A Novel Local Analysis of Objectives Approximated by Neural Network: L-Change

Pengcheng Xie, Zihao Zhou, and Zijian Zhou

Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory (P. Xie, Email: pxie@lbl.gov)
School of Mathematics and Statistics, Xi'an Jiaotong University (Z. Zhou, Email: wszzh15139520600@stu.xjtu.edu.cn)
School of Mathematical Sciences, Beijing Normal University (Z. Zhou, Email: zzj2023@bnu.edu.cn)

Abstract

We introduce a novel metric, L-Change, for characterizing the local performance and behavior of neural network approximation. Neural networks, though powerful, often suffer from unpredictable local performance, which can hinder their reliability in critical applications. L-Change addresses this issue by providing a quantifiable measure of local changes in network behavior, offering insights into the stability and performance for achieving the neural approximations. We conduct some basic properties of L-Change, demonstrating its rationality. Our numerical experiments further validate these findings, showing that L-Change can effectively identify regions where the network's performance is suboptimal, thus guiding improvements in network design and training. In addition to presenting these results, we discuss several open problems related to the application and extension of L-Change. These include the Average version of L-Change (Ave-Change), and L-Change Density function, for the objective function, and more needs related to other complex objective functions. We conclude by outlining future research directions, aiming to encourage further exploration and collaboration in this emerging area.

Introduction

Approximation is a common method in numerical algorithms. In recent years, with the enhancement of the computational power of computing devices, neural networks (NNs) have sparked an amount of interest in approximation. With the emergence, NNs, which have been in existence for over 70 years, have recently gained significant popularity due to their state-of-the-art performance in various machine learning domains and numerical mathematics including optimization, et cl. [1,2,3,4].

Definitions of L-Change

We are considering the real objective function $f : \Re \to \Re$ in the neural network approximation problem in the interval $[a, b] \subset \Re$, where $a \neq b$. Notice that the discussion is similar for the cases where f is defined in $\Re^n$ or C^n.

Definition 1. (L-Change on x of f)
For $\forall\, x \in [a + \frac{L}{2}, b - \frac{L}{2}]$, the L-Change on x of the objective function f is defined as

$$\Lambda_L(f, x) = \sup_{y_1, y_2 \in [x-\frac{L}{2}, x+\frac{L}{2}]} |f(y_1) - f(y_2)|. \tag{1}$$

Remark. If f is bounded on $[x - \frac{L}{2}, x + \frac{L}{2}]$, then $\Lambda_L(f, x)$ is bounded.

Definition 2. (L-Change Density function)
The L-Change density function of f is defined as

$$\mathcal{P}(p) = \Pr\left(\Lambda_L(f, x) = p\right), \tag{2}$$

where $\Pr(\cdot)$ is the Probability density function.

Definition 3. (Ave-Change of f)
For $\forall\, x \in [a + L, b + L]$, the L-Change on x of f is defined as

$$\Lambda_{\text{Ave}}(f, x) = \int_0^{L_{\max}} \Lambda_L(f, x) \mathrm{d}L. \tag{3}$$

Theoretical Properties of L-Change

The L-Change $\Lambda_L(f)$ has the following properties:

Theorem 1. (Affine in-variance respective to f)
Given a function f with L-Change $\Lambda_L(f, x)$ for $\forall\, x$, the L-Change of the function kf on x is

$$\Lambda_L(kf + c, x) = |k|\Lambda_L(f, x),\ \forall\, k, c \in \Re. \tag{4}$$

Theorem 2. (Uniqueness of L-Change)
L-Change of f on x is unique.

Theorem 3. (Symmetric in-variance respective to x)
If a function $f(\boldsymbol{x})$ is symmetric respective to $x = d \in \Re$, then

$$\Lambda_L(f, x) = \Lambda_L(f, 2d - x). \tag{5}$$

References

[1] Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359-366.

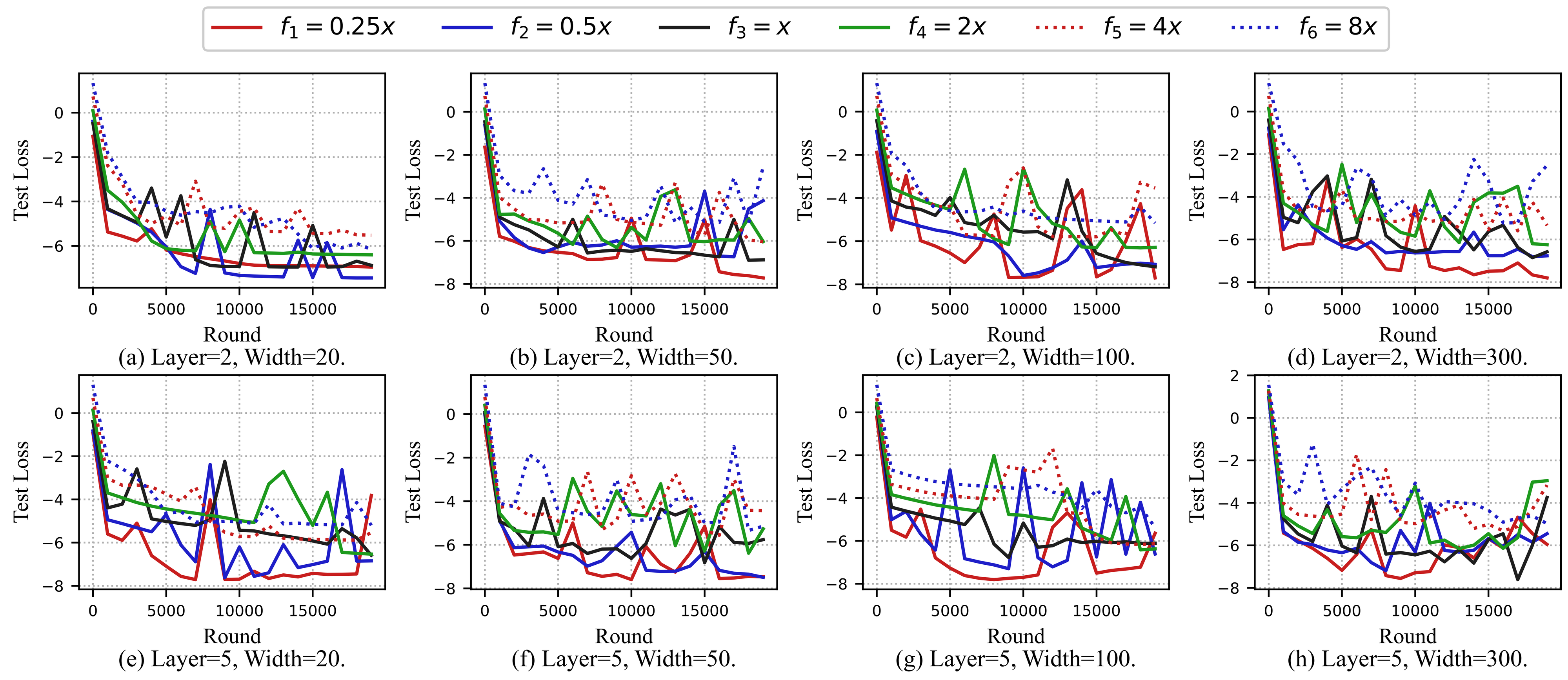
[2] Larson, J., Menickelly, M., Wild, S. M. (2019). Derivative-free optimization methods. Acta Numerica, 28, 287-404.

[3] DeVore, R., Hanin, B., Petrova, G. (2021). Neural network approximation. Acta Numerica, 30, 327-444.

[4] Xu, Z. Q. J. (2020). Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks. Communications in Computational Physics, 28(5), 1746-1767.

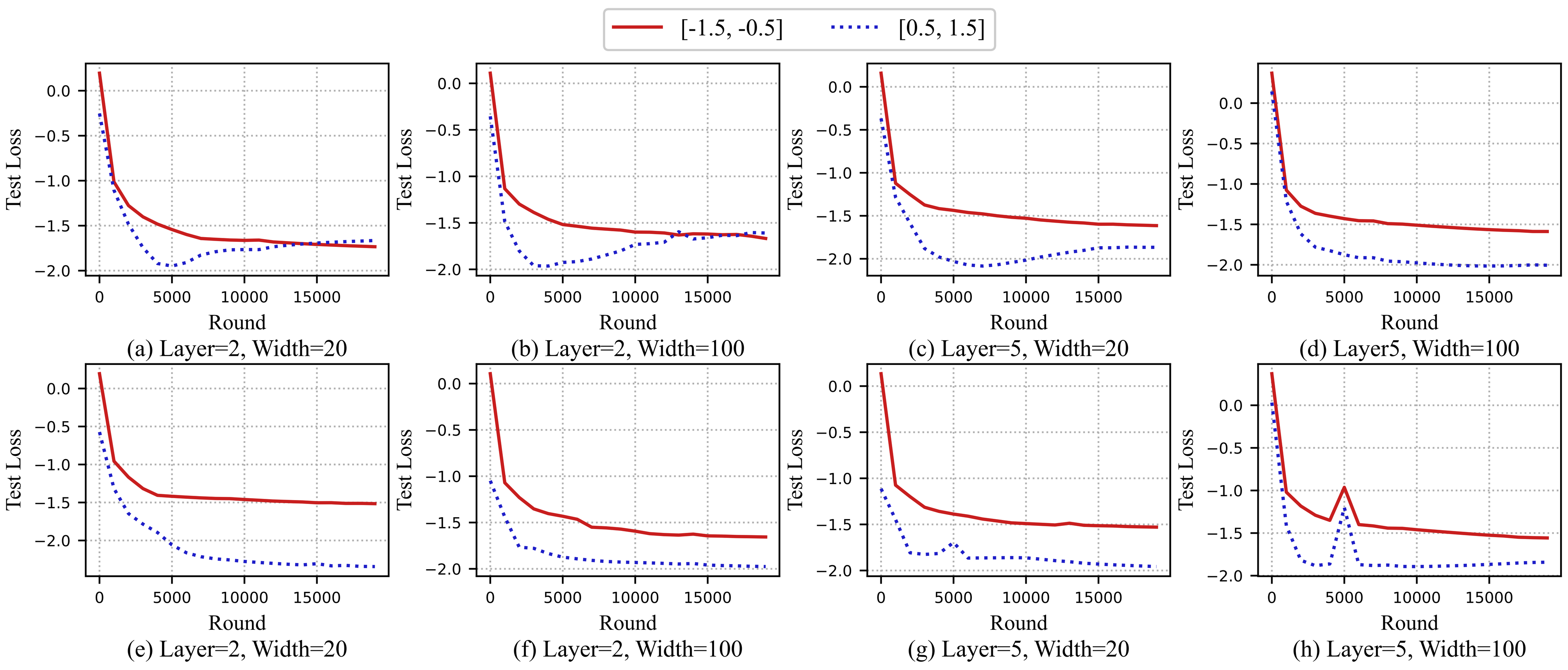
Numerical Results: L-Change $\uparrow$, Convergence of approximation $\downarrow$

Approximate to linear functions: $f(x) = f_i(x),\ \forall\, x \in [-1, 1],\ i = 1, \cdots, 6.$



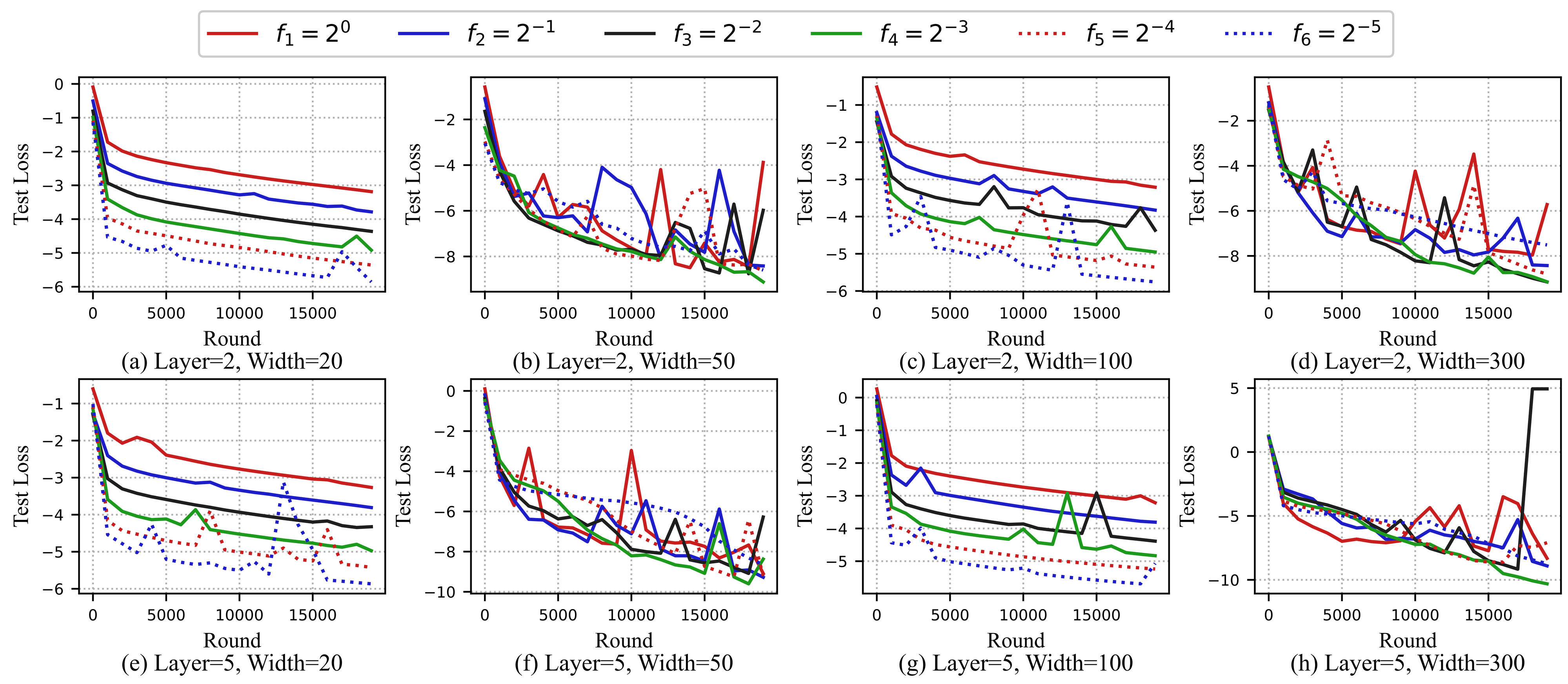
The loss function curve of the approximation to linear functions using NNs with different parameters. According to the affine in-variance, $\Lambda_L(f_1, x) = \frac{1}{2}\Lambda_L(f_2, x) = \frac{1}{4}\Lambda_L(f_3, x) = \frac{1}{8}\Lambda_L(f_4, x) = \frac{1}{16}\Lambda_L(f_5, x) = \frac{1}{32}\Lambda_L(f_6, x)$

Approximate to piecewise function: $f(x) = \begin{cases} 2x + 2,\ \forall\, x \in [-2, 0], \\ -2x + 2 \text{ or } -x + 4,\ \forall\, x \in [0, 2]. \end{cases}$



The loss function curve of the approximation to piecewise functions using NNs with different parameters.

Approximate to step function: $f(x) = \begin{cases} 0,\ \forall\, x \in [-1, 0], \\ f_i(x),\ \forall\, x \in [0, 1],\ i = 1, \cdots, 6. \end{cases}$



The loss function curve of the approximation to step functions using NNs with different parameters.