מכון ויצמן למדע

THE WEIZMANN INSTITUTE OF SCIENCE

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Vision and Robotics Seminar

Room 1 ,Ziskind Building
on Thursday, Apr 12, 2018
at 12:15

Yoav Levine
The Hebrew University of Jerusalem

Bridging Many-Body Quantum Physics and Deep Learning via Tensor Networks

Abstract:

The harnessing of modern computational abilities for many-body wave-function representations is naturally placed as a prominent avenue in contemporary condensed matter physics. Specifically, highly expressive computational schemes that are able to efficiently represent the entanglement properties which characterize many-particle quantum systems are of interest. In the seemingly unrelated field of machine learning, deep network architectures have exhibited an unprecedented ability to tractably encompass the convoluted dependencies which characterize hard learning tasks such as image classification or speech recognition. However, theory is still lagging behind these rapid empirical advancements, and key questions regarding deep learning architecture design have no adequate answers. In the presented work, we establish a Tensor Network (TN) based common language between the two disciplines, which allows us to offer bidirectional contributions. By showing that many-body wave-functions are structurally equivalent to mappings of convolutional and recurrent arithmetic circuits, we construct their TN descriptions in the form of Tree and Matrix Product State TNs, and bring forth quantum entanglement measures as natural quantifiers of dependencies modeled by such networks. Accordingly, we propose a novel entanglement based deep learning design scheme that sheds light on the success of popular architectural choices made by deep learning practitioners, and suggests new practical prescriptions. Specifically, our analysis provides prescriptions regarding connectivity (pooling geometry) and parameter allocation (layer widths) in deep convolutional networks, and allows us to establish a first of its kind theoretical assertion for the exponential enhancement in long term memory brought forth by depth in recurrent networks. In the other direction, we identify that an inherent re-use of information in state-of-the-art deep learning architectures is a key trait that distinguishes them from TN based representations. Therefore, we suggest a new TN manifestation of information re-use, which enables TN constructs of powerful architectures such as deep recurrent networks and overlapping convolutional networks. This allows us to theoretically demonstrate that the entanglement scaling supported by state-of-the-art deep learning architectures can surpass that of commonly used expressive TNs in one dimension, and can

support volume law entanglement scaling in two dimensions with an amount of parameters that is a square root of that required by Restricted Boltzmann Machines. We thus provide theoretical motivation to shift trending neural-network based wave-function representations closer to state-of-the-art deep learning architectures.

Email: hanni.naor@weizmann.ac.il * Phone: (08)9343349 * Fax: (08)9344308