

--- input url: 在这里输入pdf的链接

生成

--- output url:

Copy

输入PDF的链接

输出在线预览的链接

可以复制到剪贴板

Training behavior of deep neural network in frequency domain

v:1807.01251v6 [cs.LG] 1 Nov 2019

# Training behavior of deep neural network in frequency domain<sup>\*</sup>

Zhi-Qin John Xu<sup>\*\*1</sup>, Yaoyu Zhang<sup>2</sup>, and Yanyang Xiao<sup>2</sup>

<sup>1</sup> School of Mathematical Sciences and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China,  
xuzhiqin@sjtu.edu.cn,

<sup>2</sup> School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA,  
yaoyu@ias.edu,

<sup>3</sup> The Brain Cognition and Brain Disease Institute, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China,  
xyy82148@gmail.com

**Abstract.** Why deep neural networks (DNNs) capable of overfitting often generalize well in practice is a mystery [24]. To find a potential mechanism, we focus on the study of implicit biases underlying the training process of DNNs. In this work, for both real and synthetic datasets, we empirically find that a DNN with common settings first quickly captures the dominant low-frequency components, and then relatively slowly captures the high-frequency ones. We call this phenomenon Frequency Principle (F-Principle). The F-Principle can be observed over DNNs of various structures, activation functions, and training algorithms in our experiments. We also illustrate how the F-Principle helps understand the effect of early-stopping as well as the generalization of DNNs. This F-Principle potentially provides insight into a general principle underlying DNN optimization and generalization.

**Keywords:** Deep Neural Network · Deep Learning · Fourier analysis · Generalization.

## 1 Introduction