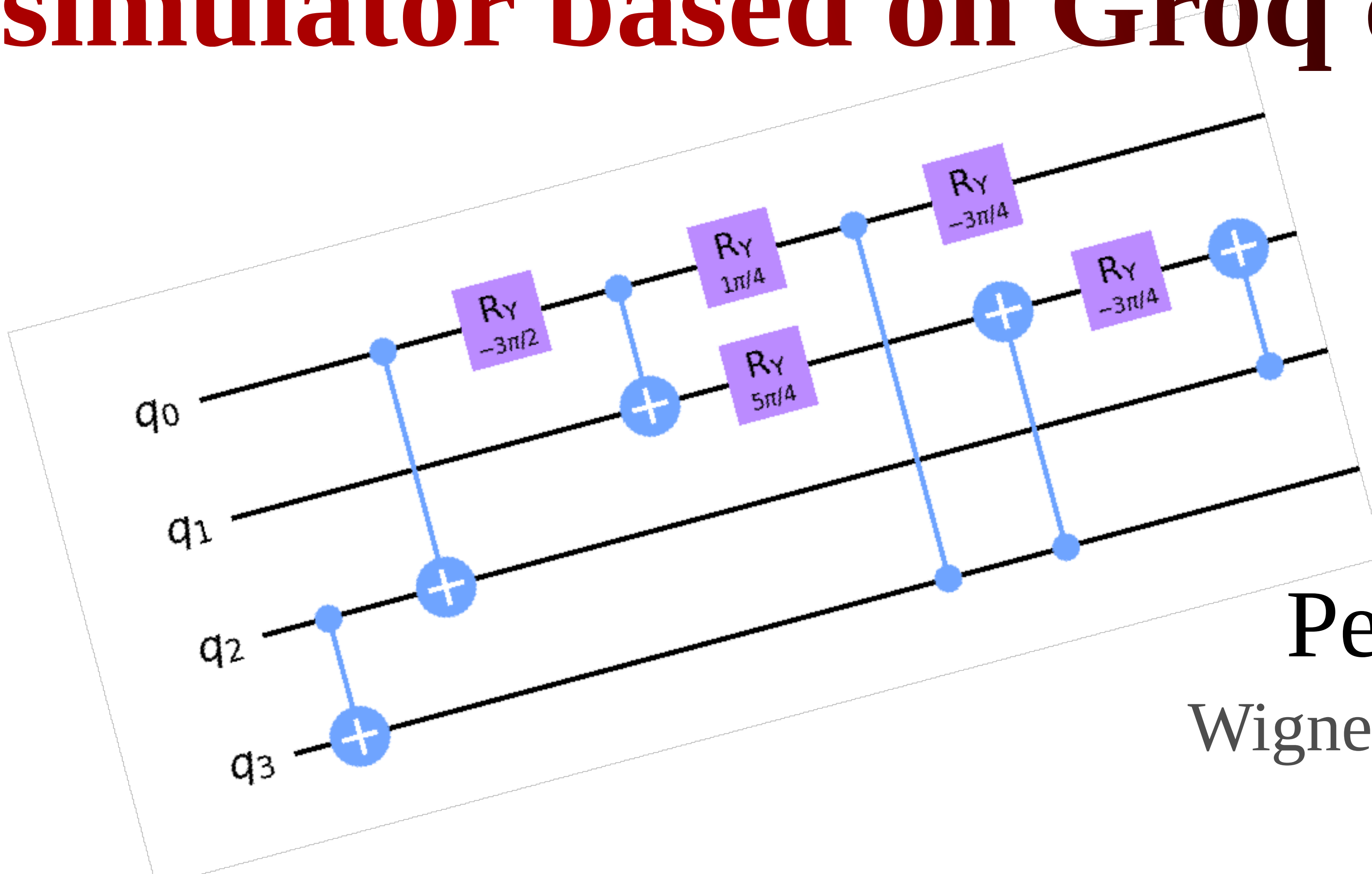
Quantum Information
National Laboratory
HUNGARY

Building a quantum computer simulator based on Groq chips



Peter Rakyta
Wigner Research Centre for Physics

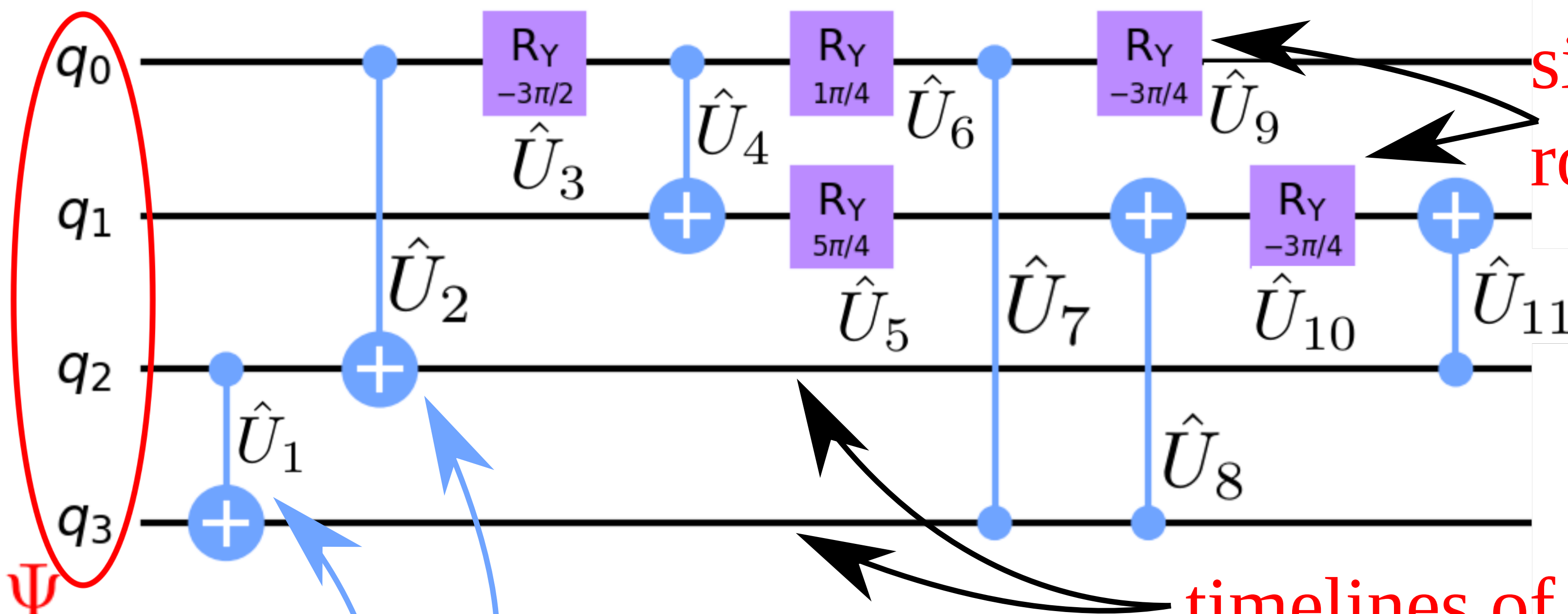
ELTE
EÖTVÖS LORÁND UNIVERSITY

Qubit based architecture

Quantum Information National Laboratory HUNGARY

quantum program (unitary)

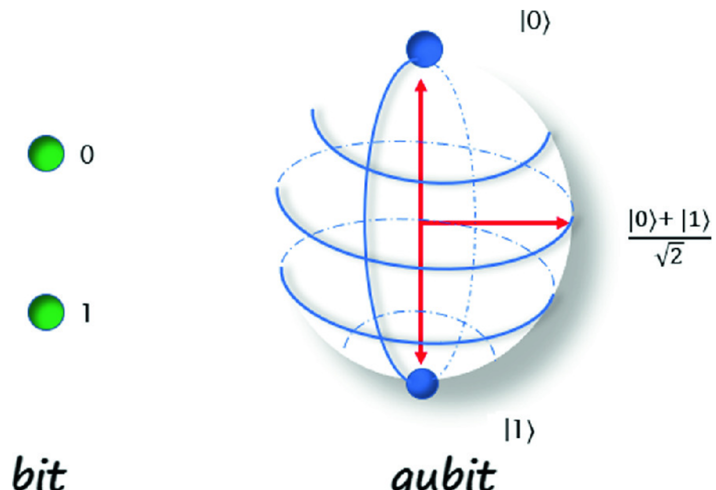
$$\hat{U} = \hat{U}_{11} \cdot \hat{U}_{10} \cdot \hat{U}_9 \cdot \hat{U}_8 \cdot \hat{U}_7 \cdot \hat{U}_6 \cdot \hat{U}_5 \cdot \hat{U}_4 \cdot \hat{U}_3 \cdot \hat{U}_2 \cdot \hat{U}_1$$



single qubit rotations

timelines of the qubits

controlled not gates



ELTE EÖTVÖS LORÁND UNIVERSITY

Wigner

Software toolkits to train parametric quantum circuits

Quantum Information
National Laboratory
HUNGARY

Gate decomposition utilities:

- Quantum Fast Approximate Synthesis Tool (QFAST)
- BQSkit: QSearch + LEAP

(Lawrence Berkeley National Laboratory)

QML utilities:

QISKIT

T|ket>: A Retargetable Compiler for NISQ Devices
(Cambridge Quantum Computing Ltd., University of Strathclyde

Qulacs: a fast and versatile quantum circuit simulator for research purposevices
(QunaSys, Osaka University, NTT, and Fujitsu)

a framework for quantum simulation with hardware acceleration

SQUANDER: toolkit to train quantum circuits

Quantum Information National Laboratory HUNGARY

- SQUANDER: Sequantial Quantum Gate Decomposer

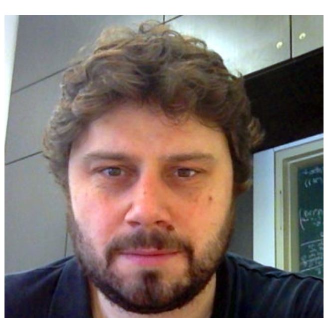
Quantum gate compilation

State preparation

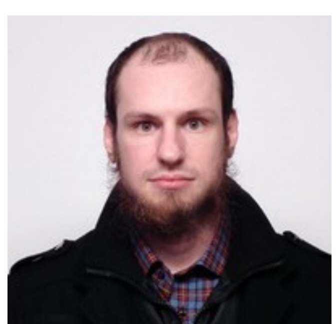
gradient descend and evolutionary optimizers

Quantum ML

GitHub https://github.com/rakytap/sequential-quantum-gate-decomposer



Zoltán Zimborás
Wigner



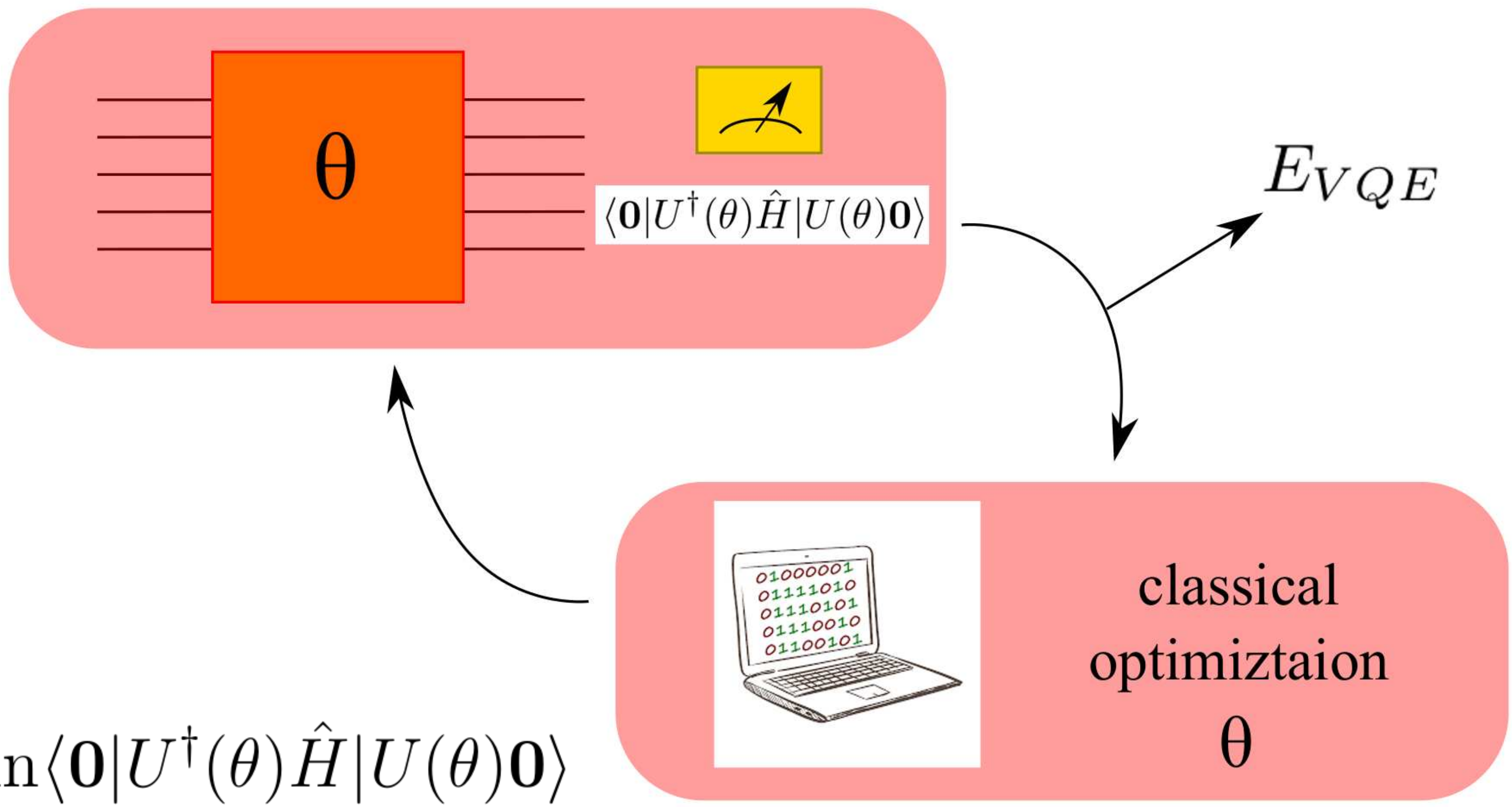
Gregory Morse

Jakab Nádori (BSc)
Barna Villám (BSc)
László Hajas (BSc)
Zita Majnay-Takács

ELTE EÖTVÖS LORÁND UNIVERSITY

Wigner

Variational quantum eigensolver



$$E_{VQE} = \min_\theta \langle \mathbf{0}|U^\dagger(\theta)\hat{H}|U(\theta)\mathbf{0}\rangle$$

$$\hat{H} = \sum_\alpha^{\mathcal{P}} w_\alpha \hat{P}_\alpha$$

Quantum Information National Laboratory HUNGARY

ELTE EÖTVÖS LORÁND UNIVERSITY

wigner

ARTICLE

DOI: 10.1038/s41467-018-07090-4 OPEN

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean[1], Sergio Boixo[1], Vadim N. Smelyanskiy[1], Ryan Babbush[1] & Hartmut Neven[1]

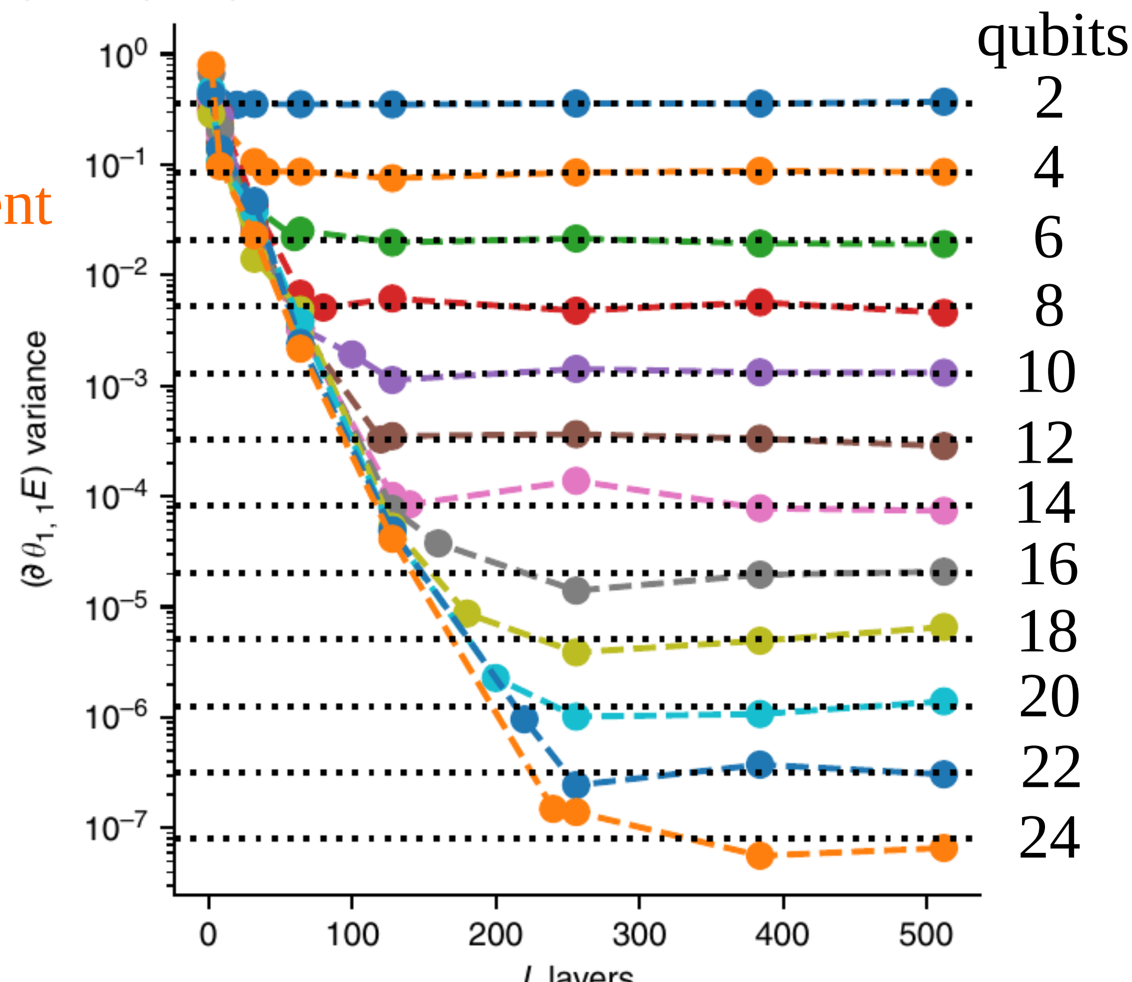
The variance of the cost function gradient

(and consequently its typical value)

vanishes exponentially

in the qubit number N

$$\mathrm{Var}[\partial_{i,l} E(\boldsymbol{\theta})] \sim \mathcal{O}\left(\frac{1}{2^{2N}}\right)$$



ELTE EÖTVÖS LORÁND UNIVERSITY

WIGNER

Barren plateau & entanglement entropy

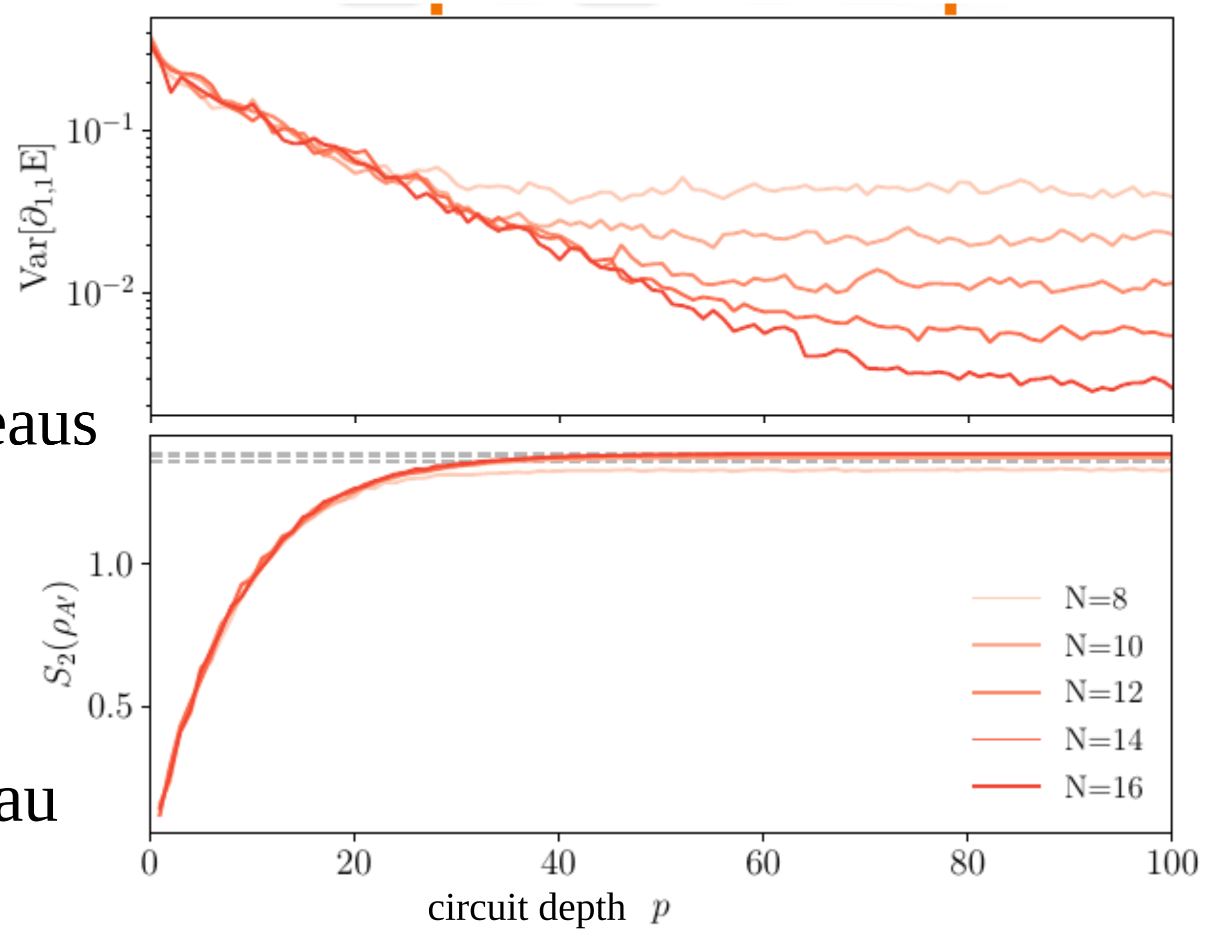
controlling entangelement
to mitigate BP?

Second Rényi entropy

$$S_2 = -\ln \mathrm{Tr}\rho_A^2$$

subsystem

Avoiding Barren Plateaus Using Classical Shadows
Stefan H. Sack, Raimel A. Medina, Alexios A. Michailidis,
Richard Kueng, and Maksym Serbyn
PRX QUANTUM 3, 020365 (2022)



Entanglement-Induced Barren Plateaus
Carlos Ortiz Marrero, Mária Kieferová,
and Nathan Wiebe
PRX Quantum 2, 040316

Entanglement devised barren plateau mitigation
Taylor L. Patti, Khadijeh Najafi, Xun Gao, and Susanne F. Yelin
Phys. Rev. Research 3, 033090

Avoiding Barren Plateaus Using Classical Shadows

Stefan H. Sack, Raimel A. Medina, Alexios A. Michailidis,
Richard Kueng, and Maksym Serbyn

PRX QUANTUM 3, 020365 (2022)

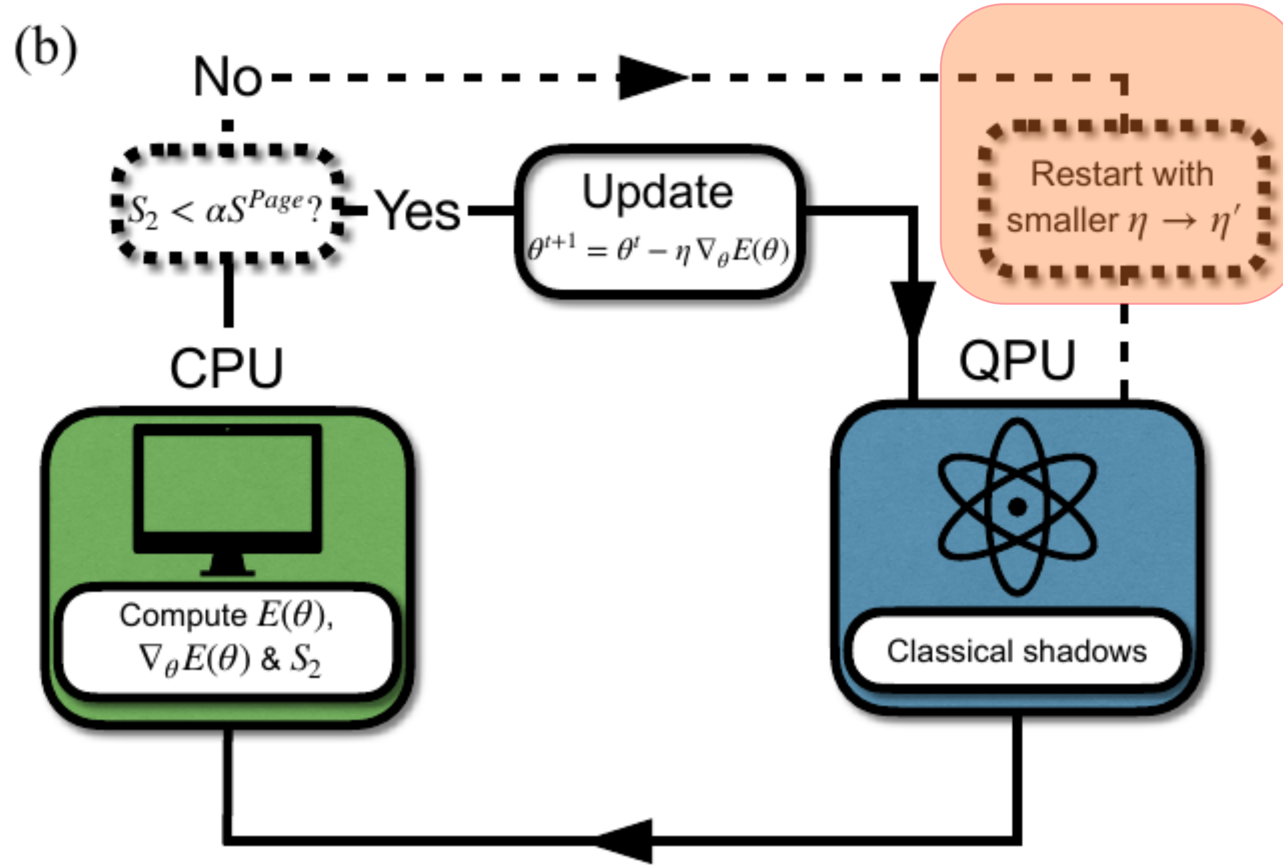
Quantum Information National Laboratory HUNGARY

- **monitor the entropy and control the lerning rate**

$$\theta^{t+1} = \theta^t - \eta \nabla_\theta E(\theta),$$

Decreasing the learning rate!!

- parameter initialization to low entangling gates
  Quantum 3, 214 (2019),
  PRX Quantum 3, 010313 (2022)
- use local cost functions
  Nature Communications 12, 1791 (2021)
- use matrix product states
  PRX Quantum 3, 010313
- layer-by-layer optimization
  Quantum Mach. Intell. 3, 5 (2021)



limiting the expressiveness of the circuit

ELTE EÖTVÖS LORÁND UNIVERSITY

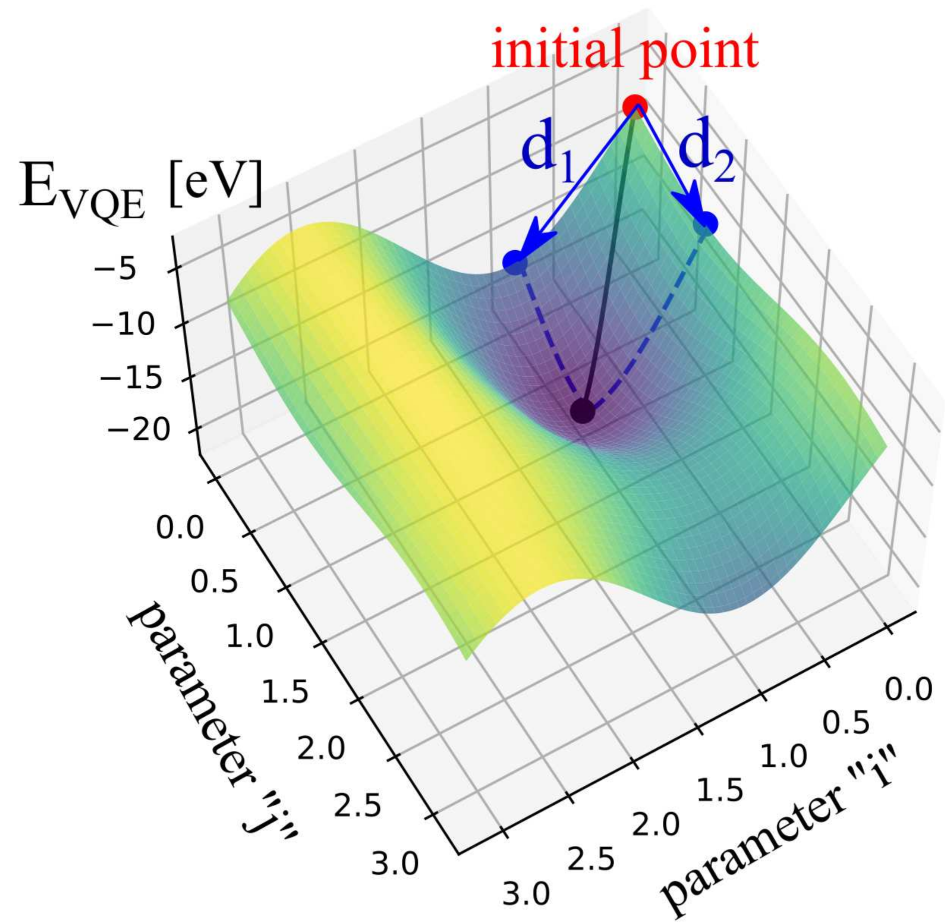
Wigner

~~Decrease~~ Increase the learning rate

- Perform line search along a well defined direction

Search direction → line search
→ Search direction → line search
→ Search direction → line search
→ Search direction → line search
→ •••

- How to determine search direction?
- What is the range of the line search?
- layer-by-layer optimization



Parameter dependence

gate containing parameter p_i: U_p

$$|\Psi\rangle = U_1 \cdot U_2 \cdot U_3 \ldots \underbrace{(A \cdot \cos(p_i) + B \cdot \sin(p_i) + C)}_{\text{matrices}} \ldots U_{K-2} \cdot U_{K-1} \cdot U_K |0\rangle$$

matrices

$$|\Psi\rangle = \cos(p_i)|a\rangle + \sin(p_i)|b\rangle + |c\rangle$$

scalars

$$E = \langle\Psi|H|\Psi\rangle = \kappa \cdot \sin(2p_i + \xi) + \gamma \cdot \sin(p_i + \phi) + C$$

knowing the constants, the line search becomes efficient

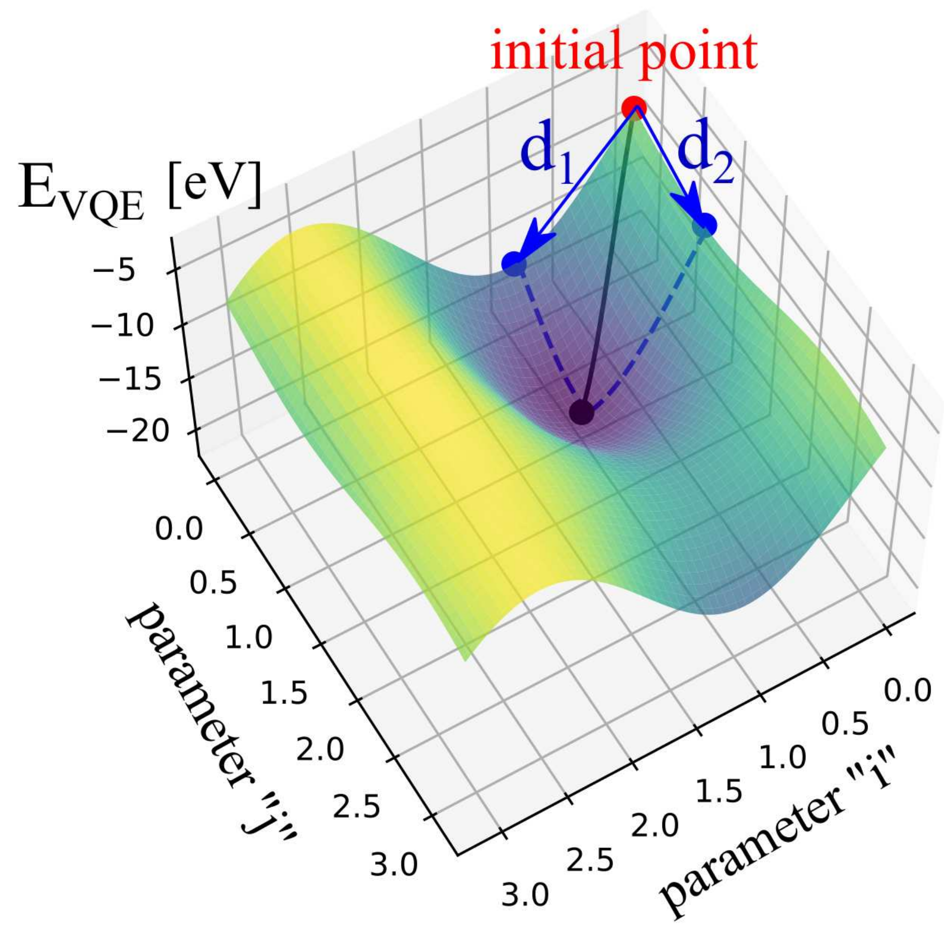
Gradient free search direction

$$E_{VQE} = \kappa \cdot \sin(2\theta_i + \xi) + C$$

- randomly select a subset Λ of the parameters
- determine the parameter-wise minimum θ_i^*
- we define the search direction

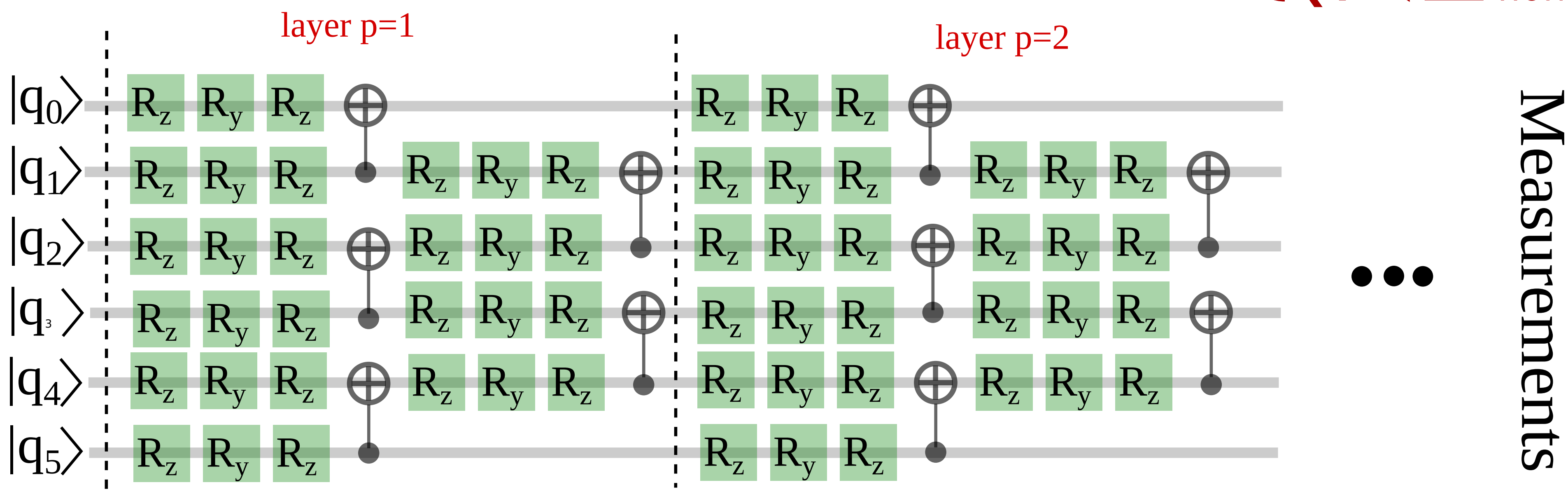
$$d_i = \begin{cases} \theta_i^* - \theta_i & \text{if } i \in \Lambda \\ 0 & \text{otherwise.} \end{cases}$$

- decreasing values in E_{VQE} are automatically associated with moderate entaglement entropy.
- There are no additional hyper-parameters in the algorithm



The studied VQE problem

Quantum Information
National Laboratory
HUNGARY



Heisenberg XXX model

$$\hat{H}_{XXX} = \sum_{i,j \in V_{\mathcal{G}}} J\left(\hat{\sigma}_i^z\hat{\sigma}_j^z + \hat{\sigma}_i^y\hat{\sigma}_j^y + \hat{\sigma}_i^x\hat{\sigma}_j^x\right) + h_z \sum_N \hat{\sigma}_i^z$$

Sachdev-Ye-Kitaev (SYC) model describing Majorana fermions

$$\hat{H}_{SYC} = \sum_{1 \le i < j < k < l \le 2N} J_{i,j,k,l}\, \chi_i \chi_j \chi_k \chi_l$$

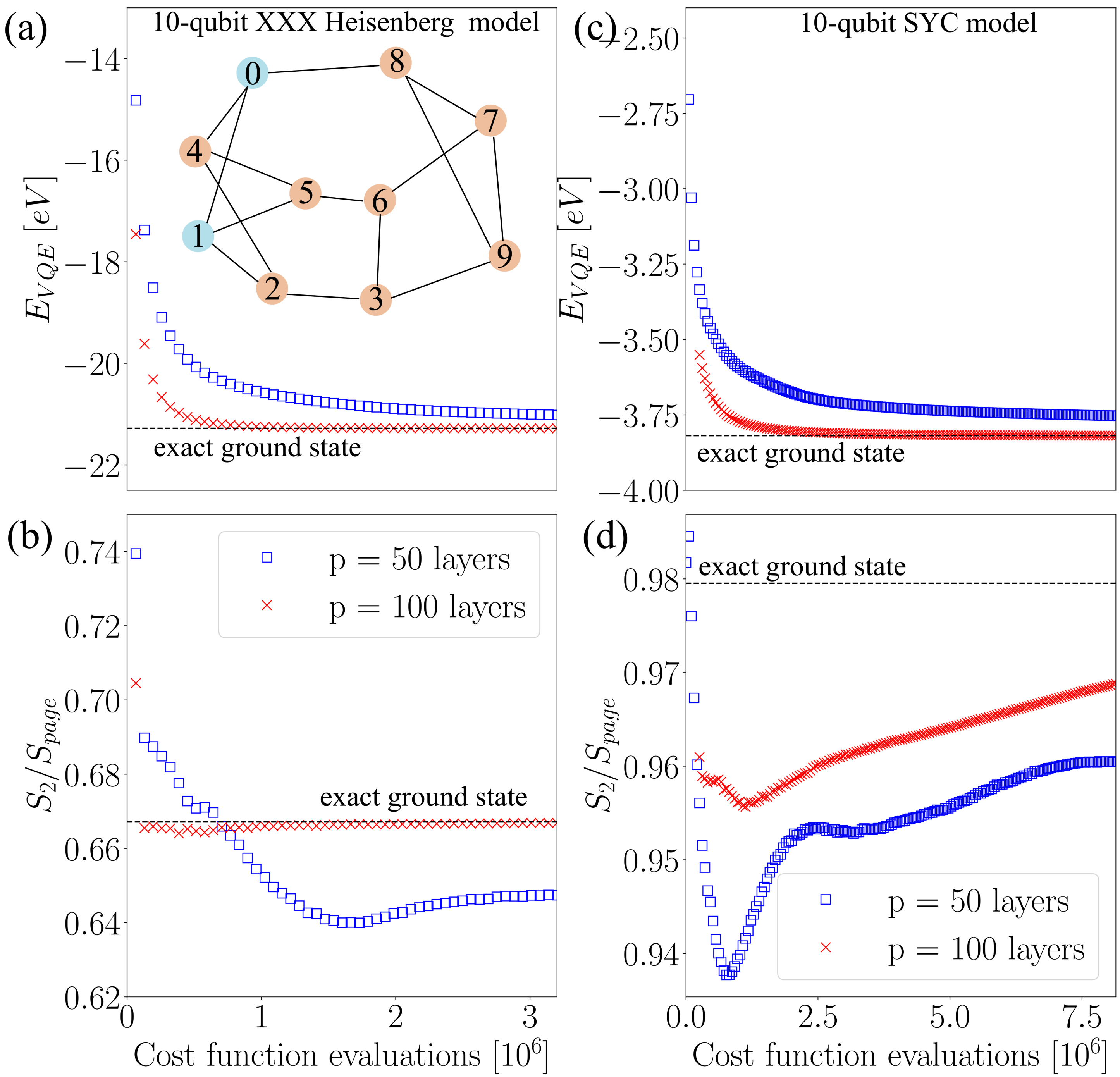
$$\{\chi_i, \chi_j\} = \delta_{i,j}$$

$J_{i,j,k,l}$ is taken from a Gaussian distribution

ELTE
EÖTVÖS LORÁND
UNIVERSITY

wigner

10-qubit system VQE



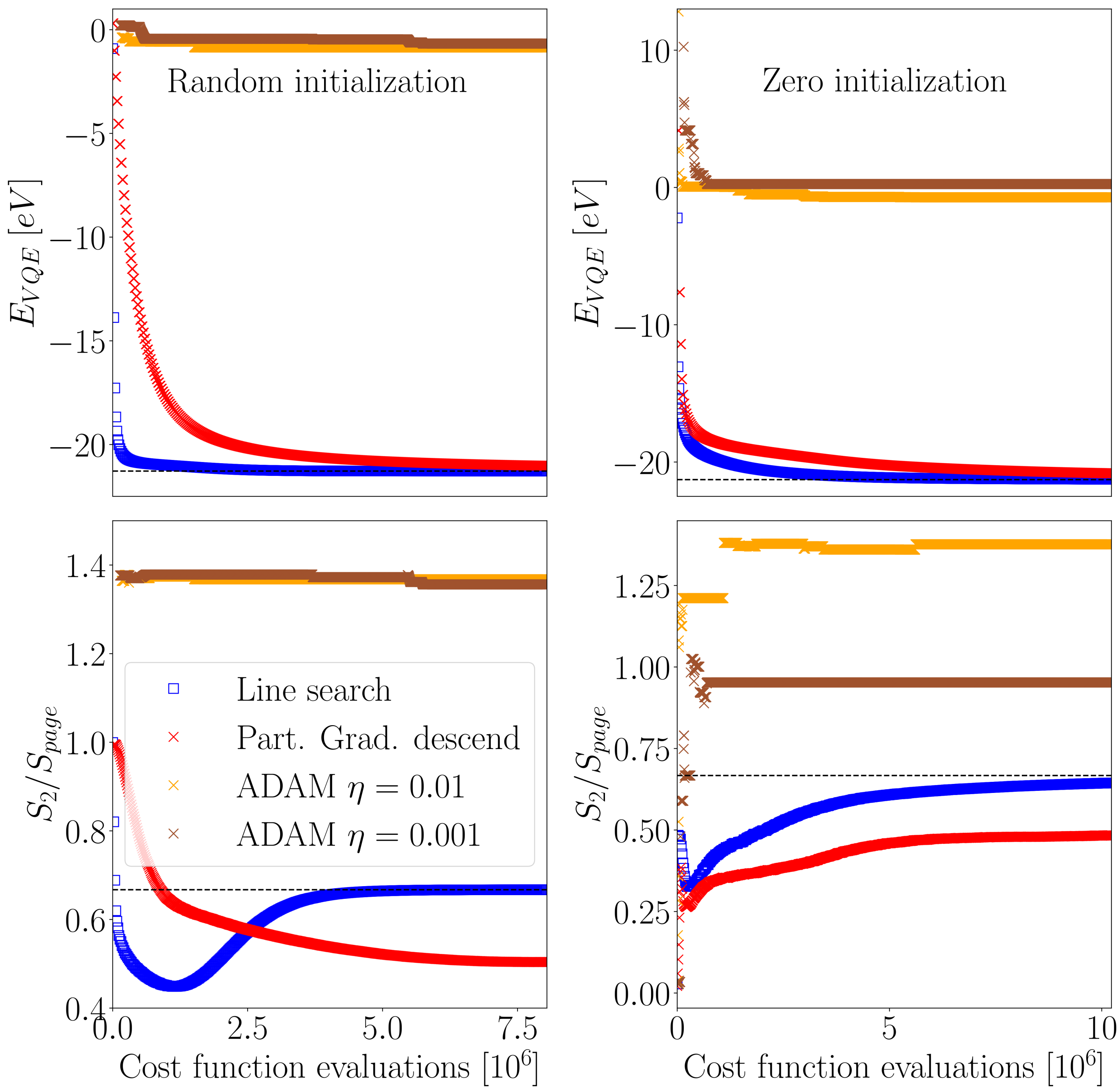
10-qubit wave-function:
2046 free parameters

circuit with 50-layers:
2700 free parameters

circuit with 100 layers:
5400 free parameters

10-qubit system VQE

Quantum Information National Laboratory HUNGARY



10-qubit wave-function:
2046 free parameters

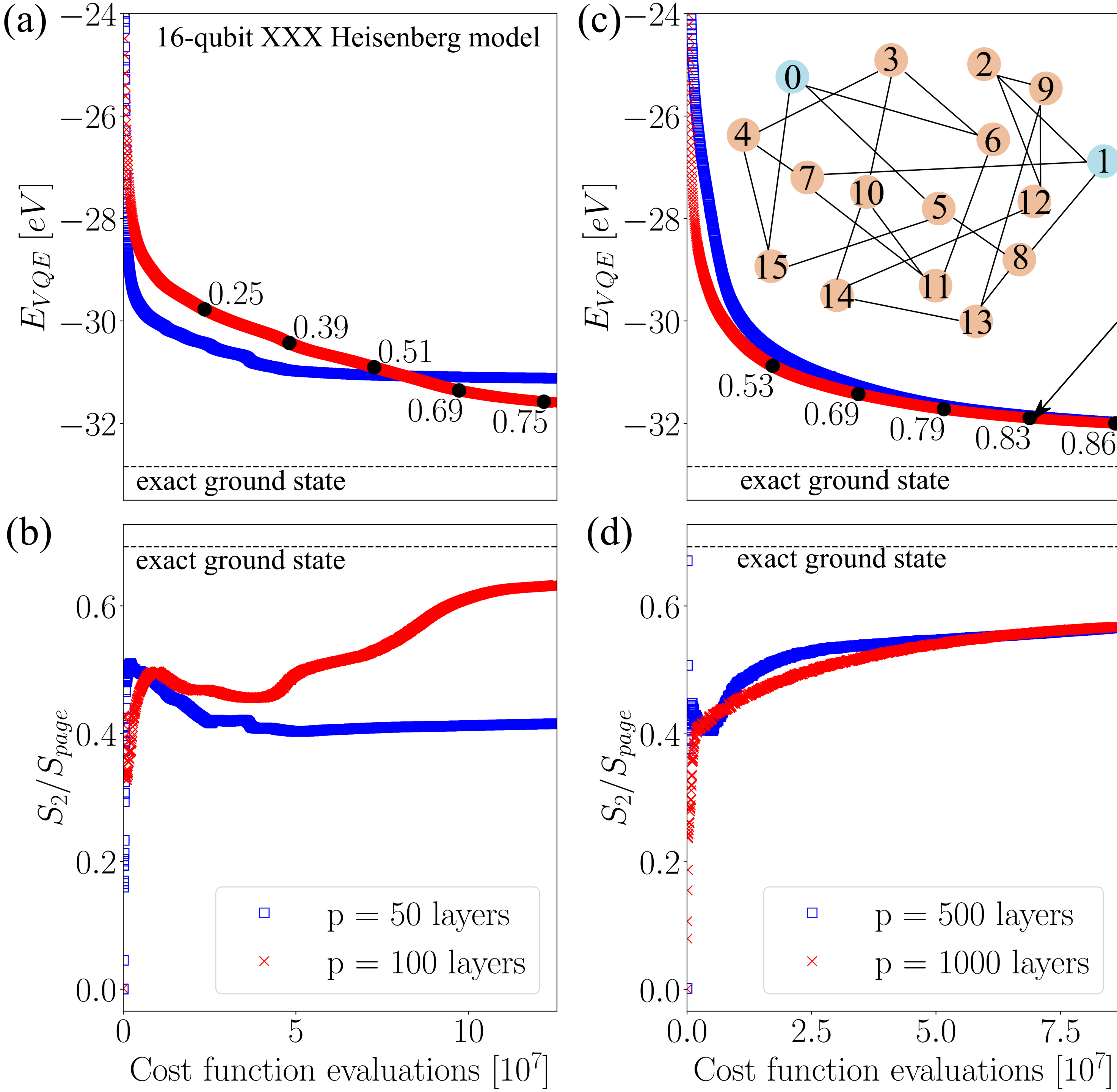
circuit with 50-layers:
2700 free parameters

circuit with 100 layers:
5400 free parameters

ELTE EÖTVÖS LORÁND UNIVERSITY

Wigner

16-qubit system VQE



$$M = |\langle \Psi_0 | \Psi(\boldsymbol{\theta}) \rangle|^2$$

16-qubit wave-function:
131070 free parameters

circuit with 100-layers:
9000 free parameters

circuit with 500-layers:
45000 free parameters

circuit with 1000-layers:
90000 free parameters

Further numerical improvements to scale up VQE experiments

- More efficient circuit ansatz adopted to
  - the coupling structure of the underlying model
  - the interaction types in the physical system
- Introducing noise into the model
- Using accelerators to evaluate the cost function

GPU

or

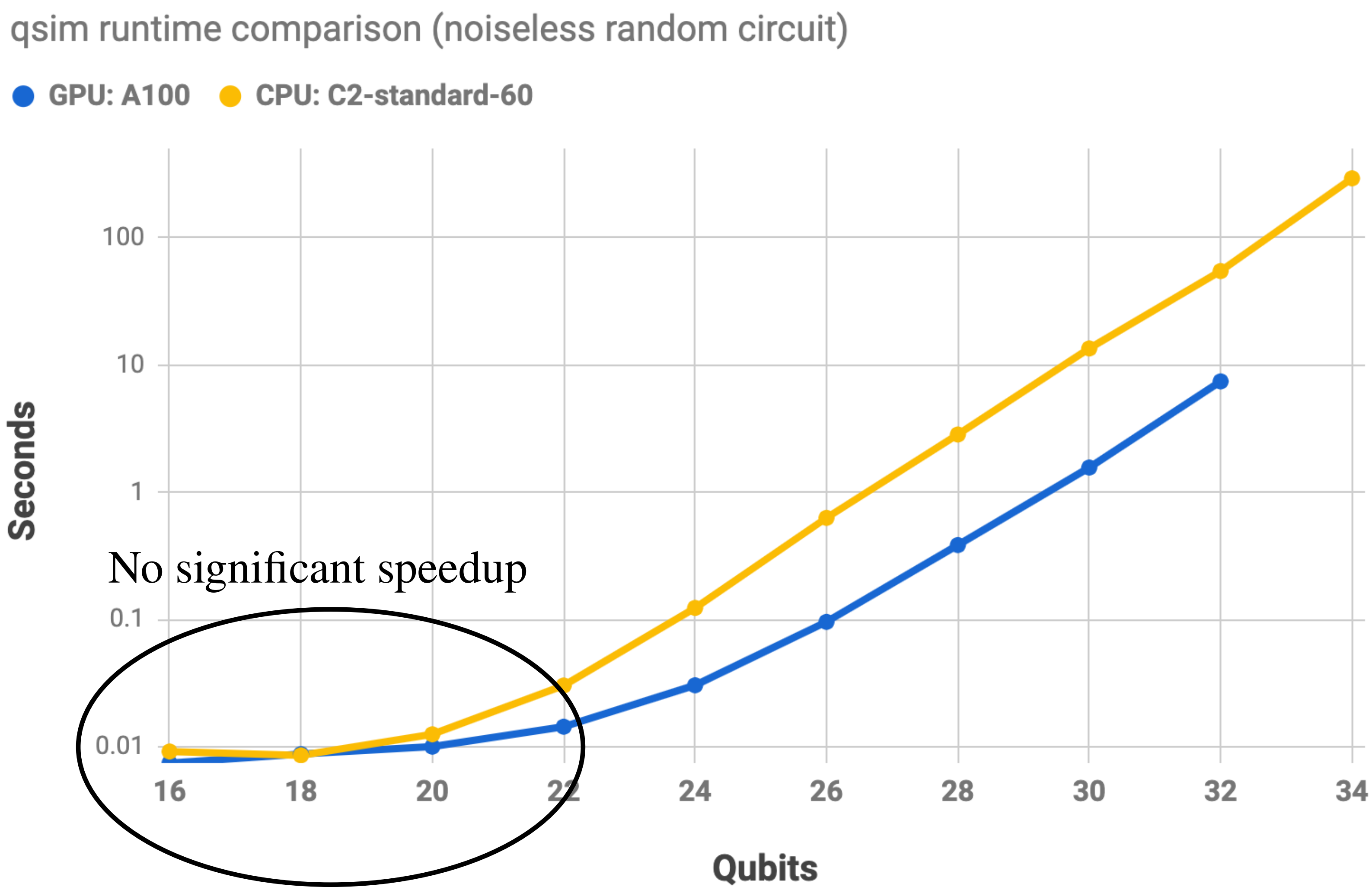
FPGA

or

AI data-flow chip

GPU QC simulation benchmark

Quantum Information National Laboratory HUNGARY



Google Quantum AI

https://quantumai.google/qsim/choose_hw

ELTE EÖTVÖS LORÁND UNIVERSITY wigner groq MAXELER Technologies Maximum Performance Computing

Performance benchmark for Groq LPU

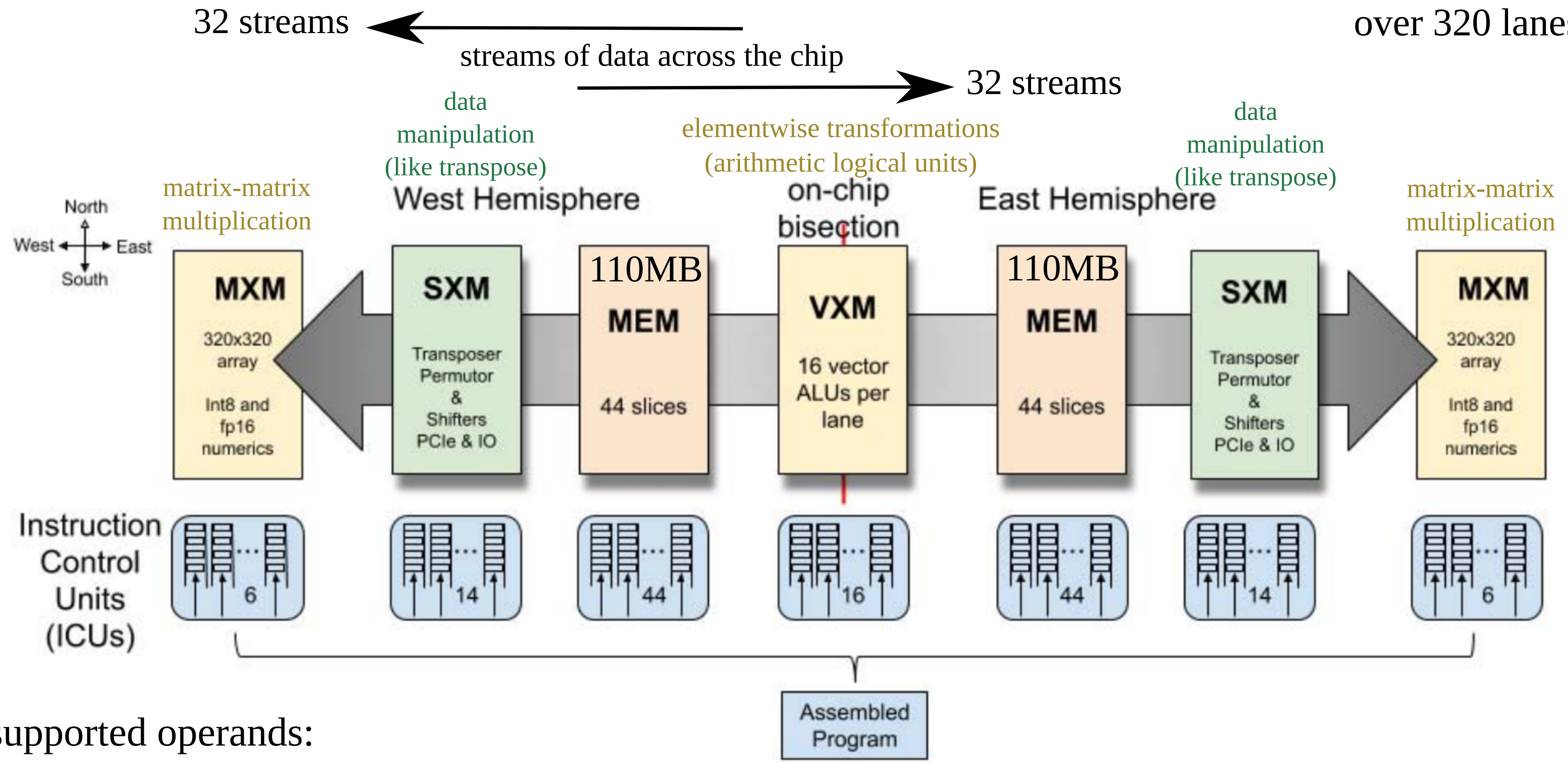
qubits	16		18		20	
performance	time	LPU speedup	time	LPU speedup	time	LPU speedup
gate count in the circuit	3601		4591		5701	
LPU implementation (estimate)	1.7 ms		8.2 ms		39 ms	
Qulacs	94 ms	55x	175 ms	21x	574 ms	14.7x
Qiskit	5784 ms	3402x	7354 ms	897x	9656 ms	247x
Squander	97 ms	57x	198 ms	24x	628 ms	16.1x

Think Fast: A Tensor Streaming Processor (TSP) for Accelerating Deep Learning Workloads

2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)

Quantum Information National Laboratory HUNGARY

each stream carries 320 vectorized bytes over 320 lanes



supported operands:

VXM: int8, int16, int32, uint8, uint16, uint32, float16, float32, bool8, bool16, bool32

MXM: int8 x int8 → int32, float16 x float16 → float32

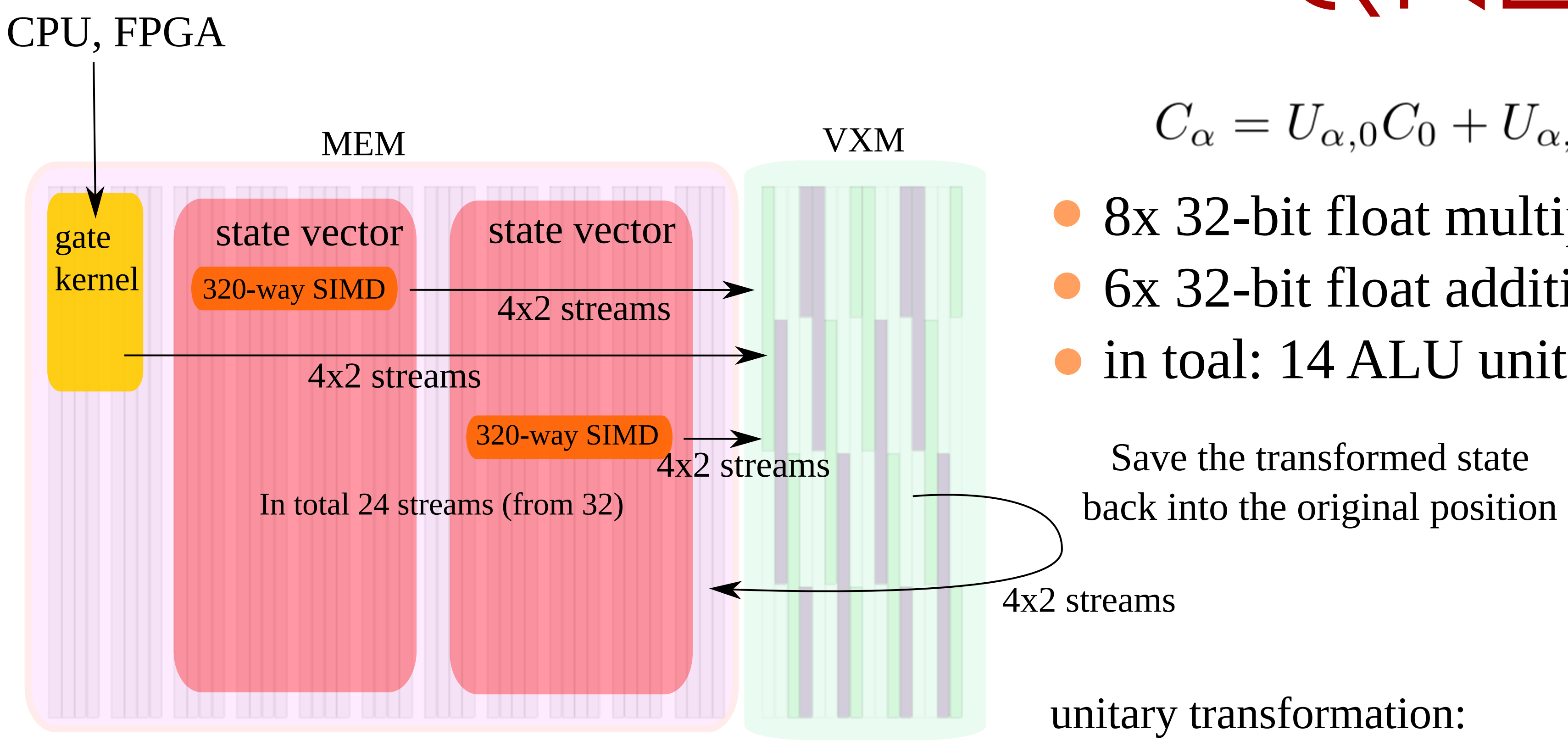
ELTE EÖTVÖS LORÁND UNIVERSITY

wigner

groq

MAXELER Technologies Maximum Performance Computing

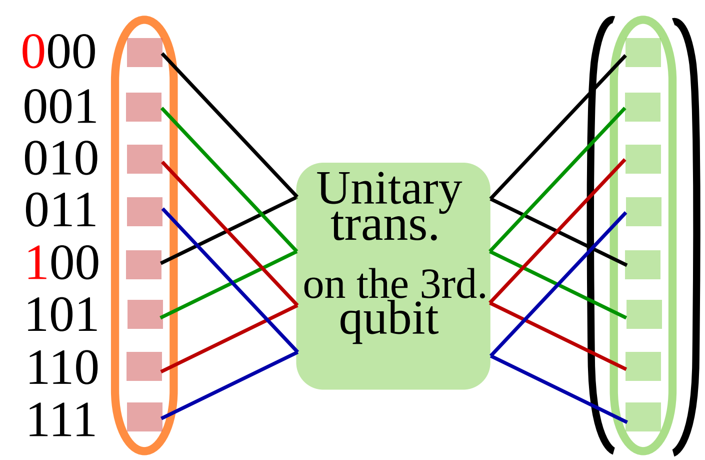
Concept of Groq QC simulator



$$C_\alpha = U_{\alpha,0}C_0 + U_{\alpha,1}C_1$$

- 8x 32-bit float multiplications
- 6x 32-bit float additions
- in toal: 14 ALU units (from 16)

unitary transformation:



transformed state

Aknowledgement

This research was supported by the Ministry of Innovation and Technology and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary and Grants No. 2022-2.1.1-NL-2022-00004, by the ÚNKP-24-5 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund, by the Hungarian Scientific Research Fund (OTKA) Grant No. K134437 and by the Hungarian Academy of Sciences through the Bolyai János Stipendium (BO/00571/22/11).

We acknowledge the computational resources provided by the Wigner Scientific Computational Laboratory (WSCLAB) (the formerWigner GPU Laboratory)

contact: Peter Rakyta, rakyta.peter@wigner.hu