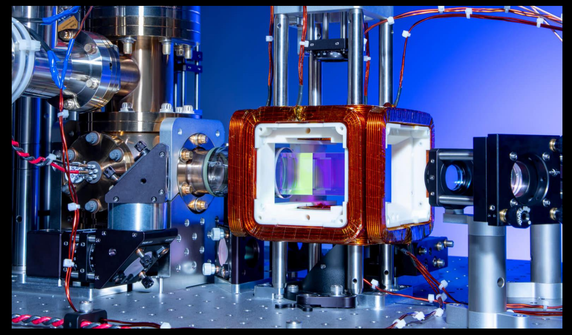
Defining the Accelerated Quantum Supercomputer

Naruhiko Tan, NVIDIA

Year of Quantum Breakthroughs

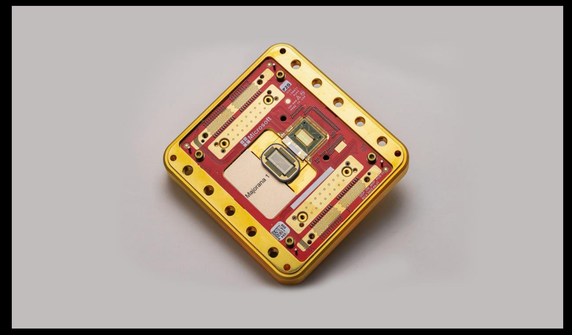
Quantum Error Correction

Google Quantum AI

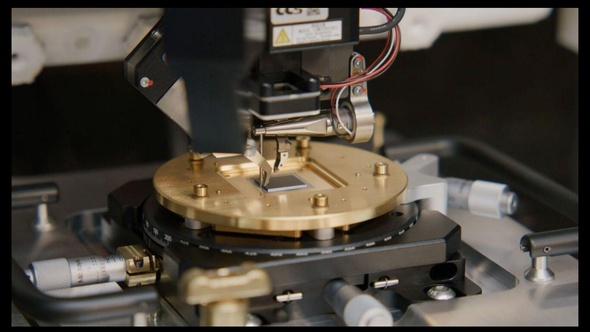


IQuEra

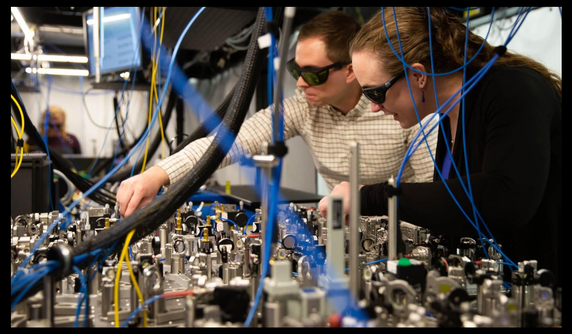
New Architectures



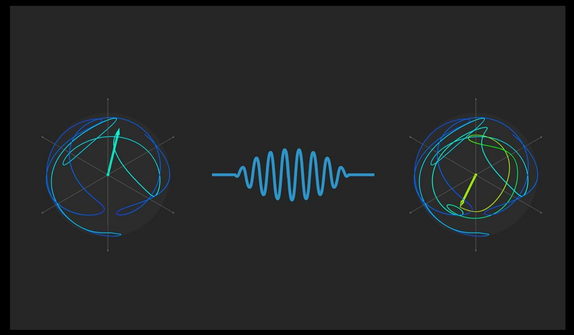
Microsoft



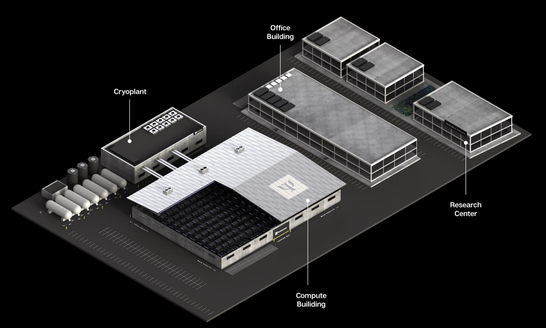
aws



QUANTINUUM



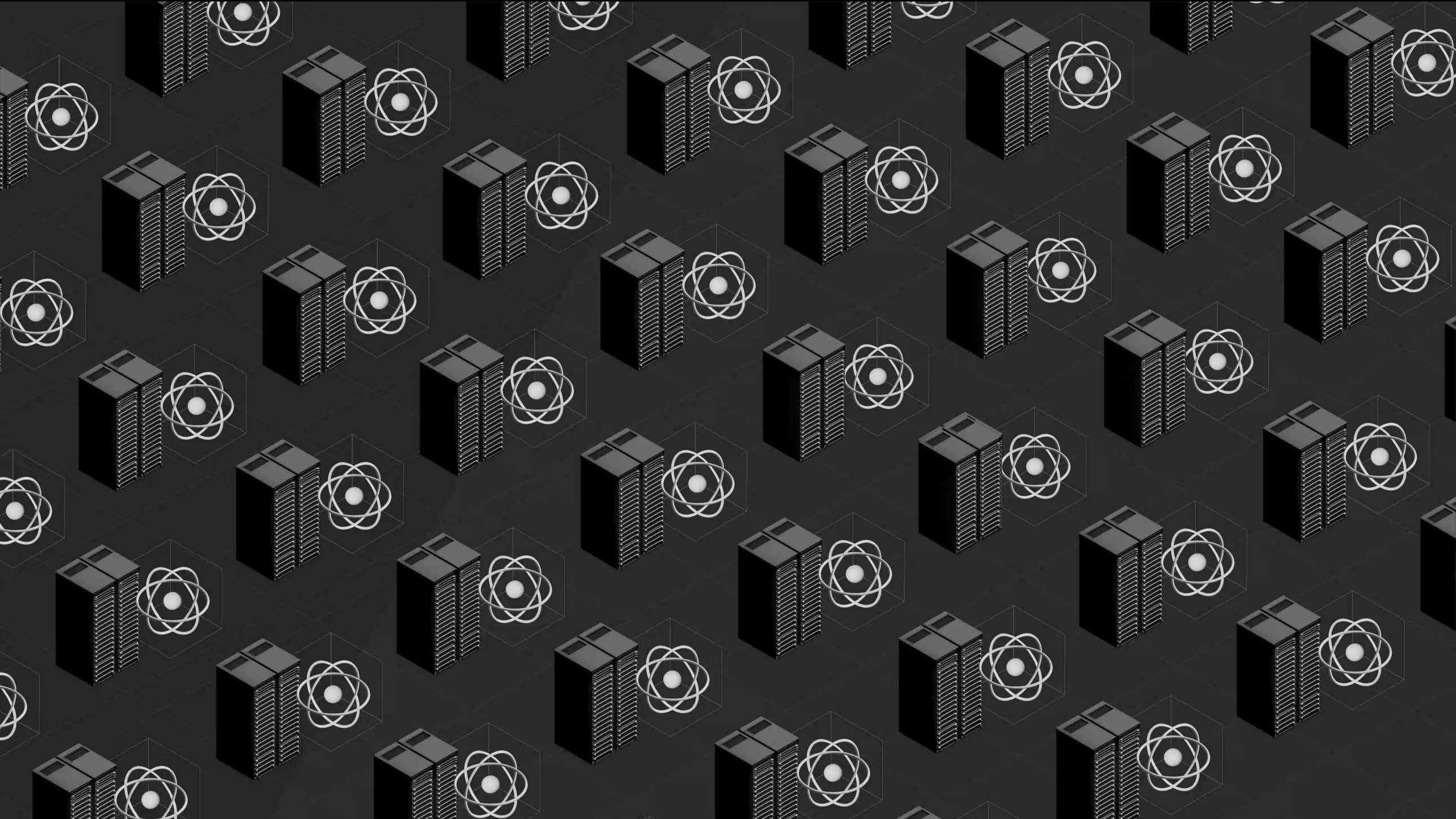
Infleqtion



PsiQuantum

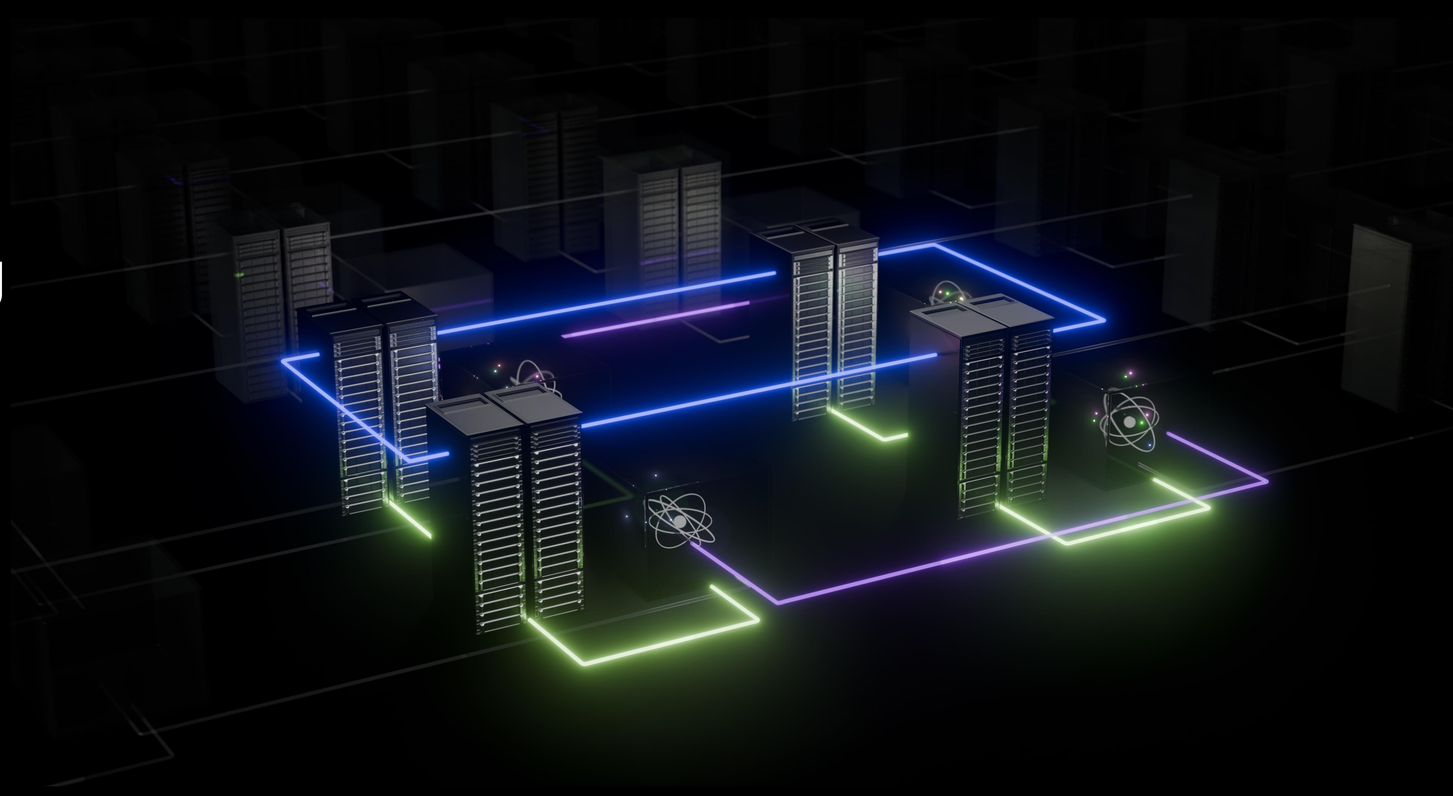
XANADU

NVIDIA



The Accelerated Quantum Supercomputer

- AI supercomputing architecture **connecting quantum hardware**
- **Hybrid applications** – combining HPC, AI and QC resources
- A **software platform** enabling domain scientists
- **Qubit-agnostic** development of control and error correction



The NVIDIA Quantum Platform

Bringing AI Supercomputing to Enable Useful Quantum Computing

CUDA-Q
Open-Source Software Platform for Accelerated Quantum Supercomputers

- Anyon Technologies
- aws
- Infleqtion
- IONQ
- IQM
- ORCA Computing
- OQC
- Pasqal
- QUANTINUUM
- QUANTUM BRILLIANCE
- QuEra Computing Inc.
- rigetti

The NVIDIA Quantum Platform

Bringing AI Supercomputing to Enable Useful Quantum Computing

CUDA-Q
Open-Source Software Platform for Accelerated Quantum Supercomputers

Pasqal — New at GTC25

QUANTINUUM — Free Access to H1-1 Through CUDA-Q

Anyon Technologies
aws
Infleqtion
IONQ
IQM
ORCA Computing
OQC
Pasqal
QUANTINUUM
QUANTUM BRILLIANCE
QuEra Computing Inc.
rigetti

The NVIDIA Quantum Platform

Bringing AI Supercomputing to Enable Useful Quantum Computing

CUDA-Q
Open-Source Software Platform for Accelerated Quantum Supercomputers

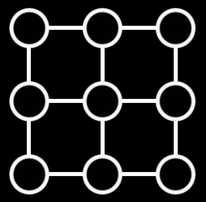
cuQuantum
Accelerated Quantum Simulations

Cirq
PENNYLANE
QIBO
Qiskit
Quest
Qulacs
TKET™
Torch Quantum

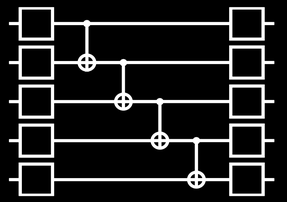
The NVIDIA Quantum Platform

Bringing AI Supercomputing to Enable Useful Quantum Computing

CUDA-QX
Turnkey Quantum Research and Application Development



CUDA-Q QEC



CUDA-Q Solvers

CUDA-Q
Open-Source Software Platform for Accelerated Quantum Supercomputers

cuQuantum
Accelerated Quantum Simulations

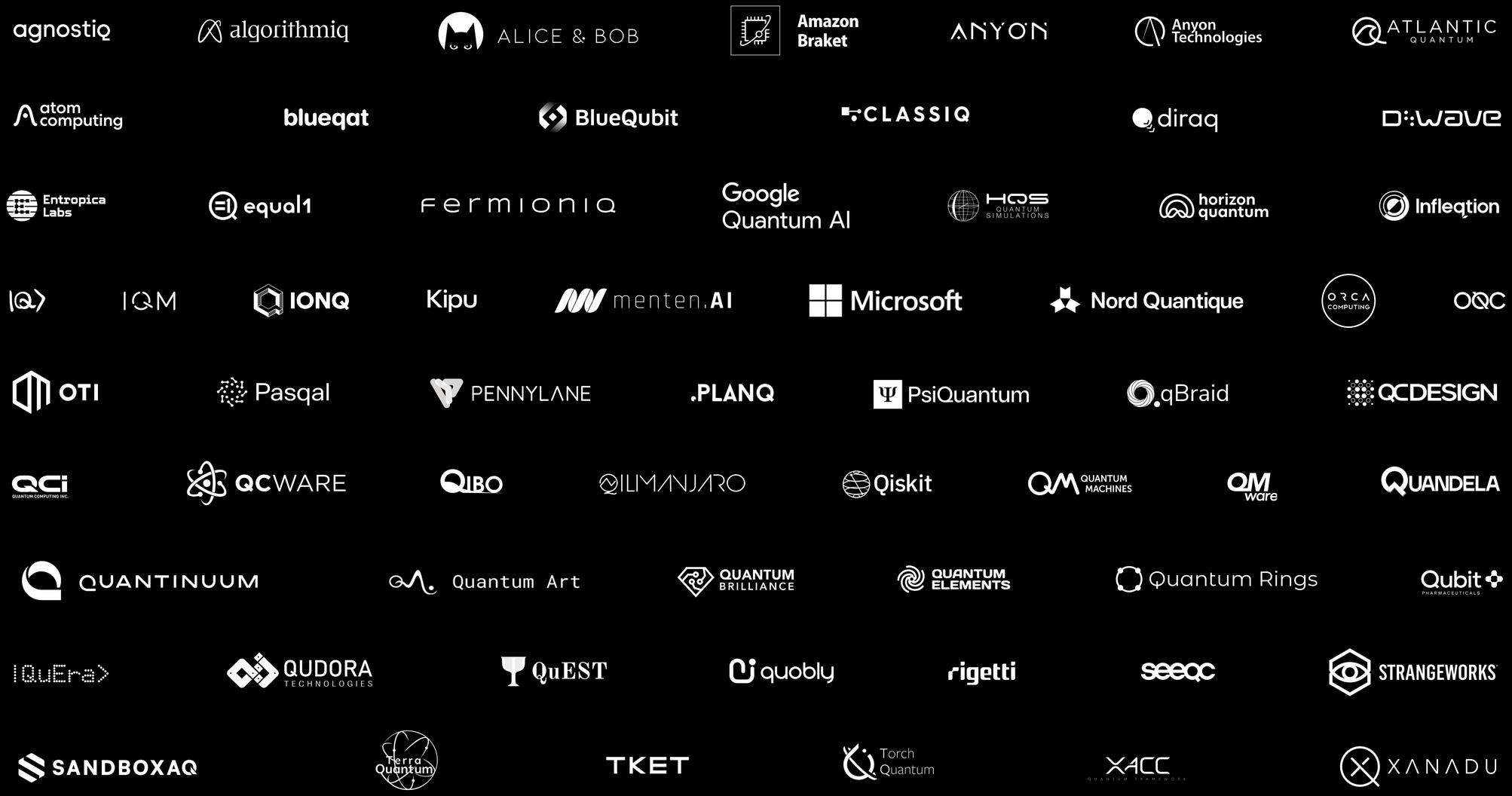
DGX Quantum
Reference Architecture for Low Latency Integration

QUANTUM MACHINES

NVIDIA

The NVIDIA Quantum Ecosystem

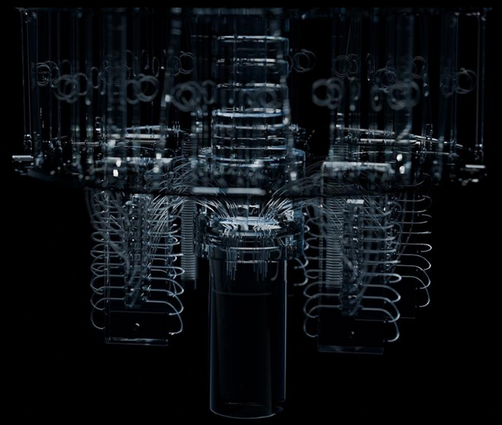
Accelerating the Quantum World

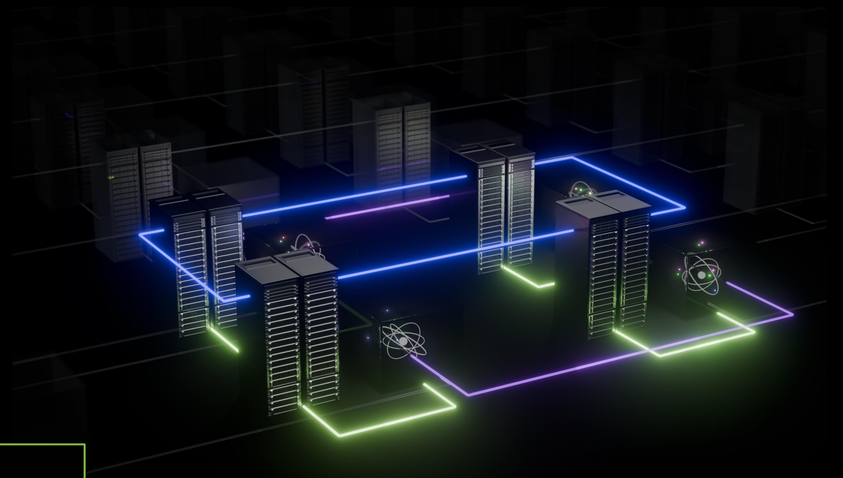


NVIDIA

Challenges Ahead

How to Turn Qubits into Accelerated Quantum Supercomputers





Qubits

Better QPUs

GPU

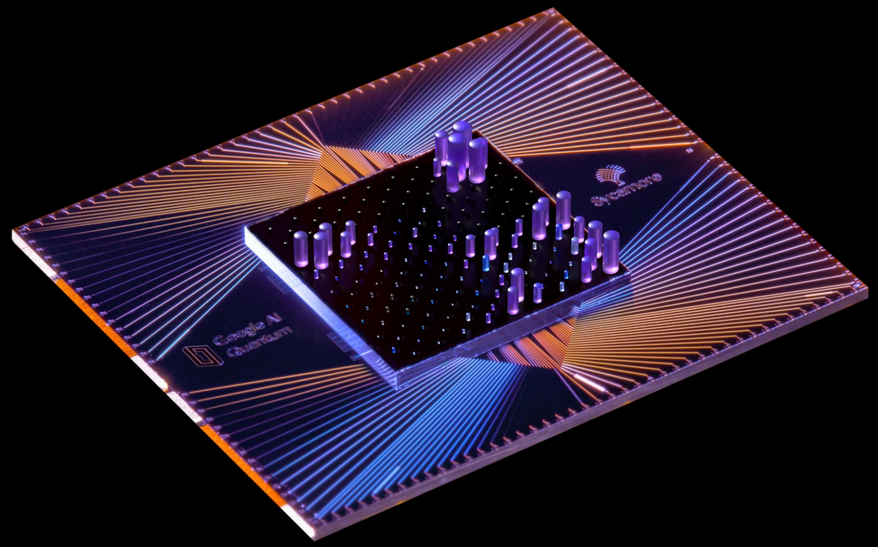
Better Algorithms

Scaling QEC

Accelerated Quantum Supercomputers

Developing Better QPUs with Simulation

NVIDIA | Google Quantum AI



Qubits

Better QPUs

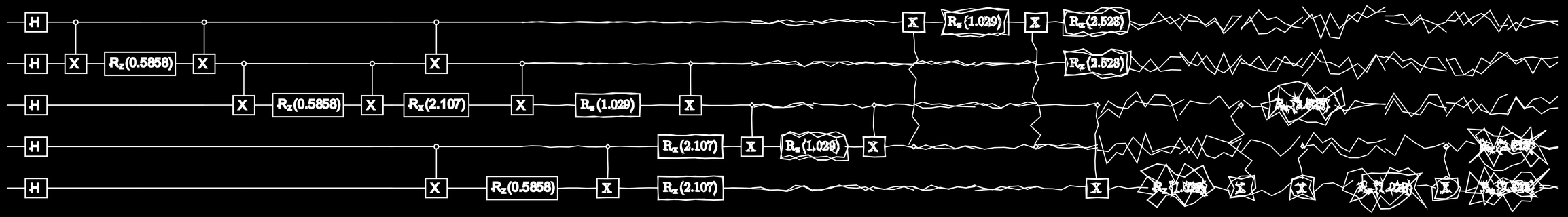
GPU

Better Algorithms

Scaling QEC

Accelerated Quantum Supercomputers

Noise Limits Today's Quantum Hardware

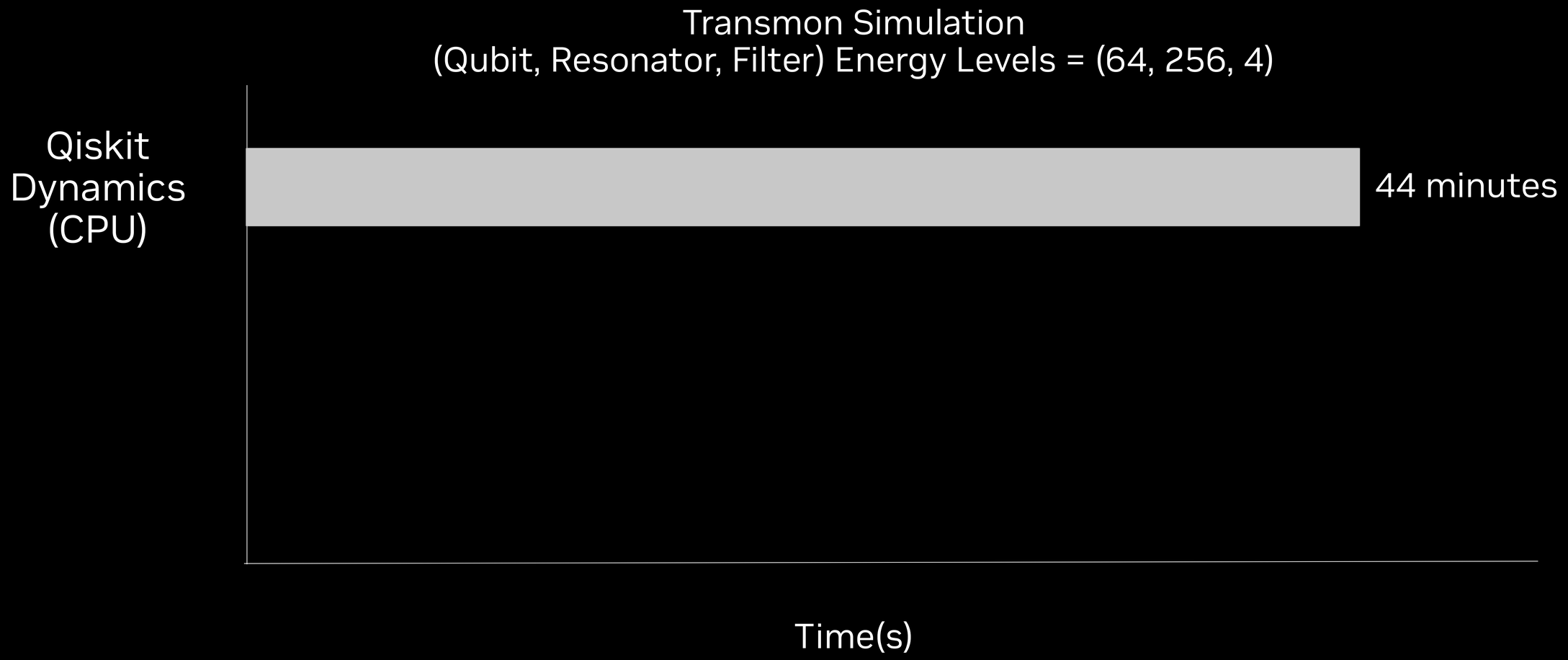


Current quantum hardware is limited to just hundreds of instructions by noise

Quantum
Processor

Dynamics in CUDA-Q

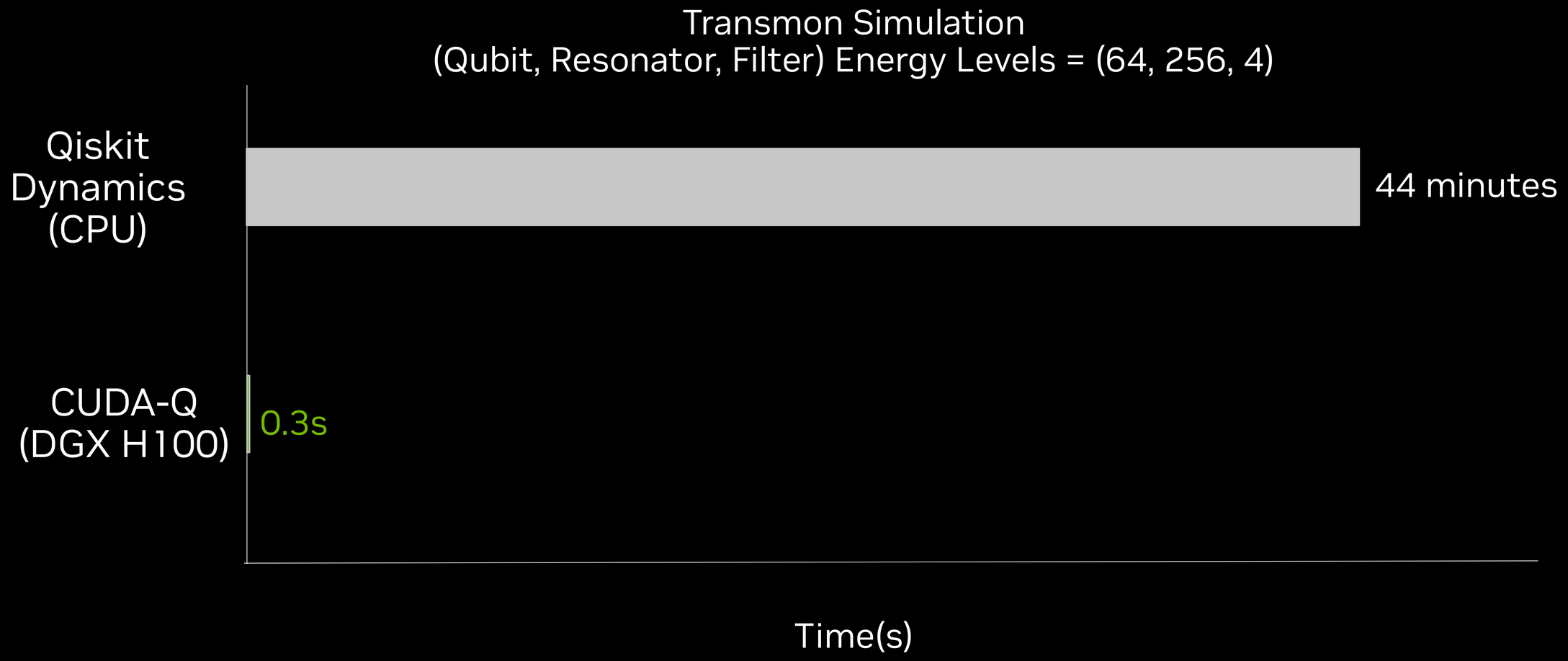
Enabling Better QPU Designs



CPU: Intel Xeon 8480CL | GPU: H100
Energy levels: 64 Transmon x 128 Resonator x 4 Purcell Filter

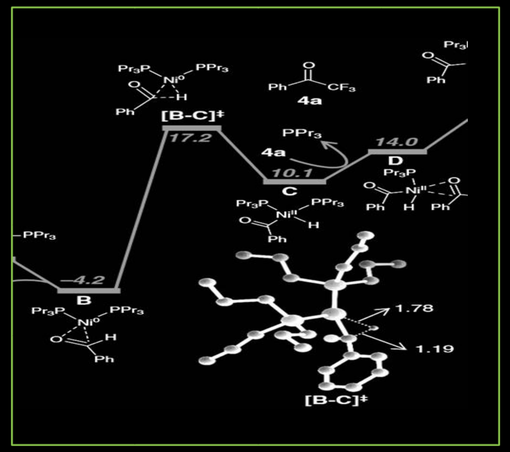
Dynamics in CUDA-Q

Enabling Better QPU Designs

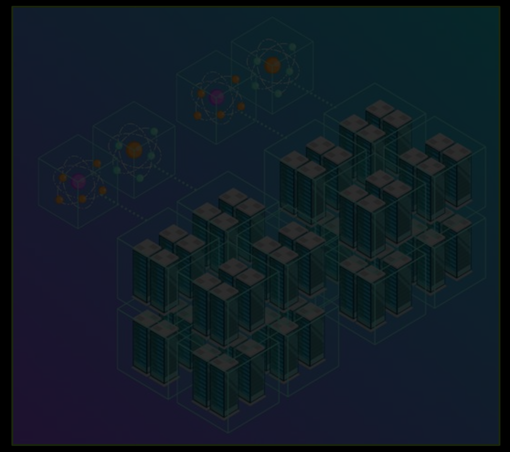


CPU: Intel Xeon 8480CL | GPU: H100
Energy levels: 64 Transmon x 128 Resonator x 4 Purcell Filter

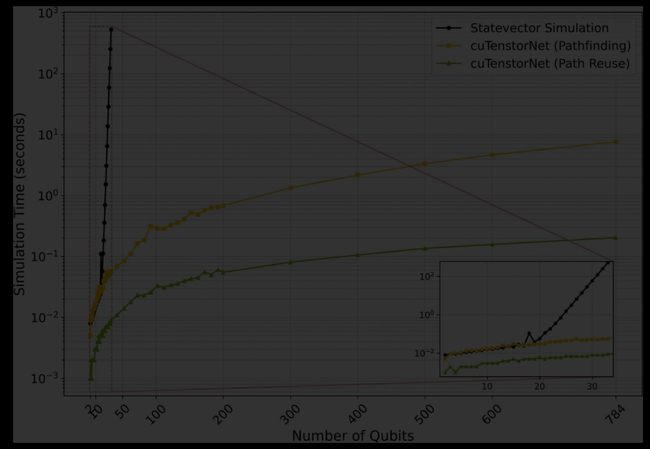
Better Algorithms with Large-Scale Simulations and AI



IonQ/AWS/AstraZeneca
AF-QMC simulation for electronic structure modeling



HPE Labs
Distributed QC with Adaptive Circuit Knitting



NCHC
784 qubit simulation to validate QML approaches

Qubits

Better QPUs

GPU

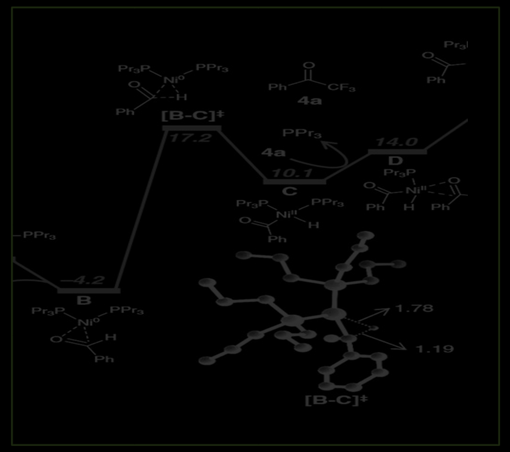
Better Algorithms

Scaling QEC

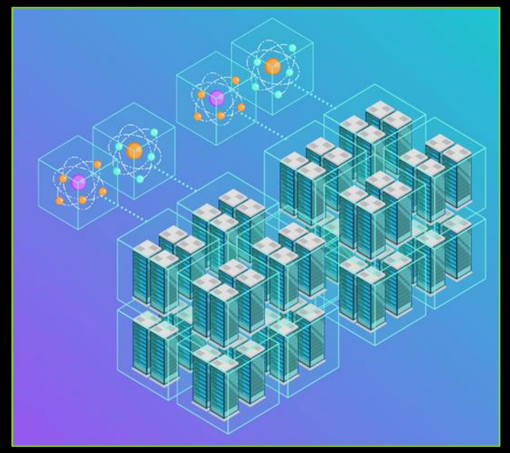
Accelerated Quantum Supercomputers

NVIDIA

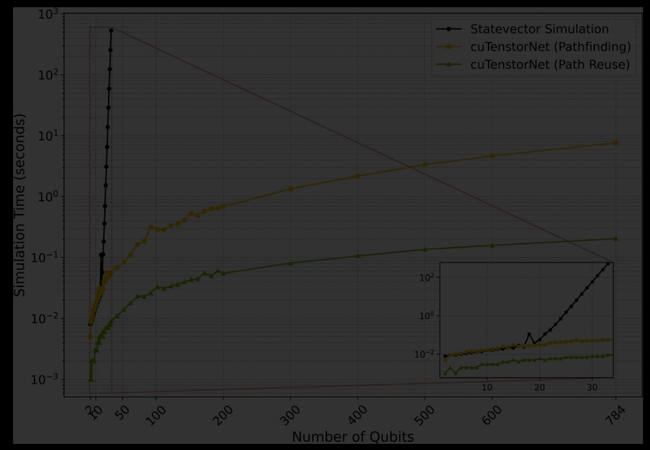
Better Algorithms with Large-Scale Simulations and AI



IonQ/AWS/AstraZeneca
AF-QMC simulation for electronic structure modeling



HPE Labs
Distributed QC with Adaptive Circuit Knitting



NCHC
784 qubit simulation to validate QML approaches

Qubits

Better QPUs

GPU

Better Algorithms

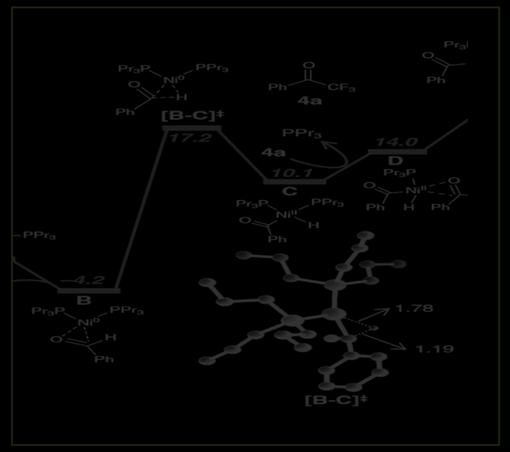
Scaling QEC

Accelerated Quantum Supercomputers

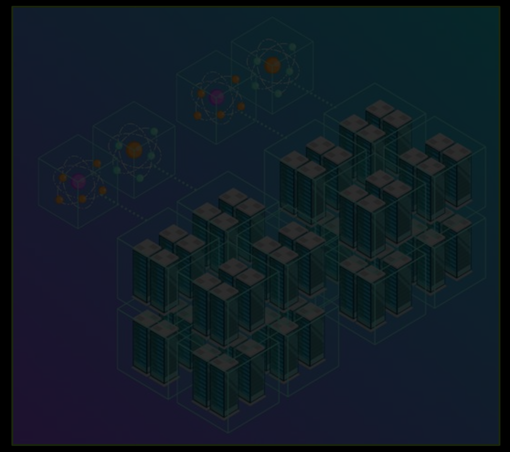
Mohseni *et. al*. How to Build a Quantum Supercomputer: Scaling Challenges and Opportunities
https://arxiv.org/html/2411.10406v1 (2024) GTC Talk DD73669

NVIDIA

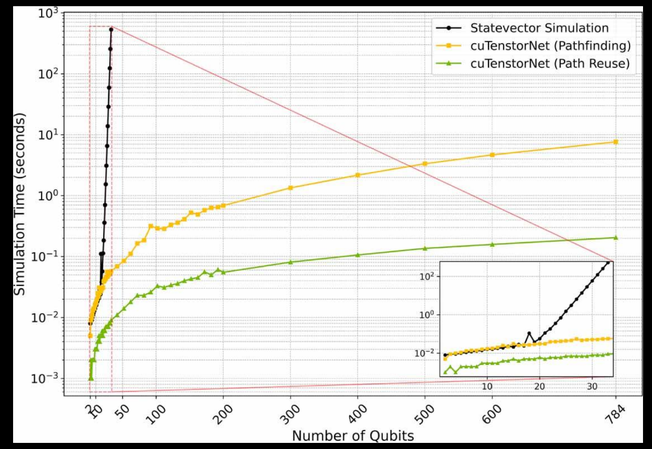
Better Algorithms with Large-Scale Simulations and AI



IonQ/AWS/AstraZeneca
AF-QMC simulation for electronic structure modeling



HPE Labs
Distributed QC with Adaptive Circuit Knitting



NCHC
784 qubit simulation to validate QML approaches

Qubits

Better QPUs

GPU

Better Algorithms

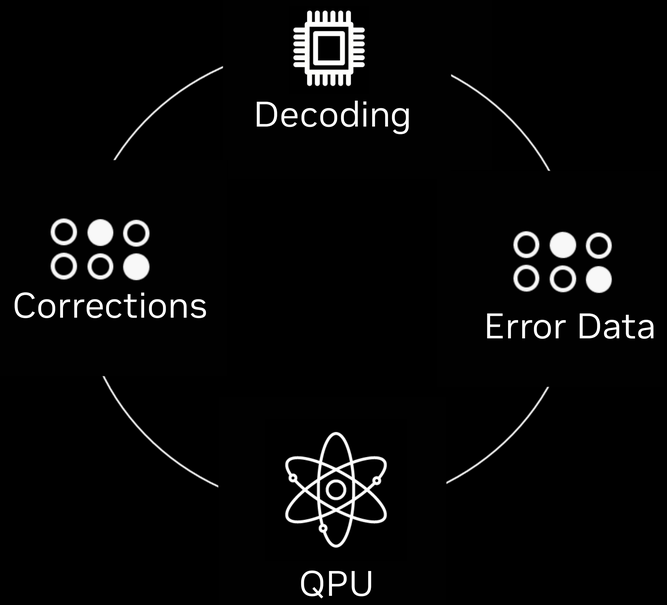
Scaling QEC

Accelerated Quantum Supercomputers

Chen *et. al*. cuTN-QSVM: cuTensorNet-accelerated Quantum Support Vector Machine with cuQuantum SDK. https://arxiv.org/html/2405.02630v2 (2024)

NVIDIA

Quantum Error Correction – The Challenge of Decoding



10^{-3}
State of the art error rates

$<10^{-10}$
Expected Error rates needed

Qubits

Better QPUs

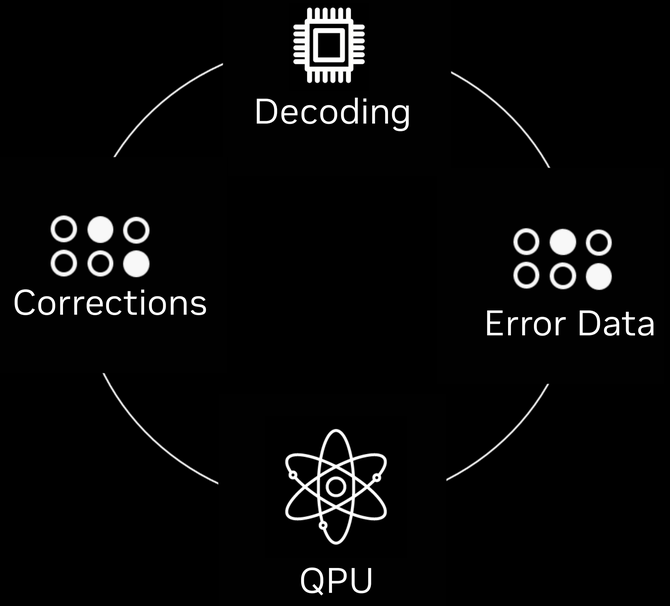
GPU

Better Algorithms

Scaling QEC

Accelerated Quantum Supercomputers

Quantum Error Correction – The Challenge of Decoding



10^{-3}
State of the art error rates

$<10^{-10}$
Expected Error rates needed

14,238
Qubits needed to factor a 2,048 bit number in 8 hours

22,325,184
Qubits needed with QEC

100TB/s
Data streaming from 1M qubits

Qubits

Better QPUs

GPU

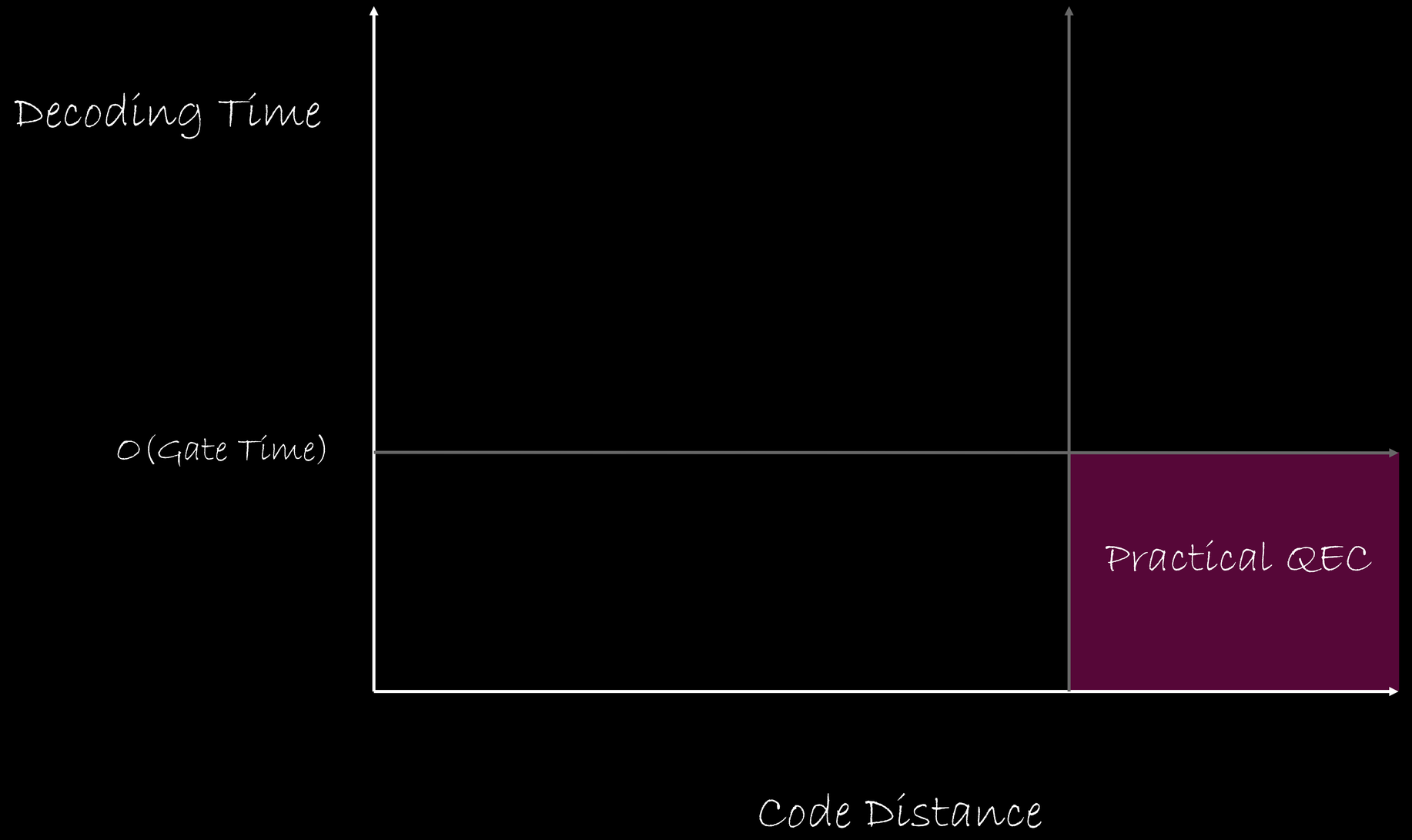
Better Algorithms

Scaling QEC

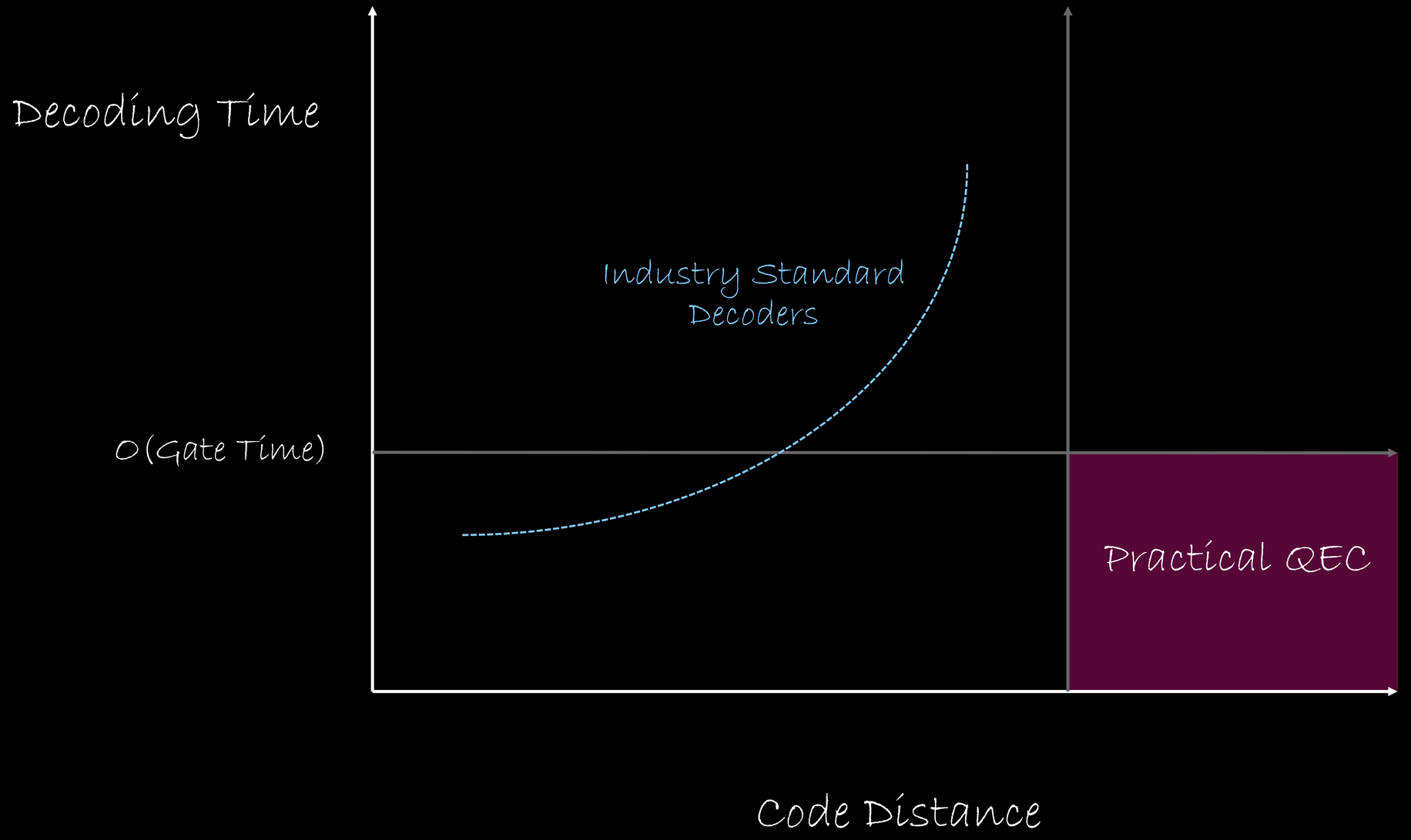
Accelerated Quantum Supercomputers

Gidney, Ekera. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. https://arxiv.org/abs/1905.09749 (2019)

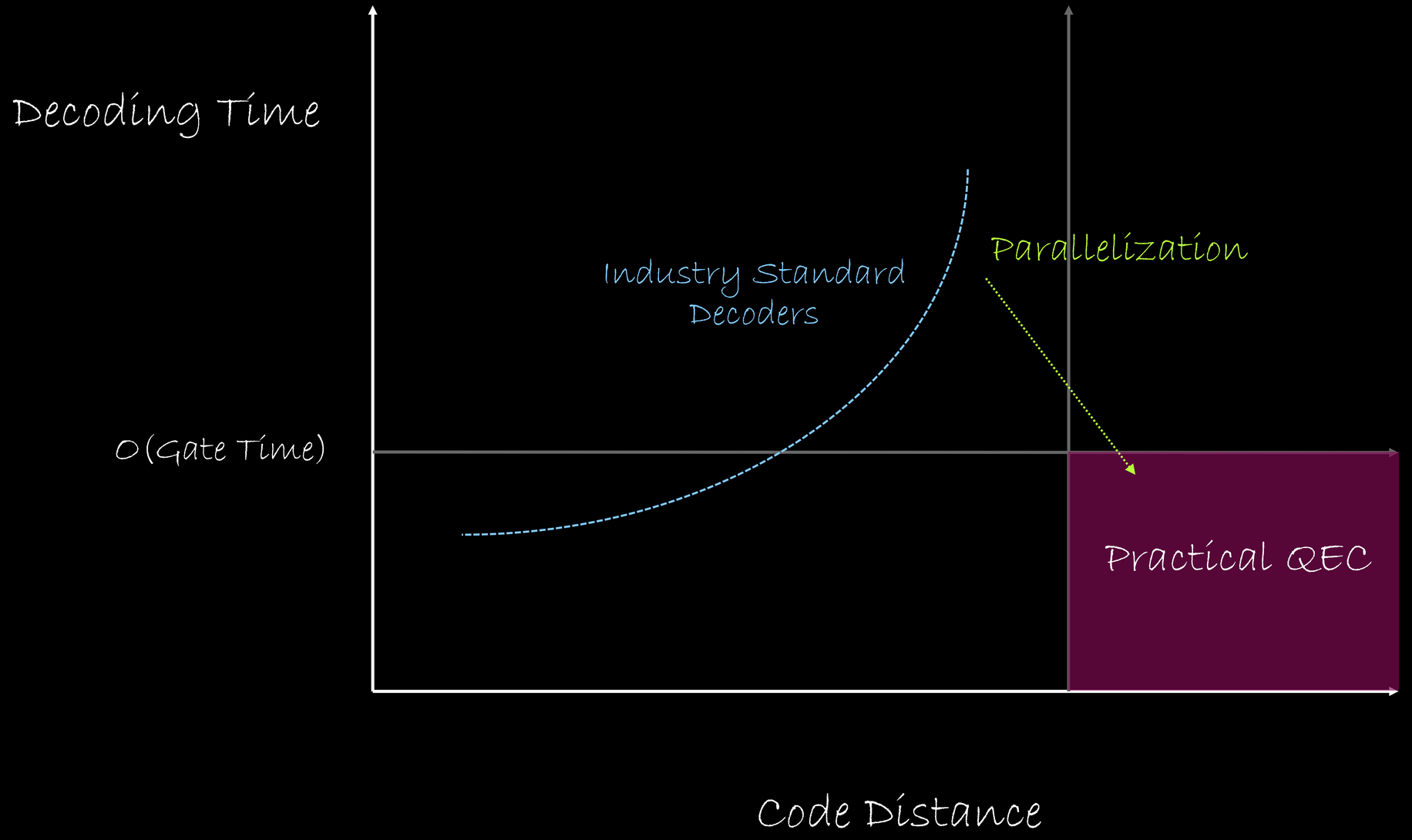
Quantum Error Correction – The Challenge of Decoding



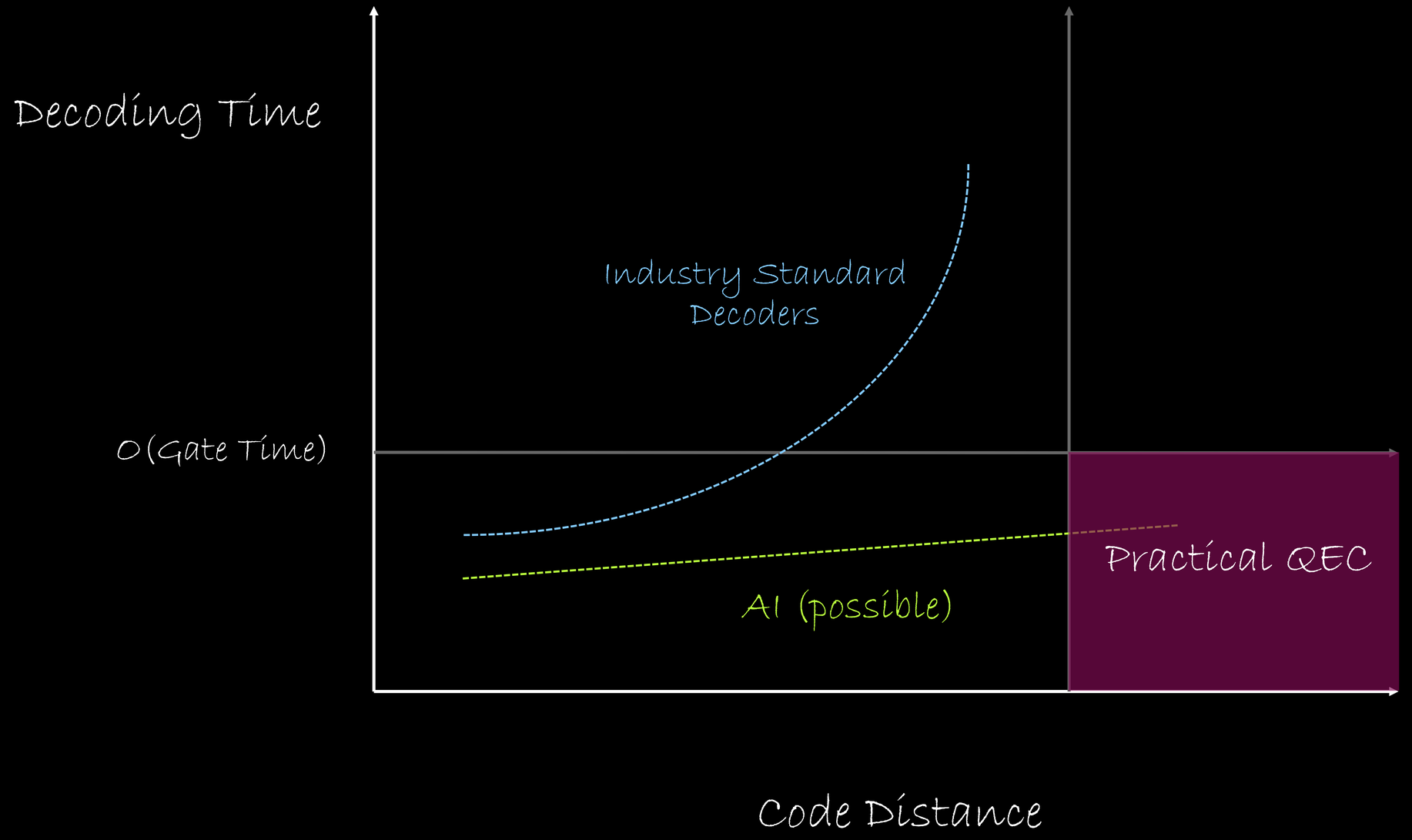
Quantum Error Correction – The Challenge of Decoding



Quantum Error Correction – The Challenge of Decoding



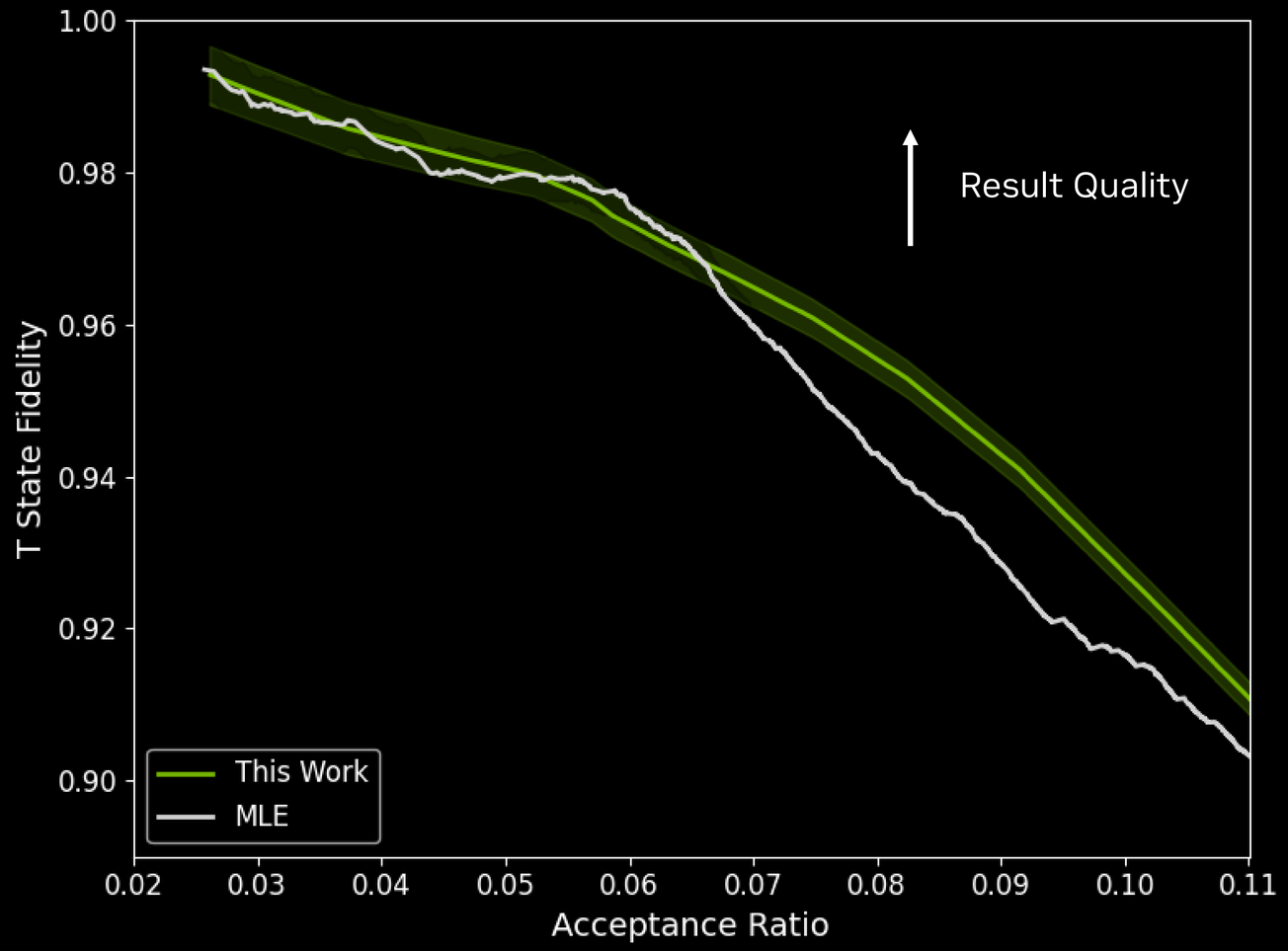
Quantum Error Correction – Why Quantum Computing Needs AI Supercomputing



Decoding QuEra's QPU with AI

Transformer-based decoder

- Trained on synthetic data with QuEra noise model
- Improved accuracy over MLE for small codes – d = 3
- Decoding time ~1ms, 50x improvement over SotA

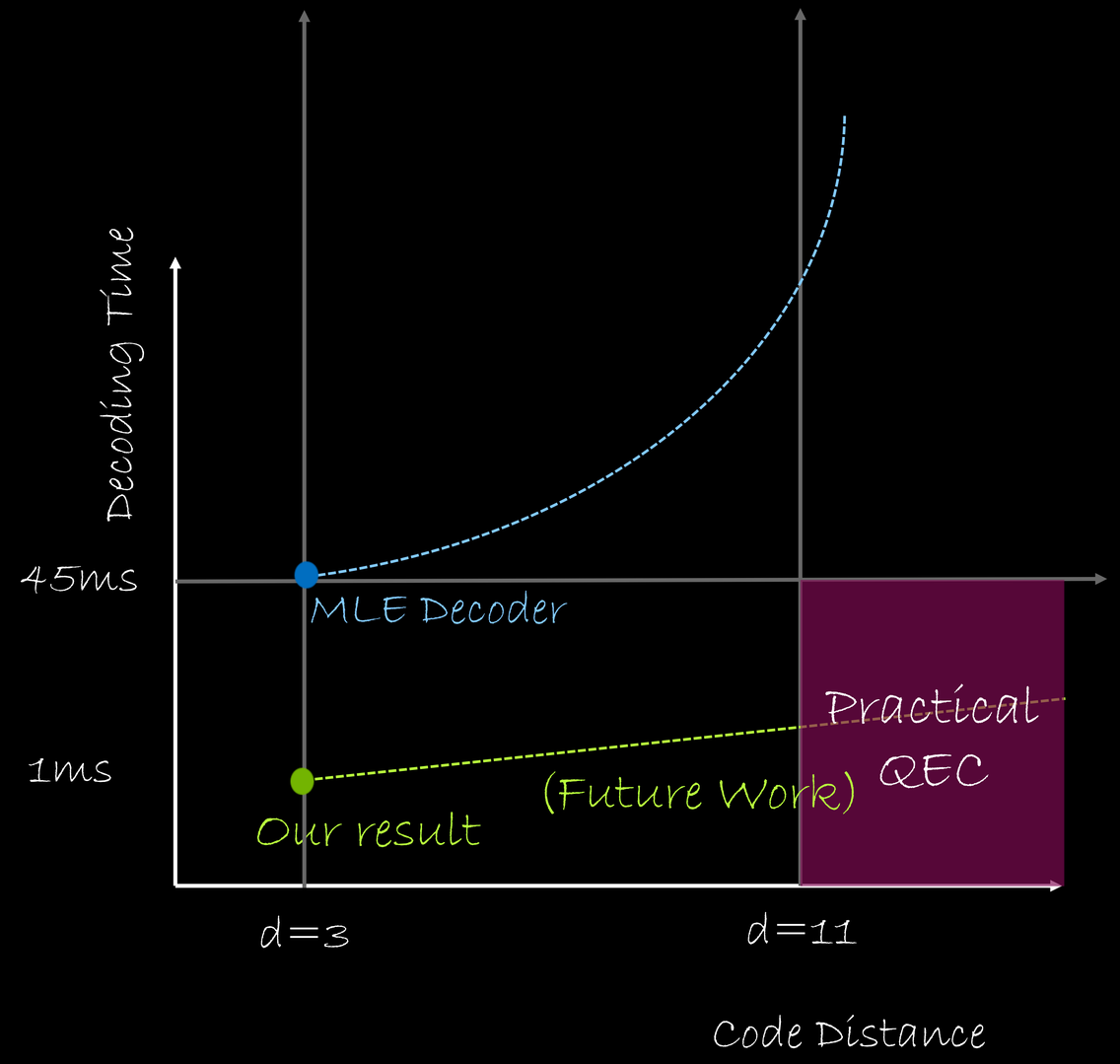


Decoding QuEra's QPU with AI

Transformer-based decoder

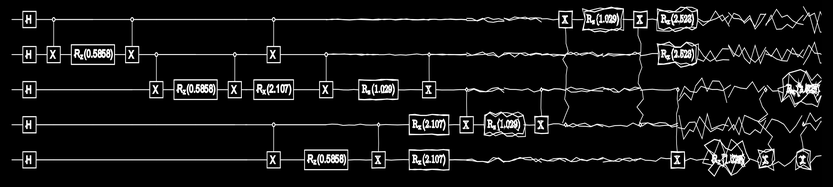
Future Work

- Larger codes – much more training data needed
- Fine-tune on QPU data for higher accuracy
- Deploy for real-time decoding



Building an AI Decoder

Custom Noise Model



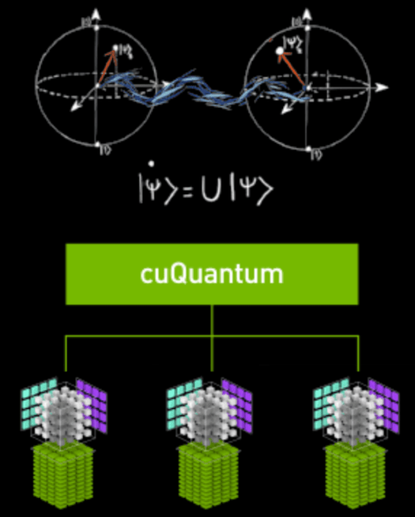
\+

Sampling Algorithm

Strategic Kraus Operator Sets

$\{K_0, \ldots, K_N\}$
$\{K_0, \ldots, K_N\}$
$\{K_0, \ldots, K_N\}$

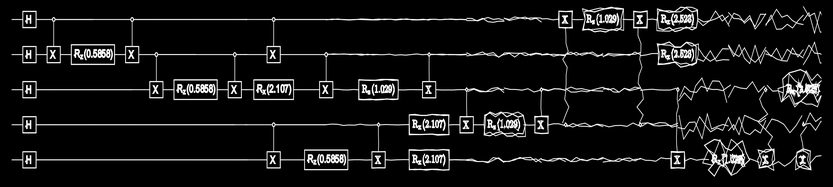
Accelerated Simulators



Synthetic Training Data

Building an AI Decoder

Custom Noise Model



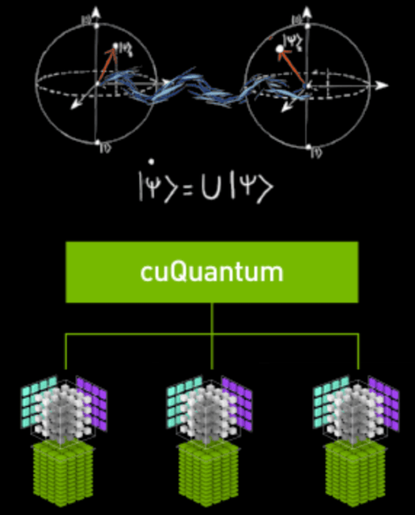
\+

Sampling Algorithm

Strategic Kraus Operator Sets

$\{K_0, \ldots, K_N\}$
$\{K_0, \ldots, K_N\}$
$\{K_0, \ldots, K_N\}$

Accelerated Simulators



Synthetic Training Data

On DGX H100 Supercomputer with 576 nodes:

- Emulated 85 qubit system – ***1 million shots*** generated in > 6 hours
- Emulated 35 qubit system – ***1 trillion shots*** generated in > 2 hours

Released open source in CUDA-Q 0.10

Announcing DGX Quantum Alpha Release

Integrated system for low-latency GPU-QPU programming

Building block for accelerated quantum supercomputer

<4us latency and real-time RL control demonstrated

Now shipping in Alpha

QUANTUM MACHINES

中央研究院 ACADEMIA SINICA

diraq

ENS DE LYON

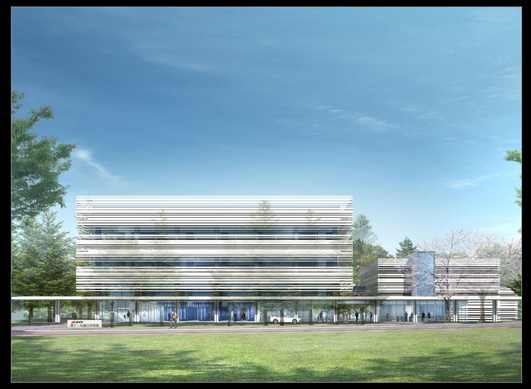
EQuS ENGINEERING QUANTUM SYSTEMS

Fraunhofer IZI

NVIDIA

Supercomputers are Driving Quantum Computing Progress

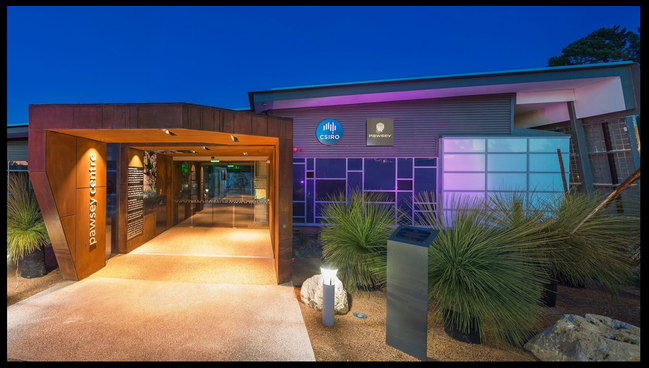
AIST - Japan
QuEra, Fujitsu, OptQC



Novo Nordisk Foundation - Denmark

Julich - Germany
IQM, D-Wave

Pawsey - Australia
Quantum Brilliance



PSNC - Poland
ORCA Computing

GENCI - France
Pasqal, Quandela

NQCC – UK
ORCA Computing, SEEQC

LRZ – Germany
AQT, IQM, planqc

Thank You