SymQNet: VQE Hamiltonian Estimation for Molecular Optimization with Quantum Computing

Significantly Faster Drug Candidate Screening

Yash V Tomar

West Lafayette Jr/Sr High School, Indiana

The Real-World Stakes: Drug Discovery is Broken

$2.6B

Average Drug Cost[1]

Average cost to bring a drug to market...much of it on wasted candidates.

12yr

Time to Market[2]

Average Development Timeline from discovery to approval, delaying life-saving treatments.

90%

Failure Rate[3]

Clinical trial failure rate, with most candidates never reaching patients.

Delayed treatment has been proven to increase mortality[4]. Faster <u>molecular optimization (MO)</u> can significantly speed up drug candidate screening.

One Reason is that Molecules are Hard to Simulate

Fundamentals Behinds MO

Born-Oppenheimer approximation implies the most **stable structure** of a structure lies at the **lowest** ground-state energy[5]

Models such as DFT attempt to minimize this energy to perform molecular optimization[5].

MO methods have limits in accuracy[5,6]

Struggle with weak *intermolecular forces* (e.g., London-Dispersion)[7]

Solution: Quantum MO (QMO)

Uses **Variational Quantum Eigensolver** (VQE) to approximate wavefunction better.[8,9]

VQE is a Solution

Thousand-year problems *potentially* solved in **days[10].**

- Hamiltonian is the operator that stores a structure's energy[11]

Calculating Hamiltonian Energy is Impossible

Mathematically impossible to diagonalize & directly calculate the energy of a structure[11]

Estimating Energy is "Next Best Bet"

VQE[8,9] uses various methods to try estimate the energy on Hamiltonian

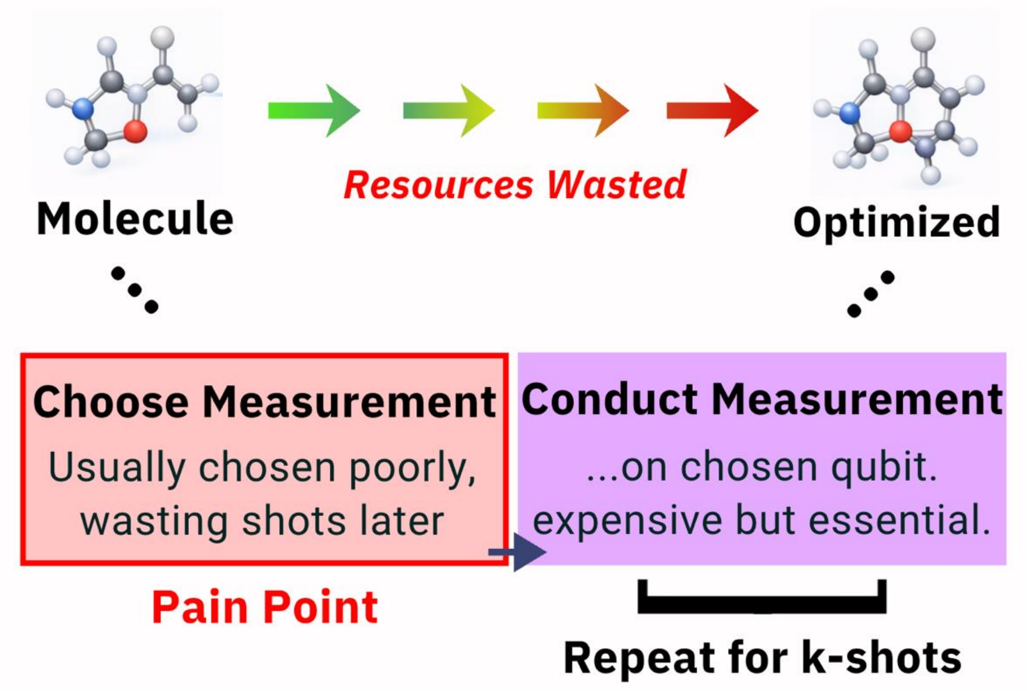
Estimation Requires Slow Measurements

To be accurate, methods must continuously measure parameters (potentially **millions of times**) to reduce variation[11,12]

***Choosing* measurements becomes important**

Optimally choosing what parameter to measure could speed up MO

Hypothesis is Based on Robust Reinforcement Learning



Hypothesis: If canonical information in each measurement shot can be maximized using reinforcement learning, then the number of estimation shots for a Hamiltonian's energy can be significantly reduced because maximizing information reduces uncertainty in the energy estimate more efficiently, allowing convergence with fewer measurements.

Objectives

1. Create **robust RL design**, validate via ablations
2. Show downstream impact on **real** molecular optimization tasks.

Information Gain Reward using Entropy

$$r_t = \max(0,\ H_{t-1} - H_t)$$

Agent Goal: Learn a **strategy** to choose which measurements **give the most information** to estimate energy **faster.**

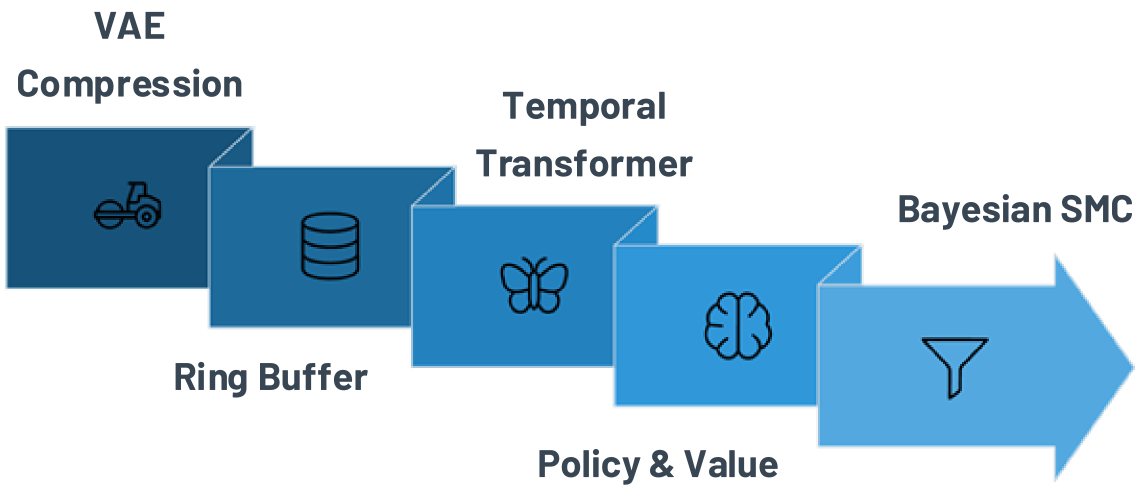
SymQNet Architecture is Based on Markov-Decision Process (MDP)

Bayesian Sequential Monte Carlo (SMC)

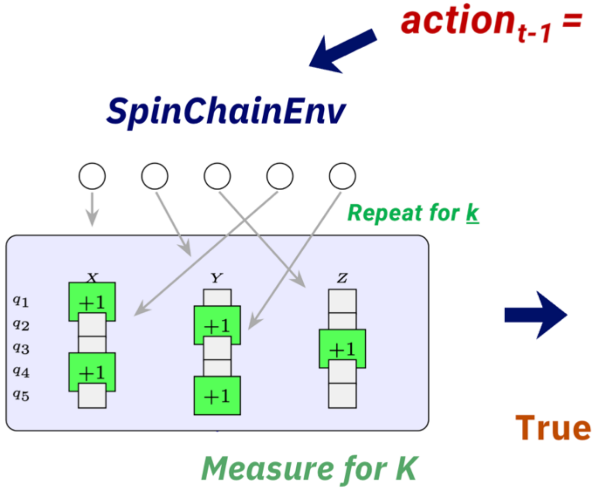
A Bayesian method that **tracks uncertainty** and **estimates** parameters[16]. SymQNet guides it using its uncertainty metric.

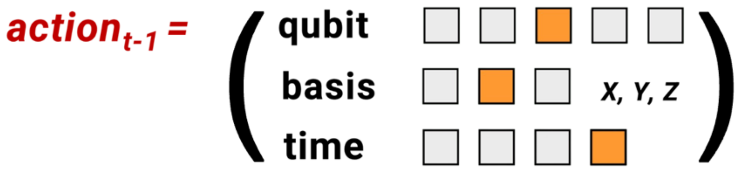
> Since measurements are SymQNet's observation, the MDP must be framed as **partially-observable**; measurements are **indirect signals** of the underlying state.

Full Architecture Pipeline



Action Space and Observation





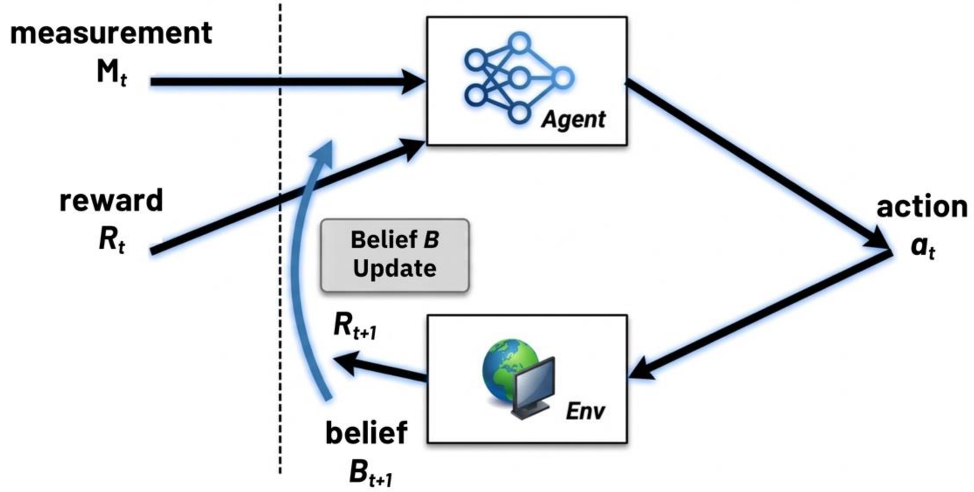
Observation= Mean Measurement=m_t

`[ 0.00, 0.00, -0.37, 0.00, 0.00 ]`

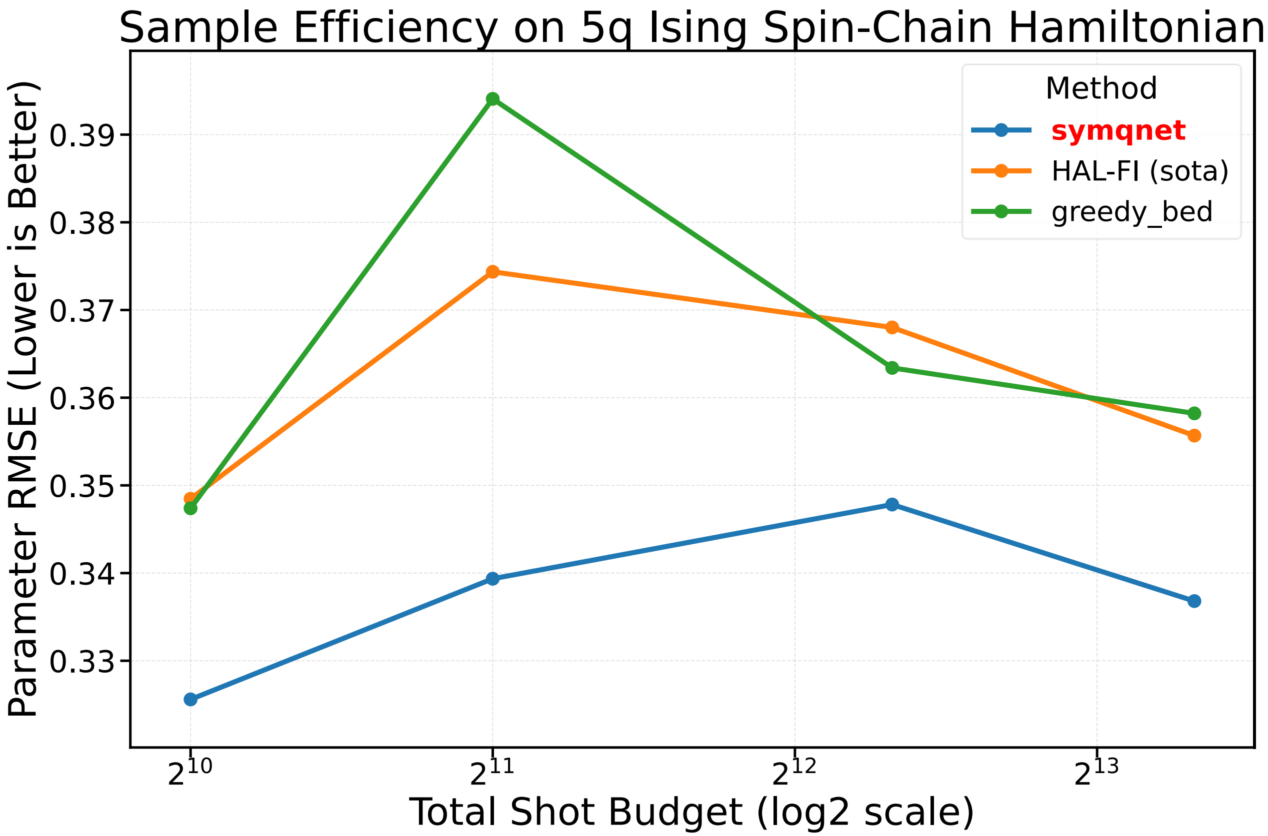
True Hamiltonian Parameters=(fixed) hidden state s_t

***Internal* Belief Update=internal transition**

Partially-Observable MDP (POMDP) Formulation



SymQNet Offers Decisive Win Over State-of-the-Art



Decisive Win Over SOTA

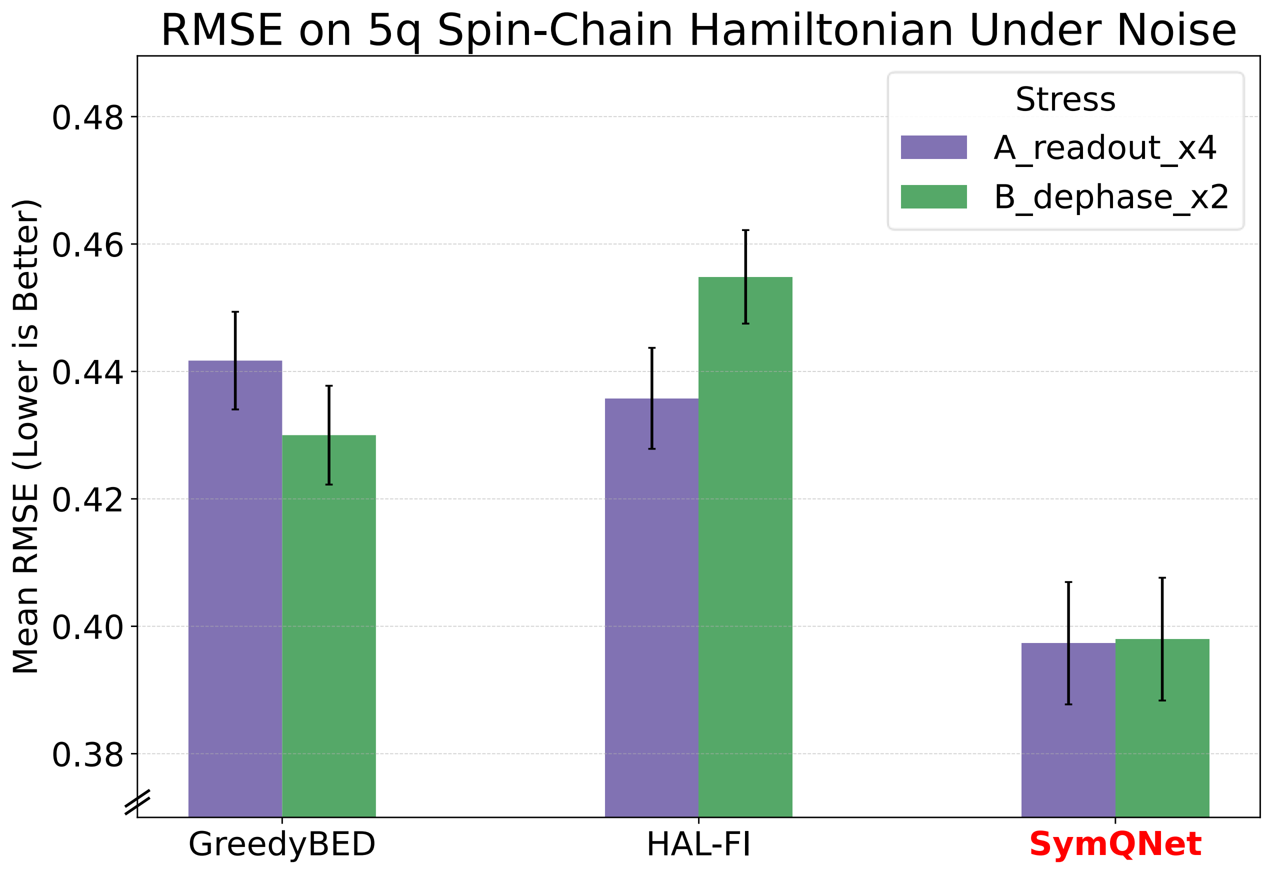
SymQNet achieves **lower Parameter RMSE** than HAL-FI (SOTA) and GreedyBED across all shot budgets (2^{10}–2^{13}).

Statistically significant: **$p < 0.05$** via paired-Wilcoxon test (Holm-corrected)

- Softens exponential error growth in downstream Hamiltonian use

Same budget. Better results.

Noise Stress-Test Supports Hypothesis



Robust Under Realistic Noise

SymQNet outperforms both GreedyBED and HAL-FI under **readout noise (4×)** and **dephasing noise (2×)**.

- Statistically significant (±2 SEM)

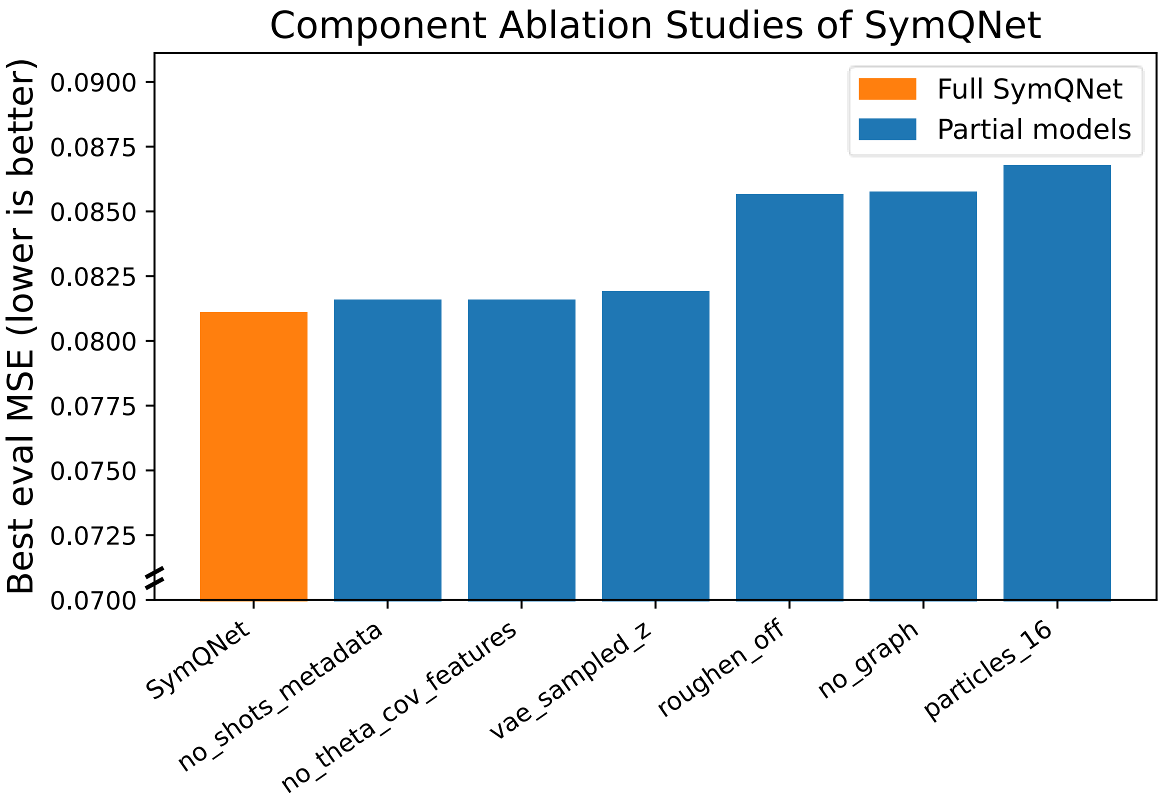
$p < 0.05$ via paired-Wilcoxon test (Holm-corrected)

Gains maintained under realistic hardware conditions: **ready for real-world deployment (see below).**

Deployment: `pip install symqnet-molopt`

Ablation Studies Show Maximal Component Utilization

*Provides empirical **evidence** SymQNet **learned**.*



$p < 0.05$ (SymQNet vs Partials) by paired-t-tests

Ablation studies systematically remove or degrade each component of SymQNet to verify its contribution. Across 30 seeds, **SymQNet (Full) achieves the lowest MSE**, confirming that every design choice (metadata, graph structure, VAE compression, particle count) contributes meaningfully to performance. Helps understand model.

No Metadata

Removing metadata injection degrades performance (context matters!).

No Graph

Removing qubit graph structure hurts as spatial relationships are informative.

Reduced Particles

Fewer SMC particles (1.6×) reduces Bayesian estimation quality.

Molecular Application Show Significant Speedup

46.7%

Shot Reduction

Fewer shots needed to reach success threshold

4/4

Molecules Won

Better solution on Pyridine, Imidazole, Urea & Ising ($p < 0.05$)

21.2%

MAE Reduction

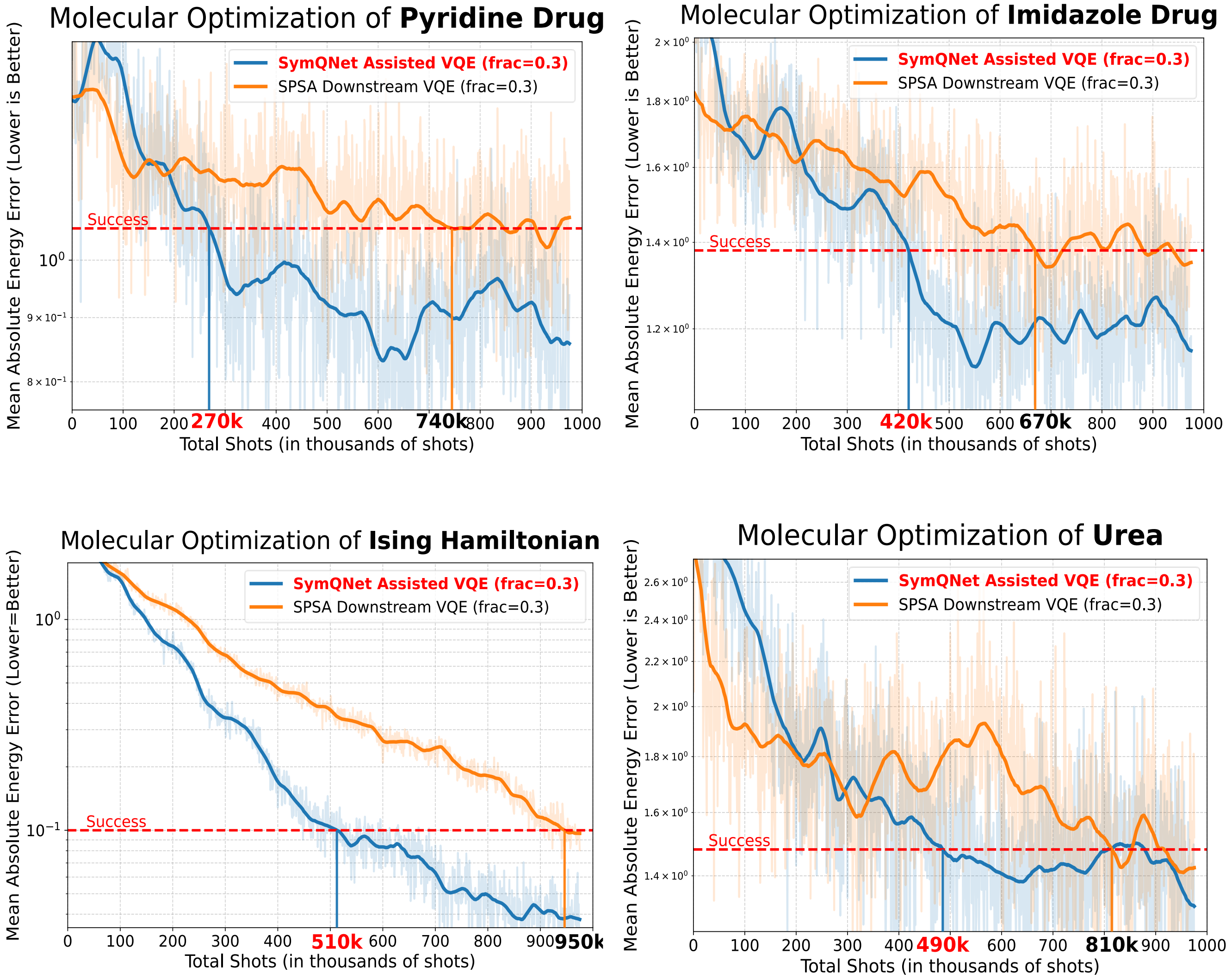
Better mean energy error reduction vs. baseline

1.88x

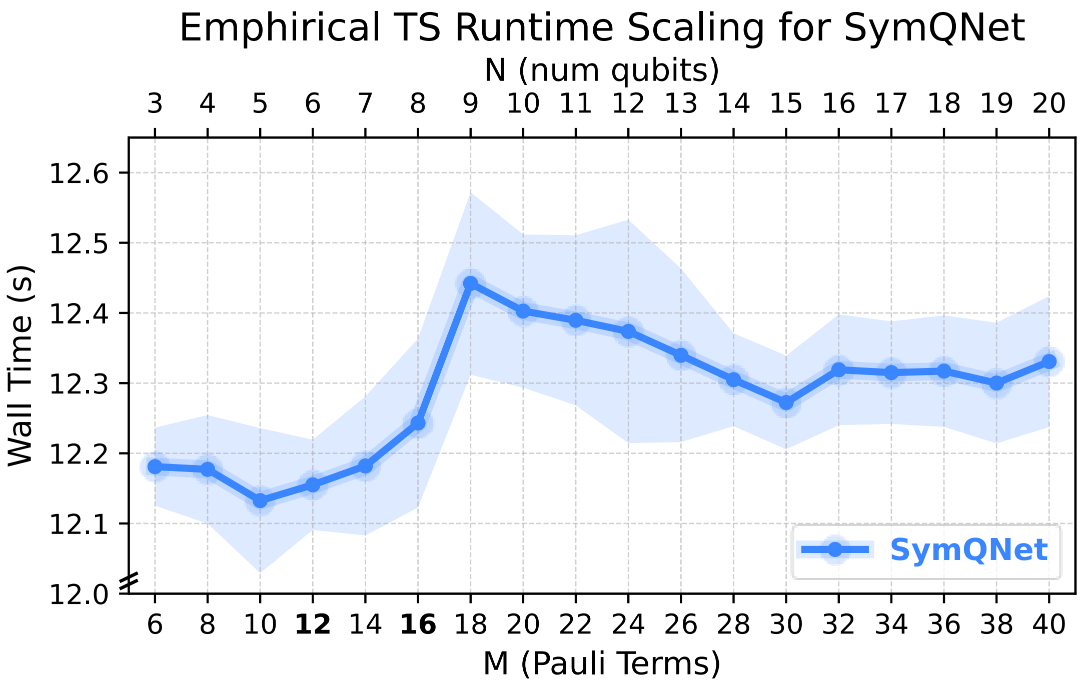
Drug Discovery Speedup

More molecules screened under the same budget

Regular VQE used standard Pauli-Term energy estimation.



Empirical Runtime Scale Tests Show O(1) Scaling



Why Should I Care?

Cost doesn't explode as molecules grow larger!

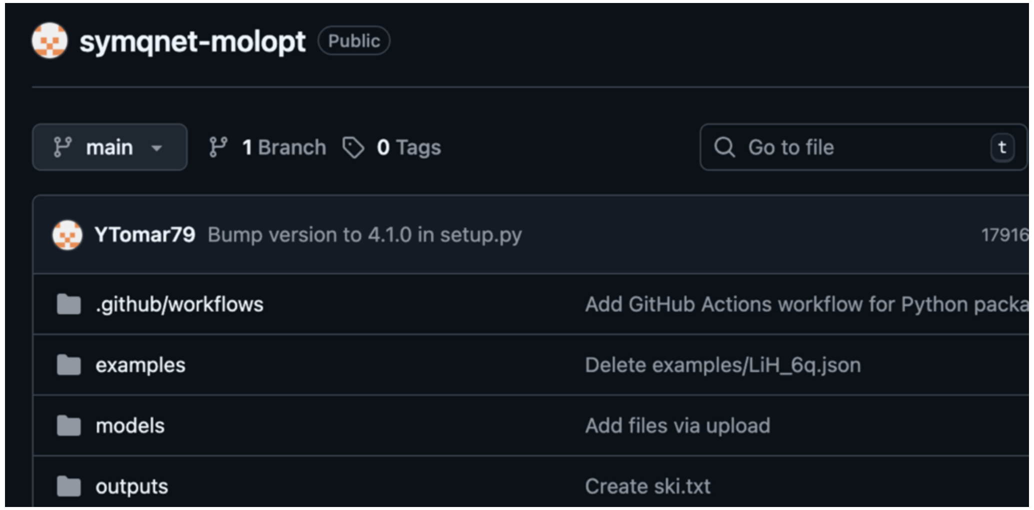
Reference[17]: Baseline methods scale as $O(N^4)$ and $O(M)$

How can SymQNet be accessible?

Accessible in One Line

```
pip install symqnet-molopt
```

Open-source Repository



SymQNet is currently **patent-pending**

Hypothesis Supported and Speedup Established

From the results, the hypothesis appears to be supported!

> *Maximizing information from each shot with RL can lead to significant speedup of MO on VQE!*

🚀 **Drug Discovery Speedup**

1.88× more molecules screened per budget. 46.7% shot reduction. 21.2% better energy error.

🏆 **Beats state-of-the-art**

Lower RMSE than HAL-FI & GreedyBED under same shot budget and under quantum noise ($p < 0.05$).

⚡ **Empirical O(1) Scaling**

Improves from $O(N^4)$, $O(M)$ to $O(1)$ (empirically confirmed).

Next Steps for SymQNet

- Train with **Higher Dimensional Hamiltonians**
- **Real-Life Deployment** of SymQNet.

My Long-Term Vision

- **Faster drug development** with lower costs
- **Better drugs** with lower clinical failure rates.

Cited References

- [1] J. A. DiMasi, H. G. Grabowski, and R. W. Hansen, “Innovation in the pharmaceutical industry: New estimates of R&D costs,” Journal of Health Economics, vol. 47, pp. 20–33, 2016.
- [2] A. Riemer and V. Freund, “Generative artificial intelligence in pharmaceutical drug development: A systematic review of time and cost efficiency across discovery, preclinical, and clinical phases,” Intelligent Pharmacy, 2025. https://doi.org/10.1016/j.ipha.2025.12.006
- [3] PK. Smietana, M. Siatkowski, and M. Møller, “Trends in clinical success rates,” Nature Reviews Drug Discovery, vol. 15, pp. 379–380, 2016. doi:10.1038/nrd.2016.85.
- [4] T. P. Hanna et al., “Mortality due to cancer treatment delay: Systematic review and meta-analysis,” BMJ, vol. 371, Art. no. m4087, 2020, doi: 10.1136/bmj.m4087.
- [5] Y. Cao et al., “Quantum chemistry in the age of quantum computing,” Chem. Rev., vol. 119, no. 19, pp. 10856–10915, 2019, doi: 10.1021/acs.chemrev.8b00803.
- [6] G. Schneider and U. Fechner, “Computer-based de novo design of drug-like molecules,” Nat. Rev. Drug Discov., vol. 4, no. 8, pp. 649–663, 2005, doi: 10.1038/nrd1799.
- [7] A. Yoshimori and J. Bajorath, “Computational analysis, alignment and extension of analogue series from medicinal chemistry,” Future Science OA, vol. 8, no. 7, 2022, doi:10.2144/fsoa-2022-0033.
- [8] A. Peruzzo et al., “A variational eigenvalue solver on a photonic quantum processor,” Nat. Commun., vol. 5, Art. no. 4213, 2014, doi: 10.1038/ncomms5213.
- [9] A. Ralli, P. J. Love, A. Tranter, and P. V. Coveney, “Implementation of measurement reduction for the variational quantum eigensolver,” Phys. Rev. Res., vol. 3, Art. no. 033195, 2021, doi: 10.1103/PhysRevResearch.3.033195.
- [10] J. Tilly *et al.*, “The variational Quantum Eigensolver: A review of methods and best practices,” *Physics Reports*, vol. 986, pp. 1–128, Nov. 2022. doi:10.1016/j.physrep.2022.08.003
- [11] A. Dutt et al., “Active learning of quantum system Hamiltonians yields query advantage,” Phys. Rev. Res., vol. 5, Art. no. 033060, 2023, doi: 10.1103/PhysRevResearch.5.033060.
- [12] K. I. O. Ben ’Attar, D. Farfurnik, and N. Bar-Gill, “Hamiltonian engineering of general two-body spin-1/2 interactions,” Phys. Rev. Res., vol. 2, Art. no. 013061, 2020, doi: 10.1103/PhysRevResearch.2.013061.
- [13] H. Singh, S. Majumder, and S. Mishra, “SHARC-VQE: Simplified Hamiltonian Approach with Refinement and Correction enabled Variational Quantum Eigensolver for Molecular Simulation,” arXiv:2407.12305 [quant-ph], Jul. 2024. doi:10.48550/arXiv.2407.12305.
- [14] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” J. Artif. Intell. Res., vol. 4, pp. 237–285, 1996, doi: 10.1613/jair.301.
- [15] J. Schulman et al., “Proximal policy optimization algorithms,” arXiv:1707.06347, 2017, doi: 10.48550/arXiv.1707.06347.
- [16] A. Doucet, N. de Freitas, and N. Gordon, Eds., Sequential Monte Carlo Methods in Practice. New York, NY, USA: Springer, 2001, doi: 10.1007/978-1-4757-3437-9.
- [17] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, “Measurement optimization in the variational quantum eigensolver using a minimum clique cover,” arXiv preprint arXiv:1907.03358, 2019.

Thank you for reading through my work! I wanted to express my passion for chemistry and CS with SymQNet. I look forward to discussions at the fair!