Multiresolution graph transformers and wavelet positional encoding for learning long-range and hierarchical structures

Nhat Khang Ngo * [1] **Truong Son Hy** * c [2] **Risi Kondor** [3]

[1]FPT Software AI Center, Vietnam

[2]University of California San Diego, USA

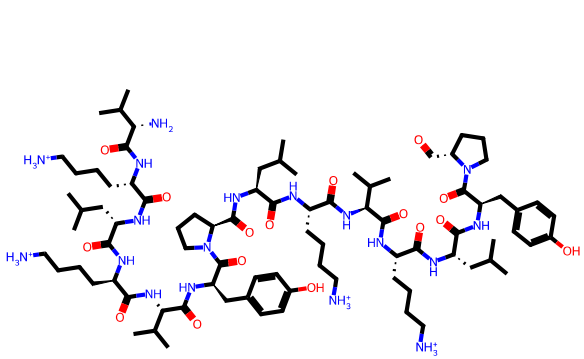
[3]University of Chicago, USA

* Equal Contribution c: Correspondent author, **tshy@ucsd.edu**

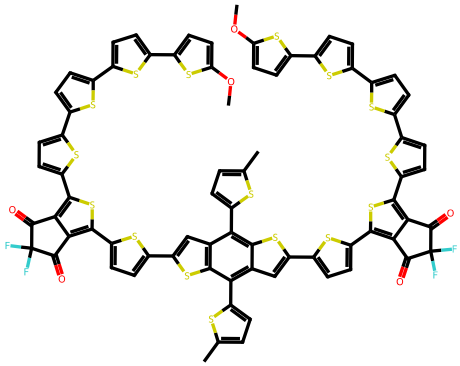
July 2023

Introduction

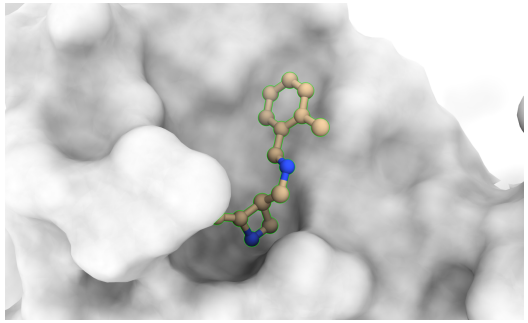
Accepted to **Journal of Chemical Physics**, Volume 159, Issue 3.



(a) Peptide



(b) Polymer



(c) Protein-Ligand

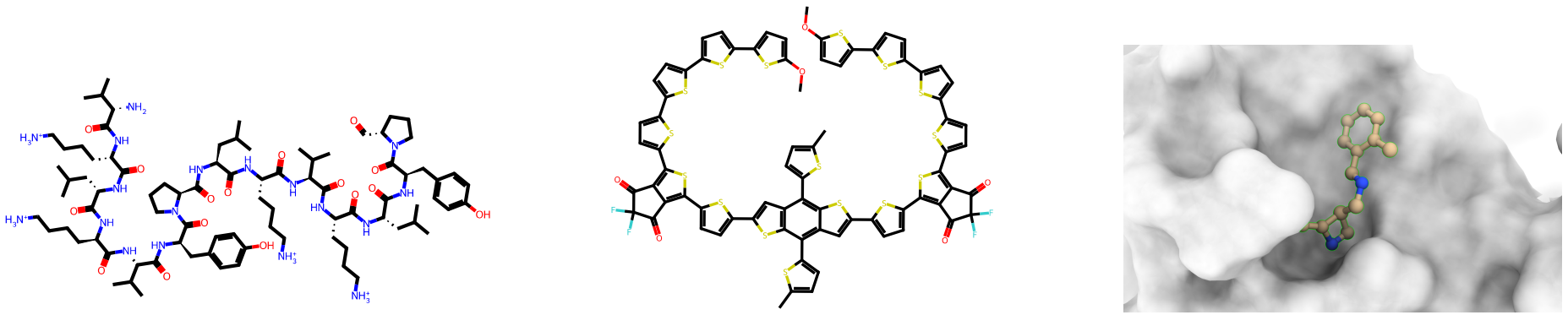
Figure: Examples of Macromolecules

- Macromolecules have **hierarchical structures** and comprise multiple **long-range** dependencies among distant atoms.

FPT Software AICenter

Introduction

Accepted to **Journal of Chemical Physics**, Volume 159, Issue 3.



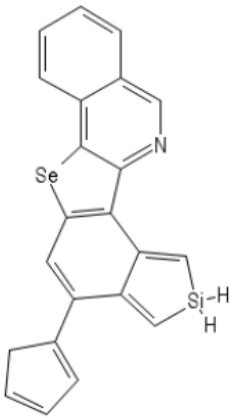
(a) Peptide (b) Polymer (c) Protein-Ligand

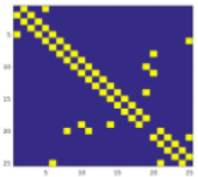
Figure: Examples of Macromolecules

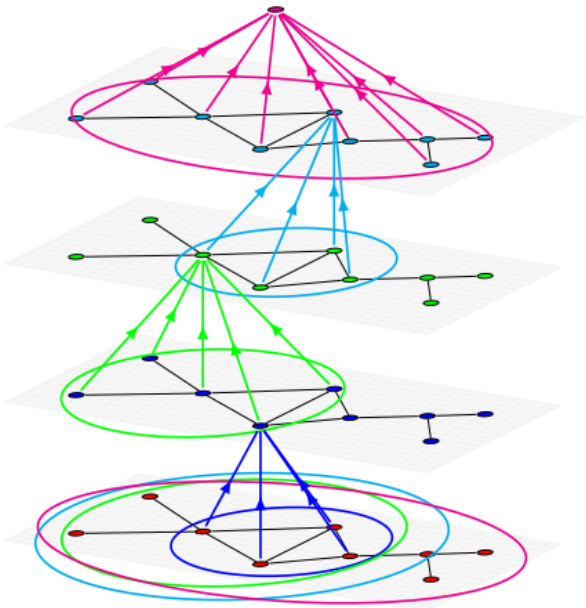
- Macromolecules have **hierarchical structures** and comprise multiple **long-range** dependencies among distant atoms.
- We want to predict functions of peptides, properties of polymers calculated from Density Functional Theory, and protein-ligand binding affinity.

FPT Software AICenter

Message passing and its limitation

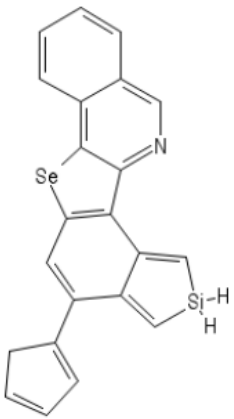


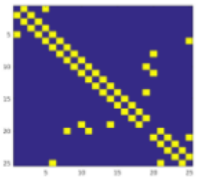


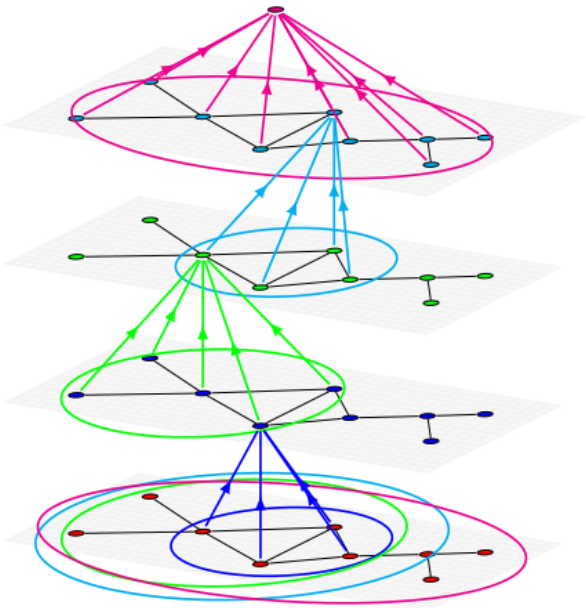


Each molecule is represented as a graph. Each node is an atom and each edge represents a bond.

Message passing and its limitation





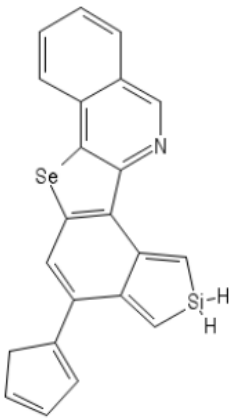


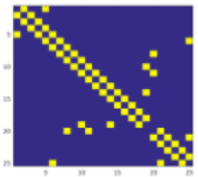
Each molecule is represented as a graph. Each node is an atom and each edge represents a bond.

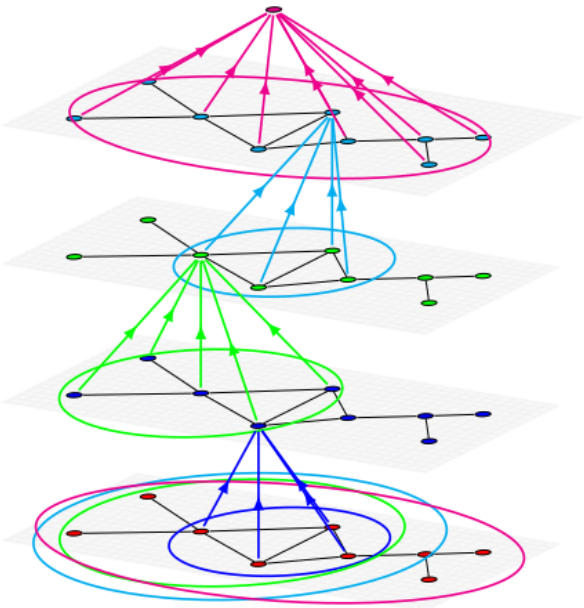
- The backbone of Graph Neural Networks is the **message passing** scheme which exchanges vectorized information among neighboring nodes.

FPT Software AICenter

Message passing and its limitation







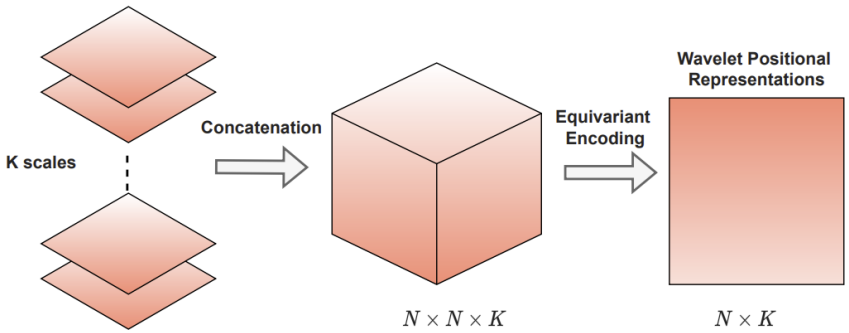
Each molecule is represented as a graph. Each node is an atom and each edge represents a bond.

- The backbone of Graph Neural Networks is the **message passing** scheme which exchanges vectorized information among neighboring nodes.
- However, message passing suffers from **over-smoothing** and **over-squashing** problems and is unable to capture the **long-range** interactions.

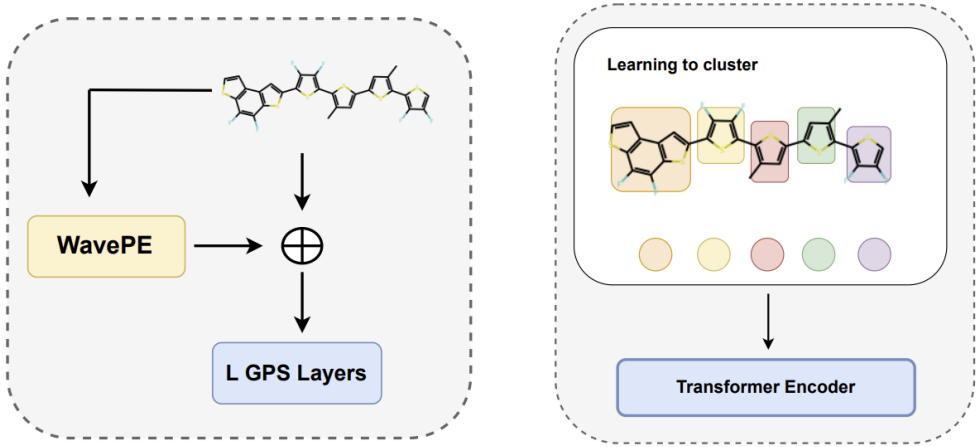
FPT Software AICenter

Our contributions

Wavelet positional encoding



Multiresolution Graph Transformer

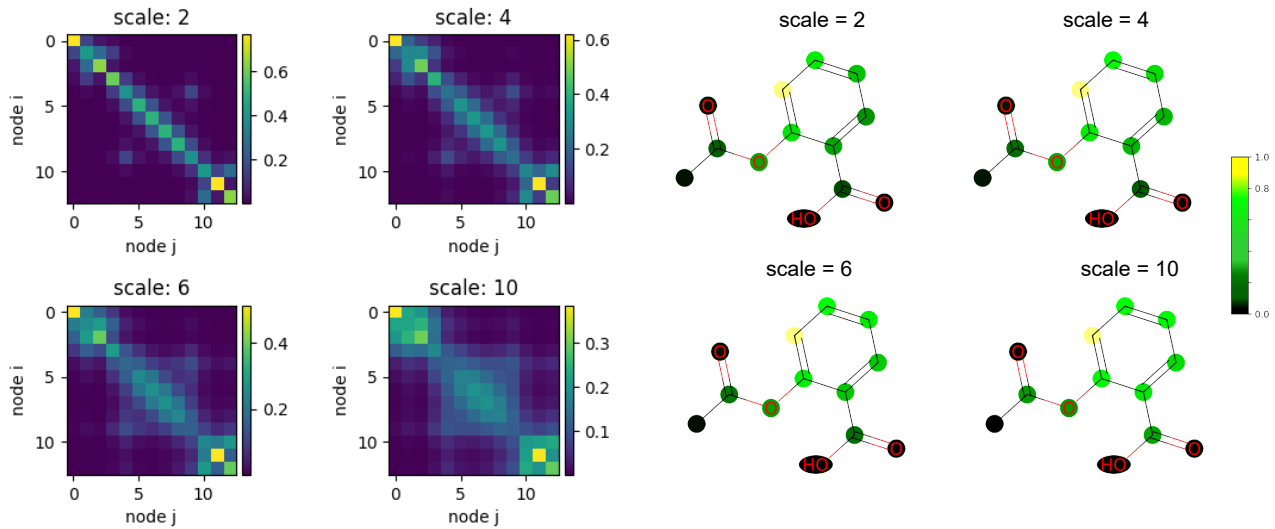


- Transformers are effective for computing the interactions between distant atoms via **self-attention** mechanisms.
- To adapt Transformer-like architectures to graphs, we need **positional encoding** schemes that embody the local structures.

FPT Software AICenter

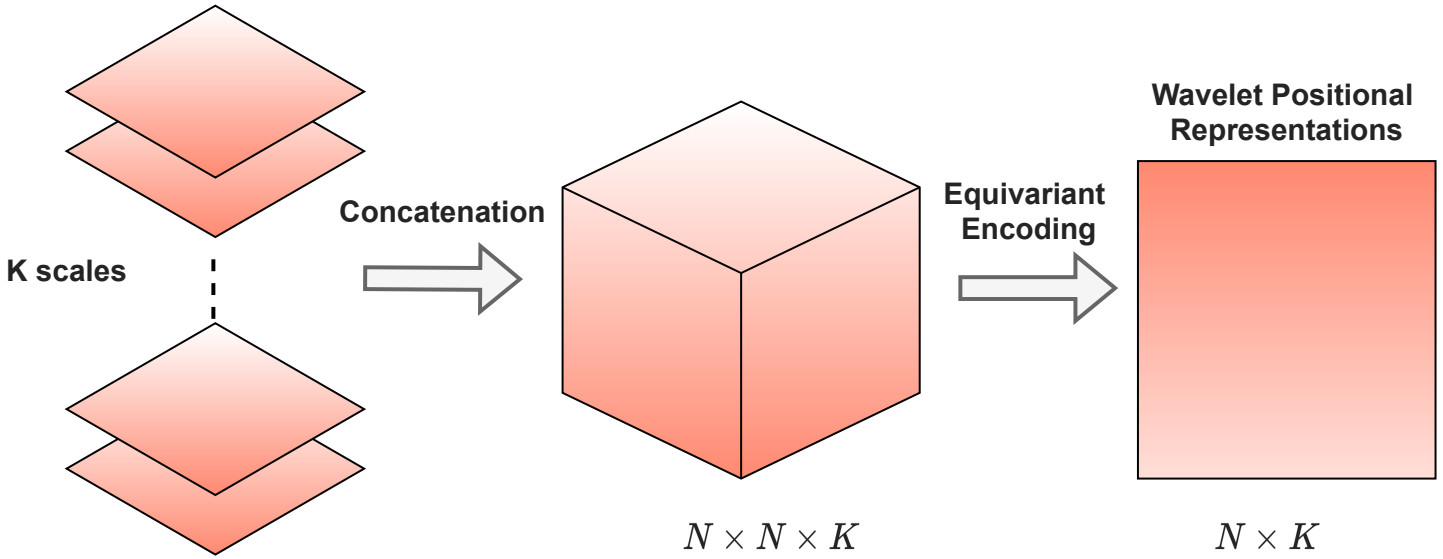
Wavelet Positional Encoding

- We compute the spectral wavelets at multiple scales.
- Low-scale wavelets are highly localized, whereas the high-scale wavelets can spread out more nodes on the molecular graphs.



Visualization of some of the wavelets with scaling parameters on the Aspirin $C_9H_8O_4$ molecular graph with 13 nodes (i.e. heavy atoms).

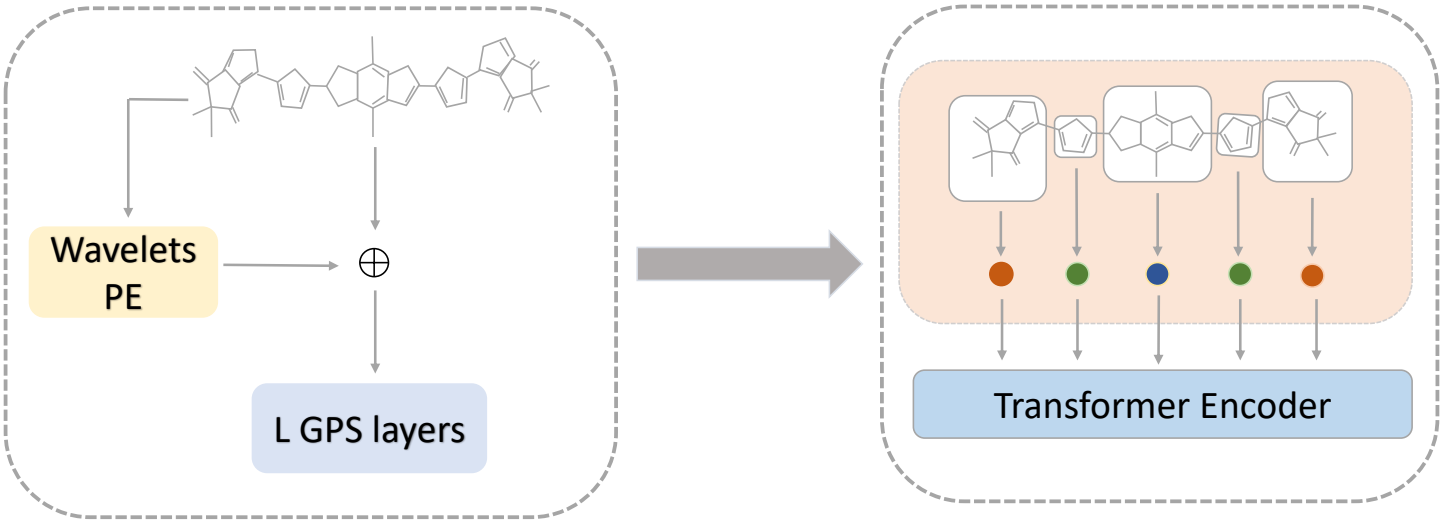
Wavelet Positional Encoding (continued)



- We concatenate the diffusion matrices at different scales together to form a tensor.
- Then, we encode this tensor to node-level positional embeddings via permutation-equivariant neural networks.

FPT Software AICenter

Multiresolution Graph Transformer (MGT)



- We propose a **learning-to-cluster** algorithm that coarsens graphs iteratively to build a multiresolution (i.e. multiple of resolutions) representation of the input graph.
- We employ the **graph transformer** model learning on each resolution and we integrate our **wavelet positional encoding**.

Peptides Property Prediction

We outperform other state-of-the-art methods in predicting peptides' functionality.

Model	Params	Peptides-struct	Peptides-func
		MAE ↓	AP ↑
GCN	508k	0.3496 ± 0.0013	0.5930 ± 0.0023
GINE	476k	0.3547 ± 0.0045	0.5498 ± 0.0079
GatedGCN	509k	0.3420 ± 0.0013	0.5864 ± 0.0077
GatedGCN + RWPE	506k	0.3357 ± 0.0006	0.6069 ± 0.0035
Transformer + LapPE	488k	0.2529 ± 0.0016	0.6326 ± 0.0126
GPS	—	0.2500 ± 0.0005	0.6535 ± 0.0041
SAN + LapPE	493k	0.2683 ± 0.0043	0.6384 ± 0.0121
SAN + RWPE	500k	0.2545 ± 0.0012	0.6562 ± 0.0075
MGT + WavePE (ours)	499k	**0.2453 ± 0.0025**	**0.6817 ± 0.0064**

Polymer Property Prediction

We **achieve the chemical accuracy** in estimating the molecular properties of polymers that are calculated from Density Functional Theory. We also outperform all other competitive baselines.

Model	Params	Property		
		GAP	HOMO	LUMO
DFT error		1.2	2.0	2.6
Chemical accuracy		0.043	0.043	0.043
GCN	527k	0.1094 ± 0.0020	0.0648 ± 0.0005	0.0864 ± 0.0014
GCN + Virtual Node	557k	0.0589 ± 0.0004	0.0458 ± 0.0007	0.0482 ± 0.0010
GINE	527k	0.1018 ± 0.0026	0.0749 ± 0.0042	0.0764 ± 0.0028
GINE + Virtual Node	557k	0.0870 ± 0.0040	0.0565 ± 0.0050	0.0524 ± 0.0010
GPS	600k	0.0467 ± 0.0010	0.0322 ± 0.0020	0.0385 ± 0.0006
Transformer + LapPE	700k	0.2949 ± 0.0481	0.1200 ± 0.0206	0.1547 ± 0.0127
MGT + LapPE (ours)	499k	**0.0378 ± 0.0004**	**0.0270 ± 0.0010**	0.0300 ± 0.0006
MGT + RWPE (ours)	499k	0.0384 ± 0.0015	0.0274 ± 0.0005	**0.0290 ± 0.0007**
MGT + WavePE (ours)	499k	0.0387 ± 0.0011	0.0283 ± 0.0004	0.0290 ± 0.0010

FPT Software AICenter

Protein-Ligand Binding Affinity Prediction

Understanding the multiscale structure of proteins is important in estimating their fitness and functionality. In this experiment, we show the effectiveness of our model in capturing the long-range and hierarchical structures of proteins. We show our competitive experimental result on ATOM3D benchmark.

Method	3D-CNN	GNN	ENN	GVP-GNN	MGT + WavePE
RMSE ↓	**1.416 ± 0.021**	1.570 ± 0.025	1.568 ± 0.012	1.594 ± 0.073	1.436 ± 0.066

Conclusion

Summary:

1. Motivated from multiresolution analysis and wavelet theory.
2. We propose novel **wavelet position encoding** and **multiresolution graph transformer** that can efficiently capture the long-range and hierarchical structures.
3. We achieve promising results for peptides, polymers and protein-ligand, with potential applications for drug discovery and material science.
4. Our software and data is available at:

   https://github.com/Fsoft-AIC/Multires-Graph-Transformer

Thank you very much for your attention!

FPT Software AICenter