Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2020-0179

Vol. **28**, No. 5, pp. 1970-2001
November 2020

Multi-Scale Deep Neural Network (MscaleDNN) for Solving Poisson-Boltzmann Equation in Complex Domains

Ziqi Liu[1], Wei Cai[2] and Zhi-Qin John Xu[3,*]

[1] *Beijing Computational Science Research Center, Beijing, 100193, P.R. China.*
[2] *Department of Mathematics, Southern Methodist University, Dallas, TX 75275, USA.*
[3] *Institute of Natural Sciences, MOE-LSC, School of Mathematical Sciences and Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai, 200240, P.R. China.*

Received 11 September 2020; Accepted (in revised version) 25 September 2020

Summary. In this paper, we propose multi-scale deep neural networks (MscaleDNNs) using the idea of radial scaling in frequency domain and activation functions with compact support. The radial scaling converts the problem of approximation of high frequency contents of PDEs' solutions to a problem of learning about lower frequency functions, and the compact support activation functions facilitate the separation of frequency contents of the target function to be approximated by corresponding DNNs. As a result, the MscaleDNNs achieve fast uniform convergence over multiple scales. The proposed MscaleDNNs are shown to be superior to traditional fully connected DNNs and be an effective mesh-less numerical method for Poisson-Boltzmann equations with ample frequency contents over complex and singular domains.

AMS subject classifications: 35Q68, 65N99, 68T07

Key words: Deep neural network, Poisson-Boltzmann equation, multi-scale, frequency principle.

---

1 Introduction

Deep neural network (DNN) has found many applications beyond its traditional applications such as image classification and speech recognition into the arena of scientific computing [10–15, 17, 22, 24, 25]. However, to apply the commonly-used DNNs to computational science and engineering problems, we are faced with several challenges. The most prominent issue is that the DNN normally only handles data with low frequency

*Corresponding author. *Email addresses:* `liuziqi@csrc.ac.cn` (Z. Liu), `cai@smu.edu` (W. Cai), `xuzhiqin@sjtu.edu.cn` (Z.-Q. J. Xu)

content well, as shown by a Frequency Principle (F-Principle) that many DNNs learn the low frequency content of the data quickly with a good generalization error, but they are inadequate when high frequency data are involved [21, 27, 28]. The fast convergence behavior of low frequency has been recently studied rigorously in theory in [2, 6, 19, 30]. As a comparison, such a behavior of DNNs is the opposite of that of the popular multi-grid methods (MGM) for solving PDEs such as the Poisson-Boltzmann (PB) equation, where the convergence is achieved first in the high frequency spectrum of the solution due to the smoothing operations employed in the MGM. Considering the potential of DNNs in handling higher dimensional solutions and approximating functions without the need of a structured mesh as in traditional finite element or finite difference method, it is of great value to extend the capability of DNN as a mesh-less PDE solver. Therefore, it is imperative to improve the convergence of DNNs for solutions with fine structures as encountered in the electrostatic potentials of complex molecules.

The electrostatic interaction of bio-molecules with ionic solvents, governed by the Poisson-Boltzmann (PB) equation within the Debye-Huckel theory [3], plays an important role in many applications including drug design and the study of disease. However, due to the complex surface structure of the bio-molecules, usually represented by a bead model, it has been a long outstanding challenging to design efficient numerical method to handle the singular molecular surface, which is either the van der Waals (vdW) surface being the sum of overlapping vdW spheres or the solvent accessible surface (SAS) generated by rolling a small ball on the vdW surface [18], and the complex distribution of the electrostatic potential over the molecular surfaces. Tradition finite element [1] and finite difference methods [29] have faced difficulties in the costly mesh generation and expensive solution of the discretized linear system. Therefore, in this paper, we will propose and investigate multi-scale DNNs, termed MscaleDNN, with the goal of approximating both low and high frequency information of a solution uniformly and developing a meshless solver for PDEs such as the PB equations in domains with complex and singular geometries.

Different learning behaviors among different frequencies are common. Leveraging this difference in designing neural network structure can benefit the learning process. In the field of computer vision, a series of works, such as image recovery [9], super-resolution [20], or classification [26], have improved the learning performance, including the generalization and training speed, by utilizing the learning difference of different image frequencies. However, it should be noted that the frequency used in the computer vision tasks, is different from the response frequency of a mapping from the input (e.g., image) to the output (e.g., label), and the former refers to the frequency within an input (i.e. an image) with respect to the spatial locations inside the image. In this work, we address different response frequencies of the mapping from the input to the output. As demonstrated in the previous work [28], the low response frequency is learned much faster than the high frequency. The main idea of the MscaleDNN is to find a way to convert the learning or approximation of high frequency data to that of a low frequency one. Similar idea has been attempted in a previous work in the development of a phase