

L2O- g^\dagger : LEARNING TO OPTIMIZE PARAMETERIZED QUANTUM CIRCUITS WITH FUBINI-STUDY METRIC TENSOR

{ YU-CHAO HUANG, HSI-SHENG GOAN } NATIONAL TAIWAN UNIVERSITY.

✉ R11222015@ntu.edu.tw

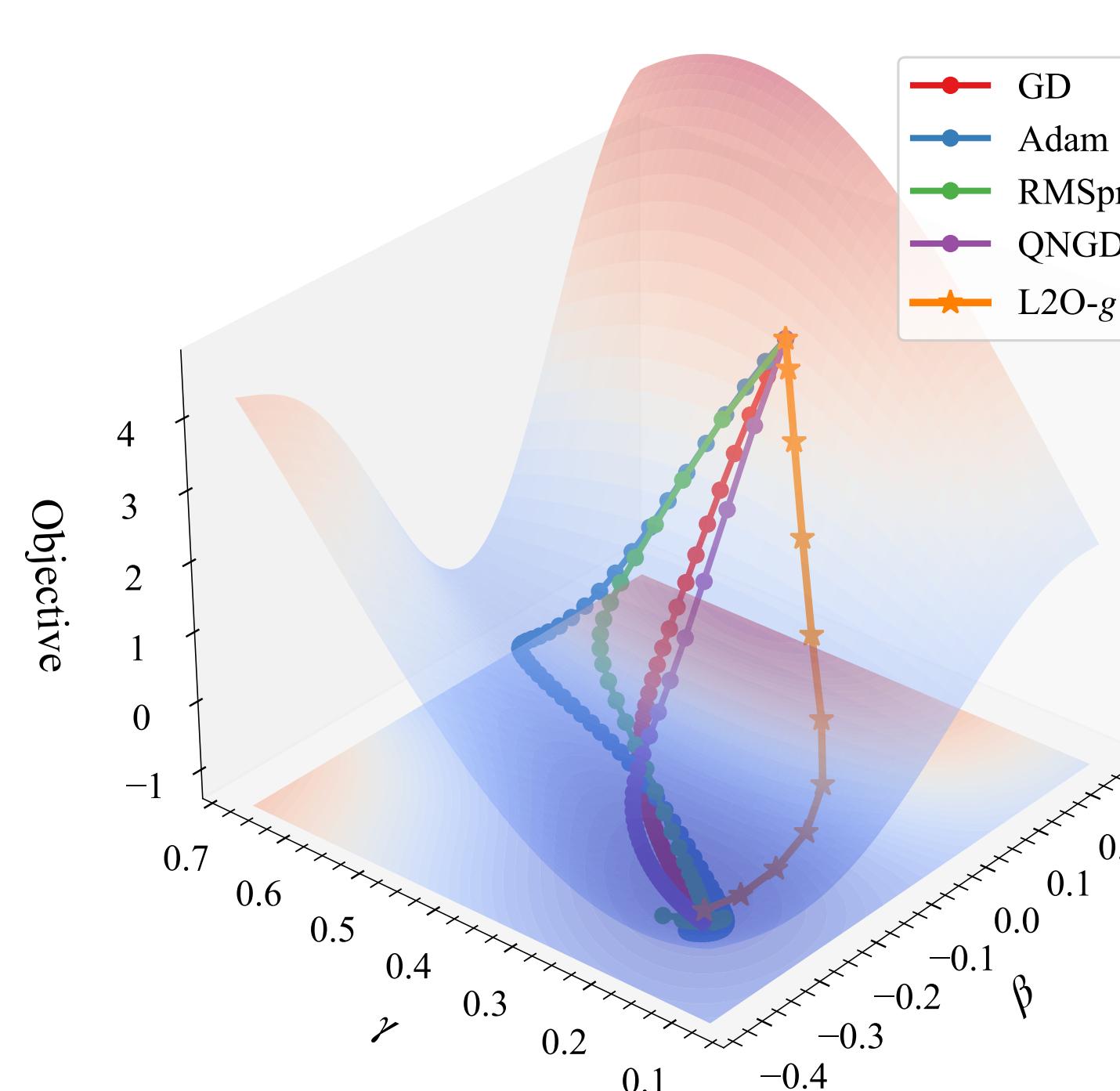
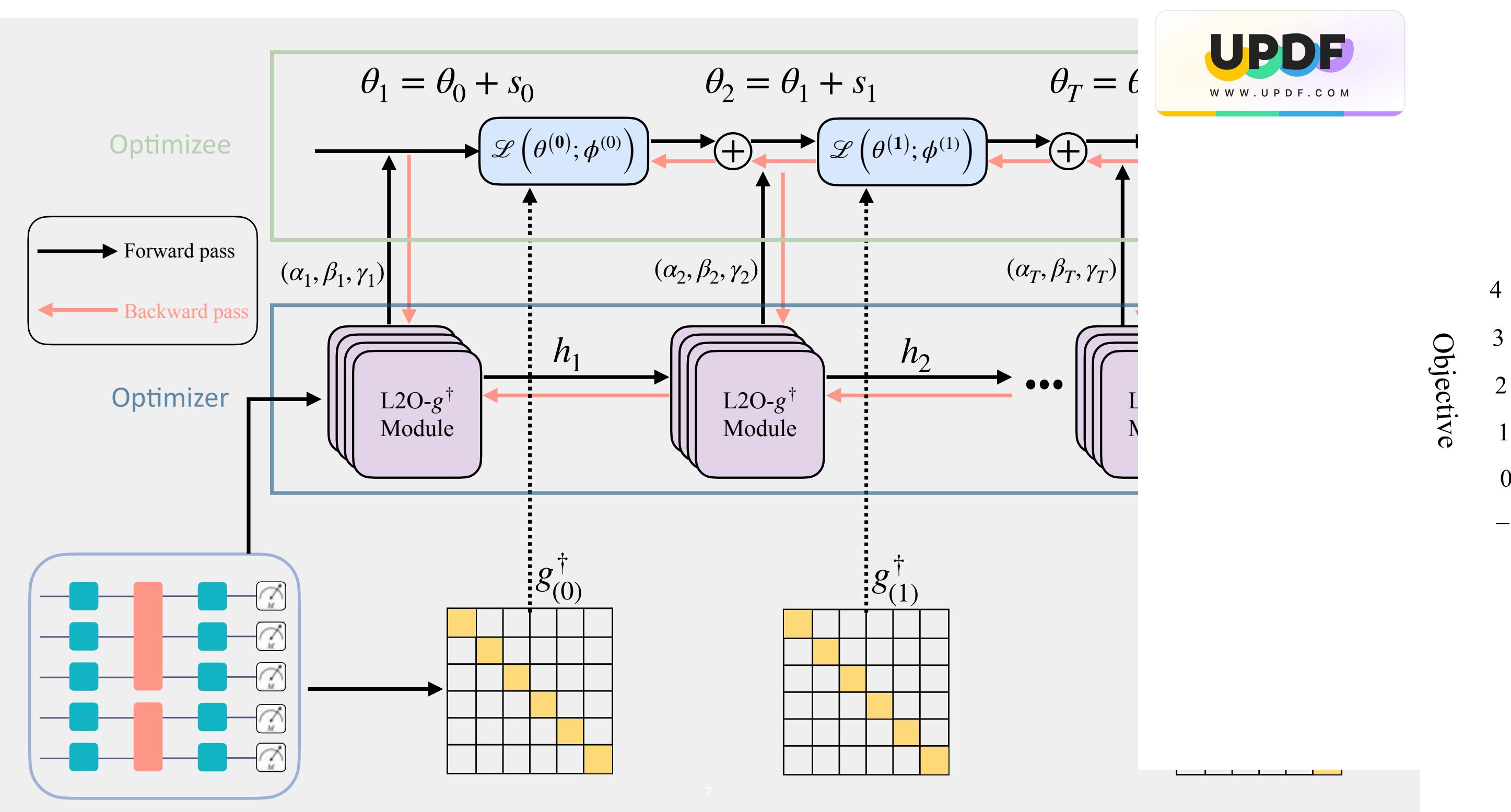
🌐 physics-morris.github.io



SUMMARY

- We present L2O- g^\dagger , a *quantum-aware* learned optimizer tailored for optimizing parameterized quantum circuits (PQCs), requiring minimal training instances.
- L2O- g^\dagger can optimize diverse VQAs like VQE, QAOA, and QNN by leveraging parameter space and distribution space optimization.
- Out-of-the-box performance can match or surpass learning rate-tuned optimizers like Adam and QNGD, as well as previous learned optimizers for PQC.
- We open source our implementation and data at: <https://github.com/Physics-Morris/L2O-g>

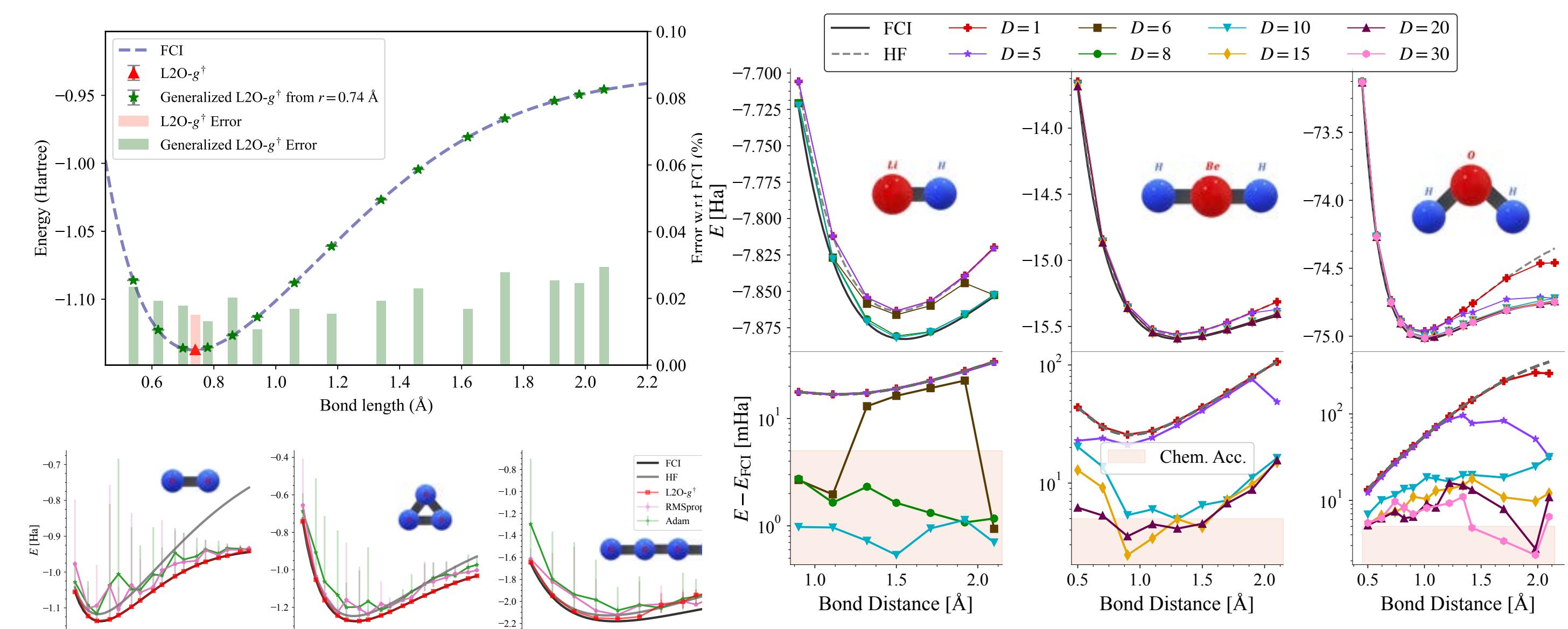
OVERALL ARCHITECTURE / TOY EXAMPLE



(Left) Given the parameters θ_t from PQC, L2O- g^\dagger balances distribution space optimization and parameter space optimization. The L2O- g^\dagger module, a recurrent neural network, outputs three vectors $(\alpha_t, \beta_t, \gamma_t)$ at each time step t in the optimizee trajectory $[x_T]$. The distribution space optimization is modified by the Fubini-Study metric tensor g^\dagger . The update step $\eta_t = \exp[\lambda_b \alpha_t] \in \mathbb{R}^N$ and the update direction $v = \lambda_a \beta_t \in \mathbb{R}^N$ update the parameters to θ_{t+1} . **(Right)** A motivating toy example of L2O- g^\dagger optimizing MaxCut problem with QAOA ($V = 5, p = 0.9, p_{\text{layer}} = 1$).

VQE FOR HEA / UCCSD / $U_{\text{ENT}}^{(1)}$ ANSATZ

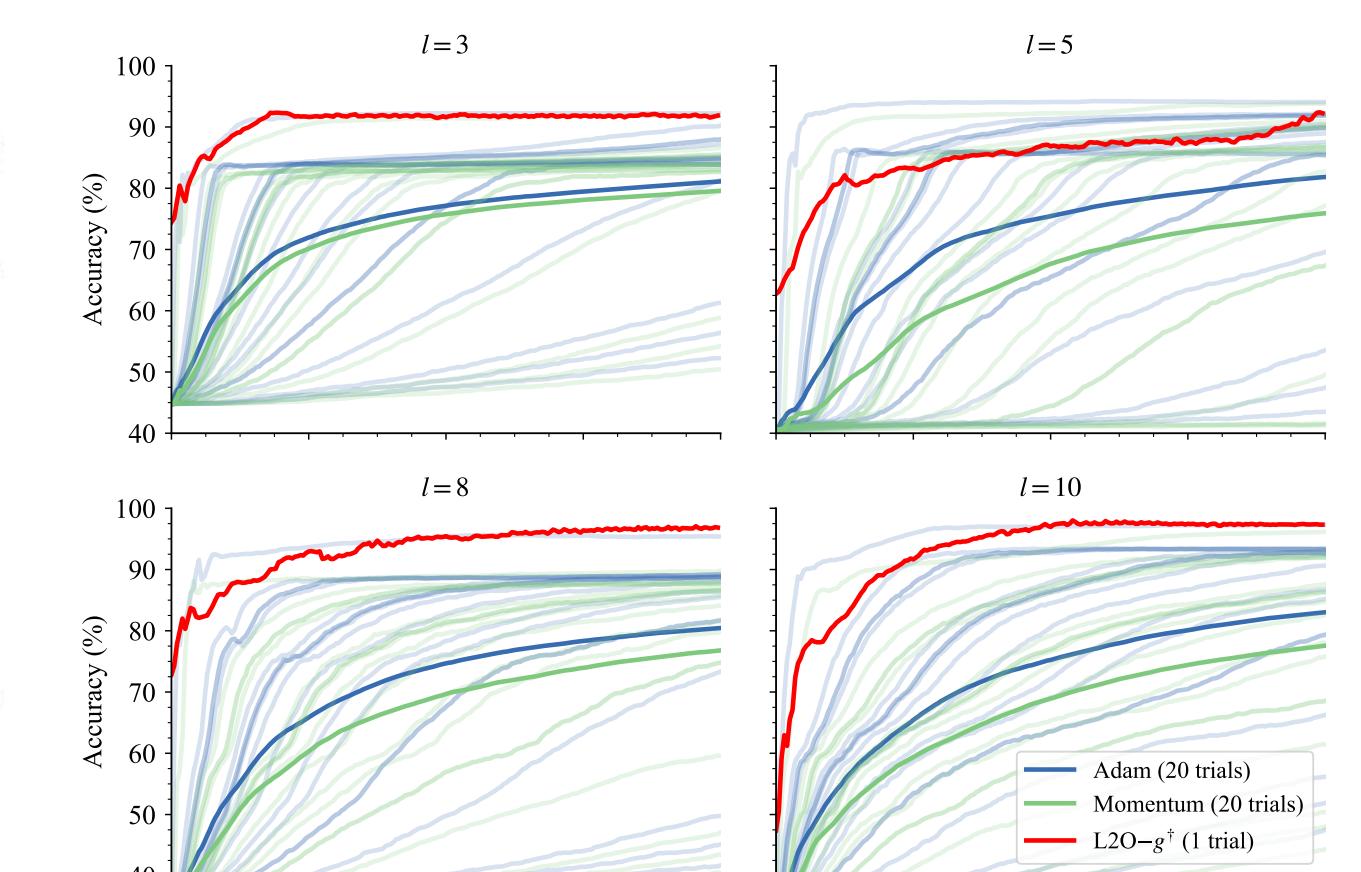
(upper left) L2O- g^\dagger trained on a single instance of the H_2 molecule to generalize to solve the ground state energy across various radii. **(down left)** Achieves higher accuracy on H_2 , H_3^+ , and H_4 compared to RMSprop and Adam optimizers for $U_{\text{ENT}}^{(1)}$ ansatz. **(right)** Achieves near chemical accuracy for more complex molecules like LiH, BeH₂, and H₂O with up to 30 entangled blocks of UCCSD ansatz.



QNN FOR DATA-REUPLOAD CLASSIFIER

For a vector x of three features, two trainable three-dimensional real vectors w and θ , and encodes them as $U(x \circ \omega + \theta)$ where $U(\phi) = e^{iZ\phi_1/2}e^{iY\phi_2/2}e^{iZ\phi_3/2}$. The cost function is $\mathcal{C}(\theta, \omega, \alpha, x_i) = \sum_{j=1}^{n_{\max}} (\alpha_{0j} F_{0j} - (1 - y_i))^2 + (\alpha_{1j} F_{1j} - y_i)^2$

LAYERS	1	3	5	8	10
L-BFGS-B	50*	85*	89*	93*	95*
VANILLA GD	44.47	49.40	48.41	43.34	47.58
MOMENTUM	56.14	79.35	82.47	77.57	79.25
ADAM	50.31	79.58	84.33	83.88	86.08
L2O-g^\dagger	50.25	90.88	91.37	95.34	95.50
C. NN			96*		
C. SVC			97*		



ABLATION: COMPARE WITH L2O-DM

Compare with L2O-DM [Andrychowicz et al., 2016], which uses the same architecture as [Verdon et al., 2019] and [Wilson et al., 2019]. The training settings are highlighted in the pink column.

PROBLEM	CONFIG.	L2O-DM [4]	L2O- g^\dagger (OURS)
RAND. PQC(\downarrow)	$(N_q, l) = (7, 5)$	-1.00 ± 0.00	-1.00 ± 0.00
	$(N_q, l) = (7, 8)$	-0.94 ± 0.06	-1.00 ± 0.00
	$(N_q, l) = (10, 5)$	-1.00 ± 0.00	-1.00 ± 0.00
VQE/HEA(\downarrow)	H_2 AT $r = 0.5$	-0.27 ± 0.31	-0.66 ± 0.48
	H_2 AT $r = 0.9$	-0.52 ± 0.11	-0.86 ± 0.21
	H_2 AT $r = 1.5$	-0.56 ± 0.06	-0.93 ± 0.05
VQE/ENT(\downarrow)	LiH AT $r = 0.9$	-7.61 ± 0.03	-7.71 ± 0.00
	BeH_2 AT $r = 0.9$	-15.23 ± 0.03	-15.34 ± 0.00
	H_2O AT $r = 0.9$	-74.85 ± 0.01	-74.95 ± 0.00
VQE/UCCSD(\downarrow)	H_2 AT $r = 0.5$	-0.87 ± 0.20	-1.06 ± 0.00
	H_2 AT $r = 0.9$	-1.07 ± 0.06	-1.12 ± 0.00
	H_3^+ AT $r = 0.5$	-0.53 ± 0.19	-0.74 ± 0.00
	H_3^+ AT $r = 0.9$	-0.79 ± 0.31	-1.27 ± 0.00
	H_4 AT $r = 0.5$	-0.15 ± 0.60	-1.47 ± 0.35
	H_4 AT $r = 0.9$	-1.24 ± 0.21	-2.07 ± 0.17
	$V = 6, p = 0.5$	34.78 ± 12.32	56.36 ± 11.90
MAXCUT(\uparrow)	$V = 6, p = 0.6$	49.06 ± 9.66	61.56 ± 6.36
	$V = 6, p = 0.7$	28.67 ± 9.95	51.53 ± 12.97
	$V = 6, p_{\text{layer}} = 1$	-0.93 ± 0.95	-0.90 ± 1.10
SK MODEL(\downarrow)	$V = 6, p_{\text{layer}} = 3$	-1.95 ± 1.65	-2.96 ± 0.06
	$V = 6, p_{\text{layer}} = 5$	-1.95 ± 2.12	-2.96 ± 0.05
	$l = 3$	66.87 ± 1.60	90.08 ± 1.57
DATA RE-UPLOAD(\uparrow)	$l = 5$	69.50 ± 7.80	91.36 ± 0.70
	$l = 8$	81.97 ± 3.60	95.34 ± 1.09

TABLE II: Ablation Study on L2O-DM [4] for Different Problems and Configurations. We report the average