ACCELERATING MANY-BODY EXPANSION THEORY THROUGH GRAPH CONVOLUTIONAL NETWORKS

YILI SHEN, *College of Software Engineering, Tongji University, Shanghai, China*; CHENGWEI JU, *Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL, USA*; JUN YI, ZHOU LIN, *Department of Chemistry, University of Massachusetts, Amherst, MA, USA*; HUI GUAN, *College of Information and Computer Sciences, University of Massachusetts, Amherst, MA, USA*.

First-principles quantum mechanical modeling can potentially interpret and predict experimentally measurable properties of large molecules or systems, such as energies, provided that the difficulty in balancing its computational efficiency and accuracy is overcome. Many-body expansion theory (MBET) has been developed to resolve this issue: it approximates the total energy of a large system through a truncated expansion of one-, two-, ..., n-body energies, but it still suffers from a computational bottleneck, expensive first-principles evaluations of all many-body energies. In the present study, we integrated the graph convolutional network (GCN), a state-of-the-art machine learning (ML) algorithm, into the existing first-principles workflow, and developed a novel scheme referred to as GCN-MBET. Operationally, we evaluated all one-body energies using conventional first-principle quantum mechanics, but obtained many-body energies based on their relationships with effortless molecular descriptors established by GCN. As the initial stage of the study, we provided a proof-of-concept of our GCN-MBET model using two- and three-body energies from representative van der Waals or hydrogen-bonded molecular aggregates, including the water cluster, the phenol cluster, and water–phenol mixture. Given sufficient configurational diversity in the training set, we successfully reproduced first-principles two- and three-body energies in the test set to the chemical accuracy (< 1 kcal/mol), but at a fractional computational cost ($\simeq 1$ %). Our results indicated that GCN-MBET provides a promising unique and powerful tool to unlock the potential of first-principles quantum mechanical modeling of large molecules or systems.