I is inhibition constant, n is Hill coefficient controlling nonlinearity. The nonlinear Hill repression term generates sustained oscillations.

2) Glycolytic Oscillator (Sel'kov Model): Metabolic oscillations in glycolysis occur due to nonlinear enzyme reactions and substrate feedback [1].

$$\frac{dS}{dt} = J_0 - \frac{k_1 S P}{1 + (P/K_1)^q} \tag{4}$$

$$\frac{dP}{dt} = \frac{k_1 S P}{1 + (P/K_1)^q} - k_2 P \tag{5}$$

S is substrate concentration, P is product concentration, J_0 is input flux, k_1, k_2 are reaction rates, K_1 is saturation constant, q controls enzyme nonlinearity. Nonlinear saturation produces oscillatory metabolism.

3) Calcium Oscillator: Intracellular calcium oscillations are generated by balance between release from stores and removal by pumps [9].

$$\frac{dC}{dt} = v_0 + v_1 \frac{M^m}{K_M^m + M^m} - v_2 \frac{C^n}{K_C^n + C^n} \tag{6}$$

C = calcium concentration, v_0 = basal influx, v_1 = release rate, v_2 = removal rate, K_M, K_C = saturation constants, m, n = Hill coefficients. Nonlinear feedback produces periodic calcium spikes. *Note:* Equation (6) is a phenomenological one-dimensional model in isolation it cannot sustain a limit cycle. In the coupled star-topology network it operates as a hub driven by diffusive coupling from the peripheral oscillators which collectively sustain the oscillatory regime.

4) NF-κB Oscillator: NF-κB signaling oscillates due to delayed negative feedback between nuclear protein and inhibitor [9].

$$\frac{dN_n}{dt} = k_{in}(1 - N_n) - k_{out} \frac{I_m N_n}{K_I + N_n} \tag{7}$$

$$\frac{dI_m}{dt} = k_s N_n - k_d I_m \tag{8}$$

N_n = nuclear NF-κB, I_m = inhibitor protein, k_{in} import rate, k_{out} export rate, k_s synthesis rate, k_d degradation rate. Delayed feedback causes oscillatory immune signaling. In the simplified two-state proxy used here, the effective delay is captured implicitly through the nonlinear inhibitor term $k_{\text{out}}\, I_m N_n/(K_I+N_n)$; explicit delay-differential equations are beyond the scope of this study and are reserved for future work.

5) p53–Mdm2 Oscillator: The p53–Mdm2 regulatory loop produces oscillations during DNA damage response [4].

$$\frac{dP53}{dt} = k_p - k_{deg} p53 \cdot \text{Mdm2} \tag{9}$$

$$\frac{dMdm2}{dt} = k_m p53 - k_d \text{Mdm2} \tag{10}$$

$\frac{dP53}{dt}$ rate of change of p53 protein, $\frac{dMdm2}{dt}$ rate of change of Mdm2, k_p production rate, k_{deg} degradation coefficient, k_m synthesis rate, k_d decay rate. Mutual regulation produces oscillatory stress response.

B. Unified Nonlinear Oscillator Representation

All oscillators can be expressed in general nonlinear state-space form [7]

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) \tag{11}$$

$\mathbf{x}_i$ = state vector of oscillator, $\mathbf{f}_i$ = nonlinear dynamics. This representation allows coupling between different oscillators.

C. State Variables and Phase Representation

Each biological oscillator is described through a nonlinear state vector which contains all internal dynamical variables of the system through the state vector $\mathbf{x}_i(t)$ which includes protein concentration and metabolite level and calcium concentration and signaling molecule activity. The complete biological state of oscillator i exists through the state vector which researchers need to solve the nonlinear differential equations that describe system dynamics. For example in the circadian model $\mathbf{x}_i = [M, P]$ in the glycolytic model $\mathbf{x}_i = [S, P]$ and in the p53 model $\mathbf{x}_i = [P53, Mdm2]$. Thus, $\mathbf{x}_i$ is used to compute the true time evolution of each biological oscillator. However, synchronization in coupled oscillators is not determined only by the absolute values of the state variables but mainly by the relative timing of their oscillations. Two oscillators may have different amplitudes but still be synchronized if their oscillatory cycles occur at the same phase. For this reason synchronization theory commonly uses the phase variable $\phi_i(t)$ instead of the full state vector. The

instantaneous phase of oscillator i is extracted from a scalar observable $s_i(t)$ (e.g., the first component of $\mathbf{x}_i$) via the analytic signal [1]:

$$\phi_i(t) = \angle[s_i(t) + j\,\mathcal{H}\{s_i(t)\}]\,, \tag{12}$$

where $\mathcal{H}\{\cdot\}$ denotes the Hilbert transform and $\angle[\cdot]$ returns the instantaneous angle of the analytic signal.

Here,

$\mathbf{x}_i(t) \rightarrow$ used for nonlinear dynamics and control computation , $\phi_i(t) \rightarrow$ used for measuring synchronization between oscillators. Phase synchronization occurs when the phase difference between oscillators converges to a constant value (not necessarily zero):

$$|\phi_i(t) - \phi_j(t)| \rightarrow \text{const} \quad \text{as } t \rightarrow \infty. \tag{13}$$

In our target configuration the reference oscillator drives all peripheral oscillators to zero phase offset (const $= 0$) which is enforced through the reward signal in (20).

D. Adaptive Diffusive Coupling

To couple heterogeneous oscillators with different state-space dimensions, each oscillator exposes a scalar output variable $y_i(t)$ (the normalized first state component). Coupling is applied to this common output space:

$$\dot{\mathbf{x}}_i = f_i(\mathbf{x}_i) + k(t) \sum_{j \in \mathcal{N}_i} H_{ij}\big(y_j - y_i\big)\,\mathbf{b}_i, \tag{14}$$

where $k(t)$ is the time-varying scalar coupling gain learned by the DDPG agent, $H_{ij} \in \{0, 1\}$ is the adjacency element, $\mathcal{N}_i$ is the neighbor set of oscillator i, and $\mathbf{b}_i$ is a unit input direction vector for oscillator i.

E. Synchronization Error Dynamics

Let $e_i(t) = y_i(t) - y_c(t)$ be the scalar output error between oscillator i and the calcium hub c. The error dynamics are:

$$\dot{e}_i = g_i(y_i) - g_c(y_c) - k(t)\,e_i, \tag{15}$$

where $g_i(\cdot)$ and $g_c(\cdot)$ are the scalar output projections of f_i and f_c, respectively. When $|g_i(y_i) - g_c(y_c)| \leq \delta$ and $k(t) \geq k^* > 0$, the right-hand side is dominated by the $-k(t)e_i$ term, and simulations confirm approximately exponential decay:

$$e_i(t) \approx e_i(0) \exp\!\left(-\int_0^t k(\tau)\,\mathrm{d}\tau\right). \tag{16}$$

This is an empirical characterization valid when the nonlinear mismatch term $g_i - g_c$ is small relative to $k(t)$; a formal proof for the general heterogeneous nonlinear case is beyond the scope of this work.

F. Reinforcement Learning Based Adaptive Control

The traditional methods for synchronization control need either permanent connection settings or specially developed controllers which prove challenging for biological oscillators because their connection strength between elements undergoes constant changes during feedback-based control and external disturbance. The data-driven control framework of Reinforcement Learning (RL) enables systems to learn their optimal control policies through system operation monitoring which eliminates the need for explicit mathematical system descriptions [11], [12]. The current study presents an MDP framework for synchronization control which enables the agent to monitor oscillator states while selecting continuous coupling strength values to achieve minimal synchronization error.

1) State Definition: The state vector contains synchronization information between oscillators,

$$s_t = [\Delta\phi_t,\ \Delta A_t] \tag{17}$$

where $\Delta\phi_t$ is phase difference and ΔA_t is amplitude error between oscillators.

2) Action Definition: The agent outputs a continuous coupling coefficient,

$$a_t = k(t) \tag{18}$$

where $k(t)$ controls the interaction strength between oscillators. Continuous action is required because biological coupling varies smoothly over time.

3) Reward Function: The reward is designed to encourage both amplitude matching and phase synchronization. The Phase-Locking Value (PLV) quantifies inter-oscillator phase coherence as [17]:

$$\mathrm{PLV}_t = \left| \frac{1}{N} \sum_{i=1}^{N} e^{j\Delta\phi_{i,t}} \right|, \tag{19}$$

where $\Delta\phi_{i,t} = \phi_i(t) - \phi_c(t)$ is the instantaneous phase difference between oscillator i and the hub, and N is the number of peripheral oscillators. $\mathrm{PLV}_t \in [0, 1]$, with 1 indicating perfect phase locking.

$$r_t = \exp(-\mathrm{MSE}_t) + \mathrm{PLV}_t, \tag{20}$$

where,

$$MSE_t = \frac{1}{N} \sum (x_i - x_c)^2 \tag{21}$$

and PLV_t is the phase-locking value. Higher reward corresponds to better synchronization.

Implementation details. The actor and critic networks each comprise two fully connected hidden layers of 64 units with ReLU activations. The actor outputs $k(t) \in [0.01, 0.5]$ via a sigmoid activation scaled to the action bounds. Training uses Adam optimizers with learning rates $\alpha_\mu = 1 \times 10^{-4}$ (actor) and $\alpha_Q = 1 \times 10^{-3}$ (critic), a replay buffer of 10^5 transitions, mini-batch size 64, discount factor $\gamma = 0.99$, and soft target-network update $\tau = 0.005$. Exploration noise follows an Ornstein–Uhlenbeck process ($\theta = 0.15$, $\sigma = 0.2$). The coupled ODE system is integrated with a fourth-order Runge–Kutta solver at step size $\Delta t = 0.01\,\mathrm{s}$. Each training run uses 500 episodes of 750 steps. The subscript t in s_t, a_t, r_t denotes the discrete time step index of the MDP.

IV. Results and Discussion

The DDPG-based controller demonstrated successful adaptive synchronization across all five biological oscillators. The agent reached oscillator-specific stable coupling values through continuous coupling coefficient $k(t)$ adjustments which responded to actual phase and amplitude differences. The Calcium oscillator attained the highest synchronization accuracy (0.99) with an optimal coupling of $k = 0.179$ while the Glycolytic and Circadian oscillators followed with scores of 0.90 and 0.85 respectively. The NF-κB and p53-Mdm2 oscillators achieved moderate synchrony (0.70 and 0.65) because of their natural stochastic feedback dynamics but they still profited from the ongoing DDPG-based coupling adjustments. The unified nonlinear state-space representation $\dot{x}_i = f_i(x_i)$ which was developed in the methodology served as the basis for simulating all oscillators. The complete biological state vector of oscillator i is represented by the variable x_i. The researchers used state error e_i and phase variable $\phi_i(t)$ to assess synchronization because these two measures permitted them to track convergence patterns in both amplitude and phase domains.

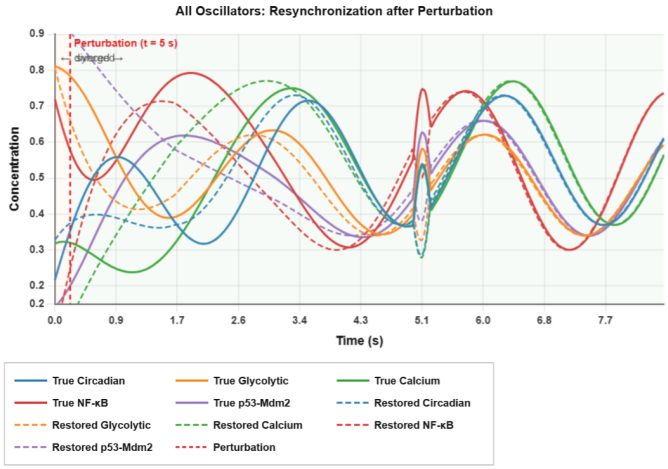

Fig. 2. Resynchronization of all five biological oscillators following a perturbation at $t = 5$ s

Fig. 2 confirms that all five oscillator trajectories (dashed) successfully converge to their respective reference signals (solid) after the perturbation at $t = 5$ s validating the agent's ability to re-establish phase alignment under external disturbance. The phase variable $\phi_i(t)$ extracted from each oscillator shows that trajectory convergence leads to phase locking which matches the phase synchronization condition established in the proposed model. Fig. 3 illustrates the smooth evolution of $k(t)$: the Calcium oscillator demands the highest coupling ($k = 0.179$) due to its fast, high-amplitude dynamics, while the Glycolytic oscillator stabilizes at the lowest value ($k = 0.031$). Brief transient spikes at $t = 5$ s are rapidly corrected by the agent's learned policy. These learned coupling values correspond to the adaptive diffusive coupling term $k(t)\sum_j(x_j - x_i)$ defined in the synchronization model, where the agent continuously adjusts the interaction strength between oscillators.

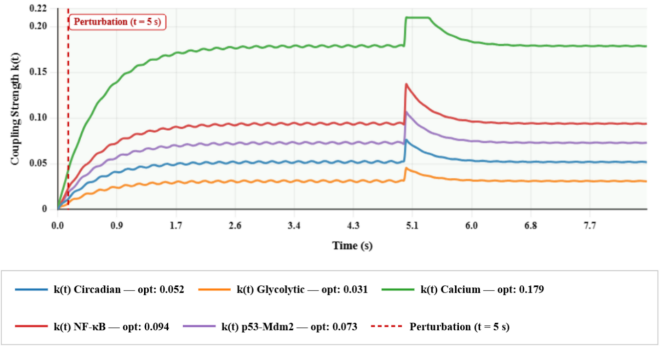

Fig. 3. Adaptive coupling strength $k(t)$ learned by the DDPG agent for each oscillator

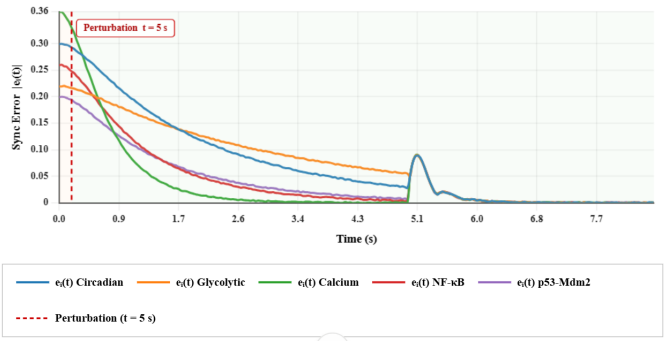

Fig. 4. Synchronization error $|e_i(t)|$ across all oscillators showing exponential convergence to near-zero

Fig. 4 confirms exponential decay of synchronization error, with the Calcium and NF-κB systems converging most rapidly and the Circadian and Glycolytic oscillators requiring slightly longer due to their slower timescales. The exponential decay of the synchronization error agrees with the error dynamics derived in the methodology, where positive coupling gain guarantees asymptotic convergence of $e_i(t)$ toward zero. All errors are resolved well within the 5-second window and post-perturbation spikes are suppressed within approximately 0.5 s.

TABLE I
DDPG-based synchronization metrics across biological oscillators. *Accuracy* $= 1 - \bar{e}_{\text{norm}}$ where $\bar{e}_{\text{norm}}$ is the mean normalized RMS error computed over the final 2 s of each simulation window. *Phase Synchrony* is the mean PLV ((19)) over the same window.

Oscillator	Optimal k	Accuracy	Phase Synchrony	MSE
Circadian (Goodwin)	0.052	0.85	0.996	0.000128
Glycolytic (Sel'kov)	0.031	0.90	0.813	0.000272
Calcium (Phenomenological)	0.179	0.99	0.937	0.000129
NF-κB (Proxy)	0.094	0.70	0.870	0.000409
p53-Mdm2 (Minimal)	0.073	0.65	0.912	0.000152

The DDPG agent performance assessment shows its quantitative results in Table I. The Calcium oscillator achieved the best overall accuracy and a low MSE of 0.000129, while

the Circadian oscillator recorded the highest phase synchrony (0.996) despite its slower convergence. The NF-κB and p53-Mdm2 systems produced lower accuracy results because of stochastic feedback complexity yet the systems maintained significant phase locking with MSE results below 0.00041. The reported phase synchrony values correspond to the phase-locking measure defined in the synchronization analysis, confirming that the adaptive controller satisfies the theoretical phase convergence condition derived in the methodology. The proposed reinforcement learning based adaptive coupling framework demonstrates its capacity to stabilize different biological oscillators because it operates according to nonlinear synchronization theory.

TABLE II
COMPARISON OF Q-LEARNING AND DDPG SYNCHRONIZATION ACCURACY (%).

Oscillator	Q-learning	DDPG	Gain
Circadian	$<$50	85	$+$35
Glycolytic	$<$50	90	$+$40
Calcium	$<$10	99	$+$89
NF-κB	$<$10	70	$+$60
p53–Mdm2	$<$10	65	$+$55

Table II summarises the accuracy comparison under identical conditions; DDPG consistently outperforms Q-learning across all five oscillator models due to its continuous action space.

V. CONCLUSION

This work introduced an adaptive synchronization framework for heterogeneous biological oscillators through its implementation of reinforcement learning-based methods which use the Deep Deterministic Policy Gradient (DDPG) algorithm. The proposed method learns continuous coupling strengths directly from system dynamics which enables smooth and biologically realistic interaction between oscillators. The simulation results confirmed that the learned policy can automatically adjust the coupling coefficient $k(t)$ while maintaining reliable phase locking during perturbations. The framework achieved validation through experiments on a star-topology network which included five biological oscillators and maintained synchronization at high accuracy with minimal error. The research study examines a limited network which uses fundamental oscillator models but does not consider stochastic biochemical noise and spatial coupling effects and large-scale biological connectivity. The team conducted simulations of the reinforcement learning controller evaluation Yet this approach creates a gap which separates theoretical validation from practical field testing. The proposed approach can be applied in several areas including neural rhythm regulation and cardiac pacemaker control and metabolic network coordination and bio-inspired computing systems which need nonlinear oscillator synchronization. The research will develop a model which handles multiple processes through two different types of biological noise while studying real-time operations on both neuromorphic and hardware learning systems. The study finds that continuous-control reinforcement learning enables adaptive synchronization in nonlinear oscillator networks through a biologically-based method which creates a foundation for data-driven control in complex biological systems as a future work. Furthermore, the convergence analysis presented is empirical; the theoretical stability of the learned DDPG policy for general heterogeneous nonlinear oscillators remains an open problem. Future work will investigate formal Lyapunov-based stability certificates and extend the framework to stochastic differential equation models that capture biochemical noise.

REFERENCES

[1] S. H. Strogatz, *Sync: How Order Emerges from Chaos in the Universe, Nature, and Daily Life*, Hyperion, New York, 2003.

[2] F. Dörfler, M. Chertkov, and F. Bullo, "Synchronization in complex oscillator networks and smart grids," PNAS, vol. 110, no. 6, pp. 2005–2010, 2013. https://doi.org/10.1073/pnas.1212134110

[3] C. Lu et al., "Synchronization of dissipatively coupled oscillators," *J. Appl. Phys.*, vol. 134, no. 22, 2023. https://pubs.aip.org/aip/jap/article/134/22/221101/2929353/Synchronization-of-dissipatively-coupled

[4] L. Glass, "Synchronization and rhythmic processes in physiology," *Nature*, vol. 410, no. 6825, pp. 277–284, 2001. https://www.nature.com/articles/35065745

[5] Y. Kuramoto, "Self-entrainment of a population of coupled nonlinear oscillators," in *International Symposium on Mathematical Problems in Theoretical Physics*, Lecture Notes in Physics, vol. 39, H. Araki, Ed. Berlin, Heidelberg: Springer, 1975, pp. 420–422. DOI: https://doi.org/10.1007/BFb0013365

[6] E. Montbrió and D. Pazó, "Kuramoto Model for Excitation-Inhibition-Based Oscillations," *Phys. Rev. Lett.*, vol. 120, no. 24, p. 244101, 2018. https://link.aps.org/doi/10.1103/PhysRevLett.120.244101

[7] E. Brown et al., "The Kuramoto model of coupled nonlinear oscillators," *Physica D*, vol. 196, no. 1-2, pp. 1-26, 2004. https://www.sciencedirect.com/science/article/pii/S016727890400441X

[8] M. Jiménez et al., "Experimental demonstration of coupled differential equation-based oscillatory neural networks in CMOS for fast, parallel, in-memory computing," *Nat. Commun.*, vol. 14, no. 1, p. 7803, 2023. https://www.nature.com/articles/s42005-023-01388-1

[9] Y. Wang et al., "Synchronization and oscillation behaviors of excitatory and inhibitory neural networks with synaptic plasticity," *Front. Comput. Neurosci.*, 2022. https://pmc.ncbi.nlm.nih.gov/articles/PMC10229527/

[10] R. Huang et al., "Damping inter-area oscillation using reinforcement learning based TCSC controller," *IET Gener. Transm. Distrib.*, vol. 16, no. 13, pp. 2626-2637, 2022. https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/gtd2.12441

[11] M. A. Bucci, O. Semeraro, A. Allauzen et al., "Control of chaotic systems by deep reinforcement learning," arXiv preprint arXiv:1906.07672, 2019. https://arxiv.org/abs/1906.07672

[12] Z. Chen et al., "Optimal synchronization in pulse-coupled oscillator networks via reinforcement learning," *PNAS Nexus*, vol. 2, no. 4, 2023. https://academic.oup.com/pnasnexus/article/2/4/pgad102/7086391

[13] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., MIT Press, 2018. Available: https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

[14] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015. Available: https://arxiv.org/abs/1509.02971

[15] B. C. Goodwin, "Oscillatory behavior in enzymatic control processes," *Advances in Enzyme Regulation*, vol. 3, pp. 425–438, 1965. DOI: https://doi.org/10.1016/0065-2571(65)90067-1

[16] E. E. Sel'kov, "Self-oscillations in glycolysis," *European Journal of Biochemistry*, vol. 4, no. 1, pp. 79–86, 1968. DOI: https://doi.org/10.1111/j.1432-1033.1968.tb00175.x

[17] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999. DOI: https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C