Smart Navigation Through a Rotating Barrier via Deep Reinforcement Learning

Mohammad Hossein Masoudi, Ali Naji

School of Nano Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Abstract

This study applies Deep Reinforcement Learning (DRL) to optimize the navigation of micro-agents in a dynamically changing environment, focusing on the size-based sorting of active microagents using a rotating Gaussian potential.

Introduction

Active matter systems, consisting of self-propelled units that convert energy into motion, show potential for applications such as targeted drug delivery and autonomous micro-robotics. This research integrates DRL to navigate and sort microagents based on size through complex, dynamic barriers [1].

Methods

This research employs a combination of computational physics models and machine learning algorithms to study the navigation and sorting of micro-agents in dynamic environments. Below, we detail the core components of our methodological approach:

Active Brownian Particles (ABP) Model:

- **Simulation Environment:** We simulate a two-dimensional environment where each agent represents an ABP capable of autonomous movement and interaction with dynamic barriers and other agents. The position and orientation of each particle are subject to stochastic fluctuations, mimicking thermal motion. [1]

Advantage Actor-Critic (A2C) Algorithm:

- **Reinforcement Learning Framework:** The A2C algorithm serves as our primary reinforcement learning model[2], where each agent learns an optimal policy for navigation based on the rewards collected from the environment. The algorithm combines policy gradient (actor) and value function (critic) approaches to stabilize learning and improve convergence rates.
- **Network Architecture:** The actor and the critic share an initial set of layers which process the state inputs, followed by separate output layers for action probabilities and state value predictions. This architecture allows for simultaneous updates to both policy and value estimates during training.

Model Overview

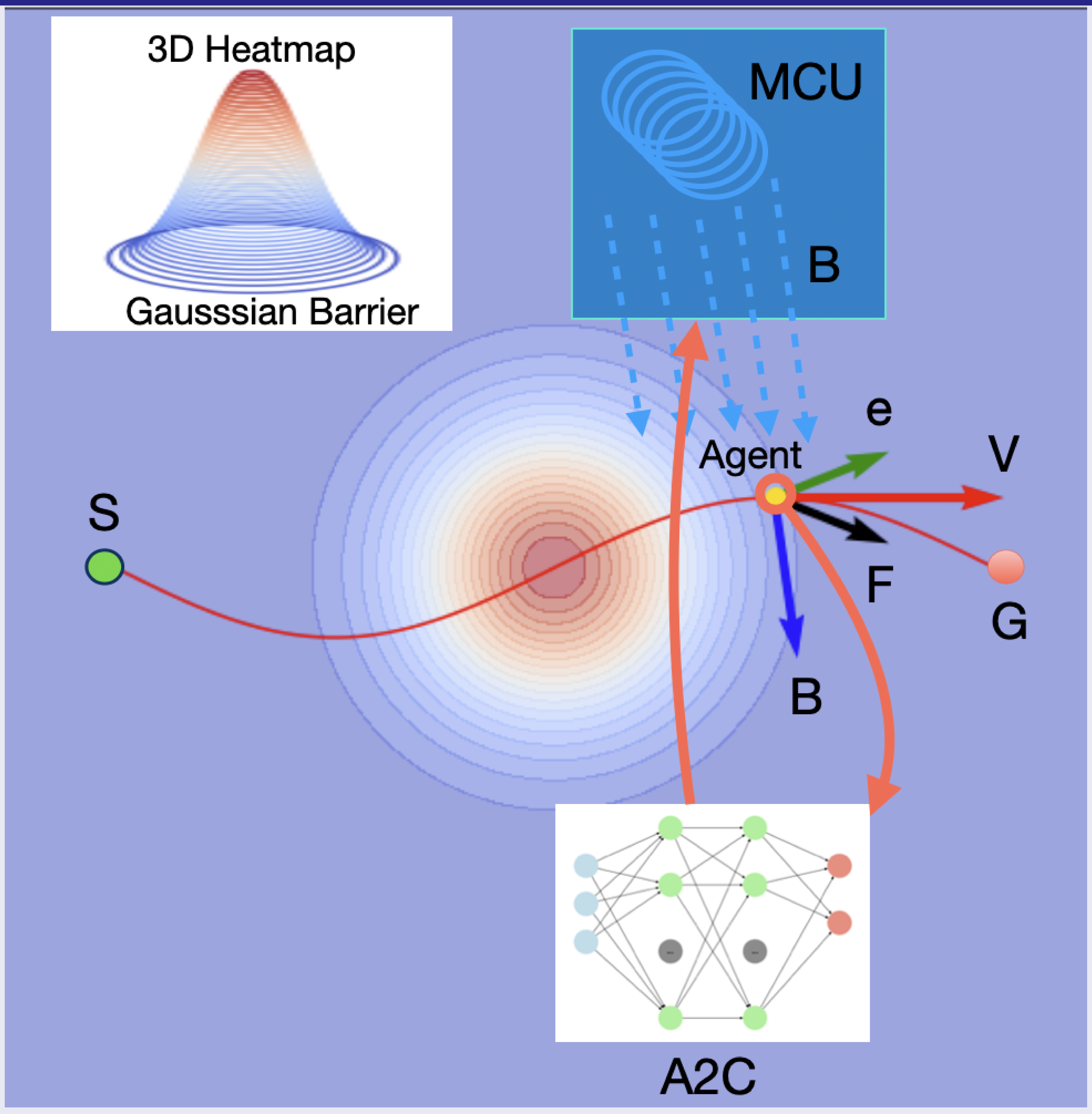

Figure 1: Schematic of the navigation model with rotating Gaussian potential.

Results: shortest-time navigation analysis

- Implementations indicate that agents can adapt to different scenarios of environmental dynamics.
- A2C vs. A* shortest path results to verify the optimality of our results.

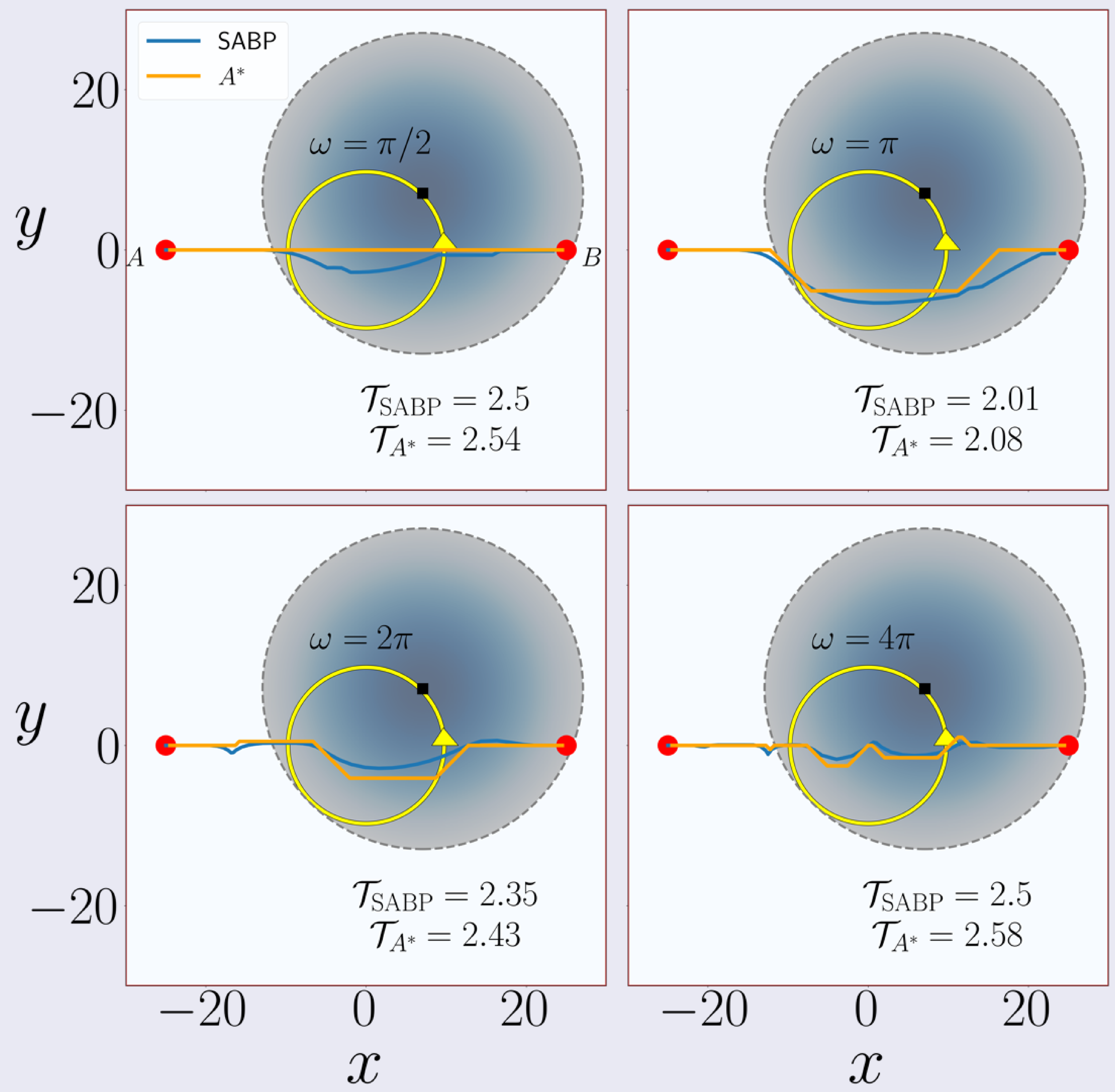

Figure 2: A2C vs. A* shortest path in different scenarios of rotating barrier

Results: Noise-induced training and its effect on the precision of our sorting mechanism

- Size-based sorting mechanism analysis in presence of thermal noises
- Providing RSM and ESM indices to measure the effectiveness of our sorting mechanism in the presence of thermal noises.

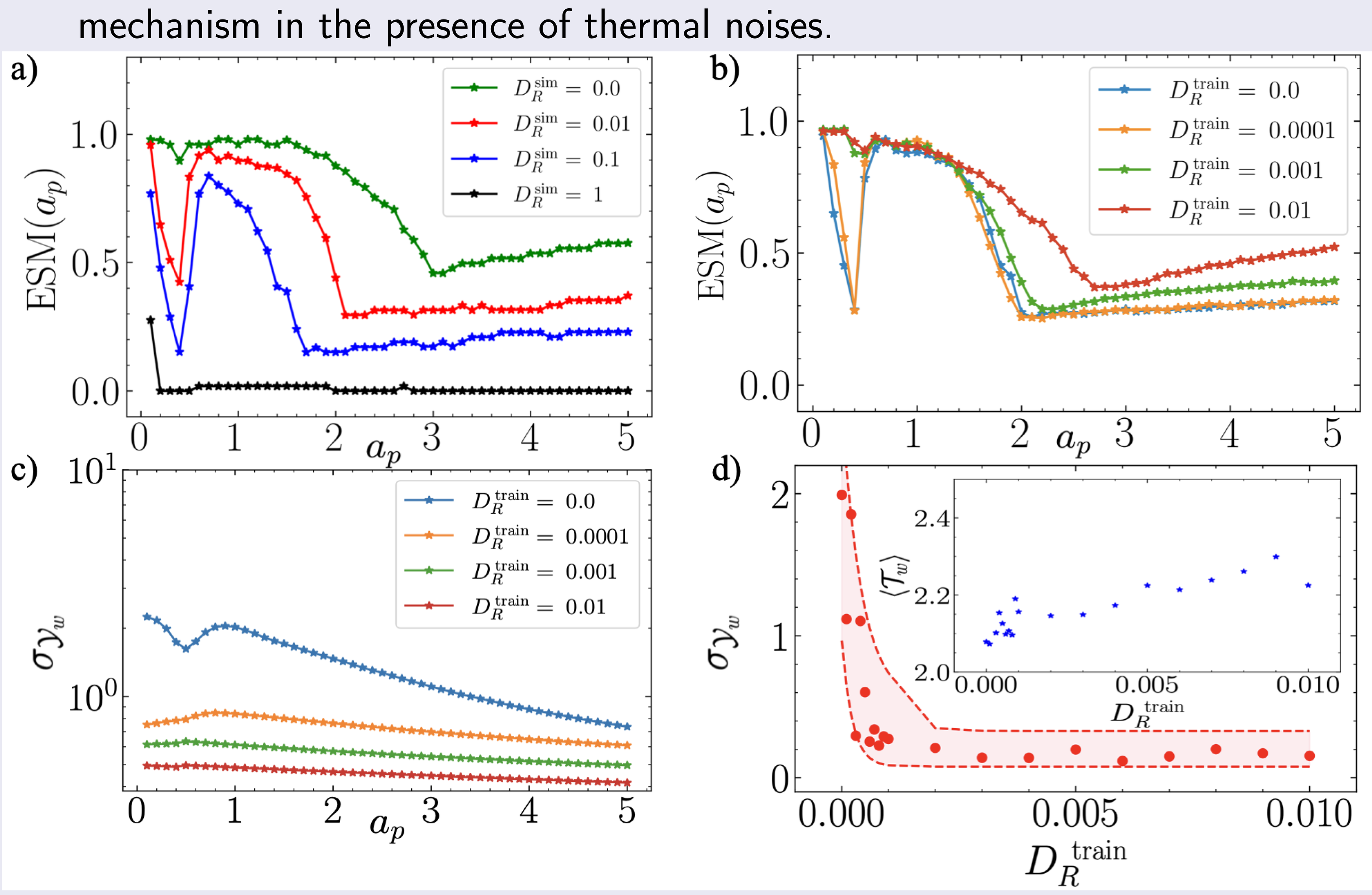

Figure 3: Noise-induced training and its effect on our sorting mechanism in the presence of thermal noises

References

[1] Clemens Bechinger et al. "Active particles in complex and crowded environments". In: *Rev. Mod. Phys.* 88 (4 Nov. 2016), p. 045006. DOI: 10.1103/RevModPhys.88.045006. URL: https://doi.org/10.1103/RevModPhys.88.045006.

[2] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT press Cambridge, 2018. URL: https://mitpress.mit.edu/9780262039246/reinforcement-learning/.