Model	BZR	COX2	MUTAG	PROTEINS	PTC_MR	PTC_MM	PTC_FM	PTC_FR
CSM (?)	84.54±0.65	79.78±1.04	87.29±1.25	OOT	58.24±2.44	63.30±1.70	63.80±1.00	65.51±9.82
HGK-SP (?)	81.99±0.30	78.16±0.00	80.90±0.48	74.53±0.35	57.26±1.41	57.52±9.98	52.41±1.79	66.91±1.46
HGK-WL (?)	81.42±0.60	78.16±0.00	75.51±1.34	74.53±0.35	59.90±4.30	67.22±5.98	64.72±1.66	67.90±1.81
WL (?)	86.16±0.97	79.67±1.32	85.75±1.96	73.06±0.47	57.97±0.49	67.28±0.97	64.80±0.85	67.64±0.74
WL-OA (?)	87.43±0.81	81.08±0.89	86.10±1.95	73.50±0.87	62.70±1.40	66.60±1.16	66.28±1.83	67.82±5.03
DGCNN (?)	79.40±1.71	79.85±2.64	85.83±1.66	75.54±0.94	58.59±2.47	62.10±14.09	60.28±6.67	65.43±11.30
GCN (?)	79.34±2.43	76.53±1.82	80.42±2.07	70.31±1.93	62.26±4.80	67.80±4.00	62.39±0.85	69.80±4.40
GIN (?)	85.60±2.00	80.30±5.17	89.39±5.60	76.16±2.76	64.60±7.00	67.18±7.35	64.19±2.43	66.97±6.17
Top-K (?)	79.40±1.20	80.30±4.21	67.61±3.36	69.60±3.50	64.70±6.80	67.51±5.96	65.88±4.26	66.28±3.71
MinCutPool (?)	82.64±5.05	80.07±3.85	79.17±1.64	76.52±2.58	64.16±3.47	N/A	N/A	N/A
DiffPool (?)	83.93±4.41	79.66±2.64	79.22±1.02	73.63±3.60	64.85±4.30	66.00±5.36	63.00±3.40	69.80±4.40
EigenGCN (?)	83.05±6.00	80.16±5.80	79.50±0.66	74.10±3.10	N/A	N/A	N/A	N/A
SAGPool (?)	82.95±4.91	79.45±2.98	76.78±2.12	71.86±0.97	69.41±4.40	66.67±8.57	67.65±3.72	65.71±10.69
HaarPool (?)	83.95±5.68	82.61±2.69	90.00±3.60	73.23±2.51	66.68±3.22	69.69±5.10	65.59±5.00	69.40±5.21
PersLay (?)	82.16±3.18	80.90±1.00	89.80±0.90	74.80±0.30	N/A	N/A	N/A	N/A
FC-V (?)	85.61±0.59	81.01±0.88	87.31±0.66	74.54±0.48	N/A	N/A	N/A	N/A
MPR (?)	N/A	N/A	84.00±8.60	75.20±2.20	66.36±6.55	68.60±6.30	63.94±5.19	64.27±3.78
SIN (?)	N/A	N/A	N/A	76.50±3.40	66.80±4.56	70.55±4.79	68.68±6.80	69.80±4.36
Wit-TopoPool (ours)	**87.80±2.44**	******87.24±3.15**	**93.16±4.11**	****80.00±3.22**	***70.57±4.43**	*****79.12±4.45**	**71.71±4.86**	*****75.00±3.51**

Table 1: Performance on molecular and chemical graphs. The best results are given in **bold** while the best performances achieved by the runner-ups are underlined.

hood for each node, hence, resulting in improvement over graph kernels. Comparing with GNN-based models, graph pooling methods such as SAGPool and HaarPool utilize the hierarchical structure of the graph and extract important geometric information on the observed graphs. Finally, PersLay, FC-V, MPR, and SIN are the state-of-the-art topological and simplicial complex-based models, specialized on extracting topological information and higher-order structures from the observed graphs. A common limitation of these approaches is that they do not simultaneously capture both local and global topological properties of the graph. Hence, it is not surprising that performance of Wit-TopoPool which systematically integrates all types of the above information on the observed graphs is substantially higher than that of the benchmark models.

Social Graphs Table 2 shows the performance comparison on 3 social graph datasets. Similarly, Table 2 indicates that our Wit-TopoPool model is always better than baselines for all social graph datasets. We find that, even compared to the baselines (which feeds neural networks with topological summaries (i.e., PersLay) or integrates higher-order structures into GNNs (i.e., SIN)), Wit-TopoPool is highly competitive, revealing that global and local topological representation learning modules can enhance the model expressiveness.

Model	IMDB-B	IMDB-M	REDDIT-B
CSM (?)	OOT	OOT	OOT
HGK-SP (?)	73.34±0.47	51.58±0.42	OOM
HGK-WL (?)	72.75±1.02	50.73±0.63	OOM
WL (?)	71.15±0.47	50.25±0.72	77.95±0.60
WL-OA (?)	74.01±0.66	49.95±0.46	87.60±0.33
DGCNN (?)	70.00±0.90	47.80±0.90	76.00±1.70
GCN (?)	66.53±2.33	48.93±0.88	89.90±1.90
GIN (?)	75.10±5.10	52.30±2.80	92.40±2.50
Top-K (?)	73.17±4.84	48.80±3.19	79.40±7.40
MinCutPool (?)	70.77±4.89	49.00±2.83	87.20±5.00
DiffPool (?)	68.60±3.10	45.70±3.40	79.00±1.10
EigenGCN (?)	70.40±3.30	47.20±3.00	N/A
SAGPool (?)	74.87±4.09	49.33±4.90	84.70±4.40
HaarPool (?)	73.29±3.40	49.98±5.70	N/A
PersLay (?)	71.20±0.70	48.80±0.60	N/A
FC-V (?)	73.84±0.36	46.80±0.37	89.41±0.24
MPR (?)	73.80±4.50	50.90±2.50	86.20±6.80
SIN (?)	75.60±3.20	52.50±3.00	92.20±1.00
Wit-TopoPool (ours)	*****78.40±1.50**	**53.33±2.47**	***92.82±1.10**

Table 2: Performance on social graphs. The best results are given in **bold** while the best performances achieved by the runner-ups are underlined.

Ablation Study

To evaluate the contributions of the different components in our Wit-TopoPool model, we perform exhaustive ablation studies on COX2, PTC_MM, and IMDB-B datasets. We use Wit-TopoPool as the baseline architecture and consider three ablated variants: (i) Wit-TopoPool without topological pooling enhanced graph convolutional layer (W/o TPGCL), (ii) Wit-TopoPool without witness complex-based topological layer (W/o Wit-TL), and (iii) Wit-TopoPool without attention mechanism (W/o Attention Mechanism). The experimental results are shown in Table 3 and we prove the validity of each component. As Table 3 suggest, we find that (i) ablating each of above component leads to the performance drops in comparison with the full Wit-TopoPool model, thereby, indicating that each of the designed components contributes to the success of Wit-TopoPool, (ii) on all three datasets, TPGCL module significantly improves the classification results, i.e., Wit-TopoPool outperforms Wit-TopoPool w/o TPGCL with an average relative gain 5.30% over three datasets – this phenomenon implies that learning both local topological information and node features are critical for successful graph learning, (iii) in comparison to Wit-TopoPool and Wit-TopoPool w/o Wit-TL, Wit-TopoPool always outperforms because the Wit-TL module enables the model to effectively incorporate more global topological information, demonstrating the significance of the proposed global topological representation learning module for graph classification, and (iv) Wit-TopoPool consistently outperforms Wit-TopoPool w/o Attention Mechanism on all 3 datasets, indicating the attention mechanism can successfully extract the most correlated information and, hence, improves the generalization of unseen graph structures. Moreover, we also compare Wit-TopoPool with VR-TopoPool (i.e., replacing witness complex in global information learning with Vietoris-Rips complex) (see Appendix B for a discussion).

	Architecture	Accuracy mean±std
COX2	**Wit-TopoPool**	***87.24±3.15**
	W/o TPGCL	82.67±3.26
	W/o Wit-TL	85.21±3.20
	W/o Attention mechanism	85.58±3.53
PTC_MM	**Wit-TopoPool**	**76.76±5.78**
	W/o TPGCL	67.38±5.33
	W/o Wit-TL	70.58±5.29
	W/o Attention mechanism	75.12±5.59
IMDB-B	**Wit-TopoPool**	****78.40±1.50**
	W/o TPGCL	73.93±1.83
	W/o Wit-TL	77.00±1.69
	W/o Attention mechanism	76.20±1.98

Table 3: Ablation study of the Wit-TopoPool architecture.

Sensitivity Analysis We perform sensitivity analysis of (i) landmark set selection and (ii) topological score function to explore the effect of above two components on our Wit-TopoPool performance. The optimal choice of landmark set selection and topological score function can be obtained via cross-validation. We first explore the effect of landmark set selection. We consider 3 types of landmark set selections, i.e., (i) randomly ($\mathfrak{L}_r$), (ii) node betweenness centrality ($\mathfrak{L}_b$), and (iii) node degree centrality ($\mathfrak{L}_d$), and report results on COX2 and PTC_MM datasets. As Table 4 shows, we observe that the landmark set selection based on either node betweenness or degree centrality helps to improve the graph classification performance, whereas the landmark set based on randomly results in the performance drop. We also explore the effect of topological score function for the importance measurement of the persistence diagram (see Eq. 2). As the results in Table 5 suggest, summing over lifespans of topological features (points) in persistence diagrams can significantly improve performance, but applying piecewise linear weighting function on topological features may result in deterioration of performance.

Computational Complexity Computational complexity of the standard persistent homology matrix reduction algorithm (**?**) (i.e., based on column operations over boundary matrix of the complex) runs in cubic time in the worst case, i.e., $\mathcal{O}(m^3)$, where m is the number of simplices in the filtration. For 0-dimensional PH, it can be computed efficiently using disjoint sets with complexity $\mathcal{O}(m\alpha^{-1}m)$, where $\alpha^{-1}(\cdot)$ is the inverse Ackermann function (**?**). Computational complexity of the witness complex construction is $\mathcal{O}(\mathfrak{L}\log{(n)})$ (where n is the number of data points and $\mathfrak{L}$ is the landmark set), involving calculating the distance between data points and landmark points.

Dataset	Landmark set	Accuracy mean±std
COX2	$\mathfrak{L}_r$	82.98±3.88
	$\mathfrak{L}_b$	85.10±2.52
	$\mathfrak{L}_d$	**87.24±3.15**
PTC_MM	$\mathfrak{L}_r$	71.53±6.17
	$\mathfrak{L}_b$	**79.12±4.45**
	$\mathfrak{L}_d$	76.76±5.78

Table 4: Sensitivity analysis with respect to the landmark set selection for Wit-TopoPool on COX2 and PTC_MM.

Dataset	Weighting function	Accuracy mean±std
COX2	$d_\rho - b_\rho$	*****87.24±3.15**
	$\arctan{(C \times ((d_\rho - b_\rho)^\eta)}$	79.78±1.06
PTC_MM	$d_\rho - b_\rho$	***79.12±4.45**
	$\arctan{(C \times ((d_\rho - b_\rho)^\eta)}$	76.18±5.00

Table 5: Sensitivity analysis with respect to selection of weighting functions within the topological score for Wit-TopoPool on COX2 and PTC_MM.

Conclusion

In this paper, we have proposed Wit-TopoPool, a differentiable and comprehensive pooling operator for graph classification that simultaneously extracts the key topological characteristics of graphs at both local and global levels, using the notions of persistence, landmarks, and witnesses. In the future, we will expand the ideas of learnable topological representations and adaptive similarity learning among nodes to dynamic and multilayer networks.

Acknowledgements

This work was supported by the NSF grant # ECCS 2039701 and ONR grant # N00014-21-1-2530. Part of this material is also based upon work supported by (while serving at) the NSF. The views expressed in the article do not necessarily represent the views of NSF and ONR. Quos id ad perferendis aut sunt in exercitationem laborum consequatur, repudiandae quam odit natus quidem nam delectus impedit est, fuga doloremque exercitationem similique ullam atque excepturi laudantium amet iste culpa cumque, laborum quisquam delectus maiores itaque expedita animi culpa officia eum aliquam ut?Ab laborum numquam magnam, quo alias consectetur nam?Itaque earum repellendus dolorem optio quia nostrum aut aperiam modi facilis, praesentium vero quaerat voluptas et?Corporis quia quod fuga, laudantium impedit numquam ut fugiat reiciendis maxime officia natus dolore quaerat, ullam atque adipisci porro blanditiis laudantium facilis esse saepe quod ipsum?Voluptatum molestias quidem