Iterated Population Based Training with Task-Agnostic Restarts

Alexander Chebykin, Tanja Alderliesten, Peter Bosman

Hyperparameter (HP) optimization

Static

- Search for fixed HPs
- Bayesian optimization (e.g., HEBO), ASHA, SMAC, …

Dynamic

- Search for a schedule of HPs
- Bandit optimization, hypergradients, **Population Based Training** (PBT) variants, …

Cowen-Rivers, Alexander I., et al. 2022. HEBO: Pushing The Limits of Sample-Efficient Hyper-parameter Optimisation. Journal of Artificial Intelligence Research 74 (July): 1269–1349.
Li, Liam, et al. 2020. A System for Massively Parallel Hyperparameter Tuning. In Proceedings of the Third Conference on Machine Learning and Systems, MLSys 2020
Lindauer, Marius, et al. 2022. SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization. Journal of Machine Learning Research 23 (54): 1–9.

Badia, Adrià Puigdomènech, et al. 2020. Agent57: Outperforming the Atari Human Benchmark. In Proceedings of the 37th International Conference on Machine Learning, ICML 2020
Baydin, Atilim Gunes, et al. 2018. Online Learning Rate Adaptation with Hypergradient Descent. In 6th International Conference on Learning Representations, ICLR 2018
Jaderberg, Max, et al. 2017. Population Based Training of Neural Networks. arXiv: 1711.09846

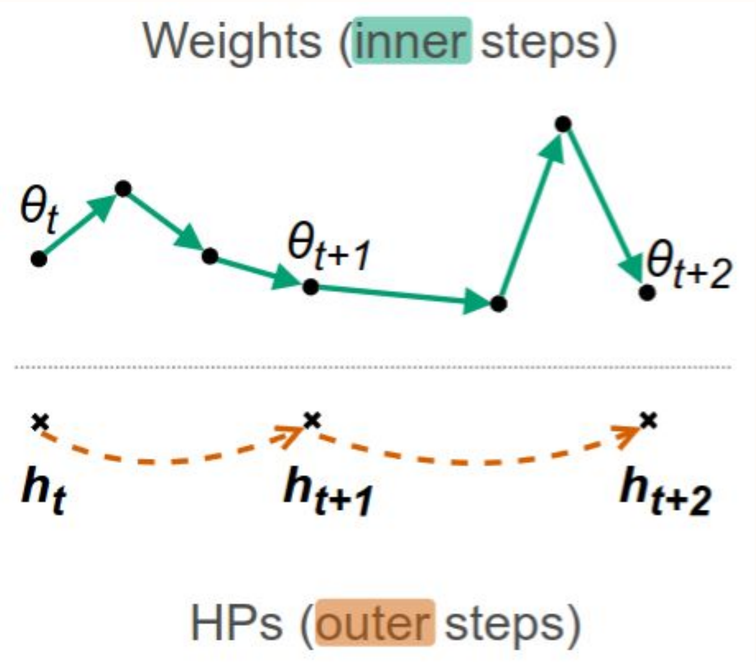
What makes PBT interesting?

- PBT optimizes a schedule of hyperparameters, not a fixed value
- PBT can run within the wall-clock time of a single training of a model
- PBT is general: not restricted to a specific hyperparameter or setting

PBT variants:
Jaderberg, Max, et al. 2017. Population Based Training of Neural Networks. arXiv: 1711.09846
Parker-Holder, Jack, et al. 2020. Provably Efficient Online Hyperparameter Optimization with Population-Based Bandits. In Advances in Neural Information Processing Systems, 33: 17200–17211
Dalibard, Valentin, & Max Jaderberg. 2021. Faster Improvement Rate Population Based Training. arXiv:2109.13800 [cs, stat]
Parker-Holder, Jack, et al. 2021. Tuning Mixed Input Hyperparameters on the Fly for Efficient Population Based AutoRL. NeurIPS 2021
Wan, Xingchen, et al. 2022. Bayesian Generational Population-Based Training. In International Conference on Automated Machine Learning, AutoML 2022
Doulazmi, Wael, et al. Multiple-Frequencies Population-Based Training. Reinforcement Learning Journal (2025).

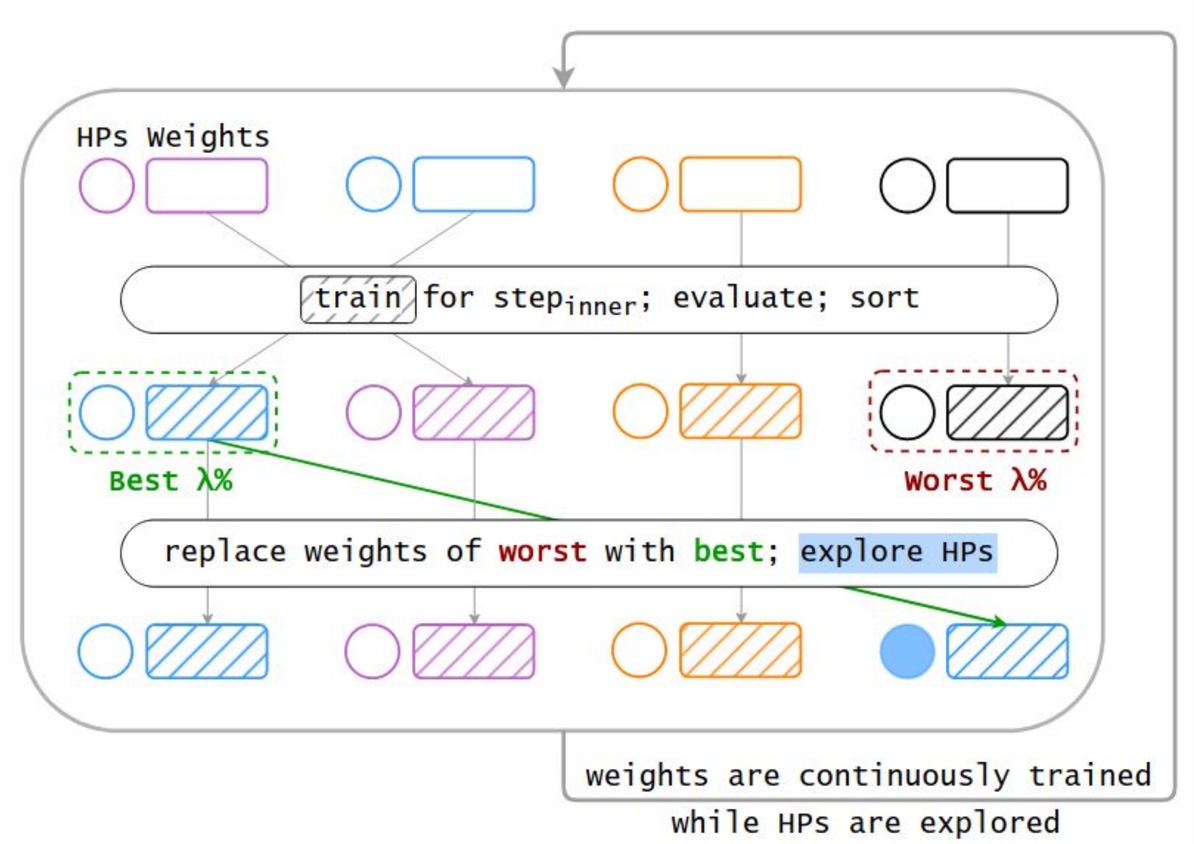
How PBT works

- Key idea 1: optimize hyperparameter during training
  - Update weights by making inner steps (e.g., SGD)
  - Update hyperparameters
  - Repeat



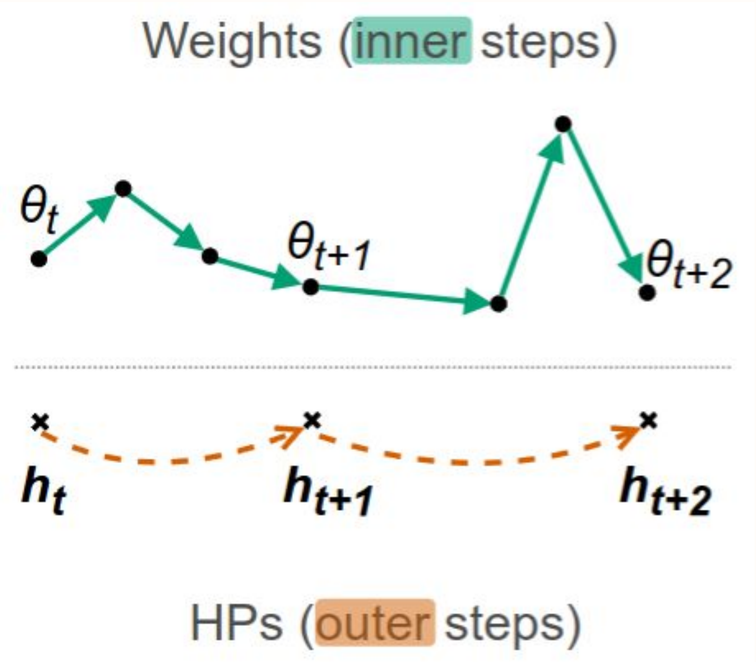
How PBT works

- Key idea 2: leverage a population
  - Many pairs of (HPs, weights) trained in parallel
  - Remove poor HPs & weights
  - Propagate good weights, explore HPs
    - via random perturbation (or Bayesian Optimization (BO) in later variants)



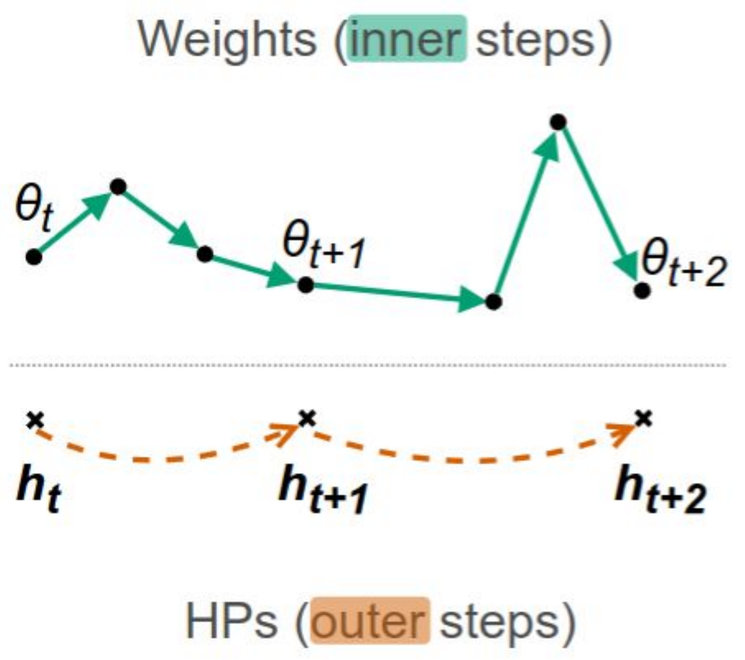
Weakness of PBT variants - setting outer step size

- Previous work: outer step size can strongly impact performance of PBT variants
- If several step sizes have to be tried, the benefits of the PBT variants diminish



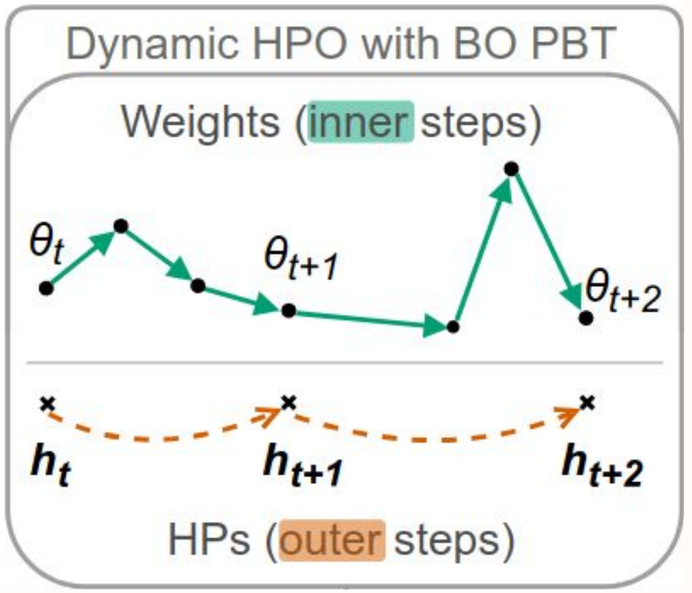
Weakness of PBT variants - setting outer step size

- Previous work: outer step size can strongly impact performance of PBT variants
- If several step sizes have to be tried, the benefits of the PBT variants diminish
- We introduce Iterated PBT (IPBT) to address this



IPBT

- Built atop the BO approach of BG-PBT

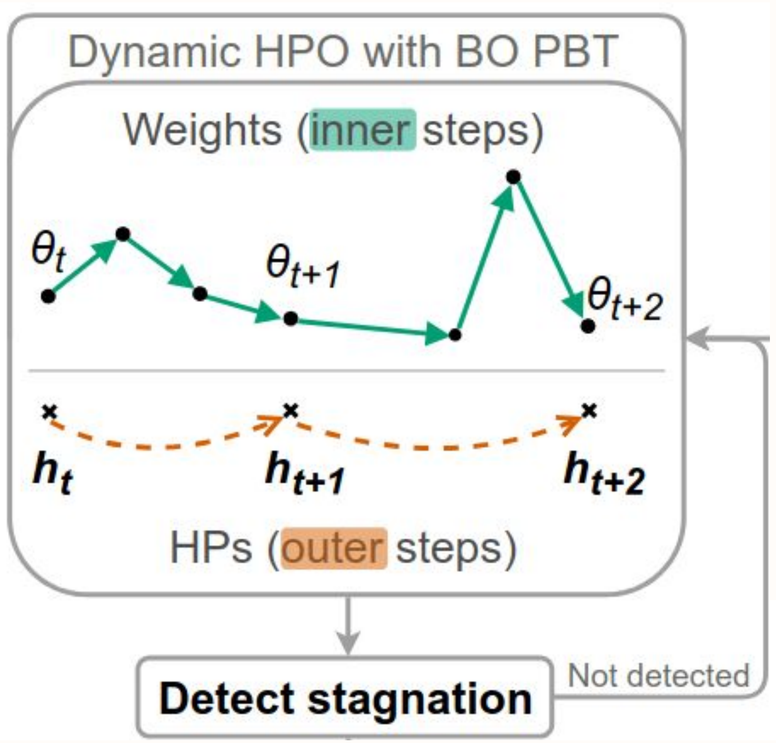


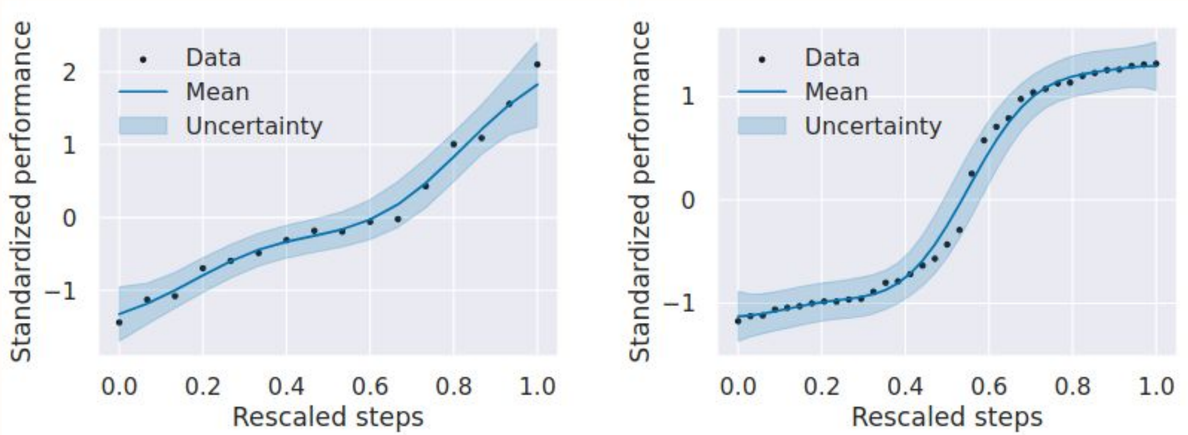
Wan, Xingchen, et al. 2022. Bayesian Generational Population-Based Training. In International Conference on Automated Machine Learning, AutoML 2022

IPBT

- Built atop the BO approach of BG-PBT

- Run until stagnation is detected
  - smooth & normalize
  - criteria to check if improvement is absent or too slow
    - (data-driven, task-agnostic)

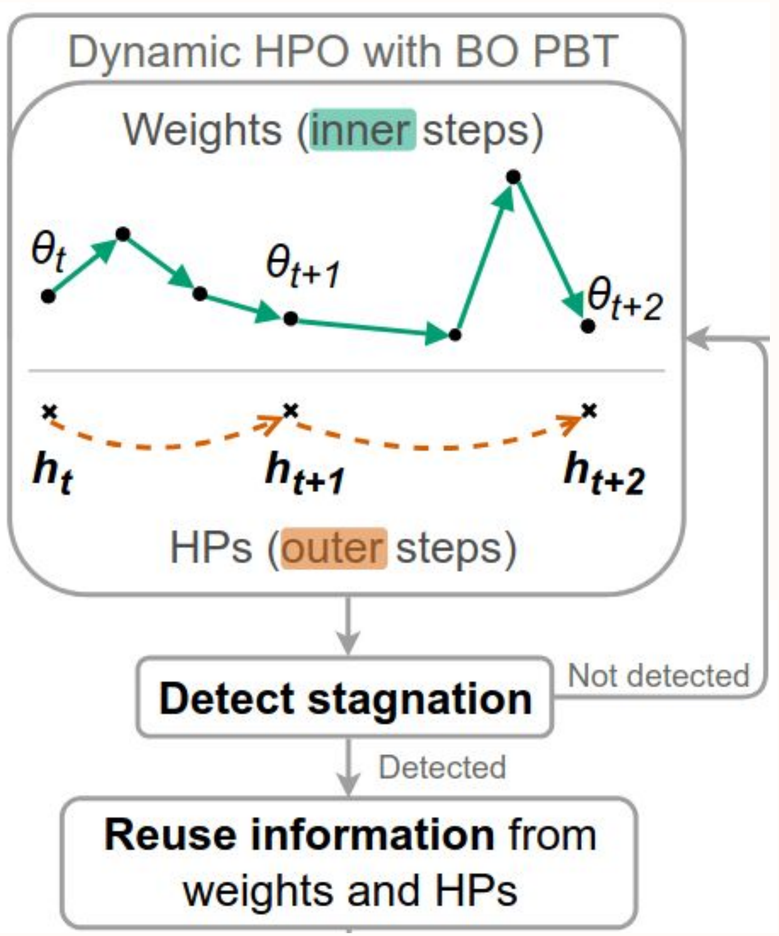




Example trajectories: (a) improving (b) stagnant

IPBT

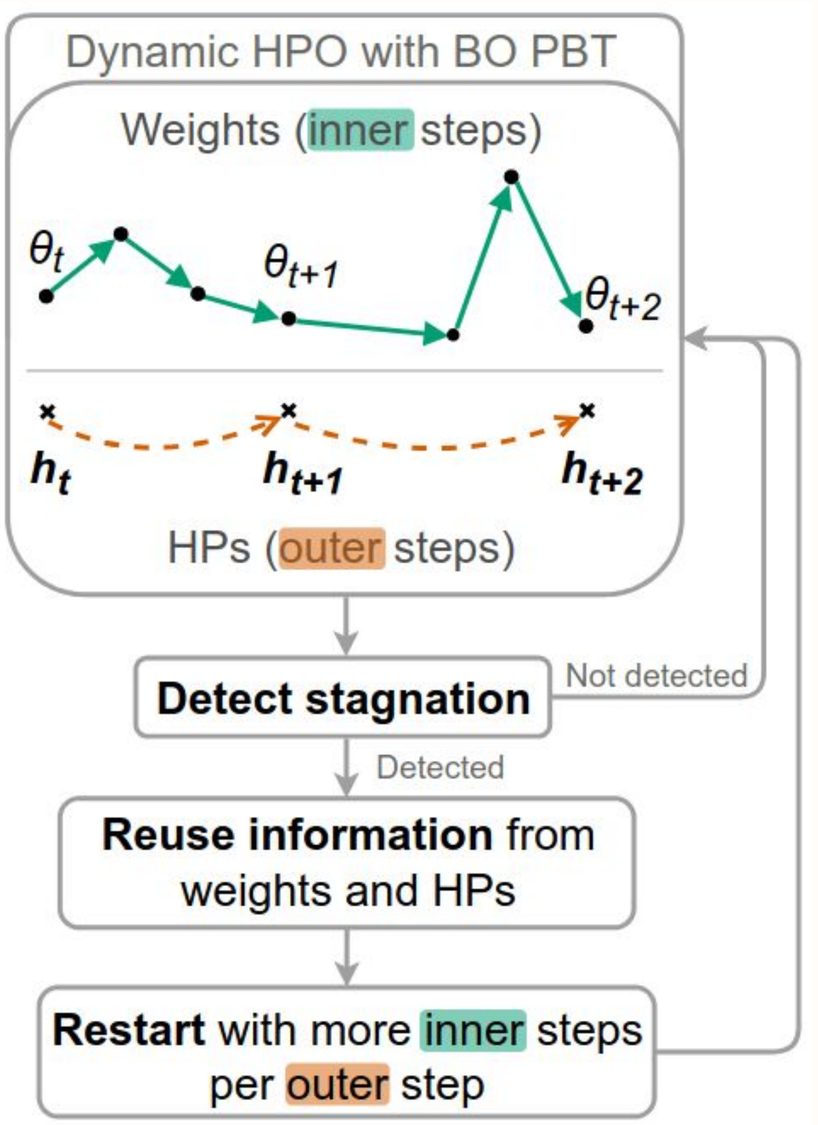
- Built atop the BO approach of BG-PBT
- Run until stagnation is detected
- Reuse information
  - Weights: shrink-perturb
  - HPs: meta-BO on the level of restarts
  - (and partial random reinit for both)



Ash, Jordan T., & Ryan P. Adams. 2020. On Warm-Starting Neural Network Training. In NeurIPS 2020

IPBT

- Built atop the BO approach of BG-PBT
- Run until stagnation is detected
- Reuse information
- Restart



Experimental setup

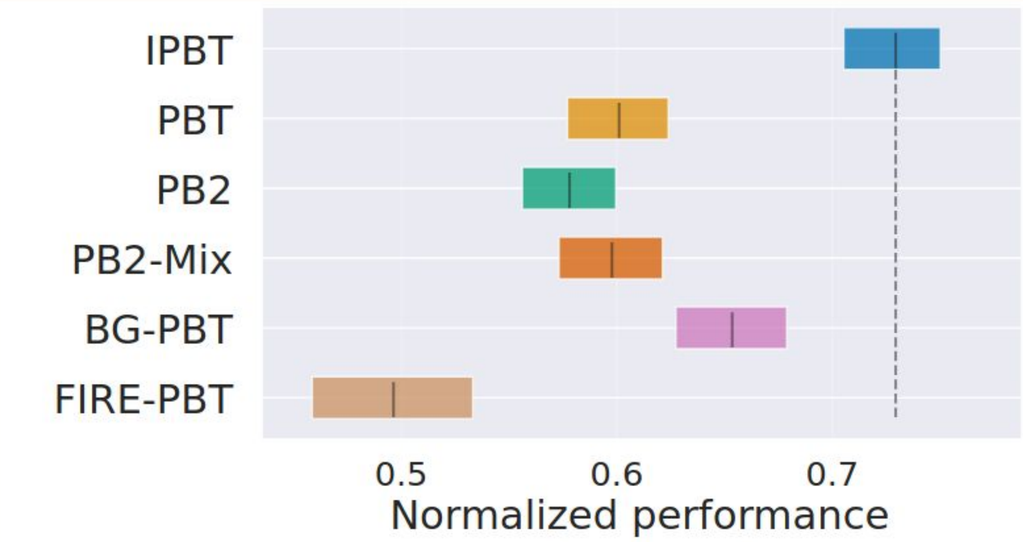
- Follow the previous setup used for benchmarking PBT variants
- Classification: 4 tasks (ResNet-12 with CIFAR-10/100, Fashion-MNIST; ConvNeXt-T with Tiny ImageNet); 5 HPs
- Reinforcement Learning (RL): 4 tasks (PPO on Brax Humanoid, Hopper, Pusher, Walker); 10 HPs
- Budget: 8 full training runs
- 8 seeds

He, Kaiming, et al. 2016. Deep Residual Learning for Image Recognition. In CVPR 2016
Liu, Zhuang, et al. 2022. A ConvNet for the 2020s. In CVPR 2022
Schulman, John, et al. 2017. Proximal Policy Optimization Algorithms. arXiv:1707.06347 [cs]
Freeman, C. Daniel, et al. 2021. Brax - A Differentiable Physics Engine for Large Scale Rigid Body Simulation. In NeurIPS Datasets and Benchmarks 2021

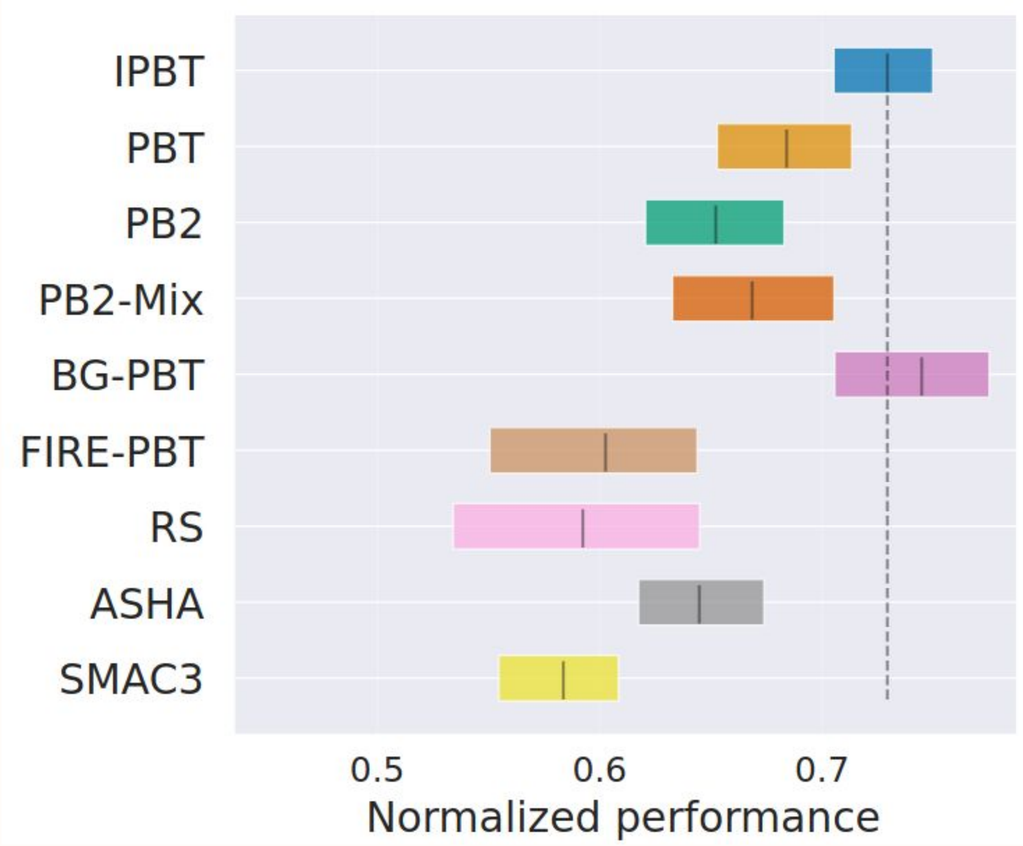
IPBT outperforms/matches baselines

- Inter-quartile mean, inter-quartile ranges across all tasks and seeds

Untuned PBT variants (same cost as IPBT):

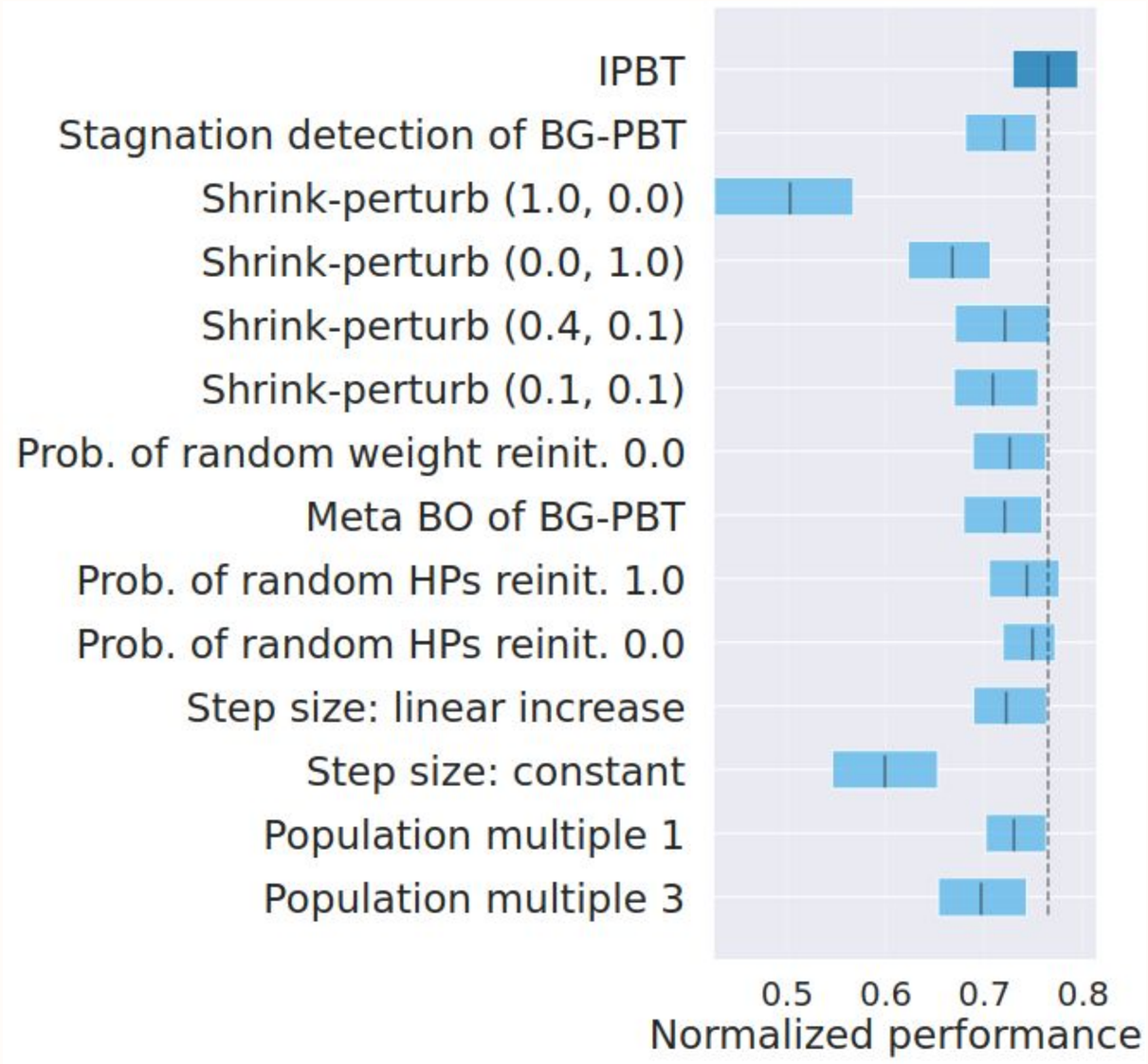


Non-PBT baselines &
tuned PBT variants (4x cost of IPBT):



Agarwal, Rishabh, et al. 2021. Deep Reinforcement Learning at the Edge of the Statistical Precipice. In NeurIPS 2021

IPBT ablations: components matter



IPBT discovers schedules for HPs beyond learning rate