TensorRL-QAS

Reinforcement learning with tensor networks for improved quantum architecture search

Akash Kundu

Stefano Mangini

NEURAL INFORMATION PROCESSING SYSTEMS

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

2025

TABLE OF CONTENTS

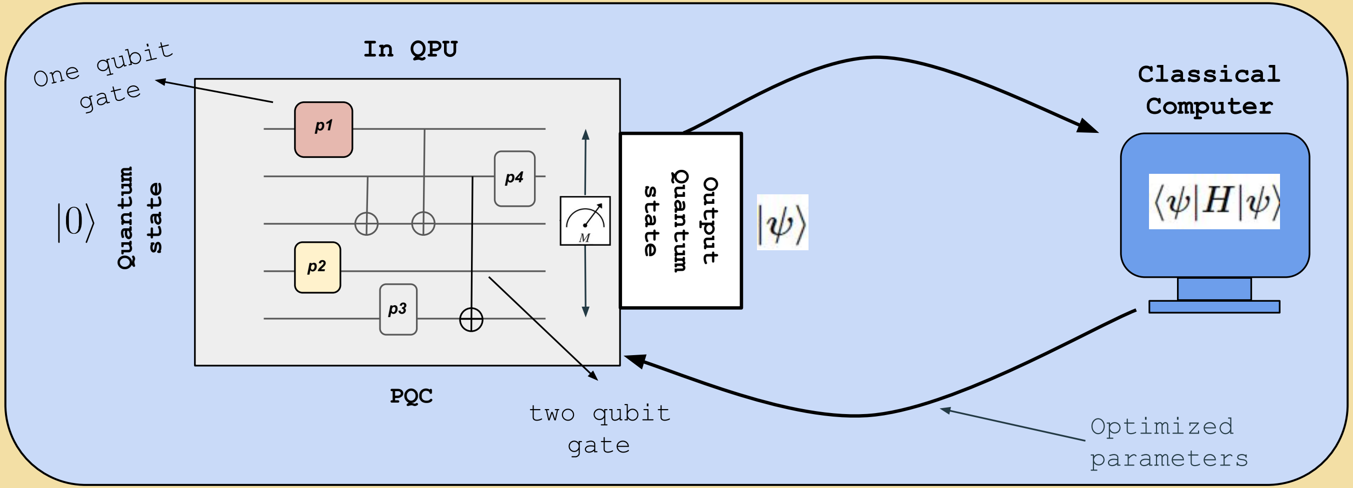
VQAs

Near-term quantum algorithms

Quantum computer (QC): Prepares quantum state (QS)

Classical computer (CC): Takes the QS and evaluate loss/cost function.

Optimization loop: Optimize cost function in CC and feed back optimized parameters to QC.



Cerezo, Marco, et al. Nature Reviews Physics 3.9 (2021): 625-644.

Bharti, Kishor, et al. Reviews of Modern Physics 94.1 (2022): 015004.

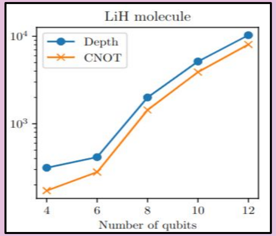
QAS: Quantum architecture search

Challenges of VQAs

- **QPU is noisy and qubit connectivity is limited**, which makes sophisticated and scalable ansatz design crucial.

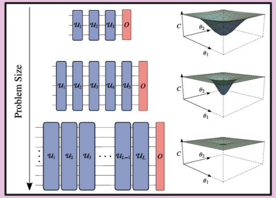
- **Manual ansatz selection** is inefficient and can have following issues:

  - **Problem inspired** (Exponential depth /gates with qubit)



Kundu, Akash. arXiv preprint arXiv:2402.13754 (2024).

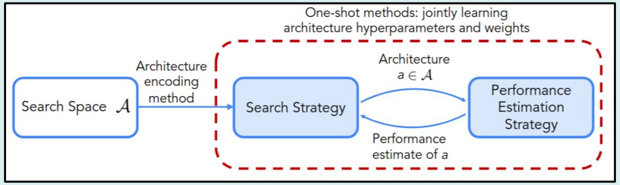
  - **Problem-agnostic** (Vanishing gradient /hinders training)



Cerezo, Marco, et al. Nature communications 12.1 (2021): 1791.

Solution: QAS

- Inspired by **neural architecture search**



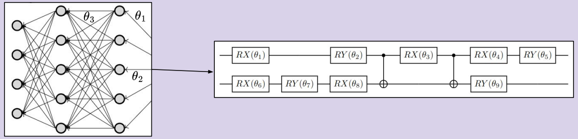
Elsken, Thomas, Jan Hendrik Metzen, and Frank Hutter. JMLR 20.55 (2019): 1-21.

- **Quantum architecture search** was introduced

Through batch sampling of circuits the good ones are evaluated where,

Search space (instead of neural net structures parametric quantum circuits)

Zhang, Shi-Xin, et al. Quantum Science and Technology 7.4 (2022): 045023.



Performance evaluation (The cost-function of the quantum optimization problem)

Search Strategy (Can be any learning model)

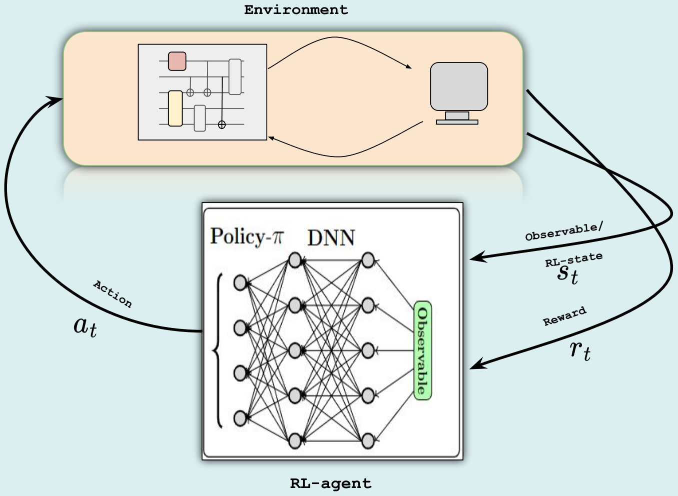
RL-QAS

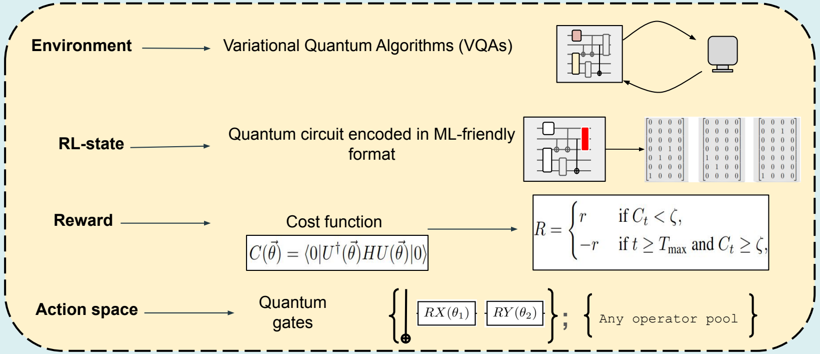
Challenges with differentiable QAS

- **Requires prior knowledge:** Valid circuit sampling requires prior knowledge for effective search.
- **Sample dependence:** Performance is highly sample dependent; invalid samples waste resources.
- **Scalability** is limited due to exponentially growing search space. **Restricted < 10 qubits**.
- **High resource scaling:** Resource consumption increases rapidly as circuit size grows.
- **QPU-friendly but not scalable**: show noise robustness mainly up to about 6–8 qubits in experiments with real or realistically simulated noise.

Solution: RL-QAS

- **No Prior Knowledge** of architecture needed
- Potential to come up with **innovative solutions**
- **Adaptability towards correcting errors** during the training process





Evaluating RL-QAS

Is it resource efficient?

- **Long time per episode:** In recent work [1] each episode (40 steps/gates per episode) execution for 4-qubit takes 25 seconds. This makes 10,000 episode to be executable in ~70 hours.
- **High cost function evaluation number by classical optimizer:** Vanilla RL [2] and curriculum RL [1] QAS requires in the order of 10^5 function evaluation to provide good approximation to find ground state of chemical Hamiltonian.

[1] Patel, Yash J., et al. ICLR (2024).

[2] Ostaszewski, Mateusz, et al. NeurIPS 34 (2021): 18182-18194.

Is it scalable?

- **QPU-friendly but not scalable**: RL-QAS shows noise robustness up to about 4-6 qubits on simulated hardware noise.
- **Not-so scalable**: Due to long time per episode, cost function evaluation and the multiple queries to the quantum state simulator evaluation of RL-QAS was restricted up to 8 qubits for various optimization problems [3, 4, 5, 6].

[3] Kundu, Akash, et al. New Journal of Physics 26.1 (2024): 013034.

[4] Kuo, En-Jui, et al. arXiv preprint arXiv:2104.07715 (2021).

[5] Fösel, Thomas, et al. arXiv preprint arXiv:2103.07585 (2021).

[6] Kundu, Akash. Machine Learning: Science and Technology (2025).

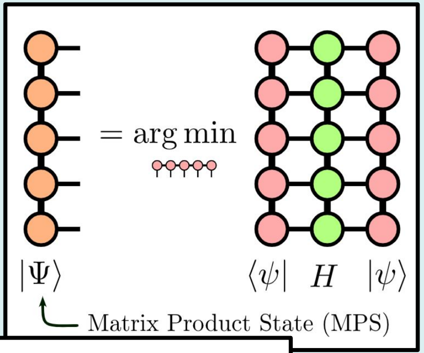
TensorRL-QAS

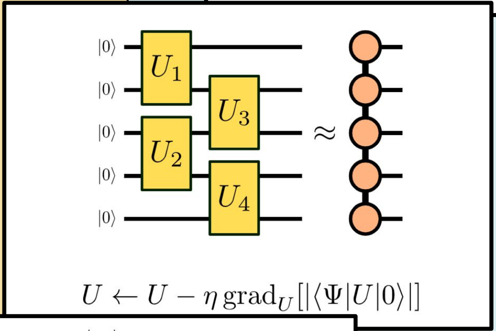
Is the proposed solution

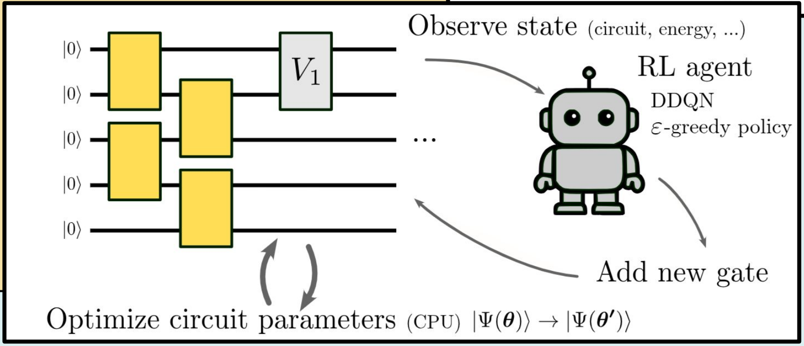
Our approach

- To address the dual challenges of action space explosion and prohibitive simulation costs, we introduce a novel framework **TensorRL-QAS**.

- **TensorRL-QAS** employs a problem-aware initialization obtained with **tensor networks** methods, **strategically initializing the search space to a physically meaningful starting point**.







TensorRL-QAS

The method

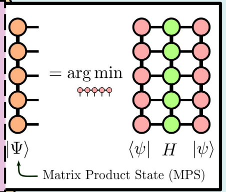
Open avenue: Considering PEPS or MERA might have the potential to enhance learning.

Step 1: Ground state approximation with DMRG

Use density matrix renormalization group (DMRG) with bond dimension χ to find a matrix product state (MPS) approximation of the ground state of the Hamiltonian H. The optimizer solves:

$$|\Psi\rangle = \arg\min_{|\psi\rangle, \||\psi\|=1} \langle\psi|H|\psi\rangle$$

$|\psi\rangle$ Is modeled as MPS, with the maximum bond distance χ



Step 2: Mapping MPS to quantum circuit

Circuit Ansatz: Construct $U = \prod_{k=1}^{m} U_k,$ where each U_k is a $2-$qubit unitary gate ($U_k \in \mathrm{U}(4)$).

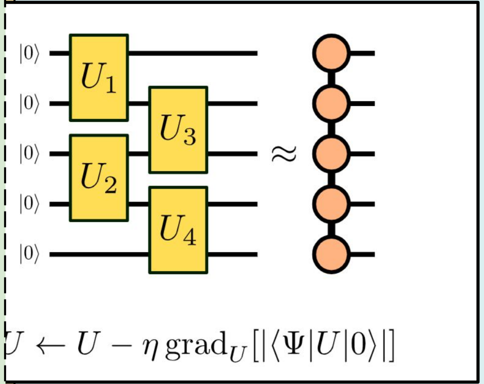
Objective Function: Optimize parameters to maximize the overlap between the circuit-prepared and target states:

$$L(U_1, U_2, \ldots) = |\langle\Psi|U|0\rangle| = \left|\langle\Psi| \prod_{k=1}^{m} U_k|0\rangle\right|$$

Gradient Computation: Build the tensor network for $\langle\Psi|U|0\rangle$ and apply automatic differentiation to compute gradients efficiently for all gate parameters.

Riemannian Optimization: Update gate parameters using Riemannian gradient steps to preserve unitarity, specifically, modifications to the Euclidean gradients are made so updates remain on the Stiefel manifold. The Cayley transform is commonly employed for these updates: $U_{\text{new}} = \text{Cayley}(U, G)$, G is Euclidean gradient.

Adam on Manifolds: Use a manifold-adapted Adam optimizer to iteratively minimize ***L*** over the Stiefel manifold, ensuring all iterative gate updates remain unitary.



TensorRL-QAS

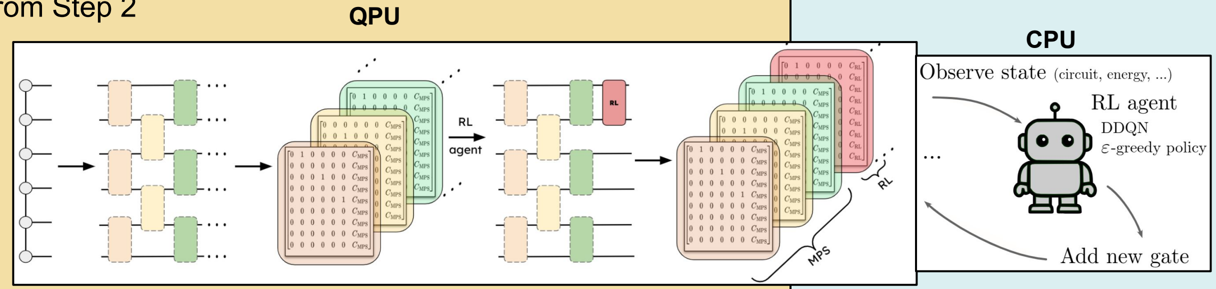
The method

Open avenue: Considering any other structure than the brickwork to map MPS to quantum circuit.

Step 3: Quantum Architecture Search by RL

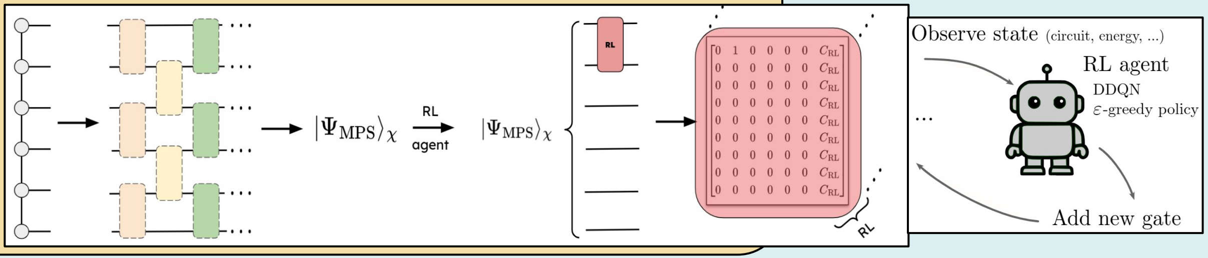
Initialize the RL agent with the circuit from Step 2 following one of the ways:

- **TensorRL (trainable, TN-init)**
  The **MPS trainable**
  (The MPS is encoded into a binary encoding to make it **visible to the RL-agent**)



Quantum circuits optimized (on **CPU**) with **COBYLA**

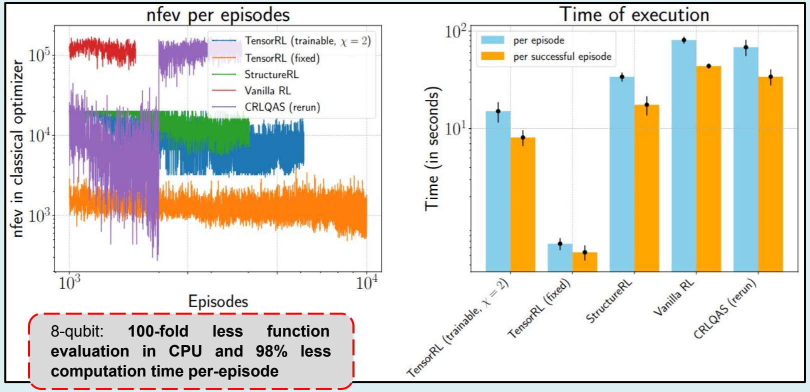
- **TensorRL (fixed, TN-init)**
  The **MPS fixed**
  (we do not encode the MPS into the RL-state, but the empty PQC is initialized with the MPS wavefunction)



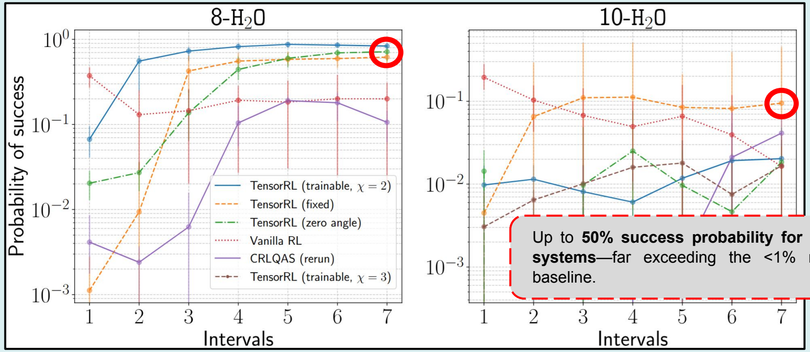
Noiseless

Results (Quantum Chemistry)

TensorRL-QAS



8-qubit: **100-fold less function evaluation in CPU and 98% less computation time per-episode**



Up to **50% success probability for 10-qubit systems**—far exceeding the <1% rates of baseline.

Molecule	Method	Error	Depth	CNOT	ROT
6-BEH_2	**TensorRL (fixed TN-init)**	6.8×10^{-5}	**7**	**5**	**12**
	TensorRL (trainable TN-init)	6.0×10^{-5}	33	22	96
	StructureRL	$\mathbf{5.9 \times 10^{-5}}$	33	21	97
	TF-QAS [42]	1.8×10^{-3}	NA	NA	57 (total gate)
	RA-QAS [37]	5.8×10^{-4}	21	26	17
	SA-QAS [39]	5.6×10^{-3}	36	45	28
	TN-VQE	5.4×10^{-3}	10	8	12
8-H_2O	**TensorRL (fixed TN-init)**	8.9×10^{-4}	**6**	**9**	**15**
	TensorRL (trainable TN-init)	2.0×10^{-4}	36	30	146
	StructureRL	$\mathbf{1.3 \times 10^{-4}}$	33	30	133
	CRLQAS [35]	1.8×10^{-4}	75	105	35
	CRLQAS (rerun)	1.7×10^{-4}	100	85	58
	Vanilla RL [21]	1.7×10^{-4}	96	117	48
	quantumDARTS [41]	3.1×10^{-4}	64	68	151
	RA-QAS	1.1×10^{-3}	56	87	41
	SA-QAS	2.6×10^{-3}	40	69	26
	TN-VQE	2.7×10^{-3}	10	14	6
10-H_2O	**TensorRL (fixed TN-init)**	4.1×10^{-4}	**17**	**15**	**17**
	TensorRL (trainable TN-init, $\chi = 3$)	6.7×10^{-4}	33	34	168
	TensorRL (trainable TN-init)	7.1×10^{-4}	35	48	173
	StructureRL	5.8×10^{-4}	37	52	169
	CRLQAS (rerun)	3.4×10^{-4}	114	140	43
	Vanilla RL	$\mathbf{2.5 \times 10^{-4}}$	73	96	32
	RA-QAS	9.9×10^{-4}	80	152	58
	SA-QAS	4.7×10^{-3}	24	42	14
	TN-VQE	1.1×10^{-3}	26	51	14
12-LiH	**TensorRL (trainable TN-init)**	$\mathbf{1.0 \times 10^{-2}}$	31	37	203
	TensorRL (fixed TN-init)	2.4×10^{-2}	**15**	**30**	**9**
	StructureRL	2.2×10^{-2}	40	53	179
	CRLQAS (rerun)	2.4×10^{-2}	32	68	23
	Vanilla RL	2.2×10^{-2}	140	321	94
	RA-QAS	2.3×10^{-2}	165	364	86
	SA-QAS	2.5×10^{-2}	178	390	88
	TN-VQE	8.6×10^{-2}	33	81	29

Gives best approximation to ground energy with **10-fold reduction in CNOT count and circuit depth compared to RL/non-RL baseline**.

Noisy

Results (Quantum Chemistry)

TensorRL-QAS

Simulated noise (8-qubit)

Noise	Method	Error	Depth	CNOT	ROT	Success prob.
Depolarizing	**TensorRL (fixed)**	$\mathbf{9.0 \times 10^{-4}}$	**7**	**5**	**12**	**100%**
	TensorRL (trainable)	2.48×10^{-3}	33	22	130	0%
	StructureRL	2.6×10^{-3}	28	22	129	0%
	CRLQAS (rerun)	1.3×10^{-3}	22	11	29	30%
Shot	**TensorRL (fixed)**	1.5×10^{-4}	**3**	**4**	**1**	100%
	TensorRL (trainable)	$\mathbf{8.7 \times 10^{-5}}$	30	28	131	100%
	StructureRL	6.8×10^{-4}	34	25	135	100%

Noise model: 1-qubit gate errors were amplified 10× (1% depolarizing), 2-qubit errors 5× (5% depolarizing), and shot-noise from 1000 samples.

TensorRL (fixed) achieved chemical accuracy with **100% success** rate under depolarizing noise, compared to CRLQAS' 30% success rate at comparable error.

On QPU (8-qubit)

Method	Noiseless error	IBMQ-Brisbane error
TensorRL (fixed, best episode)	9.0×10^{-4}	2.8×10^{-3}
TensorRL (fixed, random 1)	3.4×10^{-4}	$\mathbf{1.5 \times 10^{-3}}$
TensorRL (fixed, random 2)	1.2×10^{-3}	3.3×10^{-3}

Noiseless quantum circuits executed in **IBMQ-Brisbane device**. The best episodes refers to the episode where agent found most compact quantum circuit.

Hence **best noiseless case might not refer to the best noisy scenario**.

Noiseless

Results (Towards scalability)

TensorRL-QAS

Scalable

$$H_{\mathrm{TFIM}} = \sum_{i=1}^{n} Z_i Z_{i+1} + h \sum_{i} X_i,$$

Open avenue: With better action space representation we can scale it beyond 20-qubit.

Problem	Method	Error	Improvement
15-TFIM	TensorRL (fixed)	4.4×10^{-4}	21%
17-TFIM	TensorRL (fixed)	4.8×10^{-4}	19%
20-TFIM (bottleneck)	TensorRL (fixed TN-init)	6.7×10^{-4}	1%
20-TFIM (actions: XX, YY, ZZ, RX, RZ, RY, CX)	TensorRL (fixed)	6.1×10^{-4}	**9%**

This decreases the call to quantum statevector simulator than the baselines approaches:

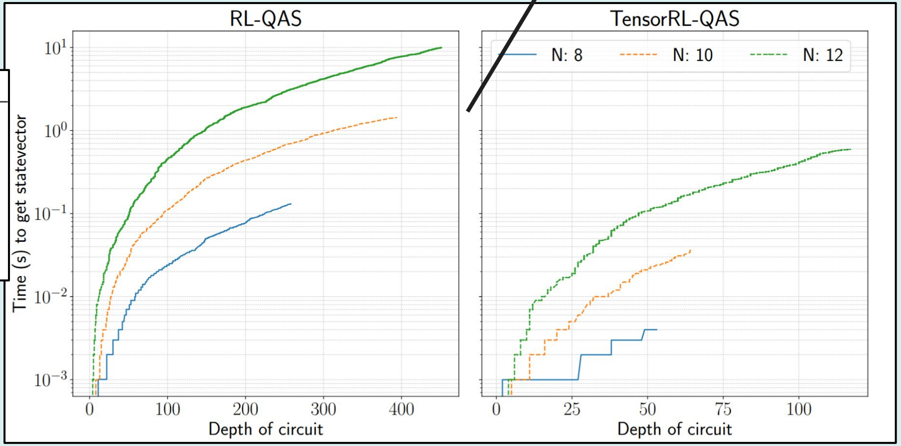
- State-of-art RL-QAS takes at least 27 hours (10 sec per episode)
- **TensorRL-QAS takens 1.5 hours (0.5 sec per episode)**.

To achieve good approximation to ground state **for 12 qubit.**

Analysis of quantum simulator

Qubits	TensorRL-QAS	CRLQAS	RAQAS
6	**20**	70	97
8	**20**	250	300
10	**20**	350	477
12	**40**	450	450
15	**60**	NA	NA
17	**60**	NA	NA
20	**80**	NA	NA

TensorRL-QAS requires 3–5× less steps per episode compared to baseline RL-QAS, across all tested system sizes.



SUMMARY

1

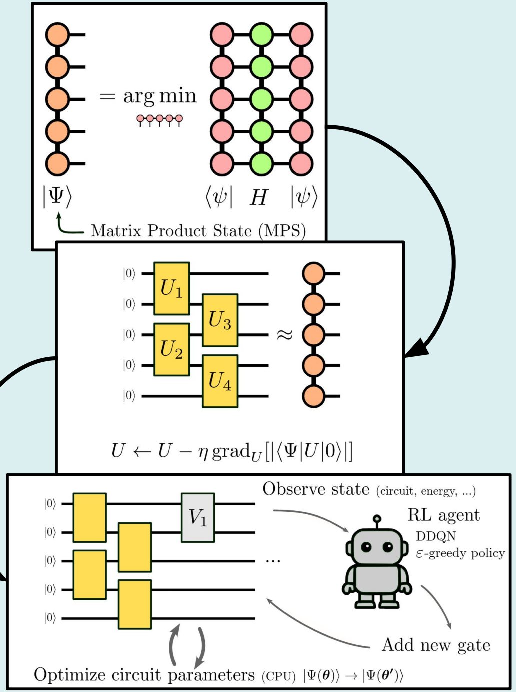
TensorRL-QAS achieves state-of-the-art results in quantum architecture search (QAS), demonstrating better ground-state energy approximations and reduced gate counts compared to traditional RL and non RL-based methods.

2

The framework enables efficient exploration of large quantum circuit spaces (up to 20 qubits), showing robust performance on both molecular and spin models under various noise conditions.

3

We open a new avenue for QAS and we believe numerous follow-up avenues will reveal the true capability of **TensorRL-QAS**.



Paper:

THANK YOU

Code: