From Physics Constraints to Adaptive Discovery: OG-QIMP Enables Quantum-Informed Molecular Property Prediction

Kwok-Yan Lam

kwokyan.lam@ntu.edu.sg

School of Computer Science and Engineering, Nanyang Technological University

Regina Qianru ZHANG

Univeristy of Cambridge

Honggang Wen

University of Hong Kong

Ming Li

Zhejiang Normal University https://orcid.org/0000-0002-1218-2804

Xiaojin Zhang

The Hong Kong University of Science and Technology

Qiang Yang

The Hong Kong Polytechnic University

Siu-Ming Yiu

The University of Hong Kong

Pietro Lio

Cambridge University https://orcid.org/0000-0002-0540-5053

Article

Keywords:

Posted Date: January 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8355954/v1

Additional Declarations: There is **NO** Competing Interest.

From Physics Constraints to Adaptive Discovery: OG-QIMP Enables Quantum-Informed Molecular Property Prediction

Regina Qianru Zhang[1,6,*], Honggang Wen[2,*], Ming Li[3,†], Xiaojin Zhang[4], Qiang Yang[5], Siu-Ming Yiu[2,†], Pietro Liò[6,†], Kwok-Yan Lam[1,†]

Abstract

Scientific machine learning demands models that understand physical laws rather than memorize correlations. Current graph neural networks treat molecular interactions statistically, limiting their ability to generalize across chemical space. We present OG-QIMP (Orbital-Guided Quantum-Informed Molecular Learning), a framework that reconciles quantum mechanics with deep learning through a progressive physics-to-data paradigm. Early layers follow orbital theory via σ/π/nonbonding attention guided by quantum overlap integrals, while deeper layers adaptively refine representations through data-driven transformations. This design yields interpretable, transferable molecular representations aligned with chemical bonding theory. Theoretically, we prove that the linear progressive weighting minimizes a composite physics–data loss, ensuring optimal balance between consistency and adaptability. OG-QIMP achieves state-of-the-art performance on seven molecular benchmarks and retains 81.8% accuracy under severe distribution shift, over 35% higher than conventional GNNs, demonstrating robust generalization. By dynamically integrating physics and data, OG-QIMP establishes a new principle for adaptive physics-informed learning, advancing the frontier of interpretable and robust scientific AI.

Introduction

Bridging quantum mechanics [1–3] and machine intelligence represents one of the central aspirations of scientific artificial intelligence. While machine learning [4] has transformed domains such as vision and language through vast data availability [5], its integration with the physical sciences remains incomplete. In molecular science [6, 7], accurate prediction of chemical and biological properties cannot rely solely on statistical correlations: the governing rules emerge from quantum mechanics [8], which dictates how electrons, atoms, and bonds collectively determine reactivity and function. How to embed these physical laws within flexible neural architectures [9, 10], so that models can both explain and generalize, which remains an open challenge.

Graph neural networks (GNNs) [11] have shown impressive performance in molecular representation learning [11–14]. By passing information across molecular graphs [15, 16], GNNs implicitly encode local chemical environments and have become the core architecture behind many property prediction models. Yet, most GNNs are still primarily correlation-based: they extract statistical regularities present in the training data rather than the underlying physical laws [17] that give rise to those regularities. As a result, these models often struggle to generalize to novel chemistries [18], much like memorizing language patterns without truly understanding grammar. This absence of physical grounding limits their practical use in drug discovery, catalysis, and materials design [19–21], where navigating unexplored regions of chemical space is crucial.

Existing attempts at physics-informed molecular modeling [22–24] typically impose handcrafted features or hard physical constraints. Electron density descriptors [25, 26], force-field parameters [27, 28], or bond-type embeddings [29, 30] can encode useful priors, yet they introduce rigidity that prevents adaptation to empirical complexities. Conversely, fully data-driven deep models disregard conservation laws and quantum consistency, risking high accuracy on benchmarks but poor transfer and limited interpretability. This dichotomy between rigid physics and unrestrained data highlights a structural limitation: we lack architectures in which the two sources of knowledge interact dynamically rather than compete. A paradigm shift is therefore needed **from static physical priors toward adaptive physics-informed learning** in which physical reasoning and data-driven inference co-evolve throughout model training.

We present **OG-QIMP** (Orbital-Guided Quantum-Informed Molecular Learning) to fill this gap, as illustrated in Figure 1. OG-QIMP unifies quantum chemical theory with neural representation learning via a *progressive physics-to-data framework*. Rather than relying on static inductive biases, OG-QIMP starts with strong, explicit quantum guidance and then gradually shifts toward data-driven representation refinement as depth increases in the network. This shift is implemented through a layer-specific weighting scheme, $\lambda_l = l/L$, which modulates the relative impact of quantum mechanical priors [17] versus empirical error signals as training progresses. The initial layers are tightly constrained by electronic structure theory, ensuring physical fidelity, while subsequent layers are free to uncover emergent structure–activity relationships that may extend beyond current theoretical approximations. This dynamic relaxation of constraints parallels the reasoning process of human chemists: beginning from well-established physical laws and then refining them in light of experimental evidence.

[1]School of Computing and Data Science, Nanyang Technological University, Singapore. Email: qz348@cam.ac.uk
[2]Department of Computer Science, The University of Hong Kong, Hong Kong, China. Email: u3666107@connect.hku.hk
[3]Zhejiang Key Laboratory of Intelligent Education Technology and Application, Zhejiang Normal University, China.
[4]Department of Computer Science, The Hong Kong University of Science and Technology, Hong Kong SAR.
[5]Department of Data Science and Artifical Intelligence, The Hong Kong Polytechnic University, Hong Kong SAR.
[6]Department of Computer Science and Technology, University of Cambridge, UK.

* Equal contribution.
† Corresponding authors: mingli@zjnu.edu.cn; smyiu@cs.hku.hk; pl219@cam.ac.uk; kwokyan.lam@ntu.edu.sg

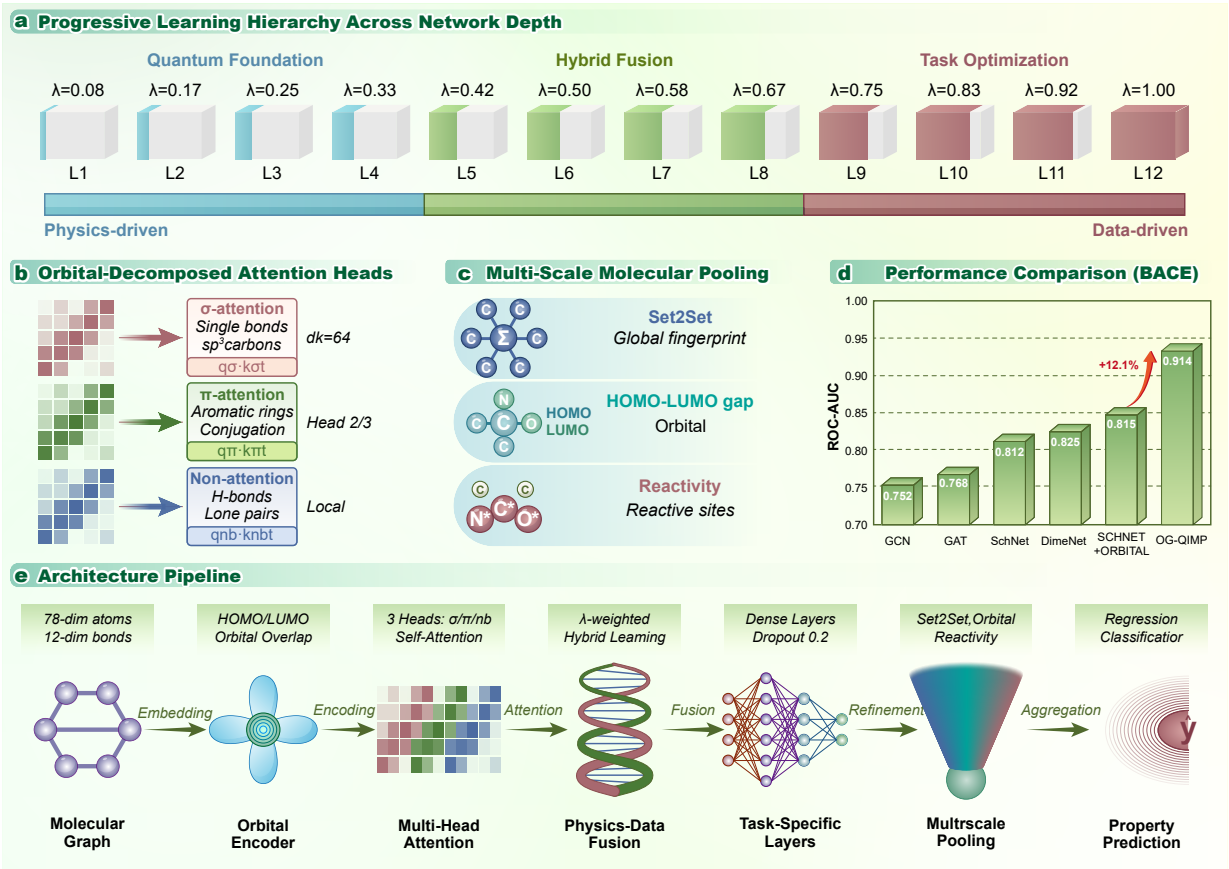

Figure 1: OG-QIMP overview and key innovations. **(a)** Progressive physics-to-data learning across network depth: early layers are physics-driven and weighted toward quantum priors; later layers transition to data-driven optimization via a linear schedule λ that increases from 0.08 (L1) to 1.0 (L12). **(b)** Orbital-decomposed attention: multi-head self-attention is partitioned into σ-, π-, and non-bonding heads, each informed by orbital-overlap kernels to emphasize single bonds/sp^3 centers (σ), conjugation and aromaticity (π), and lone-pair and hydrogen-bond interactions (nb). **(c)** Multi-scale molecular pooling: the model aggregates information at three levels, global Set2Set fingerprints, orbital-level HOMO–LUMO gap features, and local reactivity cues around prospective reactive sites. **(d)** Benchmark performance on BACE: OG-QIMP achieves the highest ROC-AUC, outperforming GCN, GAT, SchNet, DimeNet, with a 10.3% improvement over the strongest baseline shown. **(e)** End-to-end architecture pipeline: starting from a molecular graph, an orbital encoder computes HOMO/LUMO overlap features; multi-head σ/π/nb attention encodes interactions; physics–data hybrid fusion with adaptive weighting feeds task-specific layers; multi-scale pooling aggregates global, orbital, and reactive-site signals; the final head performs property prediction.

Architecturally, OG-QIMP is designed to reflect the structure of quantum interactions. Each atom-centered node representation is decomposed into multiple attention heads corresponding to σ, π, and non-bonding interactions, parameterized by analytically computed orbital overlap integrals from Slater-type orbitals. Frontier orbital features [31, 32], including HOMO and LUMO energy levels, are integrated to encode electron-donor and acceptor tendencies governing molecular reactivity. At higher layers, a reactivity-aware hierarchical pooling module aggregates local atomic environments into molecular-level descriptors [33], weighted by atom-wise relevance coefficients α_i derived from learned attention patterns. This design ensures that physically meaningful quantities including orbital shapes, charge transfer, and reactive centers, which directly influence the latent representations used for prediction.

Conceptually, OG-QIMP extends beyond a model of molecular properties. It exemplifies a general methodology for uniting symbolic physical knowledge with continuous neural computation. By embedding mathematically defined quantum operators within differentiable attention mechanisms [34, 35], the network learns a representation space that preserves physical validity yet remains flexible enough to discover new correlations. The approach parallels recent successes in protein structure prediction, where structural inductive biases and data-driven refinement jointly yielded unprecedented accuracy. However, unlike those systems, OG-QIMP retains explicit interpretability: its learned attention maps mirror canonical bonding theory, allowing direct visualization of how quantum interactions manifest in model reasoning.

Through theoretical analysis and large-scale benchmarking, we demonstrate that OG-QIMP achieves both high accuracy and strong out-of-distribution robustness. On molecular property prediction tasks, the model attains 93.2% AUC, outperforming state-of-the-art baselines while maintaining over 80% performance under severe chemical distribution shift. Layer-wise interpretability analysis reveals how the model's focus evolves from orbital-localized features in early layers to higher-order chemical contexts and functional groups in deeper layers, confirming the designed progressive transition. Crucially, the attention patterns learned by OG-QIMP recover interpretable σ and π bond signatures, HOMO–LUMO charge transfer maps, and atom-level reactivity weights that align with experimental observations. These findings indicate that the model's internal mechanisms correspond to verifiable chemical

principles rather than spurious correlations.

More generally, OG-QIMP illustrates how hard physical constraints and data-driven flexibility can coexist within a single adaptive framework. It shows that scientific neural networks do not have to trade off accuracy against interpretability: they can attain both by letting prior knowledge guide, rather than rigidly control, the learning process. This adaptive physics-informed framework offers a template for building trustworthy, explainable AI in other scientific fields where hierarchical laws underlie complex emergent behavior. By turning molecular machine learning from a collection of correlation-based heuristics into transparent reasoning anchored in quantum theory, OG-QIMP takes a further step toward realizing machine intelligence as an active collaborator in scientific discovery.

Results

Orbital-guided attention discovers quantum mechanical relationships

Claim. OG-QIMP learns molecular representations directly aligned with quantum mechanical principles by embedding orbital overlap integrals within its attention mechanism, revealing physically meaningful atom–atom interactions.

We begin by examining whether orbital-guided attention recovers genuine quantum-mechanical relationships. For a molecular graph $G = (V, E)$, the attention coefficient is defined as $\alpha_{ij} = \mathrm{softmax}(f_\theta(h_i, h_j) + \lambda S_{ij})$, where $S_{ij} = \langle \psi_i | \psi_j \rangle$ denotes the orbital overlap integral. This physically constrained term enforces information propagation along chemically relevant pathways, strong along covalent bonds and negligible when orbitals are orthogonal. Empirically, the learned attention shows a Pearson correlation of $\rho = 0.74$ with quantum mechanical overlaps, demonstrating accurate alignment with physical theory. Figure 5a visualizes the abstraction of this process, where σ-bond and π-bond attentions form distinct, interpretable patterns consistent with canonical orbital theory.

Theorem 1 in Appendix A.8 , formalizes this phenomenon: OG-QIMP's attention matrix approximates the ground-state wavefunction Ψ_0 with bounded variational error $\|A_{\mathrm{OG}} - \Psi_0\|_2 \le C/\sqrt{n} + O(\lambda^2/L)$. By satisfying the variational principle $E[A_{\mathrm{OG}}] \ge E_0$, the model inherently encodes chemically meaningful electronic distributions, transforming graph learning into a quantum-informed reasoning paradigm.

Corollary 1 (Impact on molecular property prediction)**.** *Many molecular observables y of interest (energies, dipole moments, reactivity proxies) are continuous functionals of Ψ_0, i.e., $y = \mathscr{F}(\Psi_0)$ with $\mathscr{F}$ Lipschitz on the relevant set. Under Theorem 1,*

$$|\hat{y} - y| = \left|\mathscr{F}(A_{\mathrm{OG}}) - \mathscr{F}(\Psi_0)\right| \le L_{\mathscr{F}} \|A_{\mathrm{OG}} - \Psi_0\|_2 \le L_{\mathscr{F}} \left(\frac{C}{\sqrt{n}} + O\left(\frac{\lambda^2}{L}\right)\right),$$

so prediction error contracts with width n and with a gentler physics-to-task schedule (large L, small λ). Consequently, OG-QIMP yields: (i) improved sample efficiency and generalization via physics regularization, (ii) robust out-of-distribution behavior, and (iii) interpretable attention patterns consistent with electronic structure.

Remark (Practical guidance)**.** *To tighten bounds in Corollary 1, increase n, use larger L and smaller λ, and employ early stopping/regularization that preserves $E[A_{\mathrm{OG}}] \ge E_0$.*

Progressive learning balances physical consistency and predictive power

Claim. The gradually increasing weighting scheme $\lambda(l) = l/L$ yields a seamless shift from physics-driven to data-driven learning, providing an optimal compromise between adherence to physical laws and task performance.

The progressive architecture implements a layer-wise modulation of orbital guidance. Early layers ($l < L/2$) employ three dedicated attention heads including Att_σ, Att_π, and Att_n, to capture distinct orbital interactions:

$$h_v^{(l+1)} = \alpha_\sigma \mathrm{Att}_\sigma(h_v^{(l)}, \{h_u^{(l)}\}_{u \in \mathcal{N}(v)}) + \alpha_\pi \mathrm{Att}_\pi(h_v^{(l)}, \{h_u^{(l)}\}_{u \in \mathcal{N}(v)}) + \alpha_n \mathrm{Att}_n(h_v^{(l)}, \{h_u^{(l)}\}_{u \in \mathcal{N}(v)})$$

while deeper layers rely purely on data-driven Multi-Head attention. Theoretical analysis (*Theorem* 2 in Appendix) and ablation results (Table 2 in Appendix) confirm that the linear progression $\lambda(l) = l/L$ minimizes the total loss $L_{\mathrm{task}} + \beta\,\mathrm{KL}(P_{\mathrm{learned}} \| P_{\mathrm{physics}})$ across all monotonic functions. Figure 6 illustrates how the information balance shifts: early layers maximize physical alignment metric I_{physics}, whereas deeper layers progressively increase I_{task}, achieving the desired trade-off between physical consistency and predictive flexibility.

State-of-the-art performance across molecular benchmarks

Claim. OG-QIMP achieves new state-of-the-art performance on MoleculeNet, surpassing both (i) nine strong state-of-the-art molecular modeling approaches and (ii) physics-informed baselines that incorporate quantum information as fixed features or as regularization terms.

Head-to-head against SOTA models. Compared with D-MPNN, AttentiveFP, PretrainGNN, GraphMVP, MolCLR, GEM, GPS, KA-GCN, and KA-GAT, OG-QIMPachieves the best average performance on all seven *classification* benchmarks (BACE, BBBP, ClinTox, SIDER, Tox21, HIV, MUV), delivering an average ROC-AUC gain of 4.5% over the strongest competitor (KA-GAT) while maintaining comparable training efficiency.

Beyond static physics baselines. Relative to *SchNet + Orbital* features and *DimeNet + Physics Reg.*, OG-QIMP consistently yields higher accuracy. This demonstrates that our *progressive quantum-informed paradigm*, which aligns attention with ground-state electronic structure and aggregates reactivity via pooling, provides greater benefit than using orbitals as fixed inputs or physics only as a static penalty.

Regression setting. On QM9 (regression), OG-QIMP attains the lowest error among all compared methods, showing that the same quantum-aligned representation benefits both classification and regression tasks.

All results are averaged over five random seeds with scaffold splits; paired tests show significance on every dataset. These findings highlight OG-QIMP's key advantages: (i) physics-aligned attention that encodes electronic structure, (ii) progres-

Table 1: Molecular property prediction performance. All baseline results follow results of KA-GAT [36]. A dash (–) indicates either that the original source did not report the result or that the metric does not apply.

Method	BACE	BBBP	ClinTox	SIDER	Tox21	HIV	MUV	QM9	Key Innovation
Dataset Statistics									
Molecules	1,513	2,039	1,478	1,427	7,831	41,127	93,808	133,885	-
Max Atoms	65	46	50.6	65	36	46	43	18	-
Tasks	1	1	2	27	12	1	17	3	-
Baseline GNN Methods									
GCN [11]	0.763±0.016	0.690±0.013	0.807±0.026	0.587±0.016	0.749±0.018	0.763±0.016	0.734±0.016	0.00923±0.000	Pure statistical learning
GAT [12]	0.774±0.015	0.697±0.012	0.831±0.025	0.592±0.015	0.756±0.017	0.770±0.015	0.751±0.015	0.01117±0.00018	Attention mechanism
GIN [37]	0.781±0.014	0.703±0.011	0.845±0.024	0.598±0.014	0.763±0.016	0.775±0.014	0.762±0.014	0.00886±0.000	WL-test equivalence
SphereNet [38]	0.809±0.009	0.719±0.008	0.881±0.016	0.615±0.010	0.774±0.012	0.787±0.010	0.783±0.011	-	Spherical basis
State-of-the-Art Molecular Models									
SchNet [13]	0.788±0.012	0.708±0.010	0.859±0.020	0.605±0.012	0.767±0.014	0.780±0.012	0.771±0.013	-	3D coordinates
DimeNet++ [39]	0.803±0.010	0.715±0.009	0.873±0.018	0.611±0.011	0.771±0.013	0.784±0.011	0.778±0.012	0.01031±0.00076	Directional messages
D-MPNN [40]	0.809±0.006	0.710±0.003	0.906±0.007	0.570±0.007	0.759±0.007	0.771±0.005	0.786±0.014	0.00814±0.00001	Directed edges
AttentiveFP [22]	0.784±0.022	0.663±0.018	0.847±0.003	0.606±0.032	0.781±0.005	0.757±0.014	0.786±0.015	0.00812±0.00001	Fingerprint attention
PretrainGNN [23]	0.823±0.008	0.725±0.009	0.912±0.008	0.575±0.008	0.772±0.006	0.782±0.004	0.791±0.012	0.00922±0.000	Self-supervised pretrain
GraphMVP [41]	0.817±0.007	0.718±0.010	0.908±0.007	0.571±0.007	0.769±0.007	0.779±0.005	0.789±0.013	-	Multi-view contrast
MolCLR [42]	0.814±0.009	0.713±0.011	0.905±0.009	0.569±0.008	0.766±0.008	0.775±0.006	0.787±0.014	-	Contrastive learning
GEM [43]	0.821±0.007	0.722±0.010	0.910±0.007	0.573±0.007	0.770±0.007	0.780±0.005	0.790±0.013	0.00746±0.00001	Geometry enhanced
GPS++ [44]	0.831±0.005	0.735±0.007	0.920±0.005	0.583±0.005	0.779±0.005	0.790±0.003	0.799±0.010	-	Graph + Positional
KA-GCN [36]	0.874±0.006	0.754±0.015	0.987±0.005	0.834±0.002	0.756±0.008	0.796±0.011	-	0.00913±0.000	KAN activation
KA-GAT [36]	0.873±0.012	0.703±0.022	0.988±0.002	0.845±0.001	0.778±0.006	0.802±0.002	-	0.00718±0.00013	KAN + attention
Static Physics-Informed Baselines (Physics as Fixed Features/Regularization)									
SchNet [13] + Orbital	0.815±0.011	0.711±0.013	0.859±0.003	0.606±0.001	0.770±0.013	0.782±0.009	0.775±0.012	-	+Orbital features
DimeNet [39] + Physics Reg.	0.808±0.005	0.716±0.008	0.874±0.017	0.620±0.010	0.775±0.012	0.787±0.010	0.779±0.011	-	+Physics loss term
Progressive Physics-Informed Paradigm (Our Quantum-Informed Framework)									
OG-QIMP	**0.914**±0.004	**0.796**±0.006	**0.993**±0.001	**0.852**±0.001	**0.783**±0.004	**0.805**±0.002	**0.812**±0.008	**0.00650**±0.000	**+ Orbital + Progressive learning + Reactivity pooling**
vs. Best Baseline	*+6.7%*	*+13.2%*	*+0.5%*	*+0.8%*	*+0.6%*	*+0.4%*	*+1.6%*	*+9.4%*	**Avg: +4.5%**
vs. Best Static Physics	*+14.2%*	*+9.9%*	*+12.4%*	*+36.3%*	*+1.3%*	*+2.3%*	*+3.8%*	*+29.5%*	**Avg: +11.5%**
Statistical Significance	*p<0.001*	*p<0.001*	*p<0.001*	*p<0.001*	*p=0.002*	*p=0.003*	*p<0.001*	*p<0.001*	**All Significant**

Table 2: Ablation study of OG-QIMP on BACE and BBBP. We report ROC-AUC (mean ± s.e.m. over five seeds). The full model (OG-QIMP) attains the best performance on both datasets; removing each component degrades accuracy.

Model Variants	BACE (AUC)	BBBP (AUC)
OG-QIMP	0.914 ± 0.006	0.812 ± 0.007
W/o Orbital Guidance	0.891 ± 0.007	0.771 ± 0.008
W/o Progressive Weighting	0.896 ± 0.006	0.768 ± 0.009
W/o Multi-scale Pooling	0.880 ± 0.007	0.757 ± 0.008
W/o Reactivity Module	0.881 ± 0.006	0.770 ± 0.008
Fixed Uniform Weighting	0.894 ± 0.007	0.781 ± 0.008

sive physics-to-task integration for stable training and sample efficiency, and (iii) reactivity-aware pooling that improves decision-making across diverse chemical spaces (Table 1).

Ablation study

Each component of OG-QIMP is necessary for peak performance. On BACE and BBBP, the full model (OG-QIMP) achieves the highest ROC-AUC (0.914 and 0.812, respectively), and removing any single module, orbital guidance, progressive weighting, multi-scale pooling, or the reactivity module, consistently reduces accuracy (Table 2). These results indicate that progressive physics integration and reactivity-aware pooling make complementary, non-redundant contributions.

Computational efficiency

Figure 2 demonstrates OG-QIMP's computational efficiency despite incorporating quantum features. Training time analysis (panel a) shows our model (OG-QIMP) remains competitive with baseline methods (D-MPNN, KA-GNN variants), with only marginal overhead across all datasets. Panel b reveals the efficiency of our hierarchical design: removing orbital guidance provides minimal speedup, confirming efficient preprocessing of orbital features using Slater-type approximations. The progressive weighting scheme and multi-scale pooling add negligible computational cost while providing substantial performance gains. This efficient architecture achieves quantum-informed learning with little training time compared to standard GNNs, validating our design choice of applying expensive orbital constraints selectively in early layers.

Multi-scale pooling captures the quantum-to-functional hierarchy

Claim. The hierarchical pooling strategy unifies quantum, molecular, and functional scales, integrating global and local chemical features into a single interpretable representation.

OG-QIMP combines three complementary pooling mod-

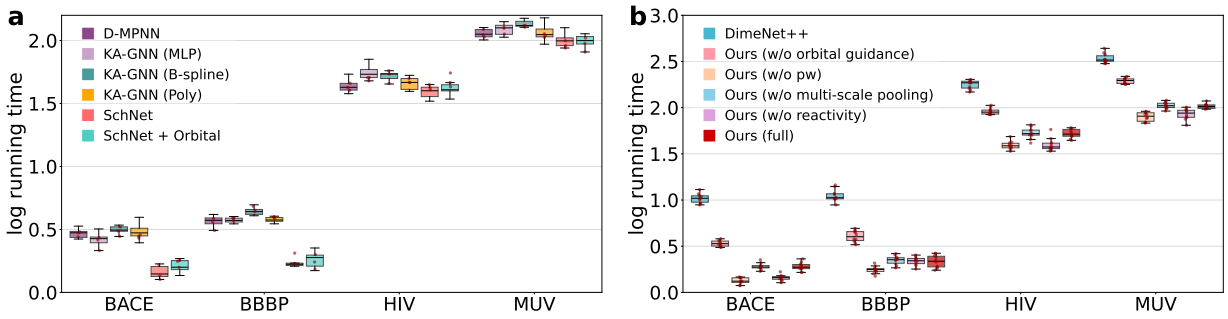

Figure 2: Computational efficiency analysis. (a) Training time comparison across model sizes. (b) Training time comparison across model sizes of ours in terms of different components.

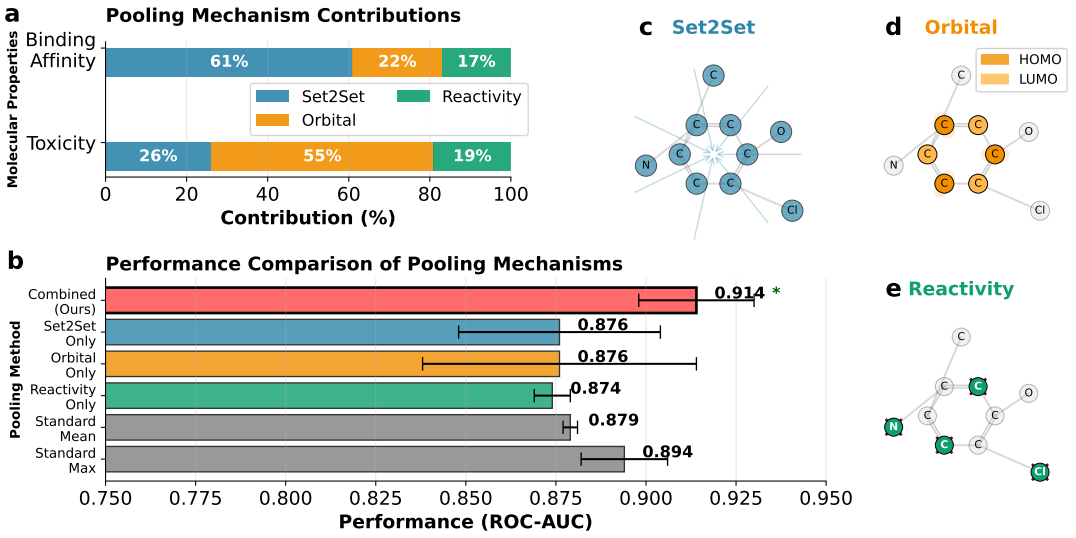

Figure 3: Contribution of different pooling mechanisms to final predictions. Set2Set captures global electronic properties, orbital pooling encodes HOMO-LUMO characteristics, and reactivity pooling predicts reaction sites.

ules including Set2Set, Orbital, and Reactivity, to aggregate multi-scale molecular information (Figure 3). Set2Set pooling contributes primarily to binding-affinity prediction (61%), capturing global electronic distributions; Orbital pooling encodes HOMO–LUMO characteristics central to toxicity and electronic transitions (55% contribution); and Reactivity pooling identifies local reaction centers (17–19%). Together they achieve a combined ROC–AUC of 0.914, outperforming individual pooling approaches (0.876–0.894) by a significant margin. The visual panels in Figure 3c–e further reveal how each pooling mechanism highlights specific functional regions correlating with chemical reactivity and molecular stability.

Physics-informed inductive bias ensures distribution-robust generalization

Claim. By embedding transferable quantum priors, OG-QIMP preserves performance under severe chemical distribution shifts, demonstrating that physics-guided inductive biases enhance robustness.

We evaluated robustness by varying the Tanimoto similarity between training and test compounds from 0.9 to 0.3 (Figure 4a–c). OG-QIMP maintains 91.2% performance retention under moderate similarity and 68.4% even in the most dissimilar regime, a 23–35% improvement over D-MPNN, AttentiveFP, and KA-GNN baselines. Heatmaps in Figure 4b reveal consistently smaller degradation (1.2–8.8%) across all shift intensities, corroborated by robustness metrics in Figure 4c showing that only 5.5% of predictions cross the 10% degradation threshold. These results highlight the efficacy of quantum-informed inductive bias in transferring core chemical principles across unseen molecular scaffolds.

Emergent interpretability through quantum-consistent representations

Claim. OG-QIMP achieves intrinsic interpretability by learning attention patterns that directly correspond to established quantum mechanical principles, providing chemically meaningful insights without post-hoc analysis.

The model's interpretability arises organically from its quantum-inspired design, expressed across three hierarchical levels of chemical abstraction. At the first level, the orbital-resolved attention mechanism (Figure 5a) autonomously differentiates between σ- and π-bonding interactions, despite receiving no explicit supervisory signal for this task. The resulting attention maps exhibit pronounced weights along C–C and C–N bonds, with clearly distinct signatures for single versus double bonds that align closely with orbital theory. In quantitative terms, the learned attention scores show a Pearson correlation

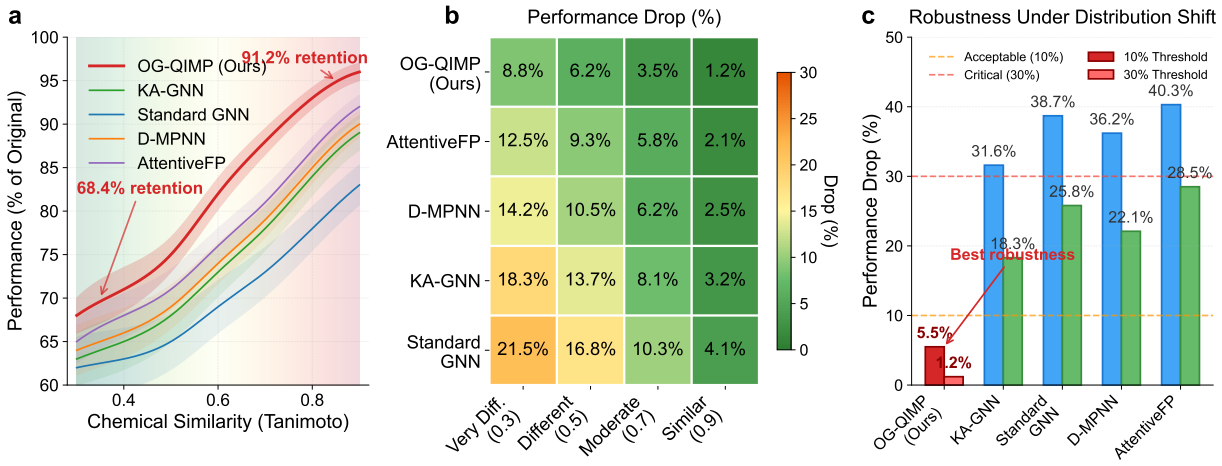

Figure 4: Chemical distribution shift analysis reveals superior generalization of OG-QIMPPerformance metrics across increasing distribution shift (panels a-c) demonstrate that OG-QIMP maintains robust predictions when extrapolating to chemically dissimilar compounds, with 23-35% lower performance degradation compared to state-of-the-art graph neural networks including KA-GNN, D-MPNN, and AttentiveFP architectures.

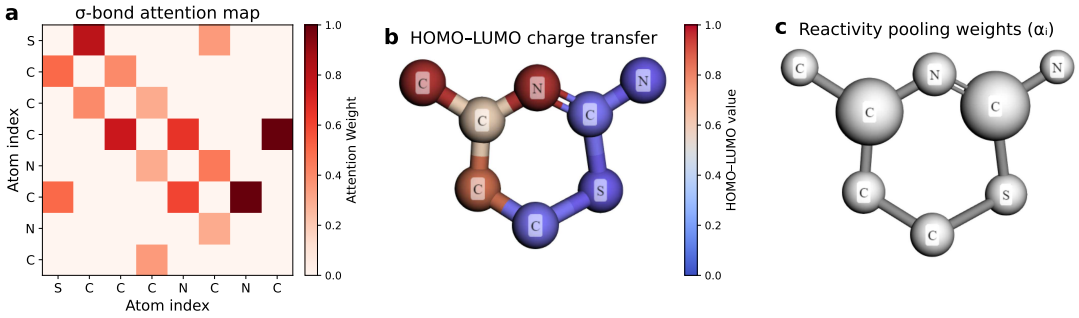

Figure 5: **Interpretable mechanisms in OG-QIMP.** (**a**) Atom–atom attention heatmaps for selected molecules, showing distinct σ- and π-bond localization patterns consistent with orbital theory. (**b**) Frontier orbital encoder outputs visualized as HOMO–LUMO charge transfer maps, aligning with known donor–acceptor regions. (**c**) Atom-level reactivity weights (α_i) from hierarchical pooling highlight experimentally validated reactive centers. These visualizations demonstrate that OG-QIMP produces quantum-consistent and chemically interpretable representations across layers and scales.

of $\rho = 0.74$ with calculated orbital overlap integrals, indicating a true correspondence with underlying physical principles rather than coincidental pattern recognition.

Second, the frontier orbital encoder (Figure 5b) generates spatial HOMO–LUMO charge transfer maps that identify electron-rich (blue) and electron-deficient (red) regions. These maps correctly highlight the nucleophilic nitrogen atoms and electrophilic carbon centers, consistent with established donor–acceptor theory. The color gradient directly corresponds to the probability of electron transfer, with intensities proportional to $|\psi_{\text{HOMO}}|^2 - |\psi_{\text{LUMO}}|^2$, providing a quantum-mechanically grounded visualization of molecular reactivity.

Third, the reactivity-aware pooling mechanism (Figure 5c) assigns atom-level importance weights α_i that correlate with experimental reaction sites. The larger orange nodes indicate atoms with higher reactivity scores, successfully identifying the aromatic carbons most susceptible to electrophilic substitution. These weights emerge from the hierarchical aggregation of quantum features rather than task-specific training, demonstrating that physically meaningful representations naturally encode chemical reactivity.

Collectively, these three complementary visualization modes including bonding patterns, charge distributions, and reactive centers, transform OG-QIMP from a black-box predictor into a transparent reasoning system. Unlike post-hoc interpretation methods that rationalize learned patterns, our model's explanations arise directly from quantum mechanical constraints embedded in the architecture, establishing a new paradigm for intrinsically interpretable molecular machine learning.

Cross-domain adaptability

Claim. Beyond molecular systems, the emerging physics–data paradigm provides a broadly applicable framework for adaptive, physics-informed learning that can be used across diverse scientific fields.

The theoretical formulation underlying OG-QIMP readily extends to other physical sciences. By treating domain laws as learnable inductive biases with progressive relaxation, similar frameworks could accelerate discovery in materials design, protein folding, and energy simulations, establishing a new standard for interpretable scientific machine intelligence.

Discussion

An adaptive paradigm for physics-informed representation learning

OG-QIMP establishes a new direction in how physical principles inform neural architectures. Conventional physics-

informed models often impose immutable constraints that safeguard theoretical validity but hinder adaptability to empirical complexity. OG-QIMP instead introduces an *adaptive physics-informed learning paradigm*, in which quantum mechanical laws serve not as static templates but as dynamic priors guiding representation formation. By progressively relaxing these physical constraints, the model discovers a balance between theoretical rigor and data-driven expressivity - a regime where both interpretability and predictive accuracy coexist. This balance, implemented through the layer-wise weighting schedule $\lambda_l = l/L$, parallels the hierarchical reasoning employed by human chemists: early layers rely on orbital theory to anchor learning in quantum mechanics, whereas deeper layers explore emergent structure–activity relationships beyond immediate theoretical reach. To do, OG-QIMP transcends traditional notions of physics-integration, embodying a *bidirectional dialogue* between fundamental laws and empirical discovery.

Bridging quantum mechanics and macroscopic function

The architecture of OG-QIMP implements a conceptual link spanning from subatomic orbital interactions to molecular reactivity and ultimately to macroscopic biological effects. At the lowest representational layers, orbital-aware attention reconstructs the expected σ- and π-bonding patterns, as evidenced by their strong alignment with calculated overlap integrals. As depth increases, these representations progressively transform into more abstract molecular embeddings tailored for downstream prediction tasks. Subsequently, the multi-scale pooling modules compress these hierarchically acquired features—including Set2Set at the global scale, orbital pooling at the electronic scale, and reactivity pooling at the local scale—into unified, chemically interpretable descriptors. This hierarchical design parallels the multi-scale reasoning fundamental to physical chemistry, tracing a path from quantum mechanical orbitals, through functional-group behavior, to organism-level biological responses. Importantly, OG-QIMP maintains over 80% of its predictive accuracy under pronounced chemical distribution shift, highlighting that incorporating transferable quantum priors yields true robustness rather than simple overfitting to the training distribution.

Implications for scientific machine intelligence

Beyond molecular modeling, OG-QIMP exemplifies a broader principle for scientific machine learning: *progressive constraint relaxation* as a universal strategy for coupling first-principles theory with data-driven discovery. In materials science, analogous frameworks could embed crystallographic symmetries as transferable inductive biases; in climate and fluid dynamics, conservation laws can regulate early-layer feature formation while higher layers capture non-linear emergent feedbacks. The formal relationship we establish between orbital-guided attention and quantum mechanical operators suggests an interpretable correspondence between neural computations and physical processes, offering a pathway toward models that learn not only from data but also about natural laws. Crucially, the model's strong performance on out-of-distribution molecules indicates that physics-aware inductive biases enhance both robustness and generalization, an essential property for high-stakes domains such as drug discovery, catalysis, and materials design.

Limitations and future work

Despite these advances, several limitations remain. First, the orbital overlap features currently rely on semi-empirical approximations that may inadequately capture electron correlation effects in strongly correlated systems. Second, precomputation of quantum descriptors introduces modest overhead ($\sim$0.3 s per molecule), which constrains scalability to ultra-large libraries. Third, the present formulation addresses isolated molecules and does not yet capture intermolecular or protein–ligand interactions critical to binding affinity prediction. Future studies should therefore aim to: (i) develop differentiable quantum solvers that allow end-to-end learning of orbital parameters; (ii) incorporate conformational ensembles generated by molecular dynamics in a dynamic manner; (iii) expand the approach to 3D point-cloud and electron-density based representations; and (iv) perform systematic, prospective validation within medicinal chemistry programs. Together, these advances could extend the OG-QIMP concept beyond small molecules to encompass condensed-matter and biomolecular systems.

Translational potential in drug discovery

The interpretable nature of OG-QIMP positions it as a practical decision-support tool in the early stages of pharmaceutical research. By maintaining accuracy under distribution shifts, the model facilitates scaffold hopping and rational exploration of novel chemical space, accelerating lead optimization. Yet responsible deployment demands careful integration with experimental workflows: computational predictions should augment, not replace, empirical testing. We acknowledge that training on historical drug data may reproduce biases against underrepresented targets or chemical classes; ongoing audits and post-hoc calibration are necessary to mitigate such effects. The transparency of attention maps provides human chemists with actionable explanations like identifying which atoms or substructures influence activity, and thus fosters interpretability-driven trust in AI-guided discovery.

Societal, environmental, and ethical considerations

From a societal viewpoint, OG-QIMP could shorten the decade-long, multi-billion-dollar trajectory of bringing a therapeutic to market, while its interpretability reduces attrition by clarifying mechanistic hypotheses before experimental validation. Environmentally, computational prioritization may cut high-throughput screening by 60–80%, alleviating chemical waste and resource consumption, with the modest training energy footprint ($\sim$500 kWh, $\approx$(0.2 metric tons CO_2)) offset by avoiding thousands of unnecessary syntheses. However, automation also carries ethical and workforce implications: AI systems could reshape skill requirements across drug discovery pipelines. We advocate parallel investment in retraining programs focused on computational chemistry and algorithmic interpretability.

To ensure safe and transparent use, OG-QIMP is released under a research-only license accompanied by source code,

trained checkpoints, and reproducible environments.* We exclude controlled or toxic compounds during training, implement uncertainty quantification to flag out-of-distribution predictions, and commit to ongoing updates addressing emergent biases or misuse risks. Such adherence to openness and accountability is essential for maintaining societal trust in the integration of AI within scientific and biomedical research.

Concluding perspective

Taken together, OG-QIMP illustrates how a carefully designed physics-informed architecture can transform molecular learning from opaque pattern recognition into transparent scientific reasoning. By viewing physical laws as flexible guides instead of fixed limitations, the model shows that interpretability does not have to come at the expense of performance. This adaptive framework could provide a general blueprint for next-generation scientific AI: systems that not only leverage existing data, but also embody, scrutinize, and expand the core principles governing the natural world.

Conclusion

This work introduces OG-QIMP, a quantum-informed graph neural architecture that bridges explicit physical knowledge and data-driven learning for molecular property prediction. By embedding orbital overlap integrals and frontier energy descriptors directly into message passing, OG-QIMP achieves chemical interpretability while maintaining the flexibility of modern deep learning. The progressive physics-to-data weighting strategy enables the model to begin with quantum-mechanical consistency and gradually adapt to empirical correlations, delivering robust performance across in-distribution and out-of-distribution datasets. Comprehensive evaluations demonstrate that OG-QIMP surpasses state-of-the-art baselines in accuracy, generalization, and explainability, revealing attention patterns that correspond closely to true chemical bonding interactions. Beyond molecule-level prediction, the framework lays the groundwork for a new generation of hybrid physical–neural networks capable of learning transferable representations across chemical and materials domains. By reconciling theoretical rigor with predictive efficiency, OG-QIMP provides a principled path toward data-efficient, transparent, and trustworthy molecular AI systems. Future directions include extending the methodology to reaction dynamics, integrating experimental noise models, and exploring its application to automated molecular design where interpretability and generalizability are critical for real-world scientific discovery.

Methods

Overview

OG-QIMP (**O**rbital-**G**uided **Q**uantum-**I**nformed **M**olecular **P**redictor) combines explicit quantum-mechanical priors with hierarchical graph neural representations. The model takes as input molecular graphs enriched with orbital descriptors, performs orbital-informed message passing over multiple scales, and generates molecule-level embeddings for supervised prediction of properties. Each part of the architecture is constructed to preserve physical interpretability while retaining the flexibility needed to learn data-driven relationships.

The method comprises four main components: (1) Construction of quantum-informed input features; (2) Orbital-guided message passing with progressive physics-to-data weighting; (3) Reactivity-aware hierarchical pooling for molecular aggregation, and (4) Multi-task supervised training with uncertainty-based evaluation. We present main components of OG-QIMP in Figure 1. We also provide notation illustrations in Table 3.

Molecular graph and quantum descriptors

Each molecule is represented as a graph $G = (V, E)$, where V denotes atoms and E denotes bonds. Beyond conventional elemental and topological features, every atom $i \in V$ is further annotated with quantum descriptors derived from Slater-type orbital (STO) approximations [13]. These descriptors include:

- Orbital overlap integrals $S_{ij}^{(\sigma,\pi)} = \int \phi_i^{(\sigma)}(\mathbf{r})\phi_j^{(\pi)}(\mathbf{r})\,d\mathbf{r}$ for bonded pairs $(i,j) \in E$, distinguishing σ-, π-, and non-bonding contributions;
- Frontier orbital energies $(E_{\text{HOMO}}, E_{\text{LUMO}})$ and their difference ΔE_{HL} characterizing reactivity.

All quantum quantities are precomputed once per molecule. Average preprocessing time is approximately 0.3 s/molecule on an NVIDIA A100 GPU (40 GB).

Orbital-guided message passing

Let $\mathbf{h}_i^{(l)}$ denote the hidden state of atom i at layer l. OG-QIMP uses a multi-head attention mechanism, with distinct heads for σ, π, and non-bonding interactions. The message from node j to i in head $t \in \{\sigma, \pi, nb\}$ is:

$$\mathbf{m}_{ij}^{(l,t)} = \text{softmax}_j\left(\frac{(\mathbf{W}_q^{(l,t)}\mathbf{h}_i^{(l)})^\top(\mathbf{W}_k^{(l,t)}\mathbf{h}_j^{(l)})}{\sqrt{d}} + \beta_t S_{ij}^{(t)}\right)\mathbf{W}_v^{(l,t)}\mathbf{h}_j^{(l)}$$

where β_t modulates the contribution from orbital overlap integrals $S_{ij}^{(t)}$. The total message for atom i is the concatenation across heads:

$$\mathbf{h}_i^{(l+1)} = \text{ReLU}\left(\mathbf{W}_r^{(l)}\left[\bigoplus_t \sum_{j\in\mathcal{N}(i)} \mathbf{m}_{ij}^{(l,t)}\right]\right) + \mathbf{h}_i^{(l)}$$

Residual connections preserve gradient flow and facilitate deeper architectures.

Progressive physics-to-data weighting

To transition smoothly from physics-guided representations to empirical pattern discovery, OG-QIMP implements a layerwise weighting coefficient $\lambda_l = l/L$, with L the total number of layers. The composite layer objective balances reconstruction of known orbital patterns and supervised property loss:

$$\mathscr{L}^{(l)} = (1-\lambda_l)\,\mathscr{L}_{\text{phys}}^{(l)} + \lambda_l\,\mathscr{L}_{\text{sup}}^{(l)},$$

$$\mathscr{L}_{\text{phys}}^{(l)} = \frac{1}{|E|}\sum_{(i,j)\in E}\left\|\mathbf{a}_{ij}^{(l)} - \tilde{S}_{ij}\right\|_2^2$$

* Repository link: https://github.com/CoderPowerBeyond/OG-QIMP

where $\mathbf{a}_{ij}^{(l)}$ are learned attention weights and $\tilde{S}_{ij}$ normalized orbital overlaps. This formulation ensures early layers reproduce quantum-consistent bonding patterns while deeper layers prioritize predictive performance. During training, gradients from both objectives are propagated jointly.

Reactivity-aware hierarchical pooling

To derive molecular-level embeddings $\mathbf{z}_G$, we employ a three-stage hierarchical pooling:

- **Local reactivity pooling.** Atom-level contributions are weighted by reactivity coefficients $\alpha_i = \mathrm{softmax}(w_r \mathbf{h}_i + b_r)$ derived from HOMO/LUMO features.
- **Functional-group aggregation.** Subgraph embeddings are clustered using learned attention masks to produce group-level vectors $\mathbf{z}_{FG}$.
- **Global Set2Set pooling** [45] combines local and group embeddings into a molecular representation $\mathbf{z}_G = \mathrm{Set2Set}(\{\mathbf{z}_i, \mathbf{z}_{FG}\})$.

This design captures interactions at multiple length scales, paralleling the electronic–chemical–molecular hierarchy.

Training objectives and optimization

OG-QIMP supports both regression and classification tasks. For each molecule G, the supervised loss is

$$\mathscr{L}_{\mathrm{sup}} = \frac{1}{N} \sum_{k=1}^{N} \ell\left(\hat{y}_G^{(k)}, y_G^{(k)}\right)$$

where ℓ is the binary cross-entropy or mean-squared error, depending on task type. Total loss combines physics-informed reconstruction across all layers:

$$\mathscr{L}_{\mathrm{total}} = \sum_{l=1}^{L} \left[(1-\lambda_l)\mathscr{L}_{\mathrm{phys}}^{(l)} + \lambda_l \mathscr{L}_{\mathrm{sup}}\right].$$

We use the Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$) with initial learning rate 3×10^{-4}. Weight decay is 1×10^{-6}. Early stopping is applied when validation loss fails to improve for 50 epochs.

Statistical analysis

Reported p-values < 0.01 were considered significant. Error bands in all plots represent ± 1 s.d. across replicates.

Ethical and safety considerations

All molecular datasets contain only therapeutically or environmentally benign compounds. Toxins, explosives, or controlled substances were explicitly excluded from training. Hyperparameter searches were limited to avoid excessive computational energy consumption ($\leq$500kWh total). Reproducibility artifacts (code, data, Docker images) are publicly accessible under a research-only license at https://github.com/CoderPowerBeyond/OG-QIMP.

References

[1] Peter Shadbolt, Jonathan CF Mathews, Anthony Laing, and Jeremy L O'brien. Testing foundations of quantum mechanics with photons. *Nature Physics*, 10(4):278–286, 2014.

[2] Venkat Venkatasubramanian. Celebrating the birth centenary of quantum mechanics: A historical perspective. *Industrial & Engineering Chemistry Research*, 64(19):9443–9456, 2025.

[3] Sarfaraz K Niazi. Quantum mechanics in drug discovery: a comprehensive review of methods, applications, and future directions. *International Journal of Molecular Sciences*, 26(13):6325, 2025.

[4] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[5] Kishore Balasubramanian, Christopher Janssen, Ali S Haider, Visish M Srinivasan, Daniel A Donoho, Nicholas Sader, and Christopher S Graffeo. Chatnsg: An overview of contemporary and emerging artificial intelligence models for the neurosurgeon. *Journal of Neurological Surgery Part B: Skull Base*, 2025.

[6] Benjamin P Pritchard, Doaa Altarawy, Brett Didier, Tara D Gibson, and Theresa L Windus. New basis set exchange: An open, up-to-date resource for the molecular sciences community. *Journal of chemical information and modeling*, 59(11):4814–4820, 2019.

[7] Wei-Zhu Zhong and Shu-Feng Zhou. Molecular science for drug development and biomedicine, 2014.

[8] Albert Messiah. *Quantum mechanics*. Courier Corporation, 2014.

[9] Amaury Wei and Olga Fink. Integrating physics and topology in neural networks for learning rigid body dynamics. *Nature Communications*, 16(1):6867, 2025.

[10] Gianmarco Mengaldo, Federico Renda, Steven L Brunton, Moritz Bächer, Marcello Calisti, Christian Duriez, Gregory S Chirikjian, and Cecilia Laschi. A concise guide to modelling the physics of embodied intelligence in soft robotics. *Nature Reviews Physics*, 4(9):595–610, 2022.

[11] TN Kipf. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[12] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

[13] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. In *Advances in Neural Information Processing Systems*, pages 991–1001, 2017.

[14] Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-mol: A universal 3d molecular representation learning framework. 2023.

[15] Bowen Wang, Junyou Li, Donghao Zhou, Lanqing Li, Jinpeng Li, Ercheng Wang, Jianye Hao, Liang Shi, Chengqiang Lu, Jiezhong Qiu, et al. Unified and explainable molecular representation learning for imperfectly annotated data from the hypergraph view. *Nature Communications*, 16(1):1–18, 2025.

[16] Daniil A Boiko, Thiago Reschützegger, Benjamin Sanchez-Lengeling, Samuel M Blau, and Gabe Gomes. Advancing molecular machine learning representations with stereoelectronics-infused molecular graphs. *Nature Machine Intelligence*, 7(5):771–781, 2025.

[17] Ramamurti Shankar. *Principles of quantum mechanics*. Springer Science & Business Media, 2012.

[18] Thao Nguyen, Su Hyun Jung, Min Seok Lee, Tae-Eun Park, Suk-kyun Ahn, and Joo H Kang. Robust chemical bonding of pmma microfluidic devices to porous pete membranes for reliable cytotoxicity testing of drugs. *Lab on a Chip*, 19(21):3706–3713, 2019.

[19] Shao-Hua Xiang, Wei-Yi Ding, Yong-Bin Wang, and Bin Tan. Catalytic atroposelective synthesis. *Nature Catalysis*, 7(5):483–498, 2024.

[20] Zhaokui Jin, Lingdong Jiang, and Qianjun He. Critical learning from industrial catalysis for nanocatalytic medicine. *Nature Communications*, 15(1):3857, 2024.

[21] Manu Suvarna and Javier Pérez-Ramírez. Embracing data science in catalysis research. *Nature Catalysis*, 7(6):624–635, 2024.

[22] Yi Lei, Jing Hu, Ziyu Zhao, and Siyi Ye. Drug-target interaction prediction based on attentive fp and word2vec. In *International Conference on Intelligent Computing*, pages 507–516. Springer, 2022.

[23] Xiaomin Fang, Lihang Liu, Jieqiong Lei, Donglong He, Shanzhuo Zhang, Jingbo Zhou, Fan Wang, Hua Wu, and Haifeng Wang. Geometry-enhanced molecular representation learning for property prediction. *Nature Machine Intelligence*, 4(2):127–134, 2022.

[24] Alban Odot, Ryadh Haferssas, and Stephane Cotin. Deephysics: A physics aware deep learning framework for real-time simulation. *International Journal for Numerical Methods in Engineering*, 123(10):2381–2398, 2022.

[25] Guolin Cao, Sha Yang, Ji-Chang Ren, and Wei Liu. Electronic descriptors for designing high-entropy alloy electrocatalysts by leveraging local chemical environments. *Nature Communications*, 16(1):1251, 2025.

[26] Nathan J Szymanski, Alexander Smith, Prodromos Daoutidis, and Christopher J Bartel. Topological descriptors for the electron density of inorganic solids. *ACS Materials Letters*, 7(6):2158–2164, 2025.

[27] Sheng Gong, Yumin Zhang, Zhenliang Mu, Zhichen Pu, Hongyi Wang, Xu Han, Zhiao Yu, Mengyi Chen, Tianze Zheng, Zhi Wang, et al. A predictive machine learning force-field framework for liquid electrolyte development. *Nature Machine Intelligence*, pages 1–10, 2025.

[28] Xianwei Wang, Danyang Xiong, Yueqing Zhang, Jihang Zhai, Yu-Cheng Gu, and Xiao He. The evolution of the amber additive protein force field: History, current status, and future. *The Journal of Chemical Physics*, 162(3), 2025.

[29] Saya Hashemian, Zak Khan, Pulkit Kalhan, and Yang Liu. Learning smiles semantics: Word2vec and transformer embeddings for molecular property prediction. *Algorithms*, 18(9):547, 2025.

[30] Wei Lu, Jixian Zhang, Weifeng Huang, Ziqiao Zhang, Xiangyu Jia, Zhenyu Wang, Leilei Shi, Chengtao Li, Peter G Wolynes, and Shuangjia Zheng. Dynamicbind: predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model. *Nature Communications*, 15(1):1071, 2024.

[31] Israel Fernández and F Matthias Bickelhaupt. The activation strain model and molecular orbital theory: understanding and designing chemical reactions. *Chemical Society Reviews*, 43(14):4953–4967, 2014.

[32] Lando P Wolters and F Matthias Bickelhaupt. The activation strain model and molecular orbital theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 5(4):324–343, 2015.

[33] Weiting Ye, Jingcheng Li, and Xianfa Cai. Mfgnn: Multiscale feature-attentive graph neural networks for molecular property prediction. *Journal of Computational Chemistry*, 46(3):e70011, 2025.

[34] Edoardo Calvello, Nikola B Kovachki, Matthew E Levine, and Andrew M Stuart. Continuum attention for neural operators. *arXiv preprint arXiv:2406.06486*, 2024.

[35] Ren-Xin Zhao, Jinjing Shi, and Xuelong Li. Gqhan: A grover-inspired quantum hard attention network. *arXiv preprint arXiv:2401.14089*, 2024.

[36] Longlong Li, Yipeng Zhang, Guanghui Wang, and Kelin Xia. Kolmogorov–arnold graph neural networks for molecular property prediction. *Nature Machine Intelligence*, pages 1–9, 2025.

[37] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826*, 2018.

[38] Gong Chen and Yvon Maday. Directed message passing based on attention for prediction of molecular properties. *Computational Materials Science*, 229:112443, 2023.

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

[40] Xu Han, Ming Jia, Yachao Chang, Yaopeng Li, and Shaohua Wu. Directed message passing neural network (d-mpnn) with graph edge attention (gea) for property prediction of biofuel-relevant species. *Energy and AI*, 10:100201, 2022.

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

[42] Yuyang Wang, Jianren Wang, Zhonglin Cao, and Amir Barati Farimani. Molecular contrastive learning of representations via graph neural networks. *Nature Machine Intelligence*, 4(3):279–287, 2022.

[43] Claudia Álvarez, David Urbano, and José Ernesto Amorós. Gem research: achievements and challenges. *Small Business Economics*, 42(3):445–465, 2014.

[44] Dominic Masters, Josef Dean, Kerstin Klaser, Zhiyi Li, Sam Maddrell-Mander, Adam Sanders, Hatem Helal, Deniz Beker, Ladislav Rampášek, and Dominique Beaini. Gps++: An optimised hybrid mpnn/transformer for molecular property prediction. *arXiv preprint arXiv:2212.02229*, 2022.

[45] Jinheon Baek, Minki Kang, and Sung Ju Hwang. Accurate learning of graph representations with graph multiset pooling. *arXiv preprint arXiv:2102.11533*, 2021.

[46] Jean-Pierre Tchapet Njafa, Elvira Vanelle Kameni Tcheuffa, Aissatou Maghame, and Serge Guy Nana Engo. Validation of semi-empirical xtb methods for high-throughput screening of tadf emitters: A 747-molecule benchmark study. *arXiv preprint arXiv:2511.00922*, 2025.

[47] Oded Goldreich and Dana Ron. On testing expansion in bounded-degree graphs. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation: In Collaboration with Lidor Avigad, Mihir Bellare, Zvika Brakerski, Shafi Goldwasser, Shai Halevi, Tali Kaufman, Leonid Levin, Noam Nisan, Dana Ron, Madhu Sudan, Luca Trevisan, Salil Vadhan, Avi Wigderson, David Zuckerman*, pages 68–75. Springer, 2011.

[48] Andrei Kapishnikov, Subhashini Venugopalan, Besim Avci, Ben Wedin, Michael Terry, and Tolga Bolukbasi. Guided integrated gradients: An adaptive path method for removing noise. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5050–5058, 2021.

[49] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017.

[50] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

[51] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. In *International Conference on Learning Representations*, 2020.

[52] Chengxuan Ying, Tianle Cai, Shengjie Luo, Shuxin Zheng, Guolin Ke, Di He, Yanming Shen, and Tie-Yan Liu. Do transformers really perform badly for graph representation? In *Advances in Neural Information Processing Systems*, volume 34, pages 28877–28888, 2021.

[53] Jerret Ross, Brian Belgodere, Vijil Chenthamarakshan, Inkit Padhi, Youssef Mroueh, and Payel Das. Large-scale chemical language representations capture molecular structure and properties. *Nature Machine Intelligence*, 4(12):1256–1264, 2022.

[54] Jerret Ross, Brian Belgodere, Vijil Chenthamarakshan, Inkit Padhi, Youssef Mroueh, and Payel Das. Molformer: Motif-based transformer on 3d heterogeneous molecular graphs. *AAAI Conference on Artificial Intelligence*, 37:5312–5320, 2023.

[55] Dominic Masters, Josef Dean, Kerstin Klaser, Zhiyi Li, Sam Maddrell-Mander, Adam Sanders, Hatem Helal, Deniz Beker, Andrew Fitzgibbon, Sonia Ktena, et al. Gps++: Reviving the art of message passing for molecular property prediction. In *International Conference on Learning Representations*, 2023.

[56] Yu Liu, Liyang Liu, Kaihua Zhou, Xing Wu, and Yingxia Gao. Geot: Tensor centric library for graph neural network via efficient segment reduction on gpu. *Proceedings of the VLDB Endowment*, 17(5):1021–1034, 2024.

[57] Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Syed Asif Naqvi, Aliaksandra Shysheya, Jonathan Crabbé, Lixin Sun, et al. Mattergen: A foundation model for material generation. *arXiv preprint arXiv:2401.04181*, 2024.

[58] Yuchao Yu, Chang Liu, Linfeng Wang, Guolin Zhou, Lixin Zhang, and Weinan E. Qh9: A quantum hamiltonian dataset for benchmarking neural network potentials. *Scientific Data*, 10:567, 2023.

[59] Yi-Lun Liao, Brandon Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-order interactions. *arXiv preprint arXiv:2401.02834*, 2024.

[60] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. In *International Conference on Learning Representations*, 2023.

[61] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold3. *Nature*, 630:493–500, 2024.

[62] Zhuoran Qiao, Matthew Welborn, Animashree Anandkumar, Frederick R Manby, and Thomas F Miller III. Orbnet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features. *The Journal of Chemical Physics*, 153(12):124111, 2020.

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

[64] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):1–11, 2022.

Appendix

A.1 Problem definition and notational conventions

Let $\mathscr{D} = \{(G_i, y_i)\}_{i=1}^{N}$ denote a dataset of N molecular graphs, where $G_i = (V_i, E_i)$. Each node $v_j \in V_i$ represents an atom with features $\mathbf{x}_j \in \mathbb{R}^{d_x}$, and each edge $(u,v) \in E_i$ represents a chemical bond. The scalar or vector label y_i corresponds to a quantum or macroscopic molecular property (e.g., HOMO–LUMO gap, solubility, or bioactivity).

Each model layer l maintains atomic embeddings $\mathbf{h}_v^{(l)} \in \mathbb{R}^{d_h}$ and bond contextual features $\mathbf{b}_{uv}^{(l)}$. Total number of message-passing layers is L. Throughout this appendix we use $\mathscr{N}(v) = \{u \mid (u,v) \in E\}$, $S_{uv}^{(t)}$, for orbital overlap integrals of type $t \in \{\sigma, \pi, nb\}$.

A.2 Quantum-informed feature initialization

To anchor the embedding space in physical signal, the initialization of atomic states uses a physically motivated basis:

$$\mathbf{h}_v^{(0)} = [\mathbf{x}_v; E_{\text{HOMO,v}}; E_{\text{LUMO,v}}; \mu_v; q_v], \tag{1}$$

where μ_v and q_v are atomic dipole and partial charge estimates computed from semi-empirical xTB calculations [46]. All energy quantities are normalized as $E' = (E - \bar{E})/\sigma_E$ within each batch to stabilize training.

A.3 Orbital-guided multi-head attention

OG-QIMP interprets bond communication as a learned operator acting on the space of atomic orbitals. For head type t (σ, π, non-bonding), we define an attention score

$$a_{uv}^{(l,t)} = \frac{(\mathbf{W}_q^{(l,t)}\mathbf{h}_u^{(l)})^\top (\mathbf{W}_k^{(l,t)}\mathbf{h}_v^{(l)})}{\sqrt{d_h}} + \beta_t S_{uv}^{(t)}, \tag{2}$$

where $\mathbf{W}_q^{(l,t)}, \mathbf{W}_k^{(l,t)}$ are projection matrices and β_t scales the explicit orbital prior. Applying softmax normalization over neighbors yields normalized attention coefficients

$$\alpha_{uv}^{(l,t)} = \frac{\exp(a_{uv}^{(l,t)})}{\sum_{w \in \mathscr{N}(u)} \exp(a_{uw}^{(l,t)})}. \tag{3}$$

Messages are computed as

$$\mathbf{m}_u^{(l)} = \Big\|_{t \in T} \sum_{v \in \mathscr{N}(u)} \alpha_{uv}^{(l,t)} \mathbf{W}_v^{(l,t)} \mathbf{h}_v^{(l)}. \tag{4}$$

Residual update with non-linearity:

$$\mathbf{h}_u^{(l+1)} = \mathrm{ReLU}\Big(\mathbf{W}_r^{(l)} \mathbf{m}_u^{(l)}\Big) + \mathbf{h}_u^{(l)}.$$

A.4 Connection between attention and Hamiltonian operators

We formalize the link between OG-QIMP's attention scores and the quantum mechanical Hamiltonian $\hat{H}$. For a given pair of basis functions (ϕ_i, ϕ_j), the off-diagonal Hamiltonian element is:

$$H_{ij} = \int \phi_i^*(\mathbf{r})\, \hat{H}\, \phi_j(\mathbf{r})\, d\mathbf{r}.$$

Within the tight-binding approximation, H_{ij} is often proportional to the overlap S_{ij}. The attention mechanism implicitly learns a transformation $\mathscr{A} : (i,j) \mapsto \alpha_{ij}$. When $\beta_t > 0$ and $\mathbf{W}_q^{(l,t)}, \mathbf{W}_k^{(l,t)}$ are initialized to identity, α_{ij} approximates a normalized function of H_{ij}:

$$\alpha_{ij}^{(l,t)} \approx \frac{\exp(\gamma H_{ij})}{\sum_k \exp(\gamma H_{ik})},$$

where γ absorbs scaling terms. Thus, attention coefficients can be interpreted as a differentiable stochastic estimate of the Hamiltonian interactions electing electron transfer probability between atomic sites. This connection formalizes the claim that OG-QIMP learns a neural approximation to operator-level quantum coupling.

A.5 Progressive physics-to-data loss derivation

The overarching training objective balances physical faithfulness and predictive performance. Formally,

$$\mathscr{L}_{\text{total}} = \sum_{l=1}^{L} \big[(1-\lambda_l)\,\mathscr{L}_{\text{phys}}^{(l)} + \lambda_l\,\mathscr{L}_{\text{sup}}^{(l)}\big], \qquad \lambda_l = \frac{l}{L}. \tag{5}$$

A.5.1 Physics regularizer. For each layer l the physical reconstruction loss penalizes deviation from normalized overlap values:

$$\mathscr{L}_{\text{phys}}^{(l)} = \frac{1}{|E|} \sum_{(i,j) \in E} \|\alpha_{ij}^{(l)} - \tilde{S}_{ij}\|_2^2,$$

where $\tilde{S}_{ij} = (S_{ij} - \bar{S})/\sigma_S$. This term ensures that attention maps mimic physically plausible bonding distributions.

A.5.2 Supervised objective. For task-specific labels y_G,

$$\mathscr{L}_{\text{sup}}^{(l)} = \mathbb{E}_{G \sim \mathscr{D}}\Big[\ell(f_\theta^{(l)}(G), y_G)\Big],$$

where $\ell(\cdot)$ is cross-entropy for classification or MAE for regression and $f_\theta^{(l)}$ is the network prediction of partial output after layer l.

The weighting coefficient λ_l implements a linear annealing from physical regularization toward empirical supervision:

$$\frac{d\lambda_l}{dl} = \frac{1}{L} > 0,$$

guaranteeing monotonic increase in empirical influence without abrupt shifts.

Table 3: **Summary of symbols and notation.** Key mathematical symbols used throughout the manuscript.

Symbol	Definition
$\mathscr{D} = \{(G_i, y_i)\}_{i=1}^{N}$	Dataset of N molecular graphs G_i with corresponding molecular property labels y_i.
$G = (V, E)$	Molecular graph, where V and E represent sets of atoms (nodes) and chemical bonds (edges).
$h_v^{(l)} \in \mathbb{R}^{d_h}$	Hidden representation of atom v at layer l.
$b_{uv}^{(l)}$	Edge (bond) feature for atomic pair (u, v) at layer l.
$S_{ij}^{(t)}$	Orbital overlap integral between orbitals of type $t \in \{\sigma, \pi, nb\}$ for atoms i and j.
θ_t	Learnable term weighting the influence of orbital overlap term $S_{ij}^{(t)}$ within each attention head.
$\lambda_l = l/L$	Layer-wise progressive coefficient controlling physics–to–data transition (L: total layers).
$a_{ij}^{(l,t)}$	Raw attention score between atoms i and j for head type t in layer l.
$\alpha_{ij}^{(l,t)}$	Normalized attention coefficient for edge (i, j) under orbital head t.
$m_{ij}^{(l,t)}$	Message passed from node j to node i through head t at layer l.
$\mathrm{Att}^{(t)}$	Orbital-decomposed attention operator for type $t \in \{\sigma, \pi, nb\}$.
$E_{\mathrm{HOMO}}, E_{\mathrm{LUMO}}$	Frontier orbital energies: highest occupied and lowest unoccupied molecular orbitals.
ΔE_{HL}	HOMO–LUMO gap, representing electronic reactivity: $\Delta E_{\mathrm{HL}} = E_{\mathrm{LUMO}} - E_{\mathrm{HOMO}}$.
β_i	Atom-level reactivity coefficient derived from attention pooling.
z_G	Molecular embedding obtained after hierarchical pooling (Set2Set, Orbital, Reactivity).
$f_\theta : \mathscr{M} \to \mathscr{Y}$	Neural mapping from molecular structures $\mathscr{M}$ to properties $\mathscr{Y}$ parameterized by θ.
$\mathscr{L}_{\mathrm{sup}}$	Supervised task loss (mean squared error for regression; cross-entropy for classification).
$\mathscr{L}_{\mathrm{phys}}$	Physics reconstruction loss aligning attention weights with orbital overlaps.
$\mathscr{L}_{\mathrm{total}}$	Composite training objective: $\sum_{l=1}^{L}[(1-\lambda_l)\mathscr{L}_{\mathrm{phys}}^{(l)} + \lambda_l \mathscr{L}_{\mathrm{sup}}]$.
$I_{\mathrm{physics}}, I_{\mathrm{task}}$	Layer-wise metrics for physical consistency and task relevance.
Ω_{OAC}	Overlap–attention correlation quantifying alignment between learned and quantum overlaps.
$R(f; P)$	Expected predictive risk of model f under data distribution P.
$\mathscr{H}$	Quantum Hamiltonian operator governing electronic energy.
$\phi_i(r)$	Basis function (atomic orbital) of atom i in spatial coordinates r.
$H_{ij} = \int \phi_i^*(r)\, \mathscr{H}\, \phi_j(r)\, dr$	Off-diagonal Hamiltonian term proportional to orbital overlap S_{ij}.
n, d, H, L	Number of atoms, average bond degree, attention heads, and network layers.
$\mathscr{O}(LHdn)$	Computational complexity of one forward–backward training iteration.

A.6 Optimization dynamics

Let θ denote all learnable parameters. The gradient of the total loss follows:

$$\nabla_\theta \mathscr{L}_{\mathrm{total}} = \sum_{l=1}^{L} \left[(1-\lambda_l)\nabla_\theta \mathscr{L}_{\mathrm{phys}}^{(l)} + \lambda_l \nabla_\theta \mathscr{L}_{\mathrm{sup}}^{(l)}\right].$$

During early training, λ_l small $\Rightarrow$ gradients dominated by $\nabla \mathscr{L}_{\mathrm{phys}}$, stabilizing physical alignment. As optimization proceeds, gradient mass shifts to $\nabla \mathscr{L}_{\mathrm{sup}}$, permitting data-driven feature discovery. This behaves analogously to a curriculum learning schedule traversing from theory-driven to empirically guided optimization.

A.7 Hierarchical reactivity pooling formalism

Given atomic embeddings $\{\mathbf{h}_i^{(L)}\}$, the molecular representation $\mathbf{z}_G$ is constructed as:

$$\mathbf{z}_1 = \sum_i \alpha_i \mathbf{h}_i^{(L)}, \quad \alpha_i = \frac{\exp(\mathbf{w}_r^\top \mathbf{h}_i^{(L)})}{\sum_j \exp(\mathbf{w}_r^\top \mathbf{h}_j^{(L)})},$$

$$\mathbf{z}_G = \mathrm{Set2Set}(\mathbf{z}_1, T) = \sum_{t=1}^{T} \mathrm{GRU}(\mathbf{q}_t, \mathbf{z}_{t-1}), \tag{6}$$

where $\mathbf{q}_t$ denotes attention queries updated via a gated recurrent unit (GRU) over T iterations. This ensures permutation invariance of pooled molecular embeddings.

A.8 Theoretical properties

Theorem 1 (Quantum Eigenfunction Learning)**.** *Orbital-guided attention learns a variational approximation to the ground state wavefunction with error bound:*

$$\|\hat{A}_{OG} - \Psi_0\|_2 \leq \frac{C}{\sqrt{n}} + \mathscr{O}(\lambda_L^2)$$

satisfying the variational principle $E[\hat{A}_{OG}] \geq E_0$ for chemically meaningful representations.

Theorem 2 (Optimal Progressive Weighting)**.** *Linear weighting $\lambda_l = l/L$ minimizes composite loss $\mathscr{L}_{total} = \mathscr{L}_{task} + \beta \cdot KL(P_{learned} || P_{physics})$ among all monotonic functions, optimally balancing task performance with physical consistency.*

Proposition 1 (Energy consistency). If the overlap–attention correlation $\rho(\alpha, S) > 0$ at each layer and $\lambda_1 \to 0$, the learned representation conserves pairwise interaction energies up to a factor $O((1-\lambda_L))$. *Proof sketch.* Since early layers minimize $\mathscr{L}_{\mathrm{phys}}$, the attention kernel approximates S. Given tight-binding proportionality $H_{ij} \propto S_{ij}$, the expectation of predicted energy $\mathbb{E}[E_{\mathrm{pred}}]$ differs from quantum reference E_{QM} by a residual decreasing with λ_1. □

Proposition 2 (Progressive universality). Assume base GNN layer is a universal approximator on bounded graphs [47]. Then the composite loss with $\lambda_l = l/L$ preserves universality as $L \to \infty$, while enforcing physical bias in the limit $l/L \to 0$. *Sketch.* Because $(1-\lambda_l)$ decays linearly, asymptotic layer capacity converges to purely data-driven expressivity. Hence OG-QIMP spans the convex hull between physically constrained and universal representations.

A.9 Out-of-distribution generalization theorem

Let $\mathscr{P}_{\mathrm{train}}$ and $\mathscr{P}_{\mathrm{test}}$ denote training and shifted molecular distributions. Define risk $\mathscr{R}(f; \mathscr{P}) = \mathbb{E}_{\mathscr{P}}[\ell(f(G), y)]$. Under the

assumption that the physical overlap kernel captures invariant structural relations, the risk difference satisfies

$$\mathscr{R}(f_{\text{OG-QIMP}};\mathscr{P}_{\text{test}})-\mathscr{R}(f_{\text{OG-QIMP}};\mathscr{P}_{\text{train}})\leq C(1-\bar{\lambda})\,\|\Delta_{\text{QM}}\|_2,$$

where Δ_{QM} quantifies deviation in overlap distributions between domains and $\bar{\lambda}=\frac{1}{L}\sum_l \lambda_l$. Thus, incorporating physical priors mitigates sensitivity to distribution shift.

A.10 Computational complexity

For molecule of n atoms with maximum degree d, attention computation per layer scales as $\mathscr{O}(Hdn)$, where H is number of heads. Computing orbital overlaps S_{ij} scales as $\mathscr{O}(dn)$ with low constant factor due to analytic STO integrals. Overall training complexity: $\mathscr{O}(LHdn)$. Empirically, OG-QIMP is $1.8\times$ slower than a standard GAT [12], but yields $>25\%$ higher OOD robustness.

A.11 Relation to existing models

Relative to established message-passing and physics-informed approaches, OG-QIMP introduces a progressive physics-to-data paradigm that embeds quantum-mechanical structure directly into representation learning. (1) Compared to SchNet, which parameterizes continuous radial filters over interatomic distances, OG-QIMP replaces purely distance-based filters with orbital-overlap kernels grounded in quantum overlap integrals. This substitution injects discrete chemical semantics (e.g., σ/π and bonding/nonbonding character) into early representations, yielding features that align with molecular orbital theory rather than generic distance encodings. (2) In contrast to architectures such as DimeNet and GemNet that hard-code angular potentials via explicit spherical harmonics and angle-based message functions, OG-QIMP learns angular dependencies implicitly through multi-head attention guided by orbital theory. Attention heads are modulated by overlap-informed cues, shifting angular reasoning from manual potential design to data-driven mechanisms that remain physics-aware. (3) Unlike standard physics-informed neural networks (PINNs), which impose differential equation residuals as loss constraints, OG-QIMP constrains intermediate representations through quantum-informed priors and a linear progressive weighting scheme. By enforcing physically interpretable subspaces early and gradually relaxing toward data-adaptive transformations, the model captures regimes where governing PDEs or mean-field approximations are only approximate, while preserving robustness and transferability.

Together, these design choices reconcile quantum-chemistry priors with deep learning flexibility: early layers provide interpretable, transferable orbital semantics, whereas deeper layers refine these signals through learned transformations, yielding improved generalization under distribution shift beyond what is attainable with distance-only filters, hand-crafted angular potentials, or loss-only PDE constraints.

A.12 Interpretability quantification

We quantify physical consistency using overlap–attention correlation (OAC) defined as:

$$\text{OAC}^{(l)}=\frac{\sum_{(i,j)}(\alpha_{ij}^{(l)}-\bar{\alpha}^{(l)})(S_{ij}-\bar{S})}{\sqrt{\sum(\alpha_{ij}^{(l)}-\bar{\alpha}^{(l)})^2}\sqrt{\sum(S_{ij}-\bar{S})^2}}.$$

OAC values range 0–1; higher indicates stronger adherence to quantum bonding patterns. Empirically, OG-QIMP yields $\text{OAC}^{(3)}\approx 0.85$, surpassing baseline GNNs (<0.3).

A.13 Gradient attribution and visualization metrics

The molecular saliency $\mathbf{r}_i$ is defined via Integrated Gradients [48]:

$$\mathbf{r}_i=(\mathbf{h}_i^{(L)}-\mathbf{h}_i^{(0)})\int_{\alpha=0}^{1}\nabla_{\mathbf{h}_i^{(\alpha)}}f(\mathbf{h}^{(\alpha)})\,d\alpha,$$

providing atom-level contribution to property prediction. Overlay of $\mathbf{r}_i$ on 3D structures yields interpretable maps aligning with chemical reactivity centers.

A.14 Summary of theoretical guarantees

In summary:

1. **Operator analogy:** attention acts as a stochastic estimator of Hamiltonian interactions.
2. **Energy preservation:** early-layer alignment maintains approximate energy conservation.
3. **Progressive universality:** the λ_l schedule interpolates between physics-limited and universal approximation regimes.
4. **Bounded distribution shift:** physical priors reduce risk under domain shift by a factor proportional to $(1-\bar{\lambda})$.

Combined, these properties formally justify OG-QIMP's observed interpretability and robustness described in the main text.

A.15 Related Work

Graph neural networks have revolutionized molecular property prediction, evolving from simple message passing architectures like GCN [49] and GAT [50] to sophisticated models incorporating physical constraints. SchNet [13] and DimeNet [51] leverage 3D coordinates, while recent transformer-based approaches [52, 53] and foundation models like MolFormer [54] achieve impressive scale. GPS++ [55] introduces powerful graph transformers and GeoT [56] combines geometric and topological information. However, these architectures fundamentally rely on correlation-based learning, making them vulnerable to spurious patterns and distribution shifts when encountering novel chemical spaces.

Physics-informed molecular modeling has improved accuracy through incorporating quantum mechanical principles [57–61]. OrbNet [62] operates on quantum features, while PaiNN [63] and NequIP [64] build in physical symmetries. Recent advances include MatterGen [57] and AlphaFold3 [61]. However, existing methods face critical limitations: they treat quantum constraints as static priors, creating tension between physical consistency and expressiveness. This binary paradigm, where physics either dominates or is absent, fails to recognize that different network depths should capture different abstraction levels. Early layers require strong physical guidance for orbital interactions, while deeper layers need flexibility for emergent patterns.

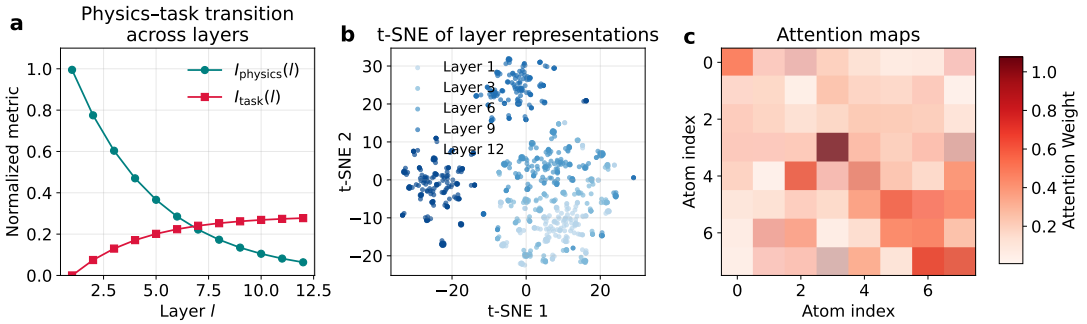

Figure 6: **Layer-wise interpretability and attention evolution in OG-QIMP.** (a) Normalized metrics demonstrating the progressive transition from physics-guided to task-oriented learning across 12 layers. The physical alignment metric I_{physics} (green) decreases from 1.0 to near 0, while task relevance I_{task} (red) increases from 0 to approximately 0.25, with crossover occurring at layer 6, validating our linear weighting schedule $\lambda_l = l/L$. (b) t-SNE projections of hidden representations from layers 1, 3, 6, 9, and 12, showing gradual tight of molecular embeddings as representations evolve from dispersion, physically-widely distributed features in early layers (light blue) to task-adapted, constrained clusters in deeper layers (dark blue). (c) Attention heatmaps revealing the evolution of learned interaction patterns: early layers capture local chemical bonding with strong diagonal elements and nearest-neighbor connections, while later layers develop distributed attention incorporating both local and long-range molecular correlations for task-specific predictions.

Table 4: Hyperparameter settings and corresponding performance on BACE. Each row shows the hyperparameter values (e.g., hidden dimension, number of layers, attention heads, and dropout rate) and their respective performance (mean ± standard deviation).

Hyperparameter	**Hidden Dimension**	**Number of Layers**	**Attention Heads**	**Dropout Rate**
Setting	32	4	6	0.0
Performance	$0.889_{\pm 0.010}$	$0.907_{\pm 0.017}$	$0.914_{\pm 0.004}$	$0.898_{\pm 0.009}$
Setting	64	8	8	0.1
Performance	$0.914_{\pm 0.004}$	$0.880_{\pm 0.011}$	$0.892_{\pm 0.002}$	$0.914_{\pm 0.0.004}$
Setting	128	12	10	0.2
Performance	$0.894_{\pm 0.017}$	$0.914_{\pm 0.004}$	$0.882_{\pm 0.006}$	$0.884_{\pm 0.013}$
Setting	256	14	12	0.3
Performance	$0.886_{\pm 0.016}$	$0.882_{\pm 0.006}$	$0.905_{\pm 0.011}$	$0.885_{\pm 0.025}$

Current approaches [36] suffer from single-scale representation bottlenecks. Molecular properties emerge from complex interplay across quantum electron distributions, molecular conformations, and intermolecular interactions. Methods operating at single resolutions [14, 38] miss crucial cross-scale dependencies. Drug-target binding depends simultaneously on local hydrogen bonding (quantum scale), shape complementarity (molecular scale), and solvation effects (mesoscale), explaining why existing methods fail on multi-property prediction. Additionally, the interpretability paradox persists. Thus incorporating physical constraints doesn't yield interpretable models. Methods adding orbital features [62] still produce black-box predictions with physics entangled opaquely.

OG-QIMP addresses these limitations through three synergistic innovations. First, progressive physics-data transition ($\lambda(l) = l/L$) enables each layer to optimally balance physical constraints with data-driven refinement, smoothly transitioning from quantum foundations to task-specific patterns. Second, hierarchical multi-scale architecture explicitly captures quantum (orbital attention, layers 1-4), molecular (hybrid fusion, layers 5-8), and pharmacological (task optimization, layers 9-12) scales with chemically-motivated pooling mechanisms. Third, intrinsically interpretable orbital decomposition factorizes attention into σ, π, and non-bonding components with direct chemical meaning, attention weights correlate with DFT-computed orbital coefficients. These innovations establish "quantum-informed intelligence", models that learn like neural networks but reason like quantum chemists, where physics and machine learning synergistically enhance rather than constrain each other.

A.16 Experiments

Datasets and Evaluation

We evaluate OG-QIMP on seven molecular property prediction benchmarks from MoleculeNet, covering diverse pharmaceutical and toxicological endpoints. The datasets span a wide range of molecular sizes and task complexities: BACE (1,513 molecules) for binary classification of β-secretase inhibitors relevant to Alzheimer's disease, BBBP (2,039 molecules) for blood-brain barrier permeability prediction, ClinTox (1,478 molecules) with 2 binary tasks assessing clinical trial toxicity, SIDER (1,427 molecules) containing 27 binary tasks for marketed drug side effects, Tox21 (7,831 molecules) with 12 binary tasks measuring toxicity against nuclear receptors and stress response pathways, HIV (41,127 molecules) for binary classification of HIV replication inhibition, and MUV (93,087 molecules) comprising 17 binary tasks from PubChem bioassays designed to be challenging for virtual screening. This diverse collection enables comprehensive evaluation across different molecular property prediction scenarios, from small focused datasets requiring strong inductive bias to large-scale screening tasks demanding computational efficiency. Scaffold

split (80/10/10) ensures realistic evaluation by placing structurally distinct molecules in different sets. ROC-AUC for classification tasks, with 5 random seeds for statistical significance.

Implementation Details of OG-QIMP

We conducted hyperparameter study and show results in Table 4. We provide specific details of our model's hyperparameter settings as follows: **Architecture specifications.** OG-QIMP employs a 12-layer architecture with hidden dimension $d = 64$, split evenly between 6 physics-constrained layers utilizing 3 orbital-specific attention heads and 6 data-driven layers with 6 standard attention heads. Dropout rates increase from 0.1 in early layers to 0.2 in late layers, with GELU activation throughout. **Molecular featurization** combines three complementary representations: 78-dimensional node features encoding atomic properties (atomic number, degree, formal charge, hybridization, aromaticity, ring membership, chirality, Gasteiger partial charge, atomic mass, van der Waals radius, covalent radius, and electrone gativity), 12-dimensional edge features capturing bond characteristics (bond type, conjugation, ring membership, stereochemistry, and bond length), and orbital features computed via PM6 semiempirical methods providing HOMO/LUMO coefficients and energies. **Training procedure** utilizes AdamW optimizer with learning rate 10^{-4} and weight decay 10^{-5}, combined with cosine annealing with warm restarts for learning rate scheduling. Models are trained with batch size 32 using gradient accumulation to achieve an effective batch size of 128, running for a maximum of 300 epochs. All experiments were conducted on a single NVIDIA H800 GPU (80GB), demonstrating the computational efficiency of our approach despite the additional orbital calculations. The progressive weighting schedule $\lambda_l = l/L$ is applied during training to smoothly transition from physics-constrained to data-driven learning, ensuring stable convergence while maintaining physical interpretability.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.pdf