LEARNING HARMONIC MOLECULAR REPRESENTATIONS ON RIEMANNIAN MANIFOLD

Anonymous authors
Paper under double-blind review

ABSTRACT

Molecular representation learning plays a crucial role in AI-assisted drug discovery research. Encoding 3D molecular structures through Euclidean neural networks has become the prevailing method in the geometric deep learning community. However, the equivariance constraints and message passing in Euclidean space may limit the network expressive power. In this work, we propose a Harmonic Molecular Representation learning (HMR) framework, which represents a molecule using the Laplace-Beltrami eigenfunctions of its molecular surface. HMR offers a multi-resolution representation of molecular geometric and chemical properties on 2D Riemannian manifold. We also introduce a harmonic message passing method to realize efficient spectral message passing over the surface manifold for better molecular encoding. Our proposed method shows comparable predictive power to current models in small molecule property prediction, and outperforms the state-of-the-art deep learning models for the rigid protein docking challenge, demonstrating its versatility in molecular representation learning.

1 INTRODUCTION

Molecular representation learning is a fundamental step in AI-assisted drug discovery. Obtaining good molecular representations is crucial for the success of downstream applications including protein function prediction (Gligorijević et al., 2021) and molecular matching, e.g., protein-protein docking (Ganea et al., 2021). In general, an ideal molecular representation should well integrate both *geometric* (e.g., 3D conformation) and *chemical* information (e.g., electrostatic potential). Additionally, such representation should capture features in various *resolutions* to accommodate different tasks, e.g., *high-level* holistic features for molecular property prediction, and *fine-grained* features for describing whether two proteins can bind together at certain interfaces.

Recently, geometric deep learning (GDL) (Bronstein et al., 2017; 2021; Monti et al., 2017) has been widely used in learning molecular representations (Atz et al., 2021; Townshend et al., 2021). GDL captures necessary information by performing neural message passing (NMP) on common structures such as 2D/3D molecular graph (Klicpera et al., 2020; Stokes et al., 2020), 3D voxel (Liu et al., 2021; Mylonas et al., 2021), and point cloud (Zhou et al., 2019; Unke et al., 2021). Specifically, GDL encodes: a) geometric features by modeling atomic positions in the *Euclidean space*, and b) chemical features by feeding atomic information into the message passing networks. High-level features could then be obtained by aggregating these atom-level features, which has shown promising results empirically.

However, we argue that current molecular representations via NMP in the Euclidean space is not necessarily the optimal solution, which suffers from several drawbacks. First, current GDL approaches need to employ equivariant networks (Thomas et al., 2018) to guarantee that the molecular representations transform accordingly upon rotation and translation (Fuchs et al., 2020), which could undermine the network expressive power (Cohen et al., 2018; Li et al., 2021). Therefore, developing a representation that could properly encode 3D molecular structure while bypassing the equivariance requirement is desirable. Second, current molecular representations in GDL are learned in a bottom-up manner, which are hardly able to provide features in different resolutions for different tasks. Specifically, NMP in Euclidean space typically achieves long-range communication between distant atoms by stacking deep layers or increasing the neighborhood radius. This would hinder the effective representation of macromolecules with tens of thousands of atoms (Liu et al., 2019;

Battiston et al., 2020; Boguna et al., 2021). To remedy this, residue-level graph representations are commonly used for large molecules (Jumper et al., 2021; Gligorijević et al., 2021). Hence designing efficient multi-resolution message passing mechanisms would be ideal for encoding molecules with distinct sizes.

Motivated by the idea of *Shape-DNA* (Reuter et al., 2006), hereby we propose Harmonic Molecular Representation learning (HMR) by utilizing the Laplace-Beltrami eigenfunctions on the molecular surface (a 2D manifold). Our representation has the following advantages: a) HMR works on 2D Riemannian manifold instead of in the 3D Euclidean space, thus the resulting molecular representation is by design roto-translation invariant. b) HMR represents a molecule in a top-down manner, and is capable of offering multi-resolution features that accommodate various target molecules (i.e., from small molecules to large proteins), thanks to the smooth nature of the Laplace-Beltrami eigenfunctions (Fig. 1). c) HMR integrates geometric and chemical features in the same embedding space. Once the molecular surface manifold is obtained, we can encode local chemical environment as learnable functions distributed on the Riemannian manifold for various applications.

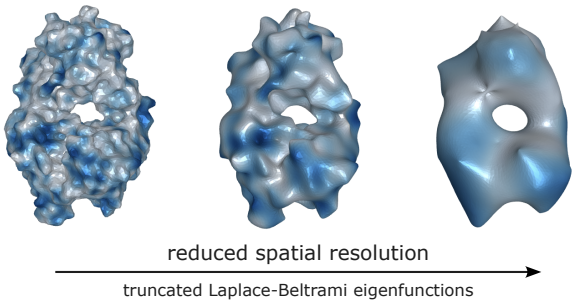

Figure 1: Multi-resolution molecular surface representation. Showing the electrostatic potential (blue being negatively charged) on the protein surface (PDB ID: 3V6F) at different resolutions under our representation.

Note that, our proposed HMR is non-trivial. To demonstrate that HMR is generally applicable to different downstream tasks including molecular property prediction and molecular matching, we propose two specific techniques for these tasks: 1) manifold harmonic message passing for realizing holistic molecular representation, and 2) learning regional functional correspondence for molecular surface matching. Without loss of generality, we apply the proposed techniques to respectively solve two drug discovery-related problems: QM9 small molecule property regression, and rigid protein docking pose prediction. Our proposed method shows comparable performance for small molecule property prediction to NMP-based models, while outperforming the state-of-the-art deep learning model in the rigid protein docking challenge.

2 Related Work

Euclidean Neural Networks Graph Neural Networks (GNNs) have been extensively used for molecular representations (Duvenaud et al., 2015; Gilmer et al., 2017). Recently, roto-translation equivariant GNNs have become more prevalent since they hard-code the equivariance condition into their architectures, hence are able to better describe 3D atomic structures (Fuchs et al., 2020; Geiger & Smidt, 2022). These models have been successfully applied to tackle the protein folding problem (Baek et al., 2021) and the rigid protein docking task (Ganea et al., 2021).

Molecular Surface Representation On the other hand, the molecular surface is a high-level representation of a molecule's shape, which has been widely used to study inter-molecular interactions (Gainza et al., 2020). Non-Euclidean convolutional neural networks (Monti et al., 2017) and point cloud-based learning models (Sverrisson et al., 2022) have been applied to encode the molecular surface for downstream prediction tasks, e.g., protein binding site prediction (Mylonas et al., 2021). However, existing methods apply filters with fixed sizes and are highly dependent on the surface mesh quality, which limits the expressive power for molecular shape representation.

Geometry Processing Our work was inspired by some pioneering work for shape encoding and shape matching in the geometry processing research community (Litman & Bronstein, 2013; Biasotti et al., 2016). The surface of a 3D object is typically discretized into a polygon mesh with vertices and faces. Intrinsic properties of the surface manifold are used to encode the shape (Sun et al., 2009; Bronstein & Kokkinos, 2010). Functional maps (Ovsjanikov et al., 2012; Litany et al., 2017b) have been proposed to establish spectral-space functional correspondence between two man-

ifolds. Recently, deep learning has been applied to learn representative features to facilitate shape recognition and matching (Litany et al., 2017a; Donati et al., 2020; Attaiki et al., 2021).

Spectral Message Passing Our proposed method decomposes surface functions as the linear combination of some basis functions (hence the name "harmonic") and realizes message passing by applying various spectral filters. Graph convolutional networks (GCNs) is closely related to our method, which operates in the graph Laplacian eigenspace (Kipf & Welling, 2016; Shen et al., 2021). The major difference is that graph is discrete by construction, while our method works with continuous Riemannian manifold (Coifman & Lafon, 2006). In other words, the underlying manifold and its spectrum remain the same with different surface discretizations, hence is a robust representation of the surface shape (Coifman et al., 2005). See "Shape-DNA" in Sec. 3.

Rigid Protein Docking The goal of rigid protein docking is to predict the protein complex structure given two individual protein structures, assuming no conformational change during the docking process. Traditional docking methods (Yan et al., 2020; de Vries et al., 2015) typically generate millions of candidate docking poses based on domain knowledge, and then apply a scoring function (e.g., energy-based) to select the top candidates for evaluation. Such approaches have shown good docking power, but at a huge computational cost due to the thorough searches for candidate poses, hindering their applications at large scale. Recently, deep learning models (Sverrisson et al., 2022; Ganea et al., 2021) have shown promising results in this challenge, which are much faster in making predictions. However, the model performance still has much room for improvement compared with traditional docking methods.

3 PRELIMINARY

The goal of this work is to propose a representation learning method using the molecular surface, which could properly encode both geometric and chemical features. Intuitively, the molecular surface defines the shape of a molecule in 3D Euclidean space (i.e., geometry), while chemical features (e.g., hydrophobicity) can be treated as functions distributed on the molecular surface. The surface and these associated functions co-determine the underlying molecular properties, e.g., whether two proteins have strong interactions. Therefore, a molecule can be represented as a set of (learned) functions defined on its surface.

Before moving on, we first explain some basic concepts behind our proposed representation learning framework. In Sec. 3.1, we introduce molecular "Shape-DNA", which is a set of basis functions inherent to the surface manifold. In Sec. 3.2, we illustrate how to use harmonic analysis to decompose a function (defined on the molecular surface) into the linear combination of the "Shape-DNA" basis , as shown in Fig. 2. These concepts enable us to represent a molecular surface as a 2D Riemannian manifold with associated functions, which form the cornerstone of our proposed framework.

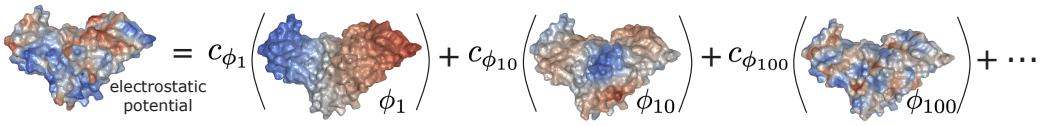

Figure 2: Illustrating harmonic analysis. Left-hand side: the electrostatic potential function on the protein surface. Right-hand side: the linear combination of Laplace-Beltrami eigenfunctions with corresponding coefficients. These eigenfunctions exhibit different spatial frequencies (resolutions).

3.1 THE SHAPE-DNA

The molecular surface [1] can be viewed as a 2D Riemannian manifold ($\mathcal{M}$), which adopts a discrete set of eigenfunctions (ϕ) that solves

$$\Delta\phi_i = \lambda_i \phi_i, \quad i = 0, 1 \ldots \tag{1}$$

Here, Δ is the Laplace-Beltrami (LB) operator acting on surface scalar fields, defined as $\Delta f = -\mathrm{div}(\nabla f)$. $\phi_0, \phi_1, \ldots$ are a set of orthonormal eigenfunctions (i.e., $\langle \phi_i, \phi_j \rangle_{\mathcal{M}} = \delta_{ij}$). And $0 =$

[1] For instance, an isosurface of its electron density field, solvent-accessible/exclusive surface, etc.

$\lambda_0 \leq \lambda_1 \leq \ldots$ are the corresponding eigenvalues. These set of eigenfunctions ϕ_i of the LB operator are called the "Shape-DNA" since each shape exhibits its unique eigenfunctions [2].

Fig. 2 shows a few selected eigenfunctions (ϕ_1, ϕ_{10}, ϕ_{100}) of a protein surface manifold on the right-hand side. These eigenfunctions are intrinsic properties of the surface manifold, which remain invariant under rigid transformations to the molecule. The eigenvalues λ_i reflect the surface Dirichlet energy (defined as $\langle \Delta f, f \rangle_{\mathcal{M}}$), which measures the smoothness of a function on manifold $\mathcal{M}$. The eigenvalues increase linearly, whose slope is roughly inversely proportional to the surface area (known as Weyl's asymptotic law,). Therefore, molecules with different surface areas exhibit distinct eigenvalues (see Appendix A).

3.2 Basics of Manifold Harmonic Analysis

Now, we introduce basic manifold harmonic analysis, a merit of using the Riemannian manifold representation. "Harmonic analysis" refers to the representation of functions as the superposition of some basic waves. Specifically in our case, given the molecular surface manifold $\mathcal{M}$ and its LB eigenfunctions $\{\phi_i\}_{i=0}^{\infty}$, any scalar-valued function f that is square-integrable on $\mathcal{M}$ can be decomposed into a generalized Fourier series:

$$f(x) = \sum_{i=0}^{\infty} \langle f, \phi_i \rangle_{\mathcal{M}} \phi_i(x) \,. \tag{2}$$

In other words, f can be represented as the linear combination of the LB eigenfunctions. The linear combination coefficient $\langle f, \phi_i \rangle_{\mathcal{M}}$ can be interpreted as the "projection" of f onto the eigenfunction ϕ_i, which reflects the contribution of this particular eigenfunction to synthesizing function f.

Interestingly, it is easy to notice that these eigenfunctions display different spatial frequencies, i.e., ϕ_1 varies slowly over the surface, while ϕ_{100} oscillates at a much higher frequency. This is analogous to the set of $\sin(kx)$ functions in the 1D case, where high frequency waves with larger k values exhibit more oscillations within the period of length 2π. Hereafter we refer to LB eigenfunctions with (relatively) small/large eigenvalues as the low/high frequency components.

Therefore, by manipulating the linear combination coefficients, we could control the contribution of different frequency components to the synthesized function f. For instance, with only low frequency components, the synthesized function will have a lower spatial resolution (i.e., more smoothed out) on the surface (e.g., Fig. 1 right panel), and vice versa. Synthesizing a new function f with some selected frequency components is commonly known as wave filtering in signal processing, which will be applied in our representation learning framework to realize multi-resolution encoding of the molecular surface. We refer the readers to some excellent review papers for more detailed introductions (Bronstein et al., 2017; Rosenberg, 1997).

4 Methodology

We realize that the aforementioned geometry processing methods are suitable for molecular surface representations. However, a clear distinction between the 3D objects used in geometry processing research and the molecular systems is that molecules are not simply shapes – their underlying atomic structures beneath the surface govern the molecular functionality. In other words, both geometry (i.e., shape) and chemistry have to be considered for molecular encoding. It has been shown that the molecular surface displays chemical and geometric patterns which fingerprint a protein's mode of interactions with other biomolecules (Gainza et al., 2020). Therefore, we formulate both geometric and chemical properties of a molecule as functions distributed on its surface manifold.

Then comes the question: how do we properly learn these geometric and chemical functions on the molecular surface? One viable solution is to emulate the message passing framework for GNNs on a discretized manifold, where we perform message passing between the surface vertices along the edges defined by the polygon mesh. However, a surface mesh typically contains tens of thousands of vertices, thus graph message passing is inefficient for macromolecules with large surface area.

[2] Two shapes may share the same LB eigendecomposition (isometries). The uniqueness of eigendecomposition up to isometries is associated with the question "Can One Hear the Shape of a Drum?" (Kac, 1966)

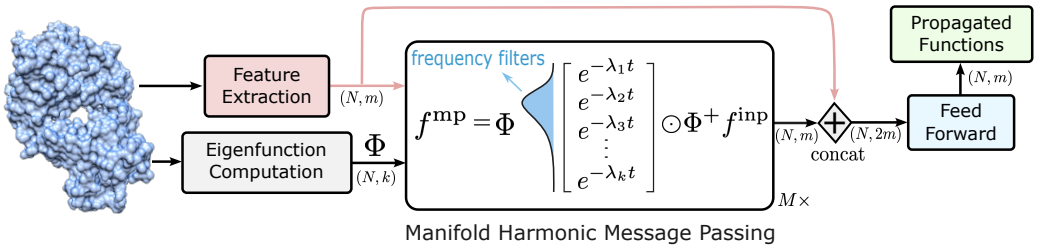

Figure 3: HMR workflow. Given a molecular surface, we first compute its Laplace-Beltrami eigenfunctions and the initial surface features. Then, we apply some neural network-learned spectral filters to propagate the initial function over the surface to achieve message passing. Multiple message passing blocks can be stacked for better representations.

In this section, we present two methods under the HMR framework. In Sec. 4.1, we introduce manifold harmonic message passing, which makes use of harmonic analysis techniques to allow efficient multi-range communication between points on the molecular surface regardless of its size. In Sec. 4.2, we propose a rigid protein docking pipeline, which is powered by the HMR module. We use this docking challenge to demonstrate the potential of our proposed representation learning method for modelling large biomolecules.

4.1 Learning Harmonic Molecular Representations

Given the 3D atomic structure of a molecule, HMR returns a set of m learned functions distributed on the molecular surface. These functions are stored as an array $\mathsf{F} \in \mathbb{R}^{N\times m}$, where N is the number of vertices in the molecular surface triangle mesh. These learned functions (i.e., features) represent the underlying molecule, and can be used for various downstream prediction tasks.

Surface Preparation We use MSMS (Ewing & Hermisson, 2010) to compute the molecular solvent-excluded surface as a triangle mesh, which includes N vertices $\{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathbb{R}^3$ and the corresponding faces. We then calculate the discrete LB operator as described in Reuter et al. (2009). We compute the first k eigenfunctions $\{\phi_i\}_{i\geq 0}^{k-1}$ with ascending eigenvalues, and stack them into an array $\Phi \in \mathbb{R}^{N\times k}$, where each column stores an eigenfunction.

Initial geometric features on the surface manifold can be directly computed given the surface triangulation. We compute the per-vertex Gaussian curvature, mean curvature, and the Heat Kernel Signatures (HKS) as described in Sun et al. (2009). The geometric features capture shape-related properties of the molecular surface, and are stored as a scalar-type array $\mathsf{F}^{\mathrm{geom}} \in \mathbb{R}^{N\times p}$, where p is the number of geometric features we used (user defined variable, same for q below).

Chemical features are projected from atoms to their neighboring surface vertices. We first obtain an initial descriptor vector $\mathbf{u} \in \mathbb{R}^r$ for each atom (e.g., atomic number, charge, etc., which is flexible and task dependent). Then, for each surface vertex $\mathbf{x}_i$, we compute its v nearest neighbor atoms centered at $\{\mathbf{a}_1^i, \dots, \mathbf{a}_v^i\}$ with features $\{\mathbf{u}_1^i, \dots, \mathbf{u}_v^i\}$. We apply a multilayer perceptron (MLP) with output dimension q to the vector $\left[\mathbf{u}_v^i, 1/\|\mathbf{x}_i - \mathbf{a}_v^i\|\right]$ for each neighboring atom, then compute the average over the neighbors to obtain the feature vector $\mathbf{f}_i^{\mathrm{chem}} \in \mathbb{R}^q$, which are assembled into the chemical features $\mathsf{F}^{\mathrm{chem}} \in \mathbb{R}^{N\times q}$.

In our implementation, we use another MLP to combine the per-vertex initial features $\mathsf{F}^{\mathrm{inp}} \leftarrow \mathrm{MLP}(\mathrm{concat}(\mathsf{F}^{\mathrm{geom}}, \mathsf{F}^{\mathrm{chem}}))$, $\mathsf{F}^{\mathrm{inp}} \in \mathbb{R}^{N\times m}$. These features reflect the local geometric and chemical environment of each surface vertex, which will be used as the input to the harmonic message passing module. The output of the surface preparation module includes the molecular surface Laplace-Beltrami eigenfunctions Φ and the per-vertex features $\mathsf{F}^{\mathrm{inp}}$.

Harmonic Message Passing Our proposed harmonic message passing mechanism is closely related to the heat diffusion process on an arbitrary surface. Joseph Fourier developed spectral analysis methods to solve the heat equation $\partial f/\partial t + \Delta f = 0$, where f is some heat distributed on the surface. This concise partial differential equation describes how a heat distribution f evolves over time, whose solution can be expressed using the heat operator $\exp(-\Delta t)$, i.e., $f(t) = \exp(-\Delta t) f_0$ for initial heat distribution f_0 at $t = 0$. Intuitively, heat will flow from hot regions to cool regions on

the surface. As time goes to infinity, the heat distribution f will converge to a constant value (i.e., the global average temperature on the surface), assuming that total energy is conserved.

In fact, heat diffusion can be thought of as a message passing process, where surface regions with different temperatures communicate with each other and propagate the initial heat distribution deterministically. The heat exchange rate is dependent on the difference in temperature, while the message passing distance is determined by the heat diffusion time t. Following this idea, we generalize the heat diffusion process by proposing a function propagation operator $\mathcal{P}$ with neural network-learned frequency filter $F_\theta(\lambda)$:

$$\mathcal{P}f = \sum_i F_\theta(\lambda_i)\langle f, \phi_i\rangle_{\mathcal{M}}\phi_i\,, \tag{3}$$

$$F_\theta(\lambda) = \exp\left(-\frac{(\lambda-\mu)^2}{\sigma^2}\right)\cdot\exp(-\lambda t) \qquad \text{where} \quad \theta = (\mu, \sigma, t)\,. \tag{4}$$

As shown in Eq. 3, the input function f is first expanded as the linear combination of the LB eigenfunctions Φ with coefficients $\langle f, \phi_i\rangle_{\mathcal{M}}$. We then learn a spectral-space frequency filter $F_\theta(\lambda)$, which is a function of the corresponding eigenvalue λ. The filter $F_\theta(\lambda)$ consists of two components (Eq. 4): a Gaussian frequency filter (with parameters μ and σ), and the heat operator part $e^{-\lambda t}$ (with parameter t). The matrix representation of the function propagation operator $\mathcal{P}$ is plotted in Fig. 3.

For each input function f, the neural network learns a unique set of parameters (μ, σ, t). In Fig. 4, we showcase a few examples of surface functions obtained by applying different frequency filters to an initial delta function.

The Gaussian frequency filter allows the network to propagate the input function along some selected eigenfunctions with eigenvalues close to μ, while the number of selected eigenfunctions is determined by the filter width σ. This is similar to a band-pass filter in signal processing, where some frequency components are eliminated. Owing to the multi-resolution nature of the LB eigenfunctions with different spatial frequencies, this filter will help capture surface functions at different resolutions. For instance, hydrogen bond is a local interaction between a few atoms, while hydrophobic interactions typically span a larger surface region.

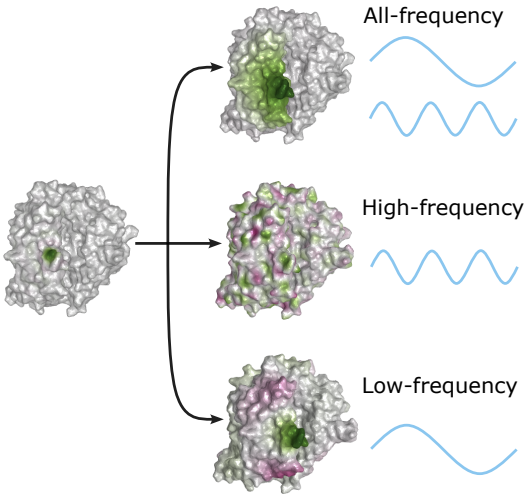

Figure 4: Examples of distinct message passing outcomes using different frequency filters.

The heat operator part governs how long we would like to propagate the function. With longer propagation time, the function f will be more averaged out towards the global mean, leading to a smoothed function. In addition, the heat operator is by definition a low-pass filter, where components with larger eigenvalues decay faster. This would make eigenfunctions with large eigenvalues contribute less significantly during message passing. Therefore, the combination of the Gaussian frequency filter and heat operator could help the network focus on some higher frequency components, which is similar in spirit to the method proposed in Aubry et al. (2011).

As demonstrated, HMR is able to represent a variety of different functions on the surface through harmonic message passing. The output of this block has the same size as the input, $\mathsf{F}^{\mathrm{mp}} \in \mathbb{R}^{N\times m}$, since each channel respectively contains the propagated version of its input function.

Feature Aggregation Finally, we use a MLP to obtain the final molecular surface representation after concatenating features from the surface preparation and harmonic message passing blocks ($\mathsf{F}^{\mathrm{out}} \leftarrow \mathrm{MLP}(\mathrm{concat}[\mathsf{F}^{\mathrm{inp}}, \mathsf{F}^{\mathrm{mp}}])$). Each surface vertex now contains m learned functions representing its neighboring geometric and chemical environment across multiple spatial scales. These per-vertex features can be used for surface property-related tasks, e.g., protein binding site prediction introduced in Sec. 4.2. In addition, molecule-level representations could be obtained through global pooling, e.g., max/average pooling, see Sec. 5.1 for a use case.

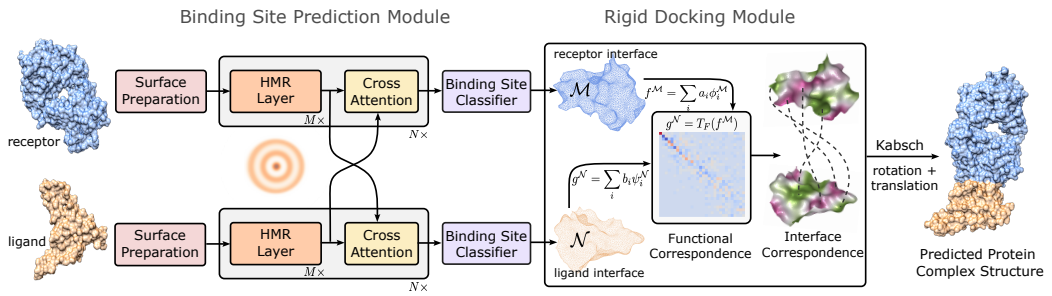

Figure 5: The rigid protein docking pipeline. Given the surface of the ligand and receptor proteins, we apply multiple HMR and cross attention layers to allow communication within and between the surfaces. The learned surface representations are then used to predict the binding interfaces and establish functional correspondence using functional maps. Rigid docking is achieve by converting the functional correspondence to rigid transformations which aligns the predicted interfaces.

4.2 Learning Surface Correspondence for Rigid Protein Docking

In this section, we demonstrate how to learn surface correspondence for molecular matching. Specifically, we introduce a surface-based rigid protein docking workflow (Fig. 5) powered by HMR and functional maps (Ovsjanikov et al., 2012).

It has been shown that protein-protein interfaces exhibit similar geometric and chemical patterns (Gainza et al., 2020). In other words, two proteins may interact if part of their surfaces display similar shapes and chemical functions (i.e., functional correspondence). Therefore, we propose to realize rigid docking in two consecutive steps: 1) given two protein surfaces, predict the region where binding might occur (i.e., binding site prediction); and 2) establish functional correspondence between the ligand/receptor binding surfaces, and convert it to real-space vertex-to-vertex correspondence. Rigid docking could then be achieved by aligning the corresponding surface vertices (e.g., through the Kabsch algorithm (Kabsch, 1976)).

Binding Site Prediction Given the protein surface meshes, we first predict the regions where they interact, which is a per-vertex binary classification problem. We iteratively apply HMR and cross-attention layers (Fig. 5) to encode the surfaces with intra- and inter-surface communications. Next, we use the learned features on each vertex to classify whether it belongs to the binding interface. Detailed descriptions of this module are available in Appendix B.

The output of the binding site prediction module includes a) the receptor (ligand) surface mesh of the binding interface, which is a submanifold of the entire protein surface, represented as $\mathcal{M}$ ($\mathcal{N}$) with M (N) vertices. b) Neural network-learned features or so-called surface functions $\mathsf{F}^{\mathcal{M}}$ and $\mathsf{G}^{\mathcal{N}}$ distributed on the binding interfaces.

Rigid Docking with Functional Maps We know that protein interfaces $\mathcal{M}$ and $\mathcal{N}$ exhibit similar geometry and also adopt a set of m corresponding functions $\mathsf{F}^{\mathcal{M}} \in \mathbb{R}^{M\times m}$ and $\mathsf{G}^{\mathcal{N}} \in \mathbb{R}^{N\times m}$ (Fig. 5). Intuitively, if we could somehow align these set of corresponding functions, then we have found a way to align the protein interfaces. To that end, we employ functional maps to "align" $\mathsf{F}^{\mathcal{M}}$ and $\mathsf{G}^{\mathcal{N}}$ in spectral domain.

Specifically, given the truncated LB eigenfunctions $\{\phi_j^{\mathcal{M}}\}_{j\geq 0}^{k}$ and $\{\phi_i^{\mathcal{N}}\}_{i\geq 0}^{k}$ of the receptor and ligand interfaces, we compute the spectral representation of learned functions as $\mathsf{A} = (\Phi^{\mathcal{M}})^{+}\mathsf{F}^{\mathcal{M}}$, and $\mathsf{B} = (\Phi^{\mathcal{N}})^{+}\mathsf{G}^{\mathcal{N}}$, where $+$ denotes the Moore-Penrose pseudo-inverse, and $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{k\times m}$. Functional correspondence can be obtained by minimizing the following:

$$\mathsf{C} = \underset{\mathsf{C}\in\mathbb{R}^{k\times k}}{\arg\min} \|\mathsf{CA} - \mathsf{B}\|_{\mathrm{F}} \,,$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. Once the functional mapping (i.e., the C matrix) is recovered, point-to-point correspondence between the ligand and receptor surfaces can be obtained by mapping indicator functions of vertices on $\mathcal{M}$ to those on $\mathcal{N}$. In practice, we adopted a slightly more complicated functional maps approach, which is illustrated in Appendix C.

Finally, we perform rigid docking by aligning the proteins according to the point-to-point surface correspondence using the Kabsch algorithm.

5 EXPERIMENTS

5.1 QM9 MOLECULAR PROPERTY REGRESSION

As a toy task to test our representation for small molecules, we perform property regression on the QM9 dataset (Ramakrishnan et al., 2014). We randomly split the dataset into 80%/10%/10% for training, validation, and test, respectively. The atomic features we used include atomic number, hybridization, and aromaticity (one-hot encoded). HMR was used to obtain the molecular surface representations, followed by an average pooling over all vertices to get the final molecular representation for the regression task.

Table 1: QM9 Mean Absolute Error.

Task Unit	α bohr^3	$\Delta\varepsilon$ meV	$\varepsilon_{\text{HOMO}}$ meV	$\varepsilon_{\text{LUMO}}$ meV	μ D	C_v cal/(mol K)
Chem. Accu.	.1	43	43	43	.1	.050
NMP	.092	69	43	38	.030	.040
SchNet	.235	63	41	34	.033	.033
Cormorant	.085	61	34	38	.038	.026
EGNN	.071	48	29	25	.029	.031
SE(3)-Tr.	.142	53	35	33	.051	.054
HMR	.141	68	42	37	.052	.053

For comparison, we show the results from both invariant and equivarient networks for 6 regression tasks in Table 1, including NMP (Gilmer et al., 2017), SchNet (Schütt et al., 2018), Cormorant (Anderson et al., 2019), EGNN (Satorras et al., 2021), and SE(3)-Transformer (Fuchs et al., 2020). Interestingly, despite HMR completely discards the bonding information between atoms and only performs massage passing over the molecular surface, it still shows comparable performance in predicting molecular properties. Indeed, from the density functional theory (Burke, 2012) point of view, it is the electron density function that ultimately determines molecular properties. This possibly explains the success of predicting these properties using only molecular surface information, since the surface is approximately an isosurface of the molecular electron density field.

5.2 RIGID PROTEIN DOCKING

We next evaluate HMR on a more challenging task: rigid protein docking.

Experimental setup HMR is trained on Database of Interacting Protein Structures (DIPS) (Townshend et al., 2019), which is modified to better represent the interactions between different proteins (see Appendix E for details); and evaluated on the golden standard Docking Benchmark 5.5 (DB5.5) (Guest et al., 2021). We compare our model with the state-of-the-art GNN-based deep learning model EQUIDOCK (Ganea et al., 2021) and two traditional docking methods, ATTRACT (de Vries et al., 2015) and HDOCK (Yan et al., 2020). To evaluate docking performance, we compute the Complex and Interface root-mean-square deviation (RMSD) following Ganea et al. (2021). In addition, we evaluate the overall quality of docking using DockQ (Basu & Wallner, 2016) and compute the success rate on whether the result achieves "Acceptable" or higher according to Lensink & Wodak (2013). See Appendix F for evaluation details.

Results Model performance are summarized in Table 2 and Fig. 6a. Our model (HMR) (Top 1) outperforms the GNN-based EQUIDOCK model under all metrics. Notably, HMR achieves a much higher success rate, indicating more test cases have result close to the ground truth structures. We note that traditional methods still exceed in terms of docking performance but at greater computational cost. Further experiments show that the harmonic message passing mechanism learns to propagate information at different scales (Fig. G.1) and is critical to the efficiency of the model (Table H.1. Our ablation study show that both the chemical and geometric information are important for rigid protein docking (Table H.2). Compared to EQUIDOCK, the novel framework of HMR and optimized training dataset collaboratively contribute to the higher performance (Table H.3).

We further examine the functions learned at predicted binding sites. As shown in Fig. 6bc, functions on the two binding sites are highly correlated, confirming the good alignment achieved using functional maps. Cases with higher docking quality show stronger function correlations between the binding sites, suggesting our HMR is able to learn complex surface interaction patterns, which also supports the claims in Gainza et al. (2020).

HMR predicts multiple binding sites for some proteins. We also assess the model performance by including candidate poses from top 3 binding site pairs, ranked by the mean probability predicted by the binding site classifer. The best scores from top 3 poses show that the performance of HMR is competitive to ATTRACT. We emphasize that multiple poses in our model are not from sampling but merely multiple predictions from the model with little computational overhead.

Table 2: Rigid Prediction Results on Docking Benchmark 5.5

	Complex RMSD ↓		Interface RMSD ↓		DockQ ↑		Success rate ↑
Model	Median	Mean	Median	Mean	Median	Mean	(≥ Acceptable)
HMR (Top 1)	12.01	12.09	10.90	11.43	0.06	0.22	0.28
HMR (Top 3)	8.44	9.68	7.50	8.70	0.10	0.26	0.33
EQUIDOCK*	15.90	17.18	14.45	14.85	0.02	0.04	0.00
ATTRACT	8.99	11.11	10.31	11.67	0.07	0.42	0.44
HDOCK	0.39	5.97	0.32	5.68	0.97	0.71	0.73

*EQUIDOCK model is trained on DIPS, provided by the authors

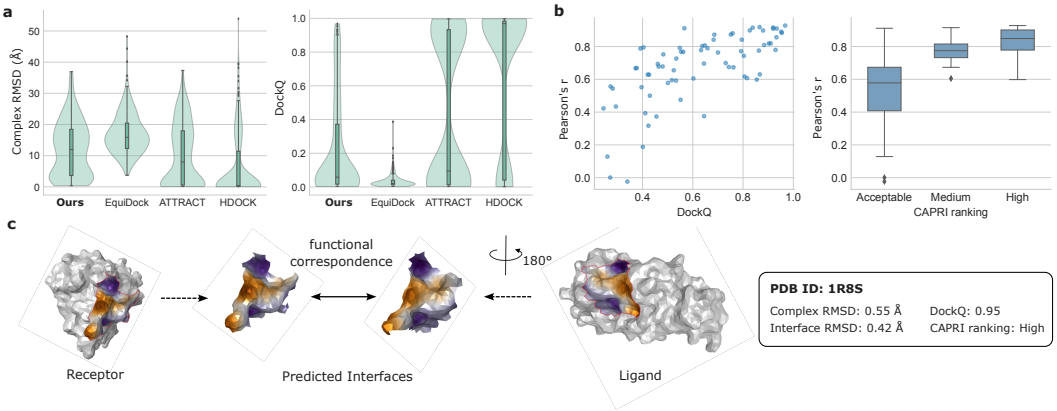

Figure 6: Rigid docking performance analysis. **a** Distribution of Complex RMSD and DockQ scores for poses predicted using HMR (ours, Top 1), EQUIDOCK, HDOCK, and ATTRACT. **b** Correlations between learned functions on the predicted ligand-receptor interfaces. For each test case, we calculate the correlation as the Pearson's r between the learned function values on ligand interface vertices and on their nearest receptor vertices (corresponding vertices), averaged over 128 hidden channels. **c** A showcase of functional correspondence between interfaces for a well-docked case.

6 CONCLUSIONS AND OUTLOOK

We presented HMR, a powerful surface manifold-based molecular representation learning framework. By integrating geometric and chemical properties as functions distributed on the molecular surface manifold, and applying harmonic message passing in the spectral domain, we achieve multi-resolution molecular representations, which shows promising performance in both molecular property prediction and molecular matching tasks. Our work highlights an important aspect of molecular "structure-activity" relationship, that is – "shape-activity" relationship. We think this is particularly important for large biomolecules, where the surface shape and chemical patterns determine some fundamental biological activities, such as protein-protein interactions.

We look forward to incorporating protein flexibility (i.e., dynamics) to molecular representations, where the protein "Shape-DNA" may reveal its unique advantages. One of the foreseeable challenges is the fast computation of the Laplace-Beltrami eigenvectors when the surface shape undergoes significant changes (e.g., the Complementarity-Determining Regions of an antibody upon binding with an antigen). We call for more research attention to surface manifold-based molecular representations and functional treatment of the chemical environment.

REPRODUCIBILITY STATEMENT

We uploaded all the code for experiments presented in this paper, the corresponding hyperparameters are presented in Table I.5. DIPS dataset can be downloaded from the following website `https://github.com/BioinfoMachineLearning/DIPS-Plus`. EQUIDOCK model and checkpoints can be downloaded from `https://github.com/octavian-ganea/equidock_public`. ATTRACT can be downloaded from `https://github.com/sjdv1982/attract`. HDOCK is implemented using its local version `HDOCKlite`, which can be downloaded from `http://huanglab.phys.hust.edu.cn/software/hdocklite/`. DockQ can be downloaded from `https://github.com/bjornwallner/DockQ/`.

REFERENCES

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

Souhaib Attaiki, Gautam Pai, and Maks Ovsjanikov. DPFM: Deep partial functional maps. In *2021 International Conference on 3D Vision (3DV)*, pp. 175–185. IEEE, 2021.

Kenneth Atz, Francesca Grisoni, and Gisbert Schneider. Geometric deep learning on molecular representations. *Nature Machine Intelligence*, 3(12):1023–1032, 2021.

Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *2011 IEEE international conference on computer vision workshops (ICCV workshops)*, pp. 1626–1633. IEEE, 2011.

Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.

Sankar Basu and Björn Wallner. DockQ: A Quality Measure for Protein-Protein Docking Models. *PLOS ONE*, 11(8):e0161879, 2016.

Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: structure and dynamics. *Physics Reports*, 874:1–92, 2020.

Silvia Biasotti, Andrea Cerri, Alex Bronstein, and Michael Bronstein. Recent trends, applications, and perspectives in 3d shape similarity assessment. In *Computer graphics forum*, volume 35, pp. 87–119. Wiley Online Library, 2016.

Marian Boguna, Ivan Bonamassa, Manlio De Domenico, Shlomo Havlin, Dmitri Krioukov, and M Serrano. Network geometry. *Nature Reviews Physics*, 3(2):114–135, 2021.

Michael M Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *2010 IEEE computer society conference on computer vision and pattern recognition*, pp. 1704–1711. IEEE, 2010.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

Kieron Burke. Perspective on density functional theory. *The Journal of chemical physics*, 136(15): 150901, 2012.

Taco S Cohen, Mario Geiger, Jonas Köhler, and Max Welling. Spherical CNNs. *arXiv preprint arXiv:1801.10130*, 2018.

Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

Ronald R Coifman, Stephane Lafon, Ann B Lee, Mauro Maggioni, Boaz Nadler, Frederick Warner, and Steven W Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the national academy of sciences*, 102(21):7426–7431, 2005.

Sjoerd J. de Vries, Christina E. M. Schindler, Isaure Chauvot de Beauchêne, and Martin Zacharias. A Web Interface for Easy Flexible Protein-Protein Docking with ATTRACT. *Biophysical Journal*, 108(3):462–465, February 2015.

Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep geometric functional maps: Robust feature learning for shape correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8592–8601, 2020.

David K Duvenaud, Dougal Maclaurin, Jorge Iparraguirre, Rafael Bombarell, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P Adams. Convolutional networks on graphs for learning molecular fingerprints. *Advances in neural information processing systems*, 28, 2015.

Gregory Ewing and Joachim Hermisson. Msms: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics*, 26(16):2064–2065, 2010.

Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. SE(3)-Transformers: 3D roto-translation equivariant attention networks. *Advances in Neural Information Processing Systems*, 33:1970–1981, 2020.

Pablo Gainza, Freyr Sverrisson, Frederico Monti, Emanuele Rodola, D Boscaini, MM Bronstein, and BE Correia. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods*, 17(2):184–192, 2020.

Octavian-Eugen Ganea, Xinyuan Huang, Charlotte Bunne, Yatao Bian, Regina Barzilay, Tommi Jaakkola, and Andreas Krause. Independent SE(3)-equivariant models for end-to-end rigid protein docking. *arXiv preprint arXiv:2111.07786*, 2021.

Mario Geiger and Tess Smidt. e3nn: Euclidean neural networks. *arXiv preprint arXiv:2207.09453*, 2022.

Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pp. 1263–1272. PMLR, 2017.

Vladimir Gligorijević, P Douglas Renfrew, Tomasz Kosciolek, Julia Koehler Leman, Daniel Berenberg, Tommi Vatanen, Chris Chandler, Bryn C Taylor, Ian M Fisk, Hera Vlamakis, et al. Structure-based protein function prediction using graph convolutional networks. *Nature communications*, 12(1):1–14, 2021.

Johnathan D. Guest, Thom Vreven, Jing Zhou, Iain Moal, Jeliazko R. Jeliazkov, Jeffrey J. Gray, Zhiping Weng, and Brian G. Pierce. An expanded benchmark for antibody-antigen docking and affinity prediction reveals insights into antibody recognition determinants. *Structure*, 29(6):606–621.e5, June 2021.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32 (5):922–923, 1976.

Mark Kac. Can one hear the shape of a drum? *The american mathematical monthly*, 73(4P2):1–23, 1966.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Johannes Klicpera, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*, 2020.

Marc F. Lensink and Shoshana J. Wodak. Docking, scoring, and affinity prediction in CAPRI. *Proteins: Structure, Function, and Bioinformatics*, 81(12):2082–2095, 2013.

Xianzhi Li, Ruihui Li, Guangyong Chen, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. A rotation-invariant framework for deep point cloud analysis. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

Or Litany, Tal Remez, Emanuele Rodola, Alex Bronstein, and Michael Bronstein. Deep functional maps: Structured prediction for dense shape correspondence. In *Proceedings of the IEEE international conference on computer vision*, pp. 5659–5667, 2017a.

Or Litany, Emanuele Rodolà, Alexander M Bronstein, and Michael M Bronstein. Fully spectral partial shape matching. In *Computer Graphics Forum*, volume 36, pp. 247–258. Wiley Online Library, 2017b.

Roee Litman and Alexander M Bronstein. Learning spectral descriptors for deformable shape correspondence. *IEEE transactions on pattern analysis and machine intelligence*, 36(1):171–180, 2013.

Qi Liu, Maximilian Nickel, and Douwe Kiela. Hyperbolic graph neural networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Qinqing Liu, Peng-Shuai Wang, Chunjiang Zhu, Blake Blumenfeld Gaines, Tan Zhu, Jinbo Bi, and Minghu Song. OctSurf: Efficient hierarchical voxel-based molecular surface representation for protein-ligand affinity prediction. *Journal of Molecular Graphics and Modelling*, 105:107865, 2021.

Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein. Geometric deep learning on graphs and manifolds using mixture model CNNs. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5115–5124, 2017.

Stelios K Mylonas, Apostolos Axenopoulos, and Petros Daras. DeepSurf: a surface-based deep learning approach for the prediction of ligand binding sites on proteins. *Bioinformatics*, 37(12): 1681–1690, 2021.

Raúl Méndez, Raphaël Leplae, Leonardo De Maria, and Shoshana J. Wodak. Assessment of blind predictions of protein–protein interactions: Current status of docking methods. *Proteins: Structure, Function, and Bioinformatics*, 52(1):51–67, 2003.

Raúl Méndez, Raphaël Leplae, Marc F. Lensink, and Shoshana J. Wodak. Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. *Proteins: Structure, Function, and Bioinformatics*, 60(2):150–169, 2005.

Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (ToG)*, 31(4):1–11, 2012.

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

Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. Continuous and orientation-preserving correspondences via functional maps. *ACM Transactions on Graphics (ToG)*, 37(6): 1–16, 2018.

Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace–beltrami spectra as 'shape-dna'of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.

Martin Reuter, Silvia Biasotti, Daniela Giorgi, Giuseppe Patanè, and Michela Spagnuolo. Discrete laplace–beltrami operators for shape analysis and segmentation. *Computers & Graphics*, 33(3): 381–390, 2009.

Steven Rosenberg. *The Laplacian on a Riemannian manifold: an introduction to analysis on manifolds*. Number 31. Cambridge University Press, 1997.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet–a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24):241722, 2018.

Yangmei Shen, Wenrui Dai, Chenglin Li, Junni Zou, and Hongkai Xiong. Multi-scale graph convolutional network with spectral graph wavelet frame. *IEEE Transactions on Signal and Information Processing over Networks*, 7:595–610, 2021.

Jonathan M Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M Donghia, Craig R MacNair, Shawn French, Lindsey A Carfrae, Zohar Bloom-Ackermann, et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702, 2020.

Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28, pp. 1383–1392. Wiley Online Library, 2009.

Freyr Sverrisson, Jean Feydy, Joshua Southern, Michael M Bronstein, and Bruno Correia. Physics-informed deep neural network for rigid-body protein docking. In *ICLR2022 Machine Learning for Drug Discovery*, 2022.

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

Raphael J. L. Townshend, Rishi Bedi, Patricia A. Suriana, and Ron O. Dror. End-to-End Learning on 3D Protein Structure for Interface Prediction. Technical report, arXiv, December 2019. arXiv:1807.01297 [cs, q-bio, stat] type: article.

Raphael JL Townshend, Stephan Eismann, Andrew M Watkins, Ramya Rangan, Maria Karelina, Rhiju Das, and Ron O Dror. Geometric deep learning of rna structure. *Science*, 373(6558): 1047–1051, 2021.

Oliver Unke, Mihail Bogojeski, Michael Gastegger, Mario Geiger, Tess Smidt, and Klaus-Robert Müller. SE(3)-equivariant prediction of molecular wavefunctions and electronic densities. *Advances in Neural Information Processing Systems*, 34:14434–14447, 2021.

Saining Xie, Jiatao Gu, Demi Guo, Charles R Qi, Leonidas Guibas, and Or Litany. Pointcontrast: Unsupervised pre-training for 3d point cloud understanding. In *European conference on computer vision*, pp. 574–591. Springer, 2020.

Yumeng Yan, Huanyu Tao, Jiahua He, and Sheng-You Huang. The HDOCK server for integrated protein–protein docking. *Nature Protocols*, 15(5):1829–1852, May 2020. doi: 10.1038/s41596-020-0312-x. Number: 5 Publisher: Nature Publishing Group.

Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N Zare, and Patrick Riley. Optimization of molecules via deep reinforcement learning. *Scientific reports*, 9(1):1–10, 2019.

Appendix

A WEYL'S ASYMPTOTIC LAW

Fig. A.1 shows the first 300 eigenvalues of a pair of ligand and receptor proteins (PDB ID: 3V6Z). The protein with smaller surface area exhibits a larger slope, which is described by Weyl's asymptotic law. Since different eigenfunctions have their unique spatial resolutions, in our protein-related experiments, we keep all eigenfunctions with eigenvalues smaller than 0.4 (determined empirically) to guarantee that proteins of different sizes have the same highest spatial resolution in their eigenfunctions.

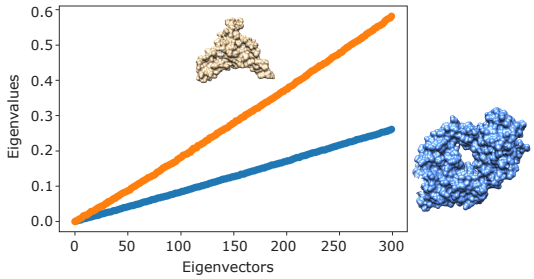

Figure A.1: Weyl's asymptotic law, where the protein with larger surface area exhibits a slower eigenvalue growth rate. Note that eigenvectors with similar eigenvalues exhibit similar spatial frequencies, since they have similar smoothness in the sense of Dirichlet energy.

B DETAILS OF THE BINDING SITE PREDICTION MODULE

First, we define the binding site as the protein surface region which is within 3 Å to its counterpart surface, and obtain the set of all corresponding surface points P as the nearest neighbor vertices between the ground truth protein interfaces.

The binding site prediction module stacks N feature propagation blocks, which consists of M HMR layers (introduced in Sec. 4.1) and a cross attention layer. Given the propagated receptor features F and ligand features G, the cross attention layer enables communication between proteins:

$$F = \text{softmax}\left(\frac{(W_Q G)(W_K F)^\top}{\sqrt{d_k}}\right)(W_V F), G = \text{softmax}\left(\frac{(W_Q F)(W_K G)^\top}{\sqrt{d_k}}\right)(W_V G)$$

where $F \in \mathbb{R}^{N_r \times d_k}$, $\mathsf{G} \in \mathbb{R}^{N_l \times d_k}$ denotes the propagated features on the receptor/ligand protein surface, d_k denotes the dimension of features, and W_Q,W_K and W_V are the parameter matrices for the key, query, and value in attention computation, respectively.

The loss function consists of two components. The first is a binary cross entropy loss, which encourages the model to correctly predict the binding site:

$$\mathcal{L}_{bce} = -[y_{\mathbf{f}_i} \log \mathbf{f}_i + (1 - y_{\mathbf{f}_i}) \log(1 - \mathbf{f}_i) + y_{\mathbf{g}_i} \log \mathbf{g}_i + (1 - y_{\mathbf{g}_i}) \log(1 - \mathbf{g}_i)]$$

where $y_{\mathbf{f}_i}$ and $y_{\mathbf{g}_i}$ are the indicate labels of whether a point i belongs to the binding site.

The second term is a PointInfoNCE loss (Xie et al., 2020), a contrastive matching loss that minimize the distance between the features of corresponding surface points and maximize the distance between unmatched points features:

$$\mathcal{L}_{match} = -\sum_{(i,j)\in P} \log \frac{\exp(\mathbf{f}_i \cdot \mathbf{g}_j / \tau)}{\sum_{(\cdot,k)\in P} \exp(\mathbf{f}_i \cdot \mathbf{g}_k / \tau)}$$

where P is the set of all corresponding surface points and τ is the temperature.

The total loss are the weighted average of two loss with weight λ

$$\mathcal{L} = \mathcal{L}_{bce} + \lambda \mathcal{L}_{match}$$

The hyper-parameters of Binding Site Prediction Module are listed in the Table I.5:

C DETAILS ON FUNCTIONAL MAPS

For completeness, we present in this appendix the details of functional maps used in Sec. 4.2.

Let us be given two manifolds $\mathcal{M}$ and $\mathcal{N}$. The aim of functional maps is to find a bijective mapping $T : \mathcal{M} \to \mathcal{N}$ to align these two manifolds. Unlike traditional methods that try to recover the point-to-point correspondence directly, functional maps lift the mapping T to a correspondence between the functional spaces on the two manifolds. Formally, let $L^2(\cdot)$ be the functional space of square integrable functions on a manifold, we infer the functional correspondence $T_F : L^2(\mathcal{M}) \to L^2(\mathcal{N})$ induced by the mapping T, and is defined by $T_F f = f \circ T^{-1}$ for any $f \in L^2(\mathcal{M})$.

To compute the functional correspondence, we need to utilize the Fourier bases on the manifolds. Actually, such functional correspondence has a concise expression in the spectral domain: given the respective truncated LB eigenfunctions $\{\phi_j^{\mathcal{M}}\}_{j\geq 0}^{k_2}$ on $\mathcal{M}$ and $\{\phi_i^{\mathcal{N}}\}_{i\geq 0}^{k_1}$ on $\mathcal{N}$, the functional correspondence T_F can be (approximately) represented as a change of basis matrix:

$$\mathsf{C} = (c_{ij})_{k_2 \times k_1} = (\langle T_F \phi_j^{\mathcal{M}}, \phi_i^{\mathcal{N}} \rangle_{\mathcal{N}})_{k_2 \times k_1}.$$

Now given a set of q corresponding functions $\{f_1, \ldots, f_q\} \subset L^2(\mathcal{M})$ and $\{g_1, \ldots, g_q\} \subset L^2(\mathcal{N})$, we denote their Fourier expansions by coefficients $\mathsf{A} = (a_{ij})_{k_1 \times q}$, where $a_{ij} = \langle f_j, \phi_i \rangle_{\mathcal{M}}$, and $\mathsf{B} = (b_{ij})_{k_2 \times q}$, where $b_{ij} = \langle g_j, \psi_i \rangle_{\mathcal{N}}$. The matrix C can be obtained by solving the following quadratic minimization problem:

$$\mathsf{C} = \operatorname{argmin}_{\mathsf{C} \in \mathbb{R}^{k_2 \times k_1}} \|\mathsf{C} \cdot \mathsf{A} - \mathsf{B}\|_{\mathrm{F}}^2 + \alpha \|\mathsf{C} \cdot \delta_{\mathcal{M}} - \delta_{\mathcal{N}} \cdot \mathsf{C}\|_{\mathrm{F}}^2 + \beta \sum_{i=1}^{q} \|\mathsf{C} \cdot \Lambda_{\mathcal{M}}^i - \Lambda_{\mathcal{N}}^i \cdot \mathsf{C}\|_{\mathrm{F}}^2, \quad (5)$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm, and the first term on the right hand side is the change of basis constraint. Two more regularization terms are introduced into the formula with tunable weights $\alpha, \beta > 0$. $\|\mathsf{C} \cdot \delta_{\mathcal{M}} - \delta_{\mathcal{N}} \cdot \mathsf{C}\|_{\mathrm{F}}$ enforces the isometry of the two manifolds where the matrix $\delta_{\mathcal{M}}$ and $\delta_{\mathcal{N}}$ are the spectral representation of the LB operators. The matrix $\Lambda_{\mathcal{M}}^i$ and $\Lambda_{\mathcal{N}}^i$ are called the orientation operator Ren et al. (2018) and are defined by the functions f_i and g_i, respectively. The commutator with the orientation operator incorporates extrinsic properties into the formulation and enforces the orientation of the functional maps.

Once the functional mapping (i.e., C) is recovered, point-to-point correspondence T between the manifold $\mathcal{M}$ and $\mathcal{N}$ can be obtained by mapping indicator functions of vertices on $\mathcal{M}$ because $T_F \delta_m = \delta_{T(m)}$ for any vertex $m \in \mathcal{M}$.

D GRADIENT ORIENTATION BLOCK

Considering that many molecules display certain symmetry or even chirality, we propose an addition gradient orientation block to differentiate symmetric points on a molecular surface. Intuitively, consider a pair of symmetric points, q and q, on the surface of a water molecule (Fig. 3a) with surface normal vectors $\vec{n}_p$ and $\vec{n}_q$. Let $\nabla h(p)$ represent the spatial gradient of function h at surface point p, then for two surface functions f and g, we have:

$$\vec{n}_p \cdot (\nabla f(p) \times \nabla g(p)) = -\vec{n}_q \cdot (\nabla f(q) \times \nabla g(q)).$$

Therefore, symmetric surface points would show opposite signs in the above dot product. We compute such dot product terms between the surface normal vector and the cross product between two learned functions for each surface vertex. The output of the gradient orientation block has the same size as the message passing block, where each channel encodes the orientation of the spatial gradient for each learned function. This block has been used in our rigid protein docking pipeline, but not in the QM9 molecular property prediction task.

E DATASET PREPARATION

The original DIPS dataset (Townshend et al., 2019) contains interacting protein chains (one vs one) extracted from experimentally determined complex structures in RCSB PDB database (https://www.rcsb.org/). While DIPS dataset is designed to represent the interactions in protein

complexes including homomeric proteins (i.e., protein complexes composed of identical proteins), we are particularly interested in interactions between different proteins in biological processes (e.g., antibody-antigen, enzyme-inhibitor, enzyme-substrates, etc.). To better resemble such interactions, we apply a modified pipeline based on DIPS, referred as DIPS-Het.

Specifically, we include the PDB entries in DIPS, as well as newly deposited entries from RCSB PDB database that 1) are determined using diffraction-based methods or electronic microscopy, 2) have $<$ 3.5 Å resolution, 3) are not hybrid protein complexes (e.g., protein-DNA complexes are excluded), and 4) have $<$ 5 protein chains. Distinguished from the original DIPS pipeline, we 1) further remove the PDB entries with only homomeric interfaces (i.e., interactions between identical proteins), and 2) adopt a one-vs-rest strategy that uses each one of the protein chains in the complex as the "ligand protein" and the rest protein chains as the "receptor protein" to form ligand-receptor pairs.

Similar to Ganea et al. (2021), we remove the cases with any protein chain sharing the same protein sequence clusters (at $>$ 30% similarity) with DB5.5. Finally, we cluster proteins based on protein chain sequence similarity (at 30%) and separate these clusters into the training and validation set. This sequence-based split encourages to select a model with better generalizability. After removing proteins failed in surface mesh generation, DIPS-Het contains 11,373 training cases and 508 validation cases.

F DOCKING EVALUATION METRICS

We evaluate the rigid protein docking results using four metrics: Complex RMSD, Interface RMSD, DockQ, and Success rate. Complex RMSD and Interface RMSD are used in Ganea et al. (2021): let $\mathbf{Z}^* \in \mathbb{R}^{3\times(n+m)}$ and $\mathbf{Z} \in \mathbb{R}^{3\times(n+m)}$ be the α-carbon coordinates of the ground truth and predicted protein complexes, respectively, after superimposing the complex using Kabsch algorithm, Complex RMSD is calculated as $\sqrt{\frac{1}{n+m}||\mathbf{Z}^* - \mathbf{Z}||_F^2}$. Similarly, Interface RMSD is calculated use the same procedure but with the α-carbon coordinates of interface residues ($<$ 8 Å to the other protein's residues). Smaller RMSD value means the predicted structure is closer the ground truth structure.

In addition, we evaluate the overall quality of docking using DockQ (Basu & Wallner, 2016) and the success rate of achieving "Acceptable" or higher according to Méndez et al. (2003; 2005). Both metrics are based on three standardized criteria used by Critical Assessment of PRedicted Interactions (CAPRI): L_{rms} is ligand (the smaller protein) RMSD calculated based on backbone atoms, after superimposing receptor's backbone atoms; I_{rms} is the backbone RMSD of interface residues, after superimposing the interface residues (residues with any atom is $<$ 10 Å to atoms in the other protein); and f_{nat} is the recall in recovering residue-residue contacts between the proteins, where two residues are "in contact" if any pair of atoms from two residues has distance $<$ 5 Å. DockQ is a continuous score between 0 and 1 (higher the better), derived from L_{rms}, I_{rms}, and f_{nat}. A docking result is considered as a "success" if it is ranked "Acceptable" or higher according to CAPRI's criteria, that is

$$f_{\text{nat}} \geq 0.1 \wedge (L_{\text{rms}} \leq 10.0 \vee I_{\text{rms}} \leq 4.0)$$
$$\text{OR}$$
$$f_{\text{nat}} \geq 0.3$$

Both DockQ score and CAPRI ranking is calculated using DockQ package (https://github.com/bjornwallner/DockQ/).

G DISTRIBUTIONS OF LEARNED PROPAGATION PARAMETER

We visualize the distribution of propagation parameters: propagation time t, frequency filter mean μ and variance σ, learned from data. The histogram of selected layers are shown in Fig. G.1.

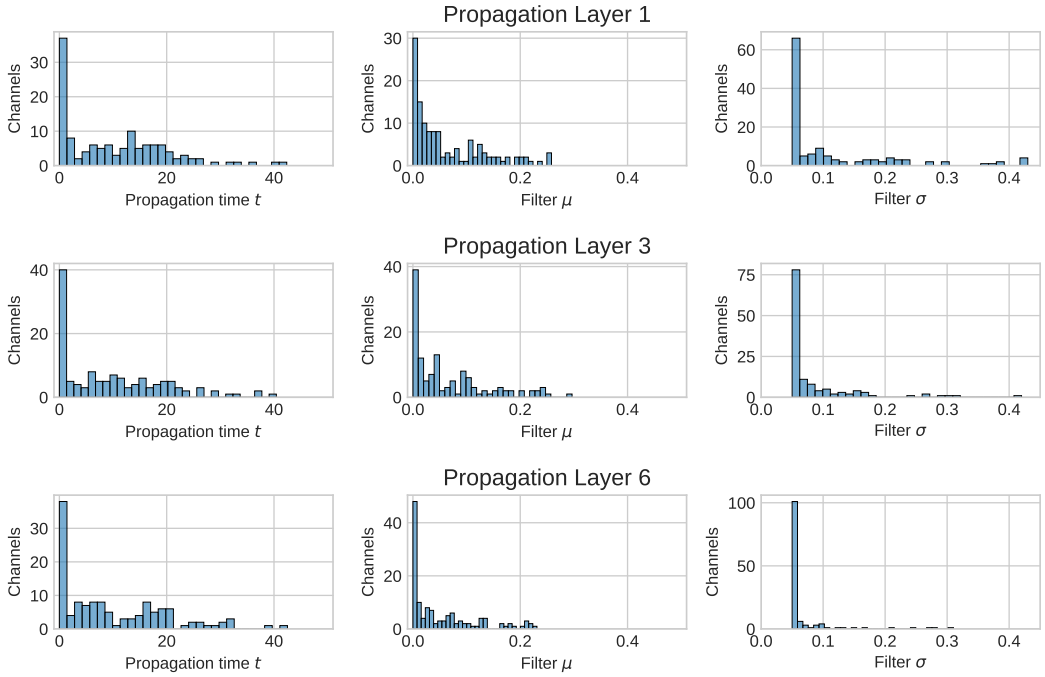

Figure G.1: Distribution of learned propagation time t and frequency filter (μ and σ) at selected propagation layers

H Additional Experiments on Rigid Protein Docking

H.1 Analysis of Different Message Passing Mechanisms

We study the effectiveness of proposed Harmonic Message Passing mechanism by comparing it with a 'Graph Message Passing' model, where Harmonic Massage Passing is replaced by a Graph Neural Network (implemented with Graph Attention Network), and a 'No Message Passing' model, where we only use the Multi-Layer Perceptron (MLP) to encode features per vertex. Experimental results show that Harmonic Message Passing model outperforms Graph Message Passing and No Message Passing models, suggesting its higher efficiency in propagating information on protein surface meshes.

Table H.1: Rigid Protein Docking Performance with Different Message Passing models

	Complex RMSD		Interface RMSD		DockQ	
Model	Median	Mean	Median	Mean	Median	Mean
Harmonic Message Passing	**12.01**	**12.09**	**10.90**	**11.43**	**0.06**	**0.22**
Graph Message Passing	15.41	15.67	13.38	14.37	0.03	0.08
No Message Passing	16.44	17.23	14.11	16.11	0.02	0.04

H.2 ABLATION STUDIES ON INPUT FEATURES

To assess the importance of chemical or geometric features for rigid protein docking, we compare the performance of HMR (Full model) to models trained with only chemical or geometric features. 'Chem only' model only includes chemically-related features (atom types, amino acid residue types, atom charge, atom radius, residue's hydrophobicity, and whether the atom is $alpha$-carbon); 'Geom only' model only contain features related to local manifold geometry, including Gaussian curvature, mean curvature, and Heat Kernel Signatures. Table H.2 shows that there are performance drop by removing either chemical or geometric features. And the chemical features seem to be critical for HMR in the rigid protein docking task.

Table H.2: Ablation studies on input features

	Complex RMSD		Interface RMSD		DockQ	
Model	Median	Mean	Median	Mean	Median	Mean
Full model	**12.01**	**12.09**	**10.90**	**11.43**	**0.06**	**0.22**
Chem only	13.96	13.95	12.73	13.33	0.04	0.17
Geom only	17.87	19.15	16.58	18.46	0.01	0.03

H.3 CROSS EXPERIMENTS WITH EQUIDOCK

Compare to EQUIDOCK, HMR differs both in the model and the dataset used for training. To understand which aspects contribute most to the better performance of HMR, we conduct cross experiments to train HMR on the original DIPS (used in EQUIDOCK) and EQUIDOCK trained on DIPS-Het (proposed by us, used in HMR). Same training and validation sets are used for the same dataset, and all models are evaluated on DB5.5. As summarized in Table H.3, model structure and dataset are both important: HMR achieved higher performance than EQUIDOCK when both models are trained on DIPS dataset; the performance of HMR further improves when it is trained on DIPS-Het dataset.

Table H.3: Compare EQUIDOCK and HMR trained on DIPS or DIPS-Het datasets

		Complex RMSD		Interface RMSD		DockQ	
Model	Training data	Median	Mean	Median	Mean	Median	Mean
HMR	DIPS	15.68	16.04	14.81	15.38	0.02	0.07
	DIPS-Het	12.01	12.09	10.90	11.43	0.06	0.22
EQUIDOCK	DIPS	17.27	18.09	15.04	16.11	0.02	0.03
	DIPS-Het	17.22	17.65	14.27	14.76	0.02	0.04

H.4 DB5.5 TEST CASES FROM EQUIDOCK

We compare HMR with EQUIDOCK on the 25 cases in DB5.5 selected as a test set in Ganea et al. (2021). For this comparison, we use the EQUIDOCK model fine-tuned on DB5.5. As shown in Table H.4 HMR still outperforms EQUIDOCK, despite not being fine-tuned on DB5.5.

Table H.4: Rigid docking performance on DB5.5 test set

	Complex RMSD		Interface RMSD		DockQ		Success rate
Model	Median	Mean	Median	Mean	Median	Mean	($\geq$ Acceptable)
HMR (Top 1)	13.10	12.42	11.05	11.68	0.07	0.15	0.20
HMR (Top 3)	11.35	11.18	10.02	10.03	0.08	0.16	0.20
EQUIDOCK	14.14	14.73	11.97	13.23	0.04	0.05	0.00

I TRAINING DETAILS

Table I.5: Hyperparameter choices of HMR and training phase

Hyperparameters	Values
QM9 Property Regression	
Number of Feature Propagation Block N	1
Number of HMR Layers M	6
Number of Attention Layers M	0
Dimension of Propagated Features d_k	256
Dropout Rate P_{drop}	0
Binding Site Prediction Module	
Number of Feature Propagation Block N	2
Number of HMR Layers M	3
Dimension of Propagated Features d_k	128
Dimension of Attention Hidden	128
Number of Attention Heads	16
Dropout Rate P_{drop}	0.1
Number of Negative Point Sampling	50
Loss Weight λ	0.1
Training	
Batch Size	4
Epoch	40
Learning Rate	5×10^{-4}
Optimizer	Adam
Learning Rate Scheduler	Cosine Annealing