# Overcoming the Spectral-bias of Neural Value approximation

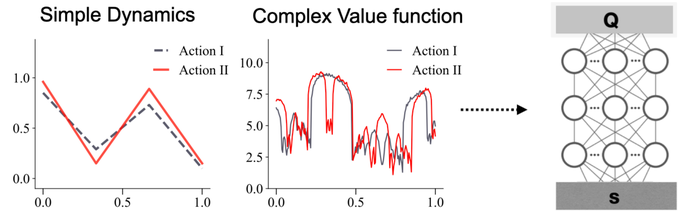
Ge Yang*, Anurag Ajay*, Pulkit Agrawal

Improbable AI Lab, NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI), MIT CSAIL (* denotes equal contribution)

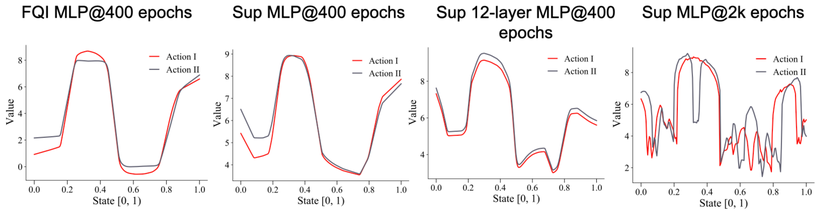
## *Spectral-bias* in MLPs slows convergence in Neural Value approximation and causes instability in off-policy RL

## A Motivating Example

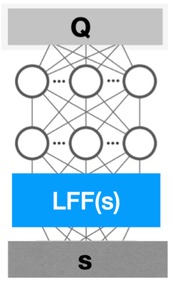
Toy MDP (Dong et al., 2020)




We learn the value function using a MLP with Fitted-Q-iteration (FQI)




MLPs have *spectral-bias* towards low-frequency function which slows down value function convergence !!




Fourier Feature Network (FFN)

Learned Fourier feature layer

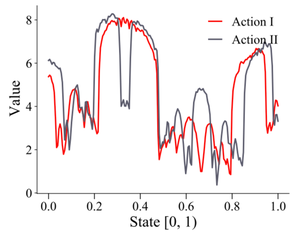
$$\mathrm{LFF}(x) = \sin(Wx + c)$$

$$W \sim \mathcal{N}\left(0, \frac{b}{\dim(x)}\right)$$

$$c \sim \mathcal{U}\left(-\pi, \pi\right)$$

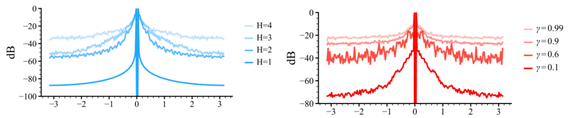
Drop-in replacement for standard MLP layer

FQI FFN@400 epochs

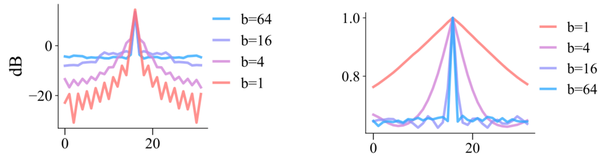



LFF layer helps MLPs evade the *spectral-bias*

## *Spectral-bias* in Value estimation




Optimal value function's spectrum for the toy MDP noticeably whitens over longer horizon (H) and larger discount factor ($\gamma$)




FFN kernel Spectrum

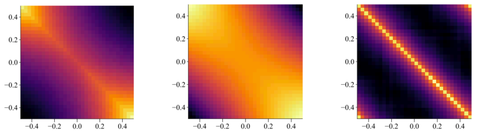
NTK cross-section

FFN kernel spectrum has larger eigenvalues on the high-frequency harmonics, thus improving convergence rate of the network

## Analyzing off-policy stability via NTK

*Cross-talks* between the gradients of near-by states, as measured by NTK, leads to off-policy divergence (Achiam et al., 2019)

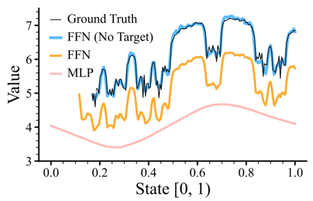
Neural Tangent Kernel (NTK) comparison



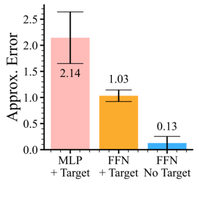
MLP+ReLU MLP+TanH FFN

The NTK of FFN is *localized* with small off-diagonal elements showing less *cross-talk*.

Neural Fitted Q Iteration

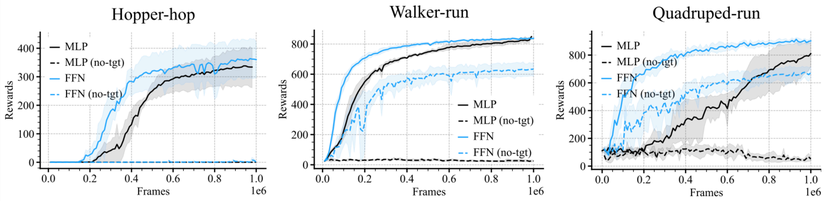



Bias & Variance

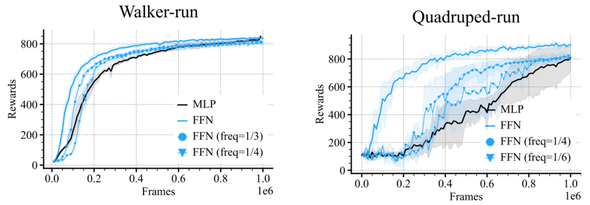



The improved learning stability allows us to *remove the target network* on a number of domains while retaining substantial performance.

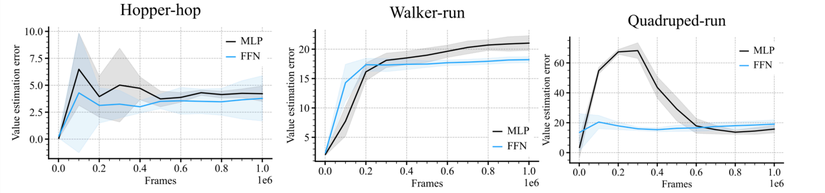
## Scaling up to Deepmind Control Suite



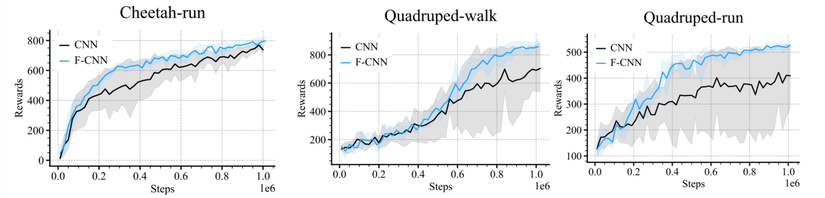

FFN leads to faster convergence than SAC and improves stability thereby retaining substantial performance on some tasks despite removing target-network




FFN improves compute efficiency and matches MLP's performance with only a fraction of compute




FFN consistently produces smaller value approximation error than MLP




FFN can be extended to image domains in form of Fourier CNN (F-CNN)