УДК 536.25

Гао Цзыхань,
Магистр
Новосибирский государственный университет, г. Новосибирск
t.gao2@g.nsu.ru

УПРАВЛЕНИЕ ТРЕХМЕРНОЙ КОНВЕКЦИЕЙ В ЗАДАЧЕ РЭЛЕЯ-БЕНАРА С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Аннотация. Конвективный теплообмен широко распространен в природе и применяется во многих отраслях промышленности. Определение эффективных стратегий управления, таких как подавление или усиление конвективного теплообмена при фиксированных внешних температурных градиентах, является важной фундаментальной и технологической задачей. В этой работе исследован подход, основанный на самых современных алгоритмах обучения с подкреплением. Этот метод значительно снижает конвективный теплообмен в двумерной системе Рэлея-Бенара за счет регулирования температуры на нижней границе системы управления, стабилизируя число Нуссельта на более низких значениях. Анализ текущих результатов исследований показывает, что управление на основе RL может эффективно управлять состоянием проводимости, когда число Рэлея Ra меньше или равно 10 000, что приводит к значительному снижению числа Нуссельта Nu по сравнению с неконтролируемым состоянием. Однако эффективность управления системой остается недостаточной, когда число Рэлея Ra достигает 100 000.

Abstract

Convective heat transfer is prevalent in nature and many industrial applications. Determining effective control strategies, such as suppressing or enhancing convective heat exchange under fixed external thermal gradients, is a prominent fundamental and technological challenge. In this work, I explore an approach based on state-of-the-art reinforcement learning algorithms. This method significantly reduces convective heat transfer in a two-dimensional Rayleigh-Bénard system by controlling the temperature at the bottom boundary of the control system, stabilizing the Nusselt number at lower values. Analysis of current research results indicates that RL-based control can effectively manage the conduction state when the Rayleigh number Ra is less than or equal to 10 000, resulting in a significant reduction in the Nusselt number Nu compared to the uncontrolled state. However, the control effectiveness for the system remains insufficient when the Rayleigh number Ra reaches 100 000.

Ключевые слова — обучение с подкреплением, тепловая конвекция, Рэлей–Бенар, управление, хаос.

Keywords—Reinforcement Learning, Thermal Convection, Rayleigh–Benard, Control, Chaos

I. INTRODUCTION

Rayleigh-Bénard convection (RBC) provides a widely studied paradigm for thermally driven flows, which are prevalent in both natural phenomena and industrial applications [1]. Buoyancy effects ultimately lead to fluid dynamical instabilities, with buoyancy effects determined by temperature gradients [2] and their influence on heat transfer. Controlling Rayleigh-Bénard convection is a fundamental research topic with scientific significance [3]. Furthermore, preventing, mitigating, or enhancing these instabilities and/or regulating heat transfer is crucial in many applications. Examples include crystal growth processes, such as silicon chip production [4]. Indeed, while these processes benefit from increased temperature gradients in terms of speed, escalating thermal gradients jeopardize the quality of outcomes due to fluid motion (i.e., flow instability). Therefore, the key question to address here is: can we control and stabilize fluid flow instabilities induced by temperature gradients?

The Rayleigh number (Ra) is a dimensionless quantity used to quantify the ratio of buoyancy driving forces to viscous forces in fluid flow. In Rayleigh-Bénard convection, when Ra is low, thermal convection within the fluid is weak and relies mainly on conductive heat transfer. As Ra increases, buoyancy effects become stronger. When Ra exceeds a critical value, thermal convection becomes significant, leading to the transition from laminar to convective flow, characterized by the formation and enhancement of convection cells.

$$\mathrm{Ra} = \frac{g\alpha(T_H - Tc)H^3}{kv}$$

The Nusselt number (Nu) characterizes the efficiency of convective heat transfer. It is defined as the ratio of total heat transfer within the boundary layer to conductive heat transfer.

$$Nu = \frac{\langle u_y T \rangle_{x,y} - k\partial_y \langle T \rangle_{x,y}}{k\Delta T \ / \ H}$$

Expressed in terms of Rayleigh and Nusselt numbers, our control problem is: Decrease or minimize the Nusselt number for a fixed Rayleigh number.

II. LITERATURE REVIEW

In recent years, various methods have been proposed to address this issue. These methods can be categorized into passive and active control methods. Passive control methods include: acceleration modulation [3, 6], oscillating shear flows [7], and oscillating boundary temperatures [8]. Active control methods include: velocity actuators [9] and thermal boundary layer perturbations [10, 11, 12]. While these methods are ingenious, they are not practical due to requirements such as a perfect understanding of the system's state or initial conditions close to the conductive state, which are difficult to establish [11]. The main challenge in controlling RBC lies in its chaotic behavior and chaotic response to control actions.

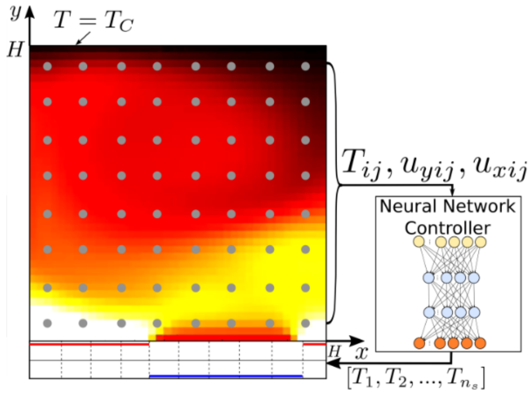
Reinforcement Learning (RL) algorithms [14] have demonstrated the ability to solve complex control problems, mastering extremely difficult challenges in high-dimensional spaces (e.g., board games [15, 16] and robotics [17]). Reinforcement Learning is a form of supervised Machine Learning (ML) [18] aimed at finding optimal control strategies. This is achieved through continuous trial and error interaction with the (simulated or real) environment, iteratively refining the initial random control policy. In fact, this is often a rather slow process, potentially requiring millions of trials to converge [19]. Reinforcement Learning is also employed in fluid dynamics, for instance, in training intelligent inertial or self-propelled particles [20, 21, 22], schooling behavior of fish [23, 24], soaring of birds and gliders in turbulent environments [25, 26], and optimal navigation in turbulent flows [27].

III. DATASET AND BASELINE

A. Dataset

FluidX3D is used for numerical simulation fluid and an environment is constructed for reinforcement learning, which acts as a dataset.

Figure 1: Schematic representation of the Reinforcement Learning control approach applied to the Rayleigh-Bénard system with the aim of reducing convective effects (i.e., Nusselt number).



The system consists of a domain with height H, aspect ratio Γ = 1, no-slip boundary conditions, constant temperature Tc at the top boundary, adiabatic side boundaries, where the temperature is controlled by n heaters $[T_1, T_2, \ldots, T_n]$. The temperatures of the heaters can vary continuously to optimize the control protocol, ensuring a constant mean value. This continuous optimization allows for finer adjustments and potentially better performance compared to discrete temperature settings. Maintain a constant mean value. Since the average temperature of the bottom plate is constant, the Rayleigh number is well-defined and constant over time. The RL method employs a neural network controller, which receives flow state information from multiple probes at fixed locations and outputs the temperature value for the bottom boundary heater. During training, the parameters of the neural network are automatically optimized by the RL method.

B. Baseline

Currently, one of the state-of-the-art active methods used to address Rayleigh-Bénard convection (RBC) is based on linear control applied to the thermal boundary layer at the bottom [13]. However, the article does not directly use the Nusselt number as a performance measure for the control effectiveness. Therefore, a comparison cannot be provided at the moment. Replication and comparison work for the baseline will be conducted in the following year. The results of the current study can be compared with the situation without RL control.

IV. METHOD

A. Proximal Policy Optimization

We employ the Proximal Policy Optimization (PPO) RL algorithm [4], which belongs to the series of policy gradient methods. Starting from random initial conditions, policy gradient methods iteratively search for the optimal (or sufficiently good) policy through gradient ascent based on locally estimated performance. Specifically, this optimization utilizes a probability distribution π(ai | si) over the action space conditioned on the instantaneous system state. At each step of the control loop, we sample and apply an action based on the distribution π(a|s). It is worth noting that sampling is crucial during training to ensure a sufficient balance between exploration and exploitation, while during testing, this stochastic approach can be turned into a deterministic one by constraining actions with the highest associated probabilities.

B. FluidX3D

FluidX3D is a fluid dynamics simulation library developed by ProjectPhysX. It provides tools for simulating various fluid phenomena, can be applied to simulate Rayleigh-Bénard convection by providing a framework for modeling fluid flow in a confined space heated from below. Users can define the geometry and boundary conditions of the system, including the temperature gradient across the boundaries.

C. MPI

To facilitate communication between FluidX3D (implemented in C++) and RL algorithms (implemented in Python), we employ the Message Passing Interface (MPI). MPI provides a standardized protocol for communication and data transfer between processes running on distributed memory systems.

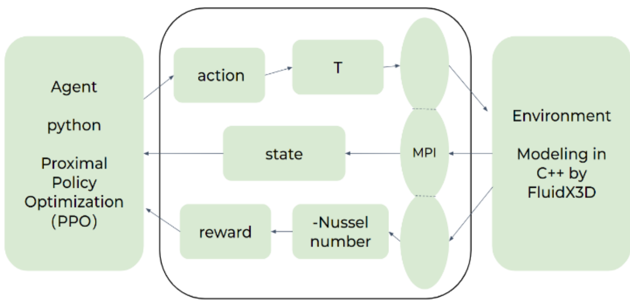


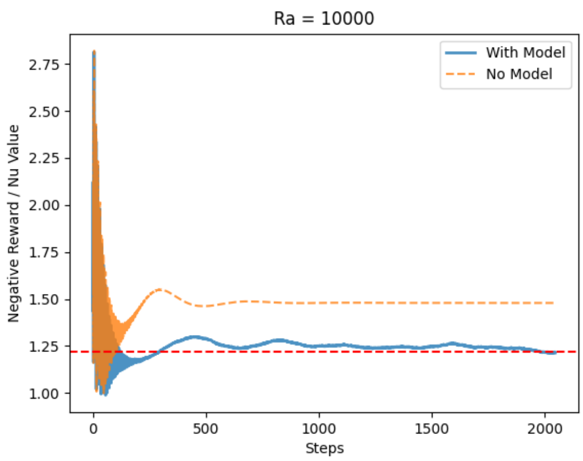
Figure 2: Using MPI to establish a communication protocol between FluidX3D and RL, data transmission occurs as follows: FluidX3D creates the environment, while RL

constructs the agent.

The environment sends the initial state to the agent, which, based on the Proximal Policy Optimization (PPO) strategy, selects an action (in this case, the temperature of the bottom heater) and transmits it back to the environment. After simulating the next phase based on the action received, the environment sends the new reward and state to the agent for learning. This iterative process enables the agent to optimize its control strategy for the Rayleigh-Bénard convection simulation.

V. Results

Figure 3: Variation of Nusselt number with time when Ra=10^4



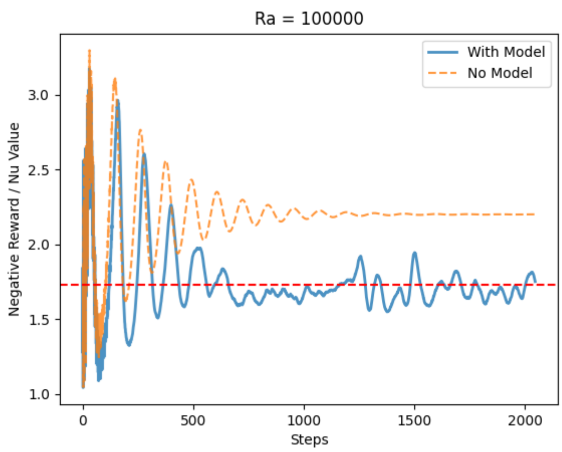


Figure 4: Variation of Nusselt number with time when Ra=10^5

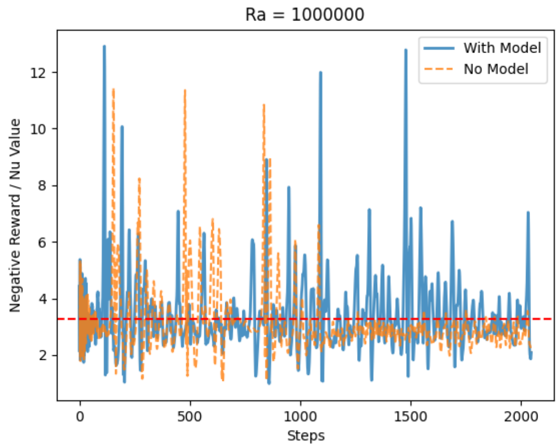


Figure 5: Variation of Nusselt number with time when Ra=10^6

The time evolution of the Nusselt number under three different Rayleigh states is observed, with control initiated at t = 0 . Each step represents a simulation conducted for a unit of time Δt (currently set to 500 units). When Ra = 10^4 ,preliminary control by RL achieves stability with the Nusselt number converging to approximately 1.24, significantly lower than the uncontrolled value of 1.5. At Ra = 10^5, RL control continues to lower the average Nusselt number, albeit with larger fluctuations and incomplete system stabilization. However, at Ra = 10^6, RL performance is subpar, with unstable Nusselt values and no significant reduction observed.

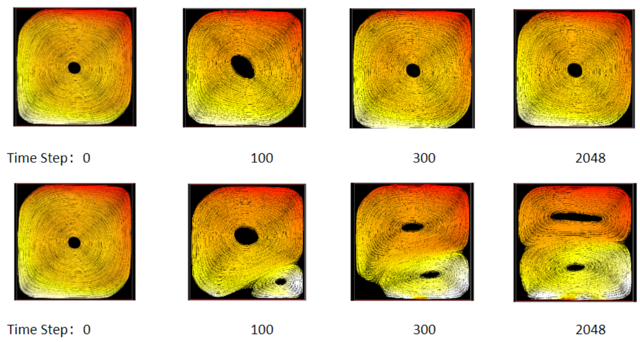


Figure 6: Temperature images of the environment without RL control (top) and with RL control (bottom) when Ra=10^5.

RL control engenders a flow regime akin to a "dual-cell pattern." In this configuration, the Nusselt number is notably lower compared to the uncontrolled scenario, as heat transfer occurs predominantly through diffusion across the interface between two cells to reach the top boundary. This "dual-cell" control strategy is autonomously crafted by RL control without any external supervision.

VI. Conclusion

Through the analysis of the results, it is evident that RL control promotes the emergence of a dual-cell pattern and demonstrates effectiveness in stabilizing and optimizing the system. Compared to the uncontrolled scenario, RL achieves a significant reduction in the Nusselt number. However, limitations of this method become apparent at higher Rayleigh states. As the Rayleigh number (Ra) increases, RL control faces challenges in maintaining system stability and further reducing the Nusselt number. Additionally, it is observed during the experiment that the method did not converge, indicating the need for further improvement.

VII. Optimization methods

Optimize the hyperparameters of the PPO algorithm. Refine the number of simulations for the FluidX3D environment at each time step. Optimize the reward function. Conduct parallel training across multiple environments to improve generalization capability.

REFERENCES

1. Welty, J. R., Wicks, C. E., Rorrer, G., & Wilson, R. E. (2009). *Fundamentals of momentum, heat, and mass transfer* (5th ed.). John Wiley & Sons.

2. Chandrasekhar, S. (1953). The instability of a layer of fluid heated below and subject to Coriolis forces. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 217*(1130), pp. 306–327.

3. Swaminathan, A., Garrett, S. L., Poese, M. E., & Smith, R. W. (2018). Dynamic stabilization of the Rayleigh-Bénard instability by acceleration modulation. *The Journal of the Acoustical Society of America, 144*(4), pp. 2334–2343.

4. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

5. Deeley, M. A., et al. (2013). Segmentation editing improves efficiency while reducing inter-expert variation and maintaining accuracy for normal brain tissues in the presence of space-occupying lesions. *Physics in Medicine and Biology, 58*(12), pp. 4071–4086.

6. Carbo, R. M., Smith, R. W., & Poese, M. E. (2014). A computational model for the dynamic stabilization of Rayleigh-Bénard convection in a cubic cavity. *The Journal of the Acoustical Society of America, 135*(2), pp. 654–668.

7. Kelly, R., & Hu, H. (1994). The effect of finite amplitude non-planar flow oscillations upon the onset of Rayleigh-Bénard convection. *International Heat Transfer Conference Digital Library*. Begel House Inc.

8. Davis, S. H. (1976). The stability of time-periodic flows. *Annual Review of Fluid Mechanics, 8*(1), pp. 57–74.

9. Tang, J., & Bau, H. H. (1994). Stabilization of the no-motion state in the Rayleigh-Bénard problem. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences, 447*(1931), pp. 587–607.

10. Singer, J., & Bau, H. H. (1991). Active control of convection. *Physics of Fluids A: Fluid Dynamics, 3*(12), pp. 2859–2865.

11. Howle, L. E. (1997). Active control of Rayleigh-Bénard convection. *Physics of Fluids, 9*(7), pp. 1861–1863.

12. Or, A. C., Cortelezzi, L., & Speyer, J. L. (2001). Robust feedback control of Rayleigh-Bénard convection. *Journal of Fluid Mechanics, 437*, pp. 175–202.

13. Remillieux, M. C., Zhao, H., & Bau, H. H. (2007). Suppression of Rayleigh-Bénard convection with proportional-derivative controller. *Physics of Fluids, 19*(1), p. 017102.

14. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

15. Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). Mastering the game of Go without human knowledge. *Nature, 550*(7676), pp. 354–359.

16. Silver, D., Hubert, T., Schrittwieser, J., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science, 362*(6419), pp. 1140–1144.

17. Andrychowicz, M., Baker, B., Chociej, M., et al. (2018). Learning dexterous in-hand manipulation. *arXiv preprint arXiv:1808.00177*.

18. Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

19. Hessel, M., Modayil, J., Van Hasselt, H., et al. (2018). Rainbow: Combining improvements in deep reinforcement learning. *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*.

20. Colabrese, S., Gustavsson, K., Celani, A., & Biferale, L. (2018). Smart inertial particles. *Physical Review Fluids, 3*(8), p. 84-91.

21. Colabrese, S., Gustavsson, K., Celani, A., & Biferale, L. (2017). Flow navigation by smart microswimmers via reinforcement learning. *Physical Review Letters, 118*(15), p. 15-26.

22. Gustavsson, K., Biferale, L., Celani, A., & Colabrese, S. (2017). Finding efficient swimming strategies in a three-dimensional chaotic flow by reinforcement learning. *The European Physical Journal E, 40*(12), p. 110.

23. Gazzola, M., Tchieu, A. A., Alexeev, D., et al. (2016). Learning to school in the presence of hydrodynamic interactions. *Journal of Fluid Mechanics, 789*, pp. 726–749.

24. Verma, S., Novati, G., & Koumoutsakos, P. (2018). Efficient collective swimming by harnessing vortices through deep reinforcement learning. *Proceedings of the National Academy of Sciences, 115*(23), pp. 5849–5854.

25. Reddy, G., Celani, A., Sejnowski, T. J., & Vergassola, M. (2016). Learning to soar in turbulent environments. *Proceedings of the National Academy of Sciences, 113*(33), pp. 4877–4884.

26. Reddy, G., Wong-Ng, J., Celani, A., et al. (2018). Glider soaring via reinforcement learning in the field. *Nature, 562*(7726), pp. 236–239.

27. Biferale, L., Bonaccorso, F., Buzzicotti, M., et al. (2019). Zermelo's problem: Optimal point-to-point navigation in 2D turbulent flows using reinforcement learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 29*(10), p. 103-138.