Can Chaos be utilized as exploration noise for locomotion learning?

Worasuchad Haomachai[1], Poramate Manoonpong[1,2*]

[1] Nanjing University of Aeronautics and Astronautics, Nanjing, China

[2]Vidyasirimedhi Institute of Science & Technology, Rayong, Thailand

haomachai@gmail.com, *poramate.m@vistec.ac.th

Introduction

- There is compelling evidence to support that chaotic patterns of behavior exist in many biological systems [1,2]. This suggests that chaotic dynamics may be involved in the biological neural control underlying spontaneous behavior.
- It also raises the question, "Can chaos be utilized in artificial neural control for robot locomotion learning?"

Objective

- This study investigates and compares the use of chaotic exploration noise and standard Gaussian noise for robot locomotion learning.

Material and Methods

- We construct a locomotion controller as a reinforcement learning framework, so that our robot (here, a gecko-like robot) has to learn to walk. The controller, configured as a neural central pattern generator (CPG) with a radial basis function (RBF)-based premotor neuron network (Fig.1) [3].
- The robot joint trajectories are encoded in the output weights connecting the CPG-RBF network to the motor neurons (dashed lines in Fig.1). The output weights are learned using a probability-based black-box optimization (BBO) approach to optimize joint trajectories with respect to robot walking performance.

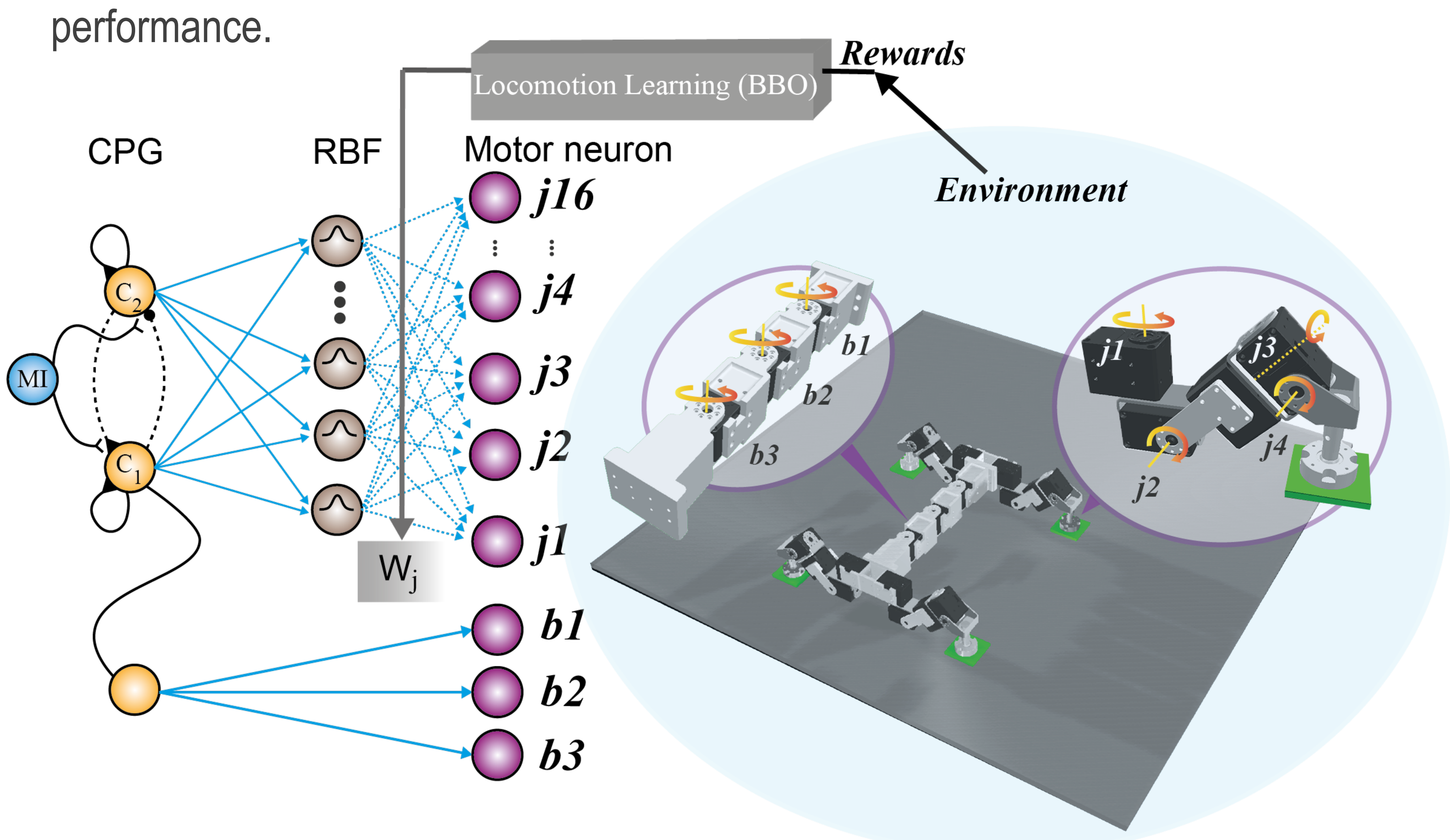

Figure 1: An overview of the CPG-RBF control network with BBO applied to the gecko-like robot.

- The BBO method, used here called Policy Improvement with Black-Box Optimization ($\mathrm{PI}^{\mathrm{BB}}$), is a parameter perturbation approach and a variant of an evolutionary algorithm. It is employed to optimize the output weights in order to generate optimal joint trajectories that maximize a reward, which basically describes how well the robot performs (Fig.2A). The pseudocode of BBO can be seen in Algorithm 1.

Algorithm 1 BBO

while *cost not converged* **do**
  for $k \in K$ **do**
    $R_k = execCPGRBFN(w_{k,j} + \varepsilon_k)$
  end
  for $k \in K$ **do**
    $S_k = e^{\lambda \cdot \frac{R_k - \min(R)}{\max(R) - \min(R)}}$
    $P_k = \frac{S_k}{\sum_{k=1}^{K} S_k}$
  end
  $\delta w_{k,j} = \sum_{k=1}^{K} P_k \cdot \varepsilon_k$
  $w_{k,j} \leftarrow w_{k,j} + \delta w_{k,j}$
  $\varepsilon_k \leftarrow \gamma \cdot \varepsilon_k$
end

Experiment and Results

- We aim to determine whether chaotic noise can be utilized as a perturbation noise to optimize control parameters (output weights) in BBO. Thus, we let the robot learn with chaotic noise (ε_k= chaotic noise) and compared its locomotion learning performance to Gaussian noise (ε_k = Gaussian noise).

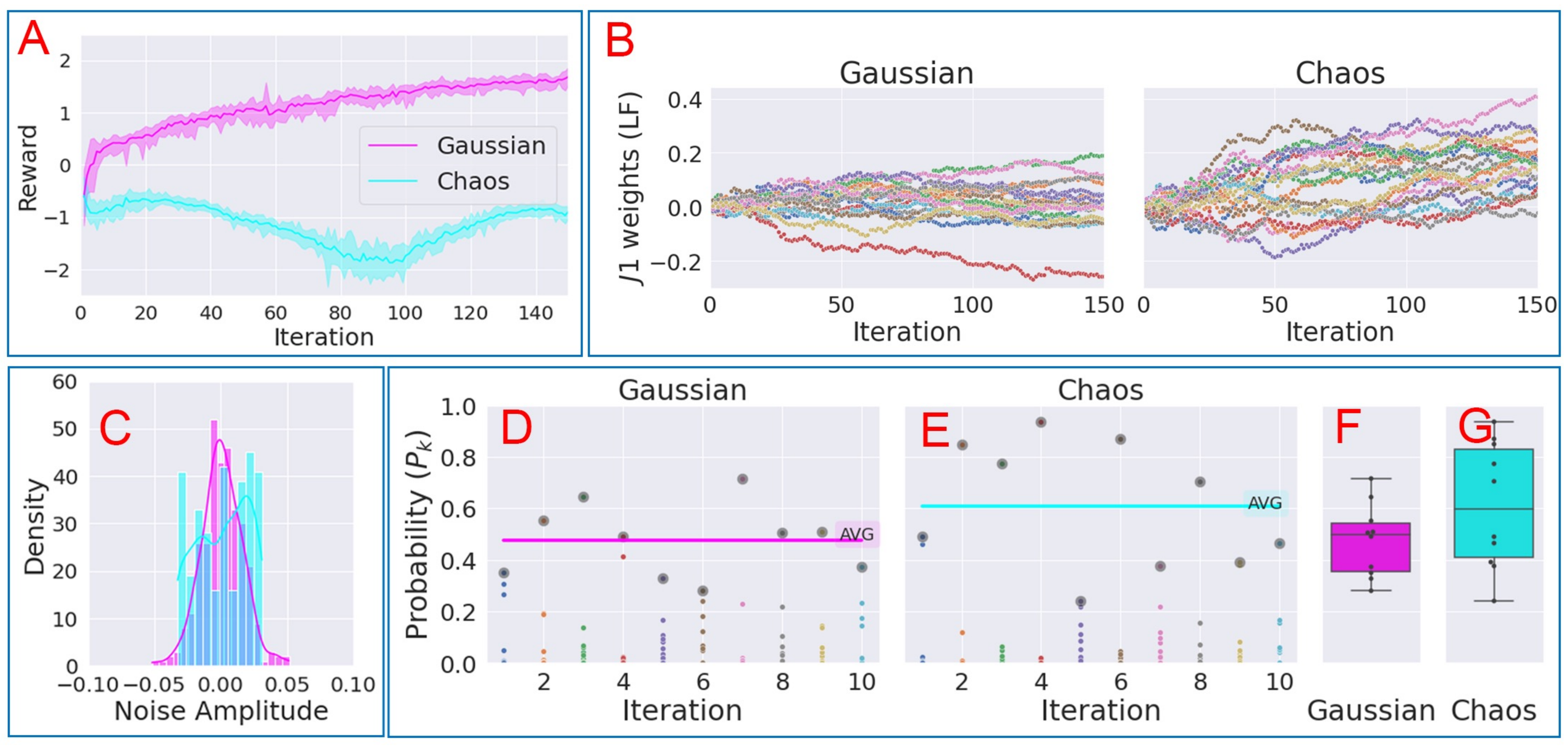

Figure 2: Learning results and analyzation of learning parameters

- Due to the asymmetric profile and chaotic dynamics (Fig.2C), the control parameters might change quickly at the beginning, subsequently converting to a certain parameter space, as observed in the weight changes of j1 of LF (Fig.2B). Based on the analysis of BBO during the first ten iterations, we observed that the highest probability P_k of chaotic noise tends to fluctuate and dominate by performing undesirable behaviors (Fig.2E). The average value of the highest P_k in each iteration was 0.60 with an SD of 0.27 (Fig.2G). Undesirable behaviors that dominate at the beginning can quickly lead the optimization process in the wrong direction. As a consequence, the optimization process could get stuck at the local optima, preventing the robot from forming a stable gait for walking forward (Fig.3, right).
- The symmetric profile of Gaussian noise can slowly adapt the parameters, leading to a balance of positive and negative parameter values with lower probability and a variant of P_k (Fig.2D and 2F). The average of the highest P_k in each iteration was 0.47 with an SD of 0.13 (Fig.2F). This results in preventing divergence (Fig.2B) where a stable gait can be formed (Fig.3, left).

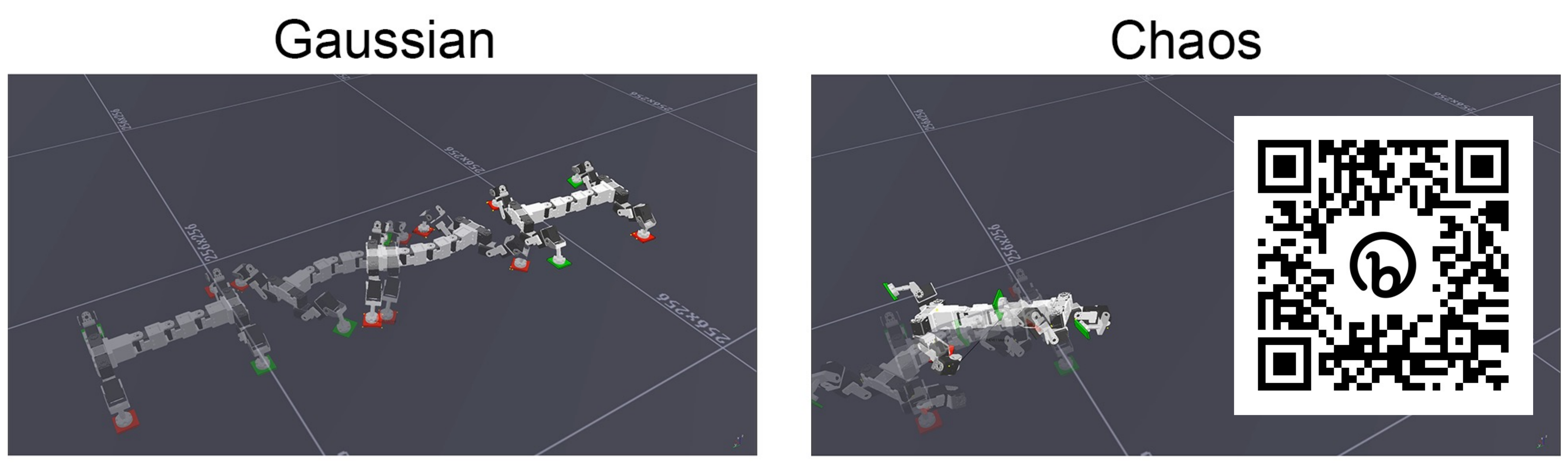

Figure 3: The robot locomotion performance after learning. Scan QR Code to see a video.

Conclusion & Future work

- Our investigation reveals that chaos cannot be directly utilized as exploration noise in BBO for locomotion learning.
- Although chaotic noise fails for locomotion learning here, it seems to facilitate learning speed (i.e., it can adapt parameters faster than Gaussian noise). Therefore, we will further explore an alternative strategy that uses chaotic dynamics to accelerate the overall optimization process of BBO with Gaussian noise for fast and stable locomotion learning.

References

1. Faure P, Korn H (2001) Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. C R Acad Sci III.
2. Maye A, Hsieh CH, Sugihara G, Brembs B. Order in spontaneous behavior. PloS one, 2(5):e443.
3. M. Thor, T. Kulvicius and P. Manoonpong, "Generic Neural Locomotion Control Framework for Legged Robots," in IEEE Transactionson Neural Networks and Learning Systems, vol. 32, no. 9, pp. 4013-4025, Sept. 2021.

Acknowledgements

This work was supported by the National Key R\&D Program of China (2020YFB1313504)[PM]. We thank the CM labs for providing Vortex.