MOAT: Graph Prompting for 3D Molecular Graphs

Qingqing Long
CNIC, CAS
UCAS
HIAS, UCAS
Beijing, China
qqlong@cnic.cn

Yuchen Yan
Peking University
Beijing, China
yanyuchen1998@gmail.com

Wentao Cui
CNIC, CAS
UCAS
HIAS, UCAS
Beijing, China
cuiwentao@cnic.cn

Wei Ju
Sichuan University
Sichuan, China
juwei@pku.edu.cn

Zhihong Zhu
Peking University
Beijing, China
zhihongzhu@stu.pku.edu.cn

Yuanchun Zhou
CNIC, CAS
UCAS
HIAS, UCAS
Beijing, China
zyc@cnic.cn

Xuezhi Wang*
CNIC, CAS
UCAS
HIAS, UCAS
Beijing, China
wxz@cnic.cn

Meng Xiao*
CNIC, CAS
UCAS
Beijing, China
shaow@cnic.cn

Abstract

Molecular property prediction stands as a cornerstone task in AI-driven drug design and discovery, wherein the atoms within a molecule serve as nodes, collectively forming a graph with bonds acting as edges. Given the crucial role of geometric structures in molecular property prediction, the integration of 3D information with various graph learning methods has been explored to enhance prediction performance. Despite the increasing adoption of the "Graph pre-training and fine-tuning" paradigm to refine molecular representations, a significant challenge persists due to the misalignment between pre-training objectives and downstream tasks. Drawing inspiration from prompt tuning techniques in Natural Language Processing (NLP), several graph prompt-based methods have emerged. However, existing approaches tend to overlook the unique properties inherent in molecular graphs. To address this gap, our paper introduces a novel approach named 3D **MO**lecul**A**rpromp**T** (**MOAT**) designed specifically for geometric molecules. Specifically, we propose atom-level prompts to capture atom distribution, geometry-level prompts tailored for molecular conformers, where different conformations have distinct chemical properties, and task-level prompts to leverage functional group properties. Results on both 3D and 2D downstream tasks demonstrate its ability to successfully bridge the data gap across diverse settings. To the best of our knowledge, this paper is the first attempt to introduce geometric graph-prompting learning for molecules.

CCS Concepts

• **Applied computing** → *Bioinformatics*; • **Computing methodologies** → *Learning paradigms*; *Machine learning approaches.*

Keywords

Graph prompting; 3D Molecules; Molecular Prompting; Molecular Fine-tuning

ACM Reference Format:
Qingqing Long, Yuchen Yan, Wentao Cui, Wei Ju, Zhihong Zhu, Yuanchun Zhou, Xuezhi Wang, and Meng Xiao. 2024. MOAT: Graph Prompting for 3D Molecular Graphs. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24), October 21–25, 2024, Boise, ID, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3627673.3679628

1 Introduction

Effective molecule representation learning is a fundamental step for various AI-driven scientific research issues in recent years, such as bioinformatics, virtual screening, drug discovery, cancer prediction [11–13, 27, 37]. Molecules can naturally be abstracted as graph-structured data, with atoms and bonds representing nodes and edges, respectively. Additionally, commonly used 3D molecules include chemical bond lengths and rotation angles, which are converted into 3D graphs where all nodes are assigned Cartesian coordinates [15, 18, 36, 38]. Due to the presence of 3D structures, 3D molecules can exhibit numerous conformers, each potentially having significantly different properties. Therefore, a key challenge in developing effective 3D molecular representations is ensuring rotational invariance [34, 41, 43].

*Corresponding Authors.

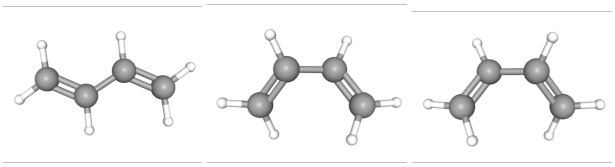
CIKM '24, October 21–25, 2024, Boise, ID, USA

ACM ISBN 979-8-4007-0436-9/24/10
https://doi.org/10.1145/3627673.3679628

Traditional approaches generally adopt the Graph Neural Networks (GNNs) architectures to model the 3D structures of molecules, we categorize these works into two classes, i.e., Equivariant GNNs and chemical GNNs. Equivariant GNNs [2, 14, 50, 56] typically operate on inputs containing absolute 3D information, such as Cartesian coordinates. While success has been achieved, each part of these models must be very carefully designed to maintain the rotation equivariance [50]. Furthermore, the performance of such models generally tends to be worse than invariant ones due to the expressivity of equivariant neural networks [34]. Another class of methods exclusively utilizes relative 3D information, including Euclidean distances of the atom Cartesian coordinates and relative polar angles. These extracted relative 3D geometries by chemical tools are naturally SE(3)-invariant [17]. In this paper, we follow the architecture of chemical GNNs. For chemical GNNs, they always adopt the end-to-end supervised learning architecture [2, 20]. In light of the constraints associated with supervised learning, significant efforts have been directed towards embracing the "pre-training and fine-tuning" framework for GNNs [53, 54, 65], drawing inspiration from the successful utilization of pre-trained models in domains such as Computer Vision (CV) and Natural Language Processing (NLP) [8, 32]. The paradigm adopted for 3D molecular generation or property prediction comprises two principal phases. Firstly, in the pre-training phase, the self-supervised learning tasks are meticulously designed specifically for unlabeled 3D molecules. These tasks aim to furnish GNNs with a broad comprehension of the inherent correlations among geometric substructures and chemical connections. For example, GNNs can amass general insights into the relationships among atoms and chemical rotation angles through tasks such as masked 3D molecule generation. Subsequently, during the fine-tuning phase, the pre-trained molecular GNNs can be adjusted to predict 3D molecular properties or geometry with a fewer number of training epochs and annotations.

Despite the diagram of "pre-training and fine-tuning" in 3D molecular GNNs achieving success in efficient fine-tuning, their performance is still hindered by the misalignment of objectives between the pre-training and fine-tuning stages. Self-supervised reconstruction objectives are adopted on the pertaining 3D molecules in the pre-training stage [2, 53], thereby facilitating the reconstruction of fundamental 3D geometric substructures. Conversely, the fine-tuning process deviates from the pre-training objective, aiming to minimize a task-specific loss driven by the quality of annotations. This misalignment poses challenges for the seamless transfer of chemical geometric knowledge acquired during pre-training, potentially resulting in catastrophic forgetting and, consequently, constraining overall performance [7, 10].

In response to these challenges, researchers made several prompt-based explorations on graph-structured data to mitigate the misalignment between the pre-training and fine-tuning stages [10, 35, 54] and achieved great success. The general idea behind these models is inspired by prompt tuning techniques in the domain of NLP [24, 26] and CV [8, 32]. They share the common principle that by reconfiguring the fine-tuning process to align with the pre-training framework, which can seamlessly transfer general knowledge to fine-tuning tasks without loss. However, all existing graph-prompting explorations are conducted on general 2D graphs, such as social networks, citation networks, and knowledge graphs, which can not directly extend to geometric 3D molecules. We summarize the specific challenges as follows,



(a) (b) (c)

Figure 1: Three types of 3D conformer for Butadiene (C4H6) [4] with varying conformational energies.

- **Gaps of atom distribution:** Molecular graphs, being a distinctive form of heterogeneous graphs, consist of diverse nodes that represent atoms [5]. These atoms exhibit notable variations in distribution and chemical attributes across different types [62], especially in different molecular datasets. However, existing methods overlook this crucial aspect by applying prompt templates uniformly to all nodes, disregarding the distinctive chemical properties inherent to individual atoms. Consequently, these approaches fail to capture the nuanced characteristics of molecular structures, potentially leading to suboptimal performance in tasks related to molecular property prediction.
- **Gaps of geometric substructures:** Molecular graphs possess a notably intricate structure, particularly in terms of their geometric arrangements [4, 60]. A real molecule is essentially an ensemble of interconverting 3D structures, referred to as conformers. While different conformers share the same 2D molecular graph, they exhibit distinct 3D arrangements [23, 31]. Typically, a molecule's conformer exists with a certain probability and may manifest unique properties. Fig. 1 illustrates this concept, showcasing three conformers of Butadiene (C4H6) with varying conformational energies. Among these conformers, the most stable configuration of 1,3-butadiene is the s-trans conformation, where the molecule lies planarly, with its two pairs of double bonds oriented in opposite directions. Another conformer, the gauche geometry, involves the twisting of double bonds from the s-cis geometry, resulting in a dihedral angle of approximately 38°. This conformer is approximately 12.0 kJ/mol (2.9 kcal/mol) higher in energy compared to the s-trans conformer. Given the significance of molecular conformers, it is imperative to identify them accurately. However, existing methods primarily focus on designing prompts for fundamental links, thereby neglecting the geometric intricacies inherent within molecular graphs.
- **Gaps of molecular properties:** Different molecular datasets can have significant molecular properties [29, 47]. Functional groups are pivotal components within molecular graphs [46, 70], as they play a fundamental role in determining molecular properties. The amalgamation of diverse functional groups profoundly impacts molecular characteristics [19]. For example, molecules with benzene rings generally

have consistent physical or chemical properties, such as solubility [44]. However, existing research works suffer from a deficiency in prompts tailored for graph motifs or functional groups, thereby constraining their capacity to preserve critical structural information.

In this paper, we propose a 3D MOleculArprompT, abbreviated as MOAT. MOAT is an elaborately designed geometric graph prompt for molecules from multi-level perspectives. To the best of our knowledge, MOAT is the first attempt to introduce geometric graph-prompting learning for 3D molecular graphs. Firstly, for the atom-level prompt, we present a novel prompting template that discriminates the various types of atoms. For the geometry-level gap, we design prompts for 3D geometrical structures through spherical Bessel function decomposition based on the relative position, which can be integrated into node representation seamlessly with message-passing, alleviating the gap of geometrical conformation in 3D molecular graphs. In addition, we design a prompt template for common functional groups in graph representation, seamlessly adapting it into node representation. The prompt-tuning process is also naturally compatible with the pre-training one through through functional group alignment tasks. Through experiments conducted on popularly used benchmark 3D datasets, we demonstrate the effectiveness of our model in bridging the data gap across diverse settings and the motivation we put forth.

2 Related Work

2.1 Pre-training and Fine-tuning Molecular-GNNs

The "Pre-training and fine-tuning" paradigm has witnessed significant advancements in the field of molecular graph representation learning. Numerous pretraining tasks have been proposed to extract valuable information from large-scale molecular datasets, typically falling into two categories. The first category, termed **Pre-training with 2D Information**, which utilizes the molecular 2D structural data for pre-training. For instance, GraphCL [67] and its various iterations [55, 61, 66] introduce an embedding-enhanced framework, distinguishing embeddings of adjacent molecular graphs and elevating embeddings of other molecules. GPT-GNN [22] utilizes an attribute graph generation task aiming at pre-training GNNs to capture both structural and semantic attributes of graphs. GraphFP [42] uses subgraph mining and fragment-based tasks to enhance GNN performance on molecular benchmarks by integrating multi-resolution structural information. However, these methods only focus on 2D pretraining structure information, disregarding 3D information. Certain molecular properties, like potential energy, are strongly correlated with molecular 3D structure and conformer [49], indicating potential enhancements in predictions with this information. The second category, **Pre-training with 3D Information**, includes approaches such as 3D Infomax [52], which maximizes mutual information between representations encoded from both 2D and 3D models, and GraphMVP [33], which incorporates 3D information into the 2D model using contrastive and generative methods. The NoisyNode [69] study applied denoising regularization to graph neural networks for pretraining equilibrium conformations, resulting in state-of-the-art performance for molecular property predictions. Nevertheless, the underlying motivation of these methods is to enhance 2D models with 3D information [25], thus their performance is still influenced by inconsistencies between the objectives of the pretraining and fine-tuning stages. 3D-GPT [59] introduces a 3D pre-training method that utilizes bond angle, bond length, and dihedral angle to optimize the effective molecular representations through a generative multi-task architecture and a total energy-based surrogate metric. In contrast, this paper endeavors to pretrain 3D models via graph prompting, focusing on their utilization in downstream 3D tasks.

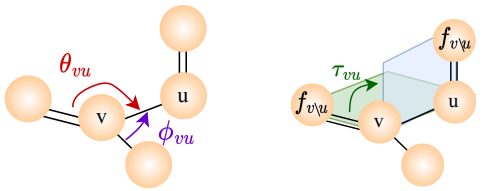
2.2 3D Molecular-GNNs

The encoding of 3D geometric structures into molecular representations is crucial for determining molecular properties, thereby driving the development of 3D graphs tailored to handle 3D information, especially in cases where pretraining models inadequately utilize such information. Within the realm of 3D molecular graph research, an important research direction is Equivariant Graph Neural Networks (EGNN), including Tensor Field Networks (TFN) [56], SE(3)-transformers [14], PaiNN [50], NequIP [2], etc. These methods typically operate on inputs containing **absolute 3D information**, such as Cartesian coordinates. While effective, each component of these models must be very carefully designed to maintain equivariance to rotations of the input graph. Furthermore, the performance of such equivariant GNNs generally tends to be worse than invariant ones [34]. Another class of methods exclusively utilizes **relative 3D information**, including Euclidean distances of the atom Cartesian coordinates and relative polar angle. These relative 3D geometries are naturally SE(3)-invariant [17]. SchNet [49] solely considers distance information, while DimeNet [17] further integrates angles between bonds. Nevertheless, both approaches incorporate incomplete 3D information, limiting network capacity. SphereNet [34] generates an approximate complete 3D representation using distance, angle, and torsion information, albeit with a complexity of $O(nk^2)$. GemNet [17], relying on quadruplet nodes, introduces even higher computational overheads. But these methods incur high computational costs. Building on this, ComENet [58] achieves the currently best performance by proposing essential rotation angles to fulfill global completeness, with a time complexity of $O(nk)$. Equiformer [31] adapts Transformer architectures with SE(3)/E(3)-equivariant features and a novel equivariant graph attention mechanism, achieving competitive results on several 3D molecular datasets. In this paper, we follow the way of relative-information-based models, avoiding the inherent limitations of equivariant 3D GNNs. We elaborately design geometric prompts for better utilizing the relative 3D characteristics.

2.3 Graph Prompt Learning

Prompt learning has been developed to address the inherent gap between pre-training and fine-tuning tasks. It has achieved significant success across various fields, including sentence classification and sentiment analysis in Natural Language Processing (NLP), as well as image recognition and segmentation in Computer Vision (CV)[8, 13, 32, 39]. The general idea behind prompt learning is to reshape data representations or fine-tuning tasks into pre-training formats. Inspired by this, many researchers naturally extend this idea into graph-structured data, for remitting the gap in graph-specific

tasks [28, 40, 54, 65]. Typically, GraphPrompt [35] improves its performance with types of task-oriented learnable readout prompt functions. All-in-One [54] designs an extra graph token structure and token insertion technique, developing the model's ability in generalization for various downstream tasks. GPF [10] integrates universal prompt templates and offers related theoretical analysis for a comprehensive understanding. HGPrompt [68] is designed for heterogeneous graphs. It proposes a dual-prompting module, which includes the feature prompt and heterogeneity prompt. The heterogeneity prompts rely on the meta-paths. MolCPT [7] extends the general idea to 2D molecules. It considers the motif structure and ignores the heterogeneous atom types and critical geometry information, which encounters the specific challenges for 3D molecular graphs. To sum up, all existing graph prompt learning methods ignore the critical and specific characteristics for distinguishing 3D molecular graphs, such as the geometric information and the 3D conformer.



(a) d_{vu}, θ_{vu} and ϕ_{vu} (b) τ_{vu}

Figure 2: Illustration for computing relative 3D molecular information $\Omega(\mathcal{G})$.

3 Preliminaries

3.1 Problem Definition

A 3D molecule can be represented as a geometric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, where $\mathcal{V}$ denotes the set of node (atom) attributes, $|\mathcal{V}|$ denotes the number of nodes (atoms), and $|\mathcal{E}|$ denotes the number of edges (bonds). The position matrix $\mathcal{P} = [p_1, p_2, ..., p_{|\mathcal{V}|}]^T \in \mathbb{R}^{|\mathcal{V}|\times 3}$, where $p_i = (x_i, y_i, z_i)$ denotes the position vector for node i in Cartesian coordinate system (CSC). Molecule properties $\mathbf{y}$ constitute the targets of the predictive task. Given a molecule $\mathcal{G}_n$, we aim to predict the properties or energy $\mathbf{y}_n$ according to its graph representation $h(\mathcal{G}_n)$, i.e., to learn a mapping function $f_\theta : \mathcal{G}_n \rightarrow \mathcal{Y}_n$.

3.2 3D Pretraining MolecluarGNNs

The 3D molecular machine learning models generally follow the GNN architectures due to their superior ability in modeling graph-structured data [52]. In this paper, we focus on the non-equivariant MolecluarGNNs [34]. They converts absolute Cartesian coordinates $\mathcal{P}$ into relative 3D expressions $\Omega(\mathcal{G})$, which is naturally $SE(3)$-invariants.

$$\Omega(\mathcal{G}) = \{(d_{vu}, \theta_{vu}, \phi_{vu}, \tau_{vu}) \in \mathbb{R}^{|\mathcal{E}|\times 4} | v \in \mathcal{V}, u \in \mathcal{N}(v)\}, \quad (1)$$

where $\mathcal{N}(v)$ denotes the connected neighbor-atoms around atom v, d_{vu} denotes the relative Euclidean distance between node u and v, θ_{vu} denotes the relative polar angle, ϕ_{vu} denotes the azimuthal angle. τ_{vu} denotes the rotation angle for edge e_{vu}, which determines the orientation of the local neighborhood. The detailed calculation of $\Omega(\mathcal{G})$ [58] is shown in Algo. 1. The 3D geometry structure guides the aggregation of local neighborhood information and leads to a more contextual representation for each node. Then a graph pooling operation is adopted to get the representation for the whole graph. Let $h_v^{(l)}$ denote the representation of node v at the l-th layer of GNN, and $\mathcal{N}(v)$ denote all the neighbor nodes of node v, the update procedure from the $(l-1)$-th layer to the l-th layer is:

$$h_v^{(l)} = \text{COMB}\left(h_v^{(l-1)}, \text{AGG}_{u\in\mathcal{N}(v)}(\{h_v^{(l-1)}, h_u^{(l-1)}, d_{vu}, \theta_{vu}, \phi_{vu}, \tau_{vu}\})\right), \quad (2)$$

where AGG denotes the aggregation function (e.g., mean or max operator). COMB combines the information of neighbours and node v (e.g., concatenation operator). After L iterations of message passing, the hidden states $h_v^{(L)}$ in the last iteration are the embeddings of v. Finally, a READOUT operation (e.g., averaging, sum or graph pooling) is adopted to get the representation h_G for the whole graph $\mathcal{G}$,

$$h_\mathcal{G} = \text{READOUT}\left(\{(h_v^{(L)} | v \in \mathcal{V}\}\right). \quad (3)$$

3.3 Pretraining and Graph-Promoting Paradigm

Let g_θ denote the GNNs pre-trained on 3D molecules $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, and $\mathcal{D}$ denote the data distribution of downstream tasks. Then the "pretraining and finetuning" paradigm [22] optimizes the parameters of the pre-trained molecular GNN g_θ and the learnable projection head g_ϕ to maximize the likelihood of correct labels $\mathcal{Y}_c$ for the downstream prediction tasks,

$$\arg\max_{\theta,\phi} P\left(\mathcal{Y} | g_\phi(g_\theta(\mathcal{X}, \mathcal{E}, \mathcal{P}))\right). \quad (4)$$

For the "pretraining and graph-promoting" paradigm [35, 54], the parameters of the pre-trained model are frozen. The target of graph promoting is to obtain the optimized learnable prompts $\mathcal{T}$ to maximize the likelihood of correct labels $\mathcal{Y}_c$. The node (atom) embeddings $\bar{\mathcal{X}}$ and edge embeddings $\bar{\mathcal{E}}$ of the downstream graphs are initialized by the pre-trained model g_θ before graph prompt learning, and the initialization process is formulated as,

$$\bar{\mathcal{X}}, \bar{\mathcal{E}}, \bar{\mathcal{P}} = g_\theta(\mathcal{X}, \mathcal{E}, \mathcal{P}). \quad (5)$$

General graph prompting templates include two learnable components, i.e., feature (node-level) prompts and adjacency (edge/motif-level) prompts. To adapt the downstream tasks, the optimization of graph prompting aims at finding the optimal prompt parameters $\mathcal{T}$ that maximize the likelihood of predicting the correct labels. The optimization process in this paper is formally defined as,

$$\begin{aligned} &\mathcal{G}^* : (\hat{\mathcal{X}}, \hat{\mathcal{P}}, \hat{\mathcal{E}}) = \mathcal{T}(\bar{\mathcal{X}}, \bar{\mathcal{E}}, \bar{\mathcal{P}}), \\ &\mathcal{T}^* = \arg\max P\left(\mathcal{Y} | \mathcal{G}^*\right). \end{aligned} \quad (6)$$

4 MODEL: MOAT

In this paper, we design a 3D MOleculArPrompT for geometric molecules from multi-level perspectives, abbreviated as MOAT. MOAT has three main components, i.e., node-level, geometry-level, and task-level graph prompts. Fig. 3 shows the framework of MOAT. Compared with the traditional graph "pre-training and fine-tuning"

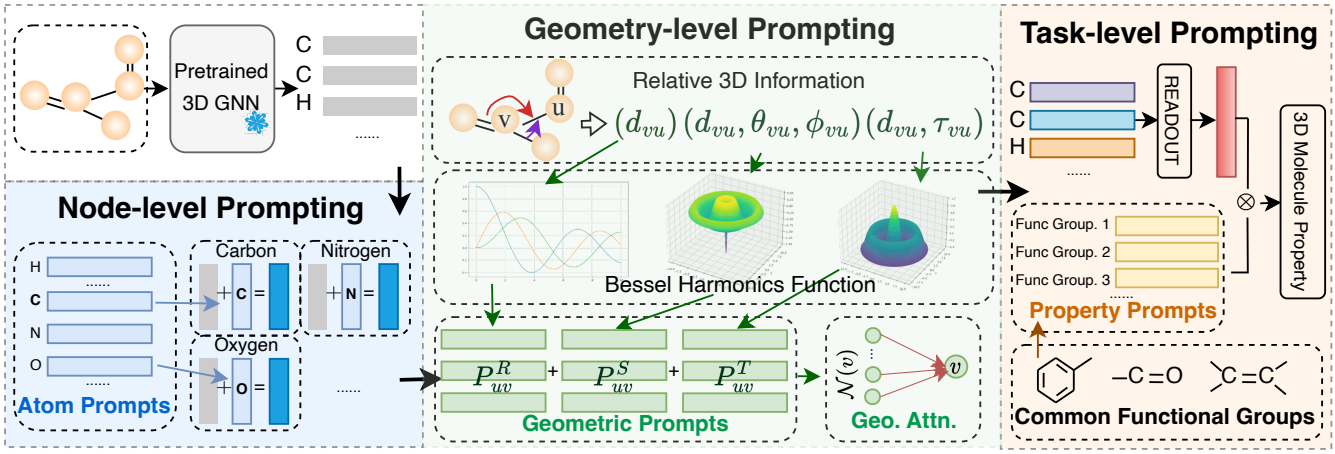


Figure 3: Overview of the proposed MOAT. a) We design an atom-level prompt to depict the atom type distribution. b) We propose a structure-level prompt for the 3D geometrical structures based on relative positions and utilize the attention mechanism to aggregate geometric information. c) Finally, we present a task-level prompt to utilize the properties of functional groups.

paradigm, graph prompt learning relieves the misalignments between pre-training objectives and downstream tasks with greater efficiency.

4.1 Node-Level Prompting.

To address the gap in atom distribution between the pre-training and fine-tuning stages, we first propose an atom-level prompting. Generally, graph prompt methods view the node-level gap as a feature perturbation and propose compensating for it by using a learnable prompt. Specifically, for node feature h_v the equation is expressed as follows,

$$h_v = \bar{h}_v + p_v, \tag{7}$$

where p_v represents the learnable prompt for node v. However, this kind of prompt template employs a learnable prompt for every node indiscriminately, ignoring the characteristics in molecular graphs. Compared with general graphs, the nodes (atoms) have different types in molecular graphs, i.e., the molecular graph is a heterogeneous graph. To capture the heterogeneity, we design a set of atom-type-related prompts to tackle the gap. The compensation process is described as the following,

$$h_v = \bar{h}_v + p_{T(v)}, \tag{8}$$

where $T(v)$ denotes the atom type of node v, $p_{T(v)}$ is a learnable vector of type $T(v)$.

4.2 Geometry-Level Prompting.

The molecular geometry structure is strongly correlated with molecular properties [4, 47], such as the orbital energy. In this paper, we follow invariant 3D GNN that takes relative 3D information, i.e., $\Omega(\mathcal{G}) = (d, \theta, \phi, \tau)$, as inputs to networks. The relative geometric information is naturally SE(3)-invariant and crucial for distinguishing the molecular 3D structure, as shown in the following theorem.

Theorem 1. Given two 3D molecular graphs, $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}, \mathcal{P}_1)$ and $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}, \mathcal{P}_2)$, the geometry transformation $\Omega(\mathcal{G})$ uniquely determined the 3D structure [58], i.e.,

$$\Omega(\mathcal{G}_1) = \Omega(\mathcal{G}_2) \Leftrightarrow \exists R \in SE(3), \mathcal{P}_1 = R(\mathcal{P}_2).$$

where $SE(3)$ denotes the three-dimensional Special Euclidean group, and R denotes a transformation that combines rotation and translation. The theorem indicates that if $\mathcal{P}_1$ and $\mathcal{P}_2$ are in the same $SE(3)$ group, then $\mathcal{G}_1$ and $\mathcal{G}_2$ would share the same 3D conformation, they are the same 3D graph.

Following previous studies [21], the relative geometric information $\Omega(\mathcal{G})$ among atoms are generally transformed into physically meaningful vectors based on quantum-based basis functions for better performance. The spherical Bessel harmonics function generally performs the best and most robustly. The basic functions of d, (d, τ) and (d, θ, ϕ) are RBF, SBF and TBF [17] respectively, which are formally defined as follows,

$$\begin{aligned}
\mathbf{e}_{\text{RBF},n}(d) &= \sqrt{\frac{2}{c}} \frac{\sin(\frac{n\pi}{c} d)}{d}, \\
\mathbf{e}_{\text{TBF},l,m,n}(d, \theta, \phi) &= \sqrt{\frac{2}{c^3 j_{l+1}^2 z_{ln}}} j_l(\frac{z_{ln}}{c} d) Y_{lm}(\theta, \phi), \\
\mathbf{e}_{\text{SBF},l,n}(d, \tau) &= \sqrt{\frac{2}{c^3 j_{l+1}^2 z_{ln}}} j_l(\frac{z_{ln}}{c} d) Y_{l0}(\tau),
\end{aligned} \tag{9}$$

where c denotes the cutoff boundary of the electron wave function, n denotes the order of the radial Bessel basis, j_l denotes the spherical Bessel function of order l. z_{ln} denotes the n-th root of the l -order Bessel function, and is precomputed numerically. Y_{lm} denotes the spherical harmonic function of order l with degree m. l, m, n are three hyper-parameters. The value of m is related to l, the range of m is $[0, l-1]$. Especially, we design a set of learnable prompts to represent each order of basis function, $V^R \in \mathbb{R}^{n \times h}$ for RBF functions, $V^T \in \mathbb{R}^{l \times n \times h}$ for TBF functions and $V^S \in \mathbb{R}^{l \times m \times n \times h}$

for SBF functions,

$$
\begin{aligned}
p_{vu}^{R} &= \sum_{n} \mathbf{e}_{\mathrm{RBF},n}(d_{vu}) V_{n}^{R}, \\
p_{vu}^{S} &= \sum_{l,m,n} \mathbf{e}_{\mathrm{SBF},l,m,n}(d_{vu}, \theta_{vu}, \phi_{vu}) V_{l,m,n}^{S}, \\
p_{vu}^{T} &= \sum_{l,n} \mathbf{e}_{\mathrm{TBF},l,n}(d_{vu}, \tau_{vu}) V_{l,n}^{T}.
\end{aligned} \tag{10}
$$

To incorporate these geometric prompts into the node representations, we design an attentive interaction layer. Specifically, three prompts are merged as edge-level auxiliary features. Then the prompt of edge e_{vu} is concatenated with the features of node v and node u. The above processes are formally defined as,

$$
\begin{aligned}
p_{vu}^{e} &= p_{vu}^{R} + p_{vu}^{S} + p_{vu}^{T}, \\
\tilde{h}_{vu} &= p_{vu}^{e} || h_v || h_u,
\end{aligned} \tag{11}
$$

where $||$ denotes the concatenation operation. Then the attention mechanism is utilized to aggregate information from different edges,

$$
\tilde{h}_v = \sum_{u \in \mathcal{N}(v)} \frac{\exp(a \cdot \tilde{h}_{vu})}{\sum_{k \in \mathcal{N}(v)} \exp(a \cdot \tilde{h}_{vk})} W \tilde{h}_{vu}, \tag{12}
$$

where $a \in \mathbb{R}^{3h}$ and $W \in \mathbb{R}^{h \times 3h}$ are two learnable parameters.

Algorithm 1 Algorithm of MOAT.

Require: Input molecules $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, pretrained 3D molecularGNN g_θ, downstream labels $\mathbf{y}_n$, such as molecular properties.
Ensure: Molecular representation $\tilde{h}_{\mathcal{G}}$,

```
Initialize atom embeddings with pretrained model X̄, Ē = g_θ(X, E);
for all u = 1,...,|V| do
    for all v ∈ N(v) do
        Compute d_uv = ||p_v − p_u||_2;
    end for
end for
for all u = 1,...,|V| do
    Find reference nodes of the center node u, i.e., f_u = arg min_{v∈N(u)} d_uv, s_u = arg min_{v∈N(u)\f_u} d_uv;
    for all v ∈ N(u) do
        Compute θ_uv = angle(f_v, v, u);
        Compute φ_uv = angle; (plane_{f_v,v,s_v}, plane_{f_v,v,s_v});
        Compute τ_uv = angle(plane_{f_{v\u},v,u}, plane_{v,u,f_{u\v}});
    end for
end for
for all u = 1,...,|V| do
    Node-level prompt learning, as shown in Eqn. (7);
end for
for all e_vu ∈ E do
    Geometry-level prompt learning, as shown in Eqn. (12);
end for
Task-level prompt learning, as shown in Eqn. (13);
Finding the optimal prompt parameters T that maximize the likelihood of predicting the correct labels with Eqn. (14).
```

4.3 Task-Level Prompting.

Except for the atom distribution and geometric substructure gap, the inconsistency of 3D molecular properties between the pre-training and fine-tuning process from different molecular datasets also needs to be considered. General graph prompting models align the fine-tuning task with the pre-training one through learnable task-level prompting. For 3D molecular graphs, we consider the gap can be relieved through the functional groups. Previous research shows that the functional groups are relative to the downstream 3D molecular property prediction analysis [63]. For example, the functional groups with impression on molecular toxicity [9], e.g., trifluoromethyl and cyanide, known as two toxic functional groups, are considered to be associated with the toxicity property of 3D molecules. Another example is that 3D molecules with benzene rings typically share analogical physical properties, such as solubility, as well as chemical characteristics like aromaticity [44]. In this paper, we consider eighty-five usual functional groups [3, 4]. These functional groups can be identified by RDKit tools based on the input *SMILES* sequence. Then the task-level prompts are formally defined as,

$$
\begin{aligned}
\tilde{h}_{\mathcal{G}} &= \mathrm{READOUT}(\{(\tilde{h}_v | v \in \mathcal{V}\}), \\
\tilde{h}_{\mathcal{G}} &= \tilde{h}_{\mathcal{G}} + \sum_{j \in S(\mathcal{G})} p_j^F,
\end{aligned} \tag{13}
$$

where $S(\mathcal{G})$ denotes the set of functional groups that appeared in the graph, $p_j \in P_f^F$ denotes the set of functional-group prompts. The dimension of p_j^F is the number of molecular properties. READOUT denotes the mean pooling operation in this paper.

The 3D downstream tasks are generally regression objectives. Finally, the **overall objective** function of MOAT is defined as,

$$
\mathcal{L} = \frac{1}{N} \sum_{n=1}^{N} ||\tilde{h}_{\mathcal{G}_n} - \mathcal{Y}_n||_2. \tag{14}
$$

4.4 Complexity analysis

The algorithmic pseudo-code of the proposed MOAT is shown in Algorithm. 1. The time complexity of MOAT is $O(N) + O(|E|)$. Compared with other 3D molecular models, such as the popular SphereNet and DimeNet++, it is efficient. The complexity of MOAT mainly comes from geometry-level prompts, as the complexity of atom-level and task-level prompting are both $O(N)$. The formulas of d_{ij}, θ_{ij}, and ϕ_{ij}, and τ_{ij} are shown in Algorithm 1 and are specified with the 1-hop local neighborhood. Thus their time complexity is $O(Nk)$, where k denotes the average degree in 3D molecular graphs, and is generally less than 100. The complexity of Eq (10) is $O(l \times m \times n \times h)$, where $l = 3, m = 2, n = 6, h = 512$ in our model. The complexity of Eq (12) is $O(|E| \times h \times 3h \times 3h) = O(|E|h^3)$. Thus complexity of geometry-level is $O(|E|h^3) = O(E)$.

The space complexity of MOAT is $O(|E|)$. Let h denote prompt dimension and t denote types of atoms, the complexity of atom-level is $O(ht)$, and t is generally less than 10. Complexity of geometry-level is $O(|E|h^2)$. The complexity of task-level is $O(|S(G)|h)$, where $|S(G)|$ is 85 and h is 64.

Table 1: The overall performance on QM9 dataset. Results with * are taken from [58] and [34].

Property	μ	α	ϵ_{HOMO}	ϵ_{LUMO}	Δ_ϵ	$< R^2 >$	ZPVE	U	U_0	H	G	c_v
Unit	D	a_0^3	meV	meV	meV	a_0^2	meV	meV	meV	meV	meV	cal/molK
PPGN*	0.0470	0.1310	40.3	32.7	60.0	0.592	3.12	36.80	36.80	36.30	36.40	0.055
SchNet*	0.0330	0.2350	41.0	34.0	63.0	**0.073**	1.7	14	19	14	14	0.033
PhysNet*	0.0529	0.0615	32.9	24.7	42.5	0.765	1.39	8.15	8.34	8.42	9.40	0.028
Cormorant*	0.1300	0.0920	36.0	36.0	60.0	0.673	1.98	28	-	-	-	0.031
MGCN*	0.0560	**0.0300**	42.1	57.4	64.2	0.110	1.12	12.9	14.4	14.3	16.2	0.038
DimeNet*	0.0286	0.0469	27.8	19.7	34.8	0.331	1.29	8.02	7.89	8.11	8.98	0.025
DimeNet++*	0.0297	0.0435	24.6	19.5	32.6	0.331	1.21	6.32	6.28	6.53	7.56	0.023
SphereNet*	0.0245	0.0449	22.8	18.9	31.1	0.268	1.12	6.26	6.36	6.33	7.78	0.022
ComENet*	0.0241	0.0404	**21.7**	18.5	**30.2**	0.239	1.20	6.18	6.40	6.13	7.98	0.024
GPF	0.0311	0.0479	27.8	20.9	36.2	0.348	1.39	8.69	8.23	8.78	9.29	0.027
GraphPrompt	0.0299	0.0475	29.9	20.8	35.4	0.355	1.35	8.21	8.47	8.57	9.67	0.026
MolCPT	0.0297	0.0470	26.4	19.3	33.7	0.349	1.34	8.33	8.28	7.76	8.56	0.025
MOAT	**0.0236**	0.0441	22.6	**18.4**	30.9	0.228	**1.04**	**6.17**	**6.34**	**6.02**	**7.55**	**0.022**

5 Experiments

5.1 Experimental Settings

5.1.1 Datasets. We examine the performance of MOAT on commonly used 3D molecular dataset QM9 [45] and OC20 dataset [6].

The QM9 dataset is widely used to predict different properties of molecules. It includes 134,000 stable small organic molecules and offers data on their geometric, energetic, electronic, and thermodynamic properties. The dataset includes 110,000 molecules in the training set, 10,000 molecules in the validation set, and 10,831 molecules in the test set. The dataset covers twelve properties, including dipole moment (μ), isotropic polarizability (α), highest occupied molecular orbital energy (ϵ_{HOMO}), lowest unoccupied molecular orbital energy (ϵ_{LUMO}), the energy gap between ϵ_{HOMO} and ϵ_{LUMO}, electronic spatial extent ($< R^2 >$), zero-point vibrational energy (ZPVE), internal energy at 0K (U_0), internal energy at 298.15K (U), enthalpy at 298.15K (H), free energy at 298.15K (G), and heat capacity at 298.15K (c_v). These properties are vital for understanding molecular behavior and have wide-ranging applications in chemistry and material science.

The OC20 dataset [6] is tailored for catalyst exploration and refinement. It encompasses 1,281,040 Density Functional Theory (DFT) relaxations, involving approximately 264,890,000 single-point evaluations, spanning a diverse array of materials, surfaces, and adsorbates. The dataset encompasses three key tasks: Structure to Energy and Forces (S2EF), Initial Structure to Relaxed Structure (IS2RS), and Initial Structure to Relaxed Energy (IS2RE). In this paper, we focus on the IS2RE task, which is a fundamental element in catalysis due to the correlation between relaxed energies and catalyst performance. The dataset for IS2RE is originally split into training/validation/test sets. There are 460,318 structures in the training set. Test labels are not publicly available. Model performance is assessed on the validation set, which includes four splits: In Domain (ID), Out of Domain Adsorbates (OOD Ads), Out of Domain Catalysts (OOD Cat), and Out of Domain Adsorbates and Catalysts (OOD Both), each split contains approximately 24,000 structures.

5.1.2 Baselines. We select the following representative baselines as our competitors:

3D Molecular Pretrain-finetuning GNNs conduct the message passing scheme to capture the 3D information for enhancing the performance of molecule embeddings. PhysNet [57], CGCNN [48], Cormorant [1], MGCN [41], PaiNN [50], and GemNet-T [51] are representative works of equivariant molecular GNNs, and they directly operate on the Cartesian coordinates. SchNet [49], PPGN [43], SphereNet [34], DimeNet [17], DimeNet++ [16], GemNet [15] and ComENet [58] are representative works of chemical GNNs, and they utilize the relative 3D information as the input of GNNs.

General Graph Prompting methods adopt the "pre-train and prompt" paradigm on 2D graphs. In this setting, their pre-trained models are frozen, and parameters of task-specific prompts are learnable to maximize the likelihood of predicting the correct downstream labels. Representative works include GPF [10] and GraphPrompt [35]. These methods are designed to relieve the gap between pre-training and downstream tasks in common graphs, such as social networks and citation networks.

Molecular Graph Prompting models utilize the "pre-train and prompt" paradigm for molecules. MolCPT [7] is the sole related work, which extends the general graph prompting idea to 2D molecules. It only considers the motif structure and ignores the heterogeneous atom types and geometry information, which encounters the specific challenges for 3D molecular graphs.

5.1.3 Experimental Settings. For molecular property prediction with 3D molecular dataset, we focus on predicting twelve properties on QM9, including dipole moment (μ), isotropic polarizability (α), etc. Additionally, we work on the Initial Structure to Relaxed Energy (IS2RE) task on OC20. All of these are regression tasks, measuring various aspects of the molecular property. For molecular property prediction with 2D molecular dataset, we use RDKit [30] to extract their 3D information as the input of our model. For a fair

Table 2: Performance on IS2RE in terms of energy MAE and the percentage of EwT of the ground truth energy. Results with * are taken from [58].

Method	Energy MAE [eV] (↓)					EwT (↑)				
	ID	OOD Ads	OOD Cat	OOD Both	Average	ID	OOD Ads	OOD Cat	OOD Both	Average
CGCNN*	0.6203	0.7426	0.6001	0.6708	0.6585	3.36%	2.11%	3.53%	2.29%	2.82%
SchNet*	0.6465	0.7074	0.6475	0.6626	0.6660	2.96%	2.22%	3.03%	2.38%	2.65%
DimeNet++*	0.5636	0.7127	0.5612	0.6492	0.6217	4.25%	2.48%	4.40%	2.56%	3.42%
GemNet-T*	0.5561	0.7342	0.5659	0.6964	0.6382	4.51%	2.24%	4.37%	2.38%	3.38%
SphereNet*	0.5632	0.6682	0.5590	0.6190	0.6023	**4.56%**	2.70%	4.59%	2.70%	3.64%
ComENet*	0.5558	0.6602	0.5491	0.5901	0.5888	4.17%	2.71%	4.53%	2.83%	3.56%
GPF	0.5743	0.7207	0.5966	0.7108	0.6506	4.20%	2.47%	4.31%	2.31%	3.32%
GraphPrompt	0.6059	0.7690	0.5665	0.6651	0.6516	3.83%	2.37%	4.09%	2.31%	3.15%
MolCPT	0.5772	0.7178	0.5733	0.6762	0.6361	4.16%	2.48%	4.15%	2.42%	3.30%
MOAT	**0.5556**	**0.6598**	**0.5488**	**0.5897**	**0.5864**	4.21%	**2.74%**	**4.63%**	**2.85%**	**3.67%**

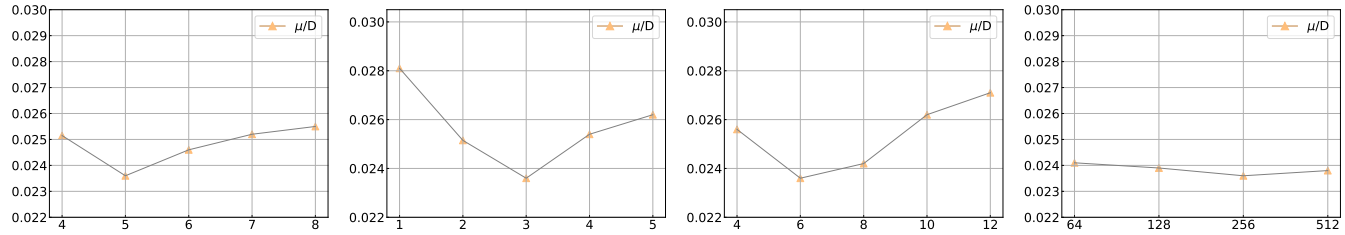
comparison, the results of some baselines are taken from the original paper. In addition, we select typical open-sourced baselines and run them in all datasets. The results are generally consistent with those reported in the original paper. The evaluated typical baselines include ComE, SphereNet, DimeNet, and DimeNet++. We use scaffold splitting to split the molecules according to their structures to mimic real-world use cases. The split for train / val / test sets is 80% : 10% : 10%. We select the ComE [58] as the pre-trained 3D molecular model on Molecule3D, which contains 3.89 billion molecules [64]. For a fair comparison with general prompting models in the 3D molecular dataset, their initialized embeddings are also predicted by the pre-trained ComE model. All models are trained on NVIDIA GeForce RTX A100 80GB GPU. The optimal hyper-parameters of MOAT are obtained on validation sets using grid search. We take Adam as the optimizer. The cutoff c is set to 5, the order l of the Bessel function is set to 3, the order n of the radial Bessel basis is set to 6, the batch size is set to 1024, and the learning rate is set to $5e^{-4}$ with a decay factor 0.5.

5.2 Overall Performance

To verify the effectiveness of MOAT, we conducted experiments on the 3D molecular datasets, i.e., QM9 and OC20 datasets, which are widely used to assess the capabilities of models in predicting properties of quantum chemistry systems. We compared MOAT with baseline models using the Mean Absolute Error (MAE) metric for each property. The comparison results are summarized in Table. 1. General graph prompting models, such as GPF and GraphPrompt, fail to leverage the coordinate information of molecules, instead treating homogeneous graphs without incorporating heterogeneous molecular information. To ensure a fair comparison between these general prompting models and MOAT, their initial embeddings were all predicted by the pre-trained ComE model. Based on the results in Table. 1, we draw the following conclusions: **(1)** Overall, Equivariant molecular GNNs exhibit poorer performance compared to chemical GNNs. For instance, in the QM9 dataset, Cormorant, MGCN, and PhysNet perform worse than DimeNet, DimeNet++, SphereNet, and ComENet. Similarly, in the OC20 dataset, while CGCNN has the lowest performance, SchNet (a chemical GNN) underperforms relative to other Equivariant molecular GNNs. **(2)** MOAT achieves competitive performance across all property prediction tasks on the QM9 dataset, achieving optimal results in 8 out of 12 tasks. At the same time, MOAT achieved the best performance in all tasks on the OC20 dataset. The overall performance of MOAT demonstrates superior accuracy, validating the effectiveness of our approach. In addition, MOAT outperforms mostly relative-information-based 3D models, which verifies the motivation we put forth. **(3)** MOAT significantly outperforms general graph prompting models (GPF and GraphPrompt). It indicates the effectiveness of incorporating both bond direction and the geometric structure of 3D molecules. Specifically, MOAT achieves notable improvements on the property prediction tasks in both 3D benchmark datasets compared with state-of-the-art methods with the "pre-train and prompt" paradigm. **(4)** MOAT also outperforms the 2D molecular prompting method MolCPT, demonstrating the importance of introducing geometry-level prompts for 3D molecules. Furthermore, compared to fine-tuning-based methods, MOAT achieves comparable or superior results, proving that the "pre-training and graph prompting" paradigm can be effectively and efficiently applied to a wide variety of molecular task scenarios.

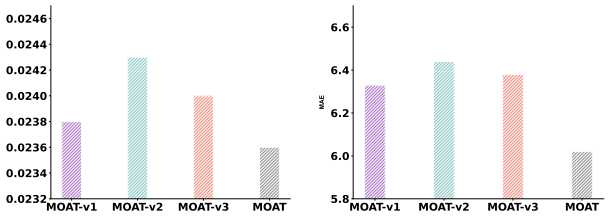
5.3 Efficiency Analysis

We conduct model efficiency analysis on MOAT and several popular and representative 3D molecular baselines, i.e. ComENet and SphereNet, in the QM9 benchmark dataset. Table 3 reports the number of parameters, average training and inference time per epoch for these models. All running time data results are tested on the same GPU with no other processes running simultaneously. Compared with other 3D molecular baselines, MOAT has the fewest parameters. The parameter count of MOAT is 25% of that in ComENet and 70% of that in SphereNet. In addition, MOAT demonstrates competitive running efficiency in both training and inference phases, attributable to its minimal parameter count. The results indicate that MOAT excels not only in parameter efficiency but also in computational performance, exhibiting scalability to large-scale 3D molecular datasets.



(a) Cutoff boundary c of the electron wave. (b) Order l of Bessel function. (c) Order n of the radial Bessel basis. (d) Embedding size

Figure 4: Parameter analysis of MOAT on the QM9 dataset.



(a) μ/D (b) H/meV

Figure 5: Performance of μ and H on QM9 dataset, with switching model components.

Model	# Parameters	Training Time / s	Inference Time/ s
ComENet	4.3M	219	9
SphereNet	1.3M	121	8
MOAT	0.9 M	92	5

Table 3: Parameters and running efficiency analysis.

5.4 Ablation Study

To assess the contribution of each component, we conducted an ablation study of MOAT using the QM9 benchmark dataset. For this evaluation, we selected the prediction of dipole moment (μ) and enthalpy at 298.15K (H) as two representative tasks. We compare MOAT with the following variants: i) **MOAT-v1**: MOAT without atom-level prompting; ii) **MOAT-v2**: MOAT without geometry-level prompting; iii) **MOAT-v3**: MOAT without task-level prompting. Results are shown in Fig. 5. Compared with other components, the geometry-level prompting objective plays the most significant role, and the atom-level prompting has the tiniest effect. This validates the crucial role of modeling 3D geometry characteristics, i.e., SE(3)-invariant, for 3D molecular property analysis, which is consistent with the motivation we put forth. Also, the performance of MOAT on all datasets surpasses that of its variants, proving the effectiveness of each component.

5.5 Parameter Analysis

We conducted a parameter analysis to evaluate the performance variability with changes in hyper-parameters, thereby validating the robustness of the model. Specifically, we examined four critical hyper-parameters in MOAT, namely the cutoff boundary of the electron wave function in Eqn. 9, the order l of the Bessel function, the order n of the radial Bessel basis, and the embedding size. We conduct experiments on the QM9 dataset, whose results are shown in Fig. 4. We make the following findings: i) The hyper-parameters in Eqn. 9 need to be chosen appropriately instead of choosing arbitrarily large values. The performance first goes good and later becomes bad with the change of parameters. ii) A larger embedding size does not necessarily lead to better performance. For MOAT, the size of 256 is the best.

6 Conclusion

Molecular property prediction is widely regarded as a cornerstone task in AI-driven drug design and discovery. Recognizing the crucial role of molecular geometry in accurate property prediction, we have integrated 3D information with advanced graph learning methods to enhance predictive performance. Despite the growing popularity of the "Graph pre-training and fine-tuning" paradigm for refining molecular representations, there remains a persistent mismatch between pre-training objectives and downstream tasks. Drawing inspiration from prompt tuning techniques in natural language processing (NLP), we propose a novel approach: 3D MOleculAr prompT, abbreviated as MOAT, specifically tailored for geometric molecules. MOAT introduces prompts at various levels to accurately depict the properties of atom distribution, geometrical conformation, and functional group distribution. Through comprehensive experiments on both 3D and 2D downstream tasks, we have demonstrated the effectiveness and robustness of our model. To the best of our knowledge, this paper marks the first attempt to introduce geometric graph-prompting learning for 3D molecular graphs. Our findings underscore the potential of MOAT to significantly advance AI-driven scientific research issues, offering a robust and versatile framework for future research and practical applications.

Acknowledgment

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences XDB38030300, the Postdoctoral Fellowship Program of CPSF GZC20232736, the China Postdoctoral Science Foundation Funded Project 2023M743565, the Special Research Assistant Funded Project of the Chinese Academy of Sciences.

References

[1] Brandon Anderson, Truong Son Hy, and Risi Kondor. 2019. Cormorant: Covariant molecular neural networks. *Advances in neural information processing systems* 32 (2019).

[2] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. 2022. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications* 13, 1 (2022), 2453.

[3] A Patrícia Bento, Anne Hersey, Eloy Félix, Greg Landrum, Anna Gaulton, Francis Atkinson, Louisa J Bellis, Marleen De Veij, and Andrew R Leach. 2020. An open source chemical structure curation pipeline using RDKit. *Journal of Cheminformatics* 12 (2020), 1–16.

[4] Theodore L Brown, HE LeMay, and Bruce E Bursten. 2006. Chemistry: The central science. UpperSaddle River.

[5] David Buterez, Jon Paul Janet, Steven J Kiddle, Dino Oglic, and Pietro Liò. 2023. Modelling local and general quantum mechanical properties with attention-based pooling. *Communications Chemistry* 6, 1 (2023), 262.

[6] Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, et al. 2021. Open catalyst 2020 (OC20) dataset and community challenges. *Acs Catalysis* 11, 10 (2021), 6059–6072.

[7] Cameron Diao, Kaixiong Zhou, Zirui Liu, Xiao Huang, and Xia Hu. 2022. Molcpt: Molecule continuous prompt tuning to generalize molecular representation learning. *arXiv preprint arXiv:2212.10614* (2022).

[8] Li Dong, Nan Yang, Wenhui Wang, Furu Wei, Xiaodong Liu, Yu Wang, Jianfeng Gao, Ming Zhou, and Hsiao-Wuen Hon. 2019. Unified language model pre-training for natural language understanding and generation. *Advances in neural information processing systems* 32 (2019).

[9] Peter Ertl, Eva Altmann, and Jeffrey M McKenna. 2020. The most common functional groups in bioactive molecules and how their popularity has evolved over time. *Journal of medicinal chemistry* 63, 15 (2020), 8408–8418.

[10] Taoran Fang, Yunchao Zhang, Yang Yang, Chunping Wang, and Lei Chen. 2022. Universal Prompt Tuning for Graph Neural Networks. *arXiv preprint arXiv:2209.15240* (2022).

[11] Yin Fang, Qiang Zhang, Ningyu Zhang, Zhuo Chen, Xiang Zhuang, Xin Shao, Xiaohui Fan, and Huajun Chen. 2023. Knowledge graph-enhanced molecular contrastive learning with functional prompt. *Nature Machine Intelligence* (2023), 1–12.

[12] Zheng Fang, Qingqing Long, Guojie Song, and Kunqing Xie. 2021. Spatial-temporal graph ode networks for traffic flow forecasting. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 364–373.

[13] Zheng Fang, Lingjun Xu, Guojie Song, Qingqing Long, and Yingxue Zhang. 2022. Polarized graph neural networks. In *Proceedings of the ACM Web Conference 2022*. 1404–1413.

[14] Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. 2020. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems* 33 (2020), 1970–1981.

[15] Johannes Gasteiger, Florian Becker, and Stephan Günnemann. 2021. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems* 34 (2021), 6790–6802.

[16] Johannes Gasteiger, Shankari Giri, Johannes T Margraf, and Stephan Günnemann. 2020. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *arXiv preprint arXiv:2011.14115* (2020).

[17] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. 2020. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123* (2020).

[18] Niklas WA Gebauer, Michael Gastegger, Stefaan SP Hessmann, Klaus-Robert Müller, and Kristof T Schütt. 2022. Inverse design of 3d molecular structures with conditional generative neural networks. *Nature communications* 13, 1 (2022), 973.

[19] Christopher J Gerry and Stuart L Schreiber. 2020. Unifying principles of bifunctional, proximity-inducing small molecules. *Nature chemical biology* 16, 4 (2020), 369–378.

[20] Jonathan Godwin, Michael Schaarschmidt, Alexander Gaunt, Alvaro Sanchez-Gonzalez, Yulia Rubanova, Petar Veličković, James Kirkpatrick, and Peter Battaglia. 2021. Simple gnn regularisation for 3d molecular property prediction & beyond. *arXiv preprint arXiv:2106.07971* (2021).

[21] Weihua Hu, Muhammed Shuaibi, Abhishek Das, Siddharth Goyal, Anuroop Sriram, Jure Leskovec, Devi Parikh, and C Lawrence Zitnick. 2021. Forcenet: A graph neural network for large-scale quantum calculations. *arXiv preprint arXiv:2103.01436* (2021).

[22] Ziniu Hu, Yuxiao Dong, Kuansan Wang, Kai-Wei Chang, and Yizhou Sun. 2020. Gpt-gnn: Generative pre-training of graph neural networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1857–1867.

[23] Ilia Igashov, Hannes Stärk, Clément Vignac, Arne Schneuing, Victor Garcia Satorras, Pascal Frossard, Max Welling, Michael Bronstein, and Bruno Correia. 2024. Equivariant 3D-conditional diffusion model for molecular linker design. *Nature Machine Intelligence* (2024), 1–11.

[24] Menglin Jia, Luming Tang, Bor-Chun Chen, Claire Cardie, Serge Belongie, Bharath Hariharan, and Ser-Nam Lim. 2022. Visual prompt tuning. In *European Conference on Computer Vision*. Springer, 709–727.

[25] Rui Jiao, Jiaqi Han, Wenbing Huang, Yu Rong, and Yang Liu. 2023. Energy-motivated equivariant pretraining for 3d molecular graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 8096–8104.

[26] Wei Ju, Zheng Fang, Yiyang Gu, Zequn Liu, Qingqing Long, Ziyue Qiao, Yifang Qin, Jianhao Shen, Fang Sun, Zhiping Xiao, et al. 2024. A comprehensive survey on deep graph representation learning. *Neural Networks* (2024), 106207.

[27] Wei Ju, Yifan Wang, Yifang Qin, Zhengyang Mao, Zhiping Xiao, Junyu Luo, Junwei Yang, Yiyang Gu, Dongjie Wang, Qingqing Long, et al. 2024. Towards Graph Contrastive Learning: A Survey and Beyond. *arXiv preprint arXiv:2405.11868* (2024).

[28] Wei Ju, Siyu Yi, Yifan Wang, Qingqing Long, Junyu Luo, Zhiping Xiao, and Ming Zhang. 2024. A survey of data-efficient graph learning. *arXiv preprint arXiv:2402.00447* (2024).

[29] Youngchun Kwon, Dongseon Lee, Youn-Suk Choi, Kyoham Shin, and Seokho Kang. 2020. Compressed graph representation for scalable molecular graph generation. *Journal of Cheminformatics* (2020).

[30] Greg Landrum et al. 2013. RDKit: A software suite for cheminformatics, computational chemistry, and predictive modeling. *Greg Landrum* 8 (2013), 31.

[31] Yi-Lun Liao and Tess Smidt. 2023. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs. *arXiv preprint arXiv:2206.11990* (2023).

[32] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *Comput. Surveys* 55, 9 (2023), 1–35.

[33] Shengchao Liu, Hanchen Wang, Weiyang Liu, Joan Lasenby, Hongyu Guo, and Jian Tang. 2021. Pre-training molecular graph representation with 3d geometry. *arXiv preprint arXiv:2110.07728* (2021).

[34] Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. 2022. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations (ICLR)*.

[35] Zemin Liu, Xingtong Yu, Yuan Fang, and Xinming Zhang. 2023. Graphprompt: Unifying pre-training and downstream tasks for graph neural networks. In *Proceedings of the ACM Web Conference 2023*. 417–428.

[36] Qingqing Long, Zheng Fang, Chen Fang, Chong Chen, Pengfei Wang, and Yuanchun Zhou. 2024. Unveiling Delay Effects in Traffic Forecasting: A Perspective from Spatial-Temporal Delay Differential Equations. In *Proceedings of the ACM on Web Conference 2024*. 1035–1044.

[37] Qingqing Long, Yilun Jin, Guojie Song, Yi Li, and Wei Lin. 2020. Graph structural-topic neural network. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1065–1073.

[38] Qingqing Long, Yilun Jin, Yi Wu, and Guojie Song. 2021. Theoretically improving graph neural networks via anonymous walk graph kernels. In *Proceedings of the Web Conference 2021*. 1204–1214.

[39] Qingqing Long, Yiming Wang, Lun Du, Guojie Song, Yilun Jin, and Wei Lin. 2019. Hierarchical community structure preserving network embedding: A subspace approach. In *Proceedings of the 28th ACM international conference on information and knowledge management*. 409–418.

[40] Qingqing Long, Lingjun Xu, Zheng Fang, and Guojie Song. 2021. HGK-GNN: Heterogeneous Graph Kernel based Graph Neural Networks. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 1129–1138.

[41] Chengqiang Lu, Qi Liu, Chao Wang, Zhenya Huang, Peize Lin, and Lixin He. 2019. Molecular property prediction: A multilevel quantum interactions modeling perspective. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 1052–1060.

[42] Kha-Dinh Luong and Ambuj K Singh. 2024. Fragment-based Pretraining and Finetuning on Molecular Graphs. *Advances in Neural Information Processing Systems* 36 (2024).

[43] Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, and Yaron Lipman. 2019. Provably powerful graph networks. *Advances in neural information processing systems* 32 (2019).

[44] John E McMurry. 2014. *Organic chemistry with biological applications*. Cengage Learning.

[45] Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. 2014. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data* 1, 1 (2014), 1–7.

[46] Yu Rong, Yatao Bian, Tingyang Xu, Weiyang Xie, Ying Wei, Wenbing Huang, and Junzhou Huang. 2020. Self-supervised graph transformer on large-scale molecular data. *Advances in Neural Information Processing Systems* 33 (2020), 12559–12571.

[47] Gerta Rücker and Christoph Rücker. 2001. Substructure, subgraph, and walk counts as measures of the complexity of graphs and molecules. *Journal of Chemical Information and Computer Sciences* 41, 6 (2001), 1457–1462.

[48] Soumya Sanyal, Janakiraman Balachandran, Naganand Yadati, Abhishek Kumar, Padmini Rajagopalan, Suchismita Sanyal, and Partha Talukdar. 2018. Mt-cgcnn:

Integrating crystal graph convolutional neural network with multitask learning for material property prediction. *arXiv preprint arXiv:1811.05660* (2018).

[49] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. 2017. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems* 30 (2017).

[50] Kristof Schütt, Oliver Unke, and Michael Gastegger. 2021. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*. PMLR, 9377–9388.

[51] Anuroop Sriram, Abhishek Das, Brandon M Wood, Siddharth Goyal, and C Lawrence Zitnick. 2022. Towards training billion parameter graph neural networks for atomic simulations. *arXiv preprint arXiv:2203.09697* (2022).

[52] Hannes Stärk, Dominique Beaini, Gabriele Corso, Prudencio Tossou, Christian Dallago, Stephan Günnemann, and Pietro Liò. 2022. 3d infomax improves gnns for molecular property prediction. In *International Conference on Machine Learning*. PMLR, 20479–20502.

[53] Mingchen Sun, Kaixiong Zhou, Xin He, Ying Wang, and Xin Wang. 2022. Gppt: Graph pre-training and prompt tuning to generalize graph neural networks. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1717–1727.

[54] Xiangguo Sun, Hong Cheng, Jia Li, Bo Liu, and Jihong Guan. 2023. All in one: Multi-task prompting for graph neural networks. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2120–2131.

[55] Susheel Suresh, Pan Li, Cong Hao, and Jennifer Neville. 2021. Adversarial graph augmentation to improve graph contrastive learning. *Advances in Neural Information Processing Systems* 34 (2021), 15920–15933.

[56] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. 2018. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219* (2018).

[57] Oliver T Unke and Markus Meuwly. 2019. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *Journal of chemical theory and computation* 15, 6 (2019), 3678–3693.

[58] Limei Wang, Yi Liu, Yuchao Lin, Haoran Liu, and Shuiwang Ji. 2022. ComENet: Towards complete and efficient message passing for 3D molecular graphs. *Advances in Neural Information Processing Systems* 35 (2022), 650–664.

[59] Xu Wang, Huan Zhao, Wei-wei Tu, and Quanming Yao. 2023. Automated 3d pre-training for molecular property prediction. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2419–2430.

[60] Michael R Wasielewski, Malcolm DE Forbes, Natia L Frank, Karol Kowalski, Gregory D Scholes, Joel Yuen-Zhou, Marc A Baldo, Danna E Freedman, Randall H Goldsmith, Theodore Goodson III, et al. 2020. Exploiting chemistry and molecular systems for quantum information science. *Nature Reviews Chemistry* 4, 9 (2020), 490–504.

[61] Jun Xia, Lirong Wu, Jintao Chen, Bozhen Hu, and Stan Z Li. 2022. Simgrace: A simple framework for graph contrastive learning without data augmentation. In *Proceedings of the ACM Web Conference 2022*. 1070–1079.

[62] Jun Xia, Chengshuai Zhao, Bozhen Hu, Zhangyang Gao, Cheng Tan, Yue Liu, Siyuan Li, and Stan Z Li. 2022. Mole-bert: Rethinking pre-training graph neural networks for molecules. In *The Eleventh International Conference on Learning Representations*.

[63] Minghao Xu, Hang Wang, Bingbing Ni, Hongyu Guo, and Jian Tang. 2021. Self-supervised graph-level representation learning with local and global structure. In *International Conference on Machine Learning*. PMLR, 11548–11558.

[64] Zhao Xu, Youzhi Luo, Xuan Zhang, Xinyi Xu, Yaochen Xie, Meng Liu, Kaleb Dickerson, Cheng Deng, Maho Nakata, and Shuiwang Ji. 2021. Molecule3d: A benchmark for predicting 3d geometries from molecular graphs. *arXiv preprint arXiv:2110.01717* (2021).

[65] Yuchen Yan, Peiyan Zhang, Zheng Fang, and Qingqing Long. 2024. Inductive Graph Alignment Prompt: Bridging the Gap between Graph Pre-training and Inductive Fine-tuning From Spectral Perspective. In *Proceedings of the ACM on Web Conference 2024*. 4328–4339.

[66] Yuning You, Tianlong Chen, Yang Shen, and Zhangyang Wang. 2021. Graph contrastive learning automated. In *International Conference on Machine Learning*. PMLR, 12121–12132.

[67] Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. 2020. Graph contrastive learning with augmentations. *Advances in neural information processing systems* 33 (2020), 5812–5823.

[68] Xingtong Yu, Yuan Fang, Zemin Liu, and Xinming Zhang. 2023. HGPROMPT: Bridging Homogeneous and Heterogeneous Graphs for Few-shot Prompt Learning. *ArXiv* abs/2312.01878 (2023).

[69] Sheheryar Zaidi, Michael Schaarschmidt, James Martens, Hyunjik Kim, Yee Whye Teh, Alvaro Sanchez-Gonzalez, Peter Battaglia, Razvan Pascanu, and Jonathan Godwin. 2022. Pre-training via denoising for molecular property prediction. *arXiv preprint arXiv:2206.00133* (2022).

[70] Zaixi Zhang, Qi Liu, Hao Wang, Chengqiang Lu, and Chee-Kong Lee. 2021. Motif-based graph self-supervised learning for molecular property prediction. *Advances in Neural Information Processing Systems* 34 (2021), 15870–15882.