PFPN: Continuous Control of Physically Simulated Characters using Particle Filtering Policy Network

Pei Xu, Ioannis Karamouzas
Clemson University

Introduction

Contribution: We propose *Particle Filtering Policy Network* (PFPN) as a substitute to Gaussian-based action policy network for improving control for physically simulated characters using reinforcement learning.

PFPN

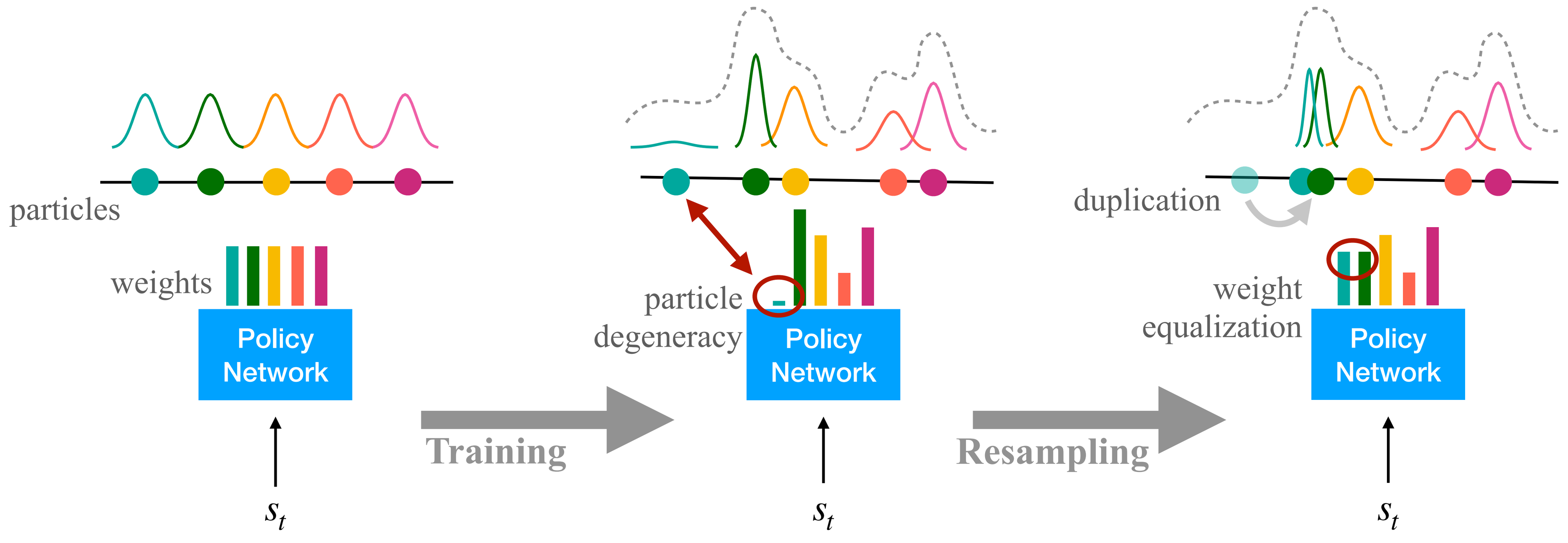
— can help better action space exploration, leading to sampling efficient training and exhibiting better robustness when applied to control tasks for highly-articulated characters with many degrees of freedom.

— is a general approach without changing the underlying architecture of training models or learning algorithms, and applicable to both on-policy and off-policy reinforcement algorithms.

Approach

PFPN employs a fixed number of particles (color dots) on each action dimension. Particles have learnable locations optimized through sampling from unimodal Gaussians. The policy network learns the weights through which the state-independent components represented by particles are mixed together.

During training, particles can move along their action dimensions towards different direction and provide an expressive, multi-modal distribution over the continuous space.

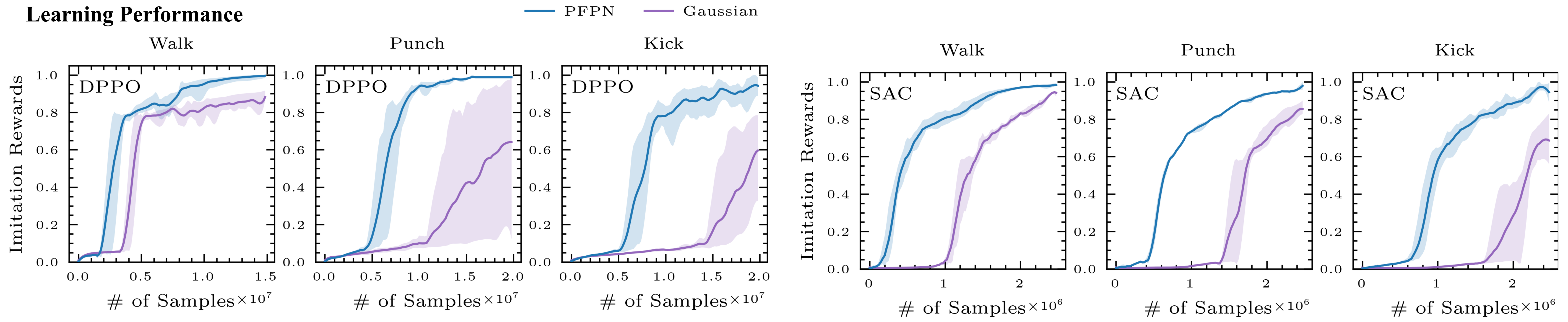


Resampling. To address the issue of particle degeneracy, where the associated weight of a particle decrease to near zero and thus have no chance to be activated any more, we adopt the resampling idea from particle filtering literature, and reactivate the degenerate particle by duplicating a target particle, which is chosen by resampling from those alive ones.

Our resampling strategy adds small noise when performing duplication and, therefore, brings about more diversity for better action space exploration. Meanwhile, by weight equalization, we guarantee that the overall policy distribution would not change too much before and after resampling, thus guaranteeing the training stability.

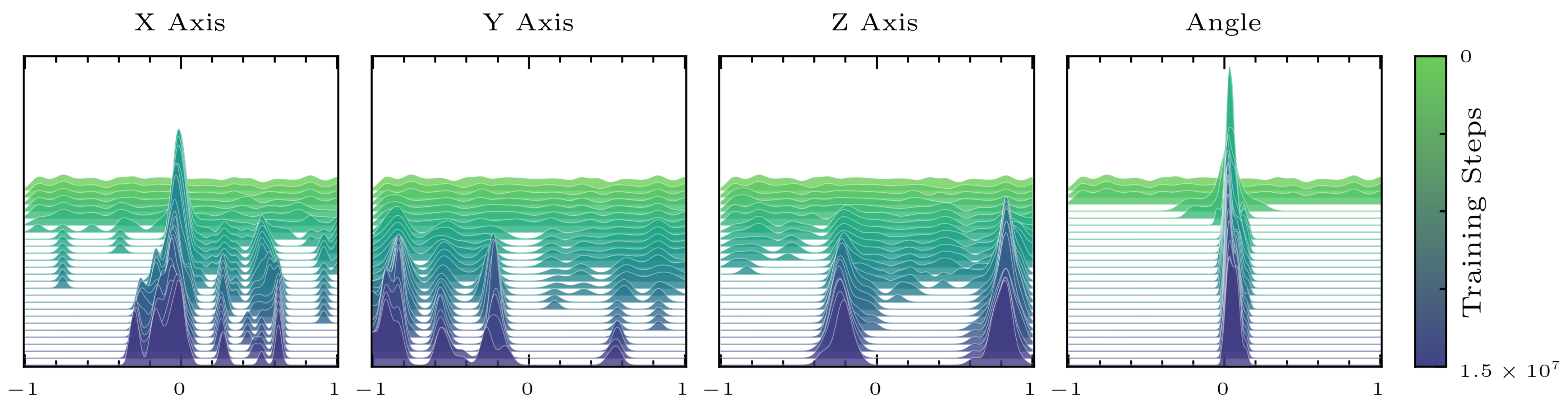
Results

Learning Performance

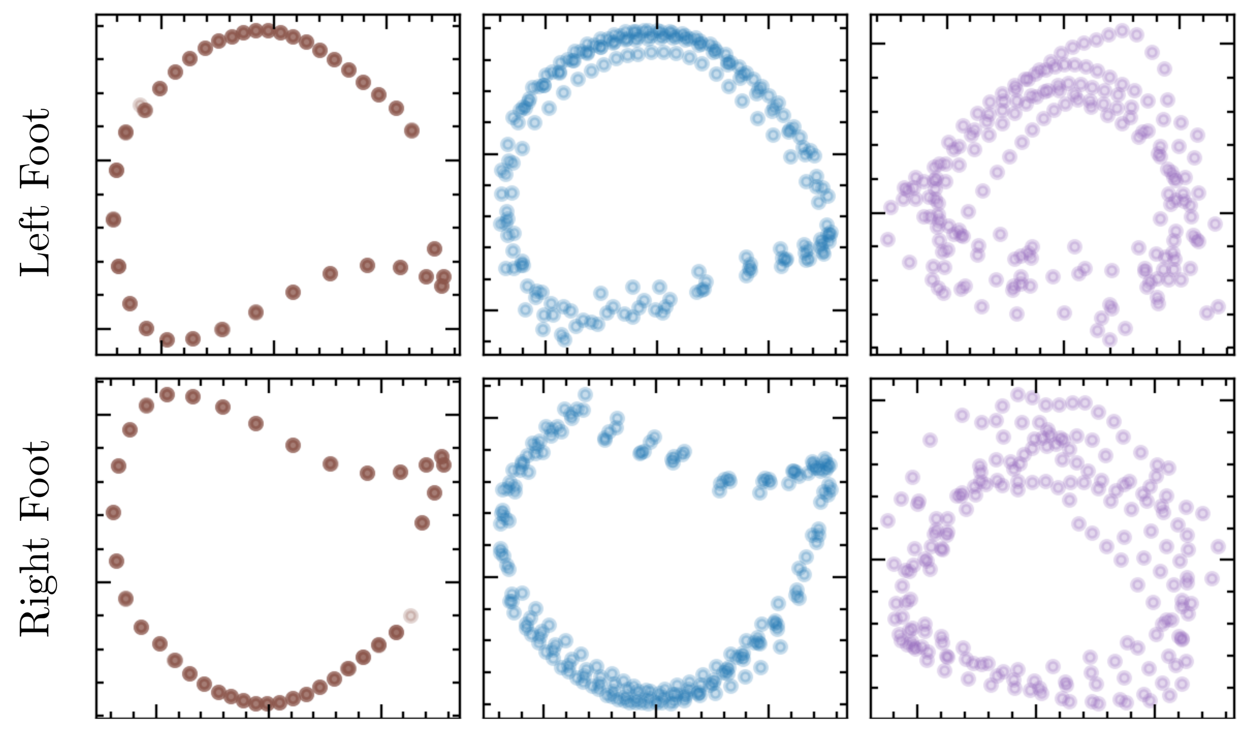


Particle Evolution during training

Motion: Walk Joint: Right Hip using Axis-Angle Representation

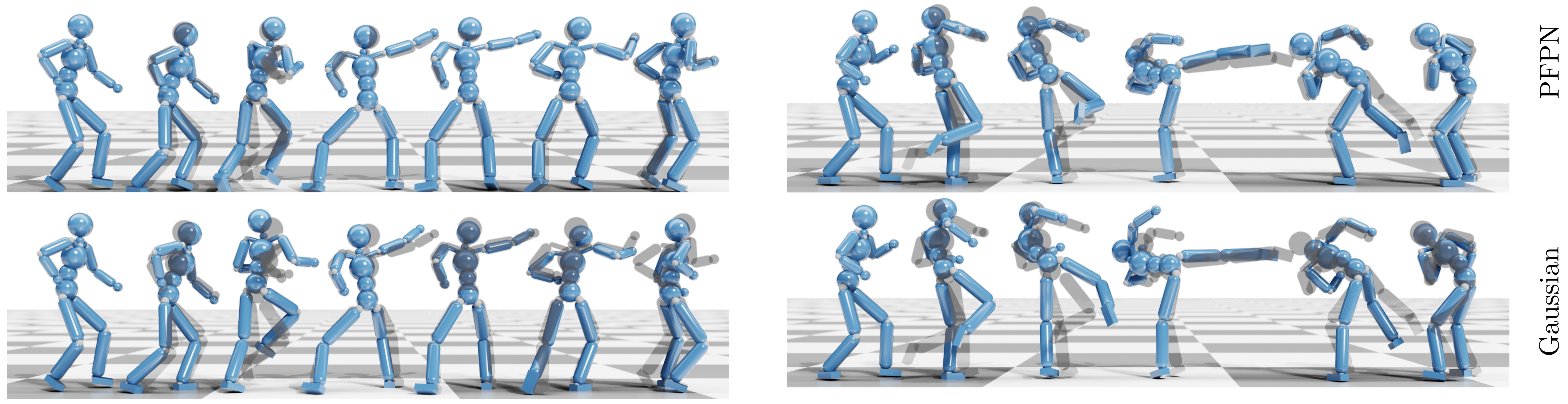


Foot Trajectories during walking



Reference PFPN Gaussian

Qualitative Comparison



Benchmark: DeepMimic DRL Algorithm: DPPO Shadow characters identify the reference motion.

Conclusion

PFPN is a great alternative to Gaussian-based action policies for physically simulated character control tasks. It exhibits better learning performance and higher motion quality in both on-policy and off-policy algorithms and various tasks.

Code Available on

https://motion-lab.github.io/PFPN

Acknowledgement

This work was supported in part by the National Science Foundation under Grant No. IIS-2047632.

NEURAL INFORMATION PROCESSING SYSTEMS