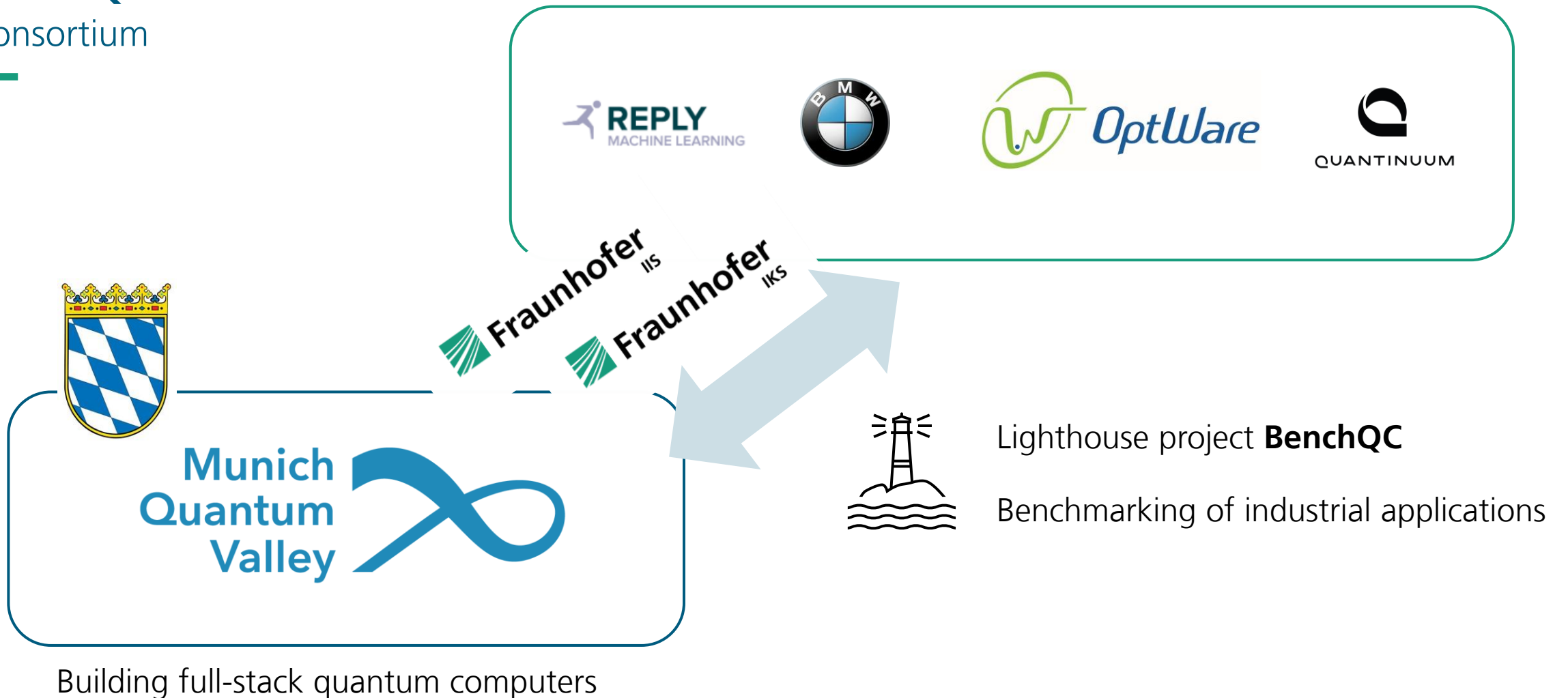


BenchQC – Scalable and modular benchmarking of industrial quantum computing applications

Speaker: **Dr. Florian Geissler** (Fraunhofer IKS)
QUEST-IS, Palaiseau, November 2025



Overview paper: <https://arxiv.org/abs/2504.11204>

Quantum computing benchmarking

An overview



Q Search docs

THE ZOO

Presentation

Acknowledgments

Contact us

Terminology

FIGURE OF MERIT (FOM)

Fidelities and errors

Application-level FOM

COMPONENT-LEVEL BENCHMARK

Qubit characterization

Welcome to Quantum Benchmark Zoo !

Quantum Benchmark Zoo aims to give an overview of the protocols and studies established to evaluate the **performance of quantum computers**. This is a living survey, so feel free to [contribute](#). Should you find any error or omission, either propose a pull request on our [GitHub](#) or [contact us](#). Any kind of help is welcome and will be [acknowledged](#).

Our concerns

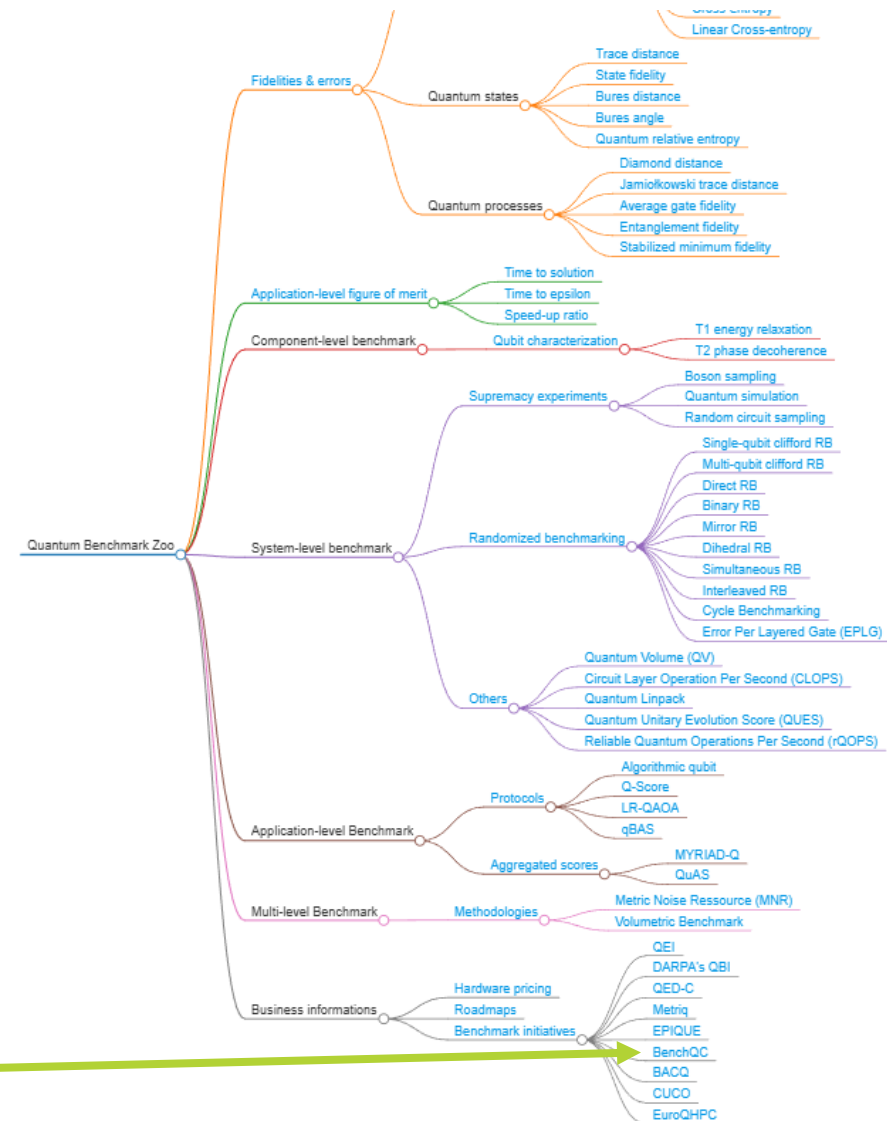
At the time of writing, quantum benchmarking protocols and experiments are flourishing, with several scientific papers published daily. This website aims to objectively describe each protocol to build a trustworthy catalog of benchmarking approaches. The second aim is to carefully report benchmark results using these protocols, verifying that the experiment complies with the protocol definition and reporting any variation in the protocol.

Map of the zoo

Click the node to reduce or expand it.

BenchQC:

- Application-centric (incl. QML)
- Tooling (QUARK)
- Holistic metrics



<https://quantumbenchmarkzoo.org/>

Emerging technology:

Quantum computing

With more powerful hardware emerging, quantum computing (QC) has the potential to lead to disruptive changes in many industrial areas:

- **Simulation** of quantum mechanical systems (development of new drugs, chemical sector with battery development,...)
- **Optimization** problems (Logistics, production, pharma,...)
- **Quantum machine learning** (Computer vision, mobility,...)

But for which applications will QC really offer an advantage in practice?

→ Application-driven benchmarking procedure required!

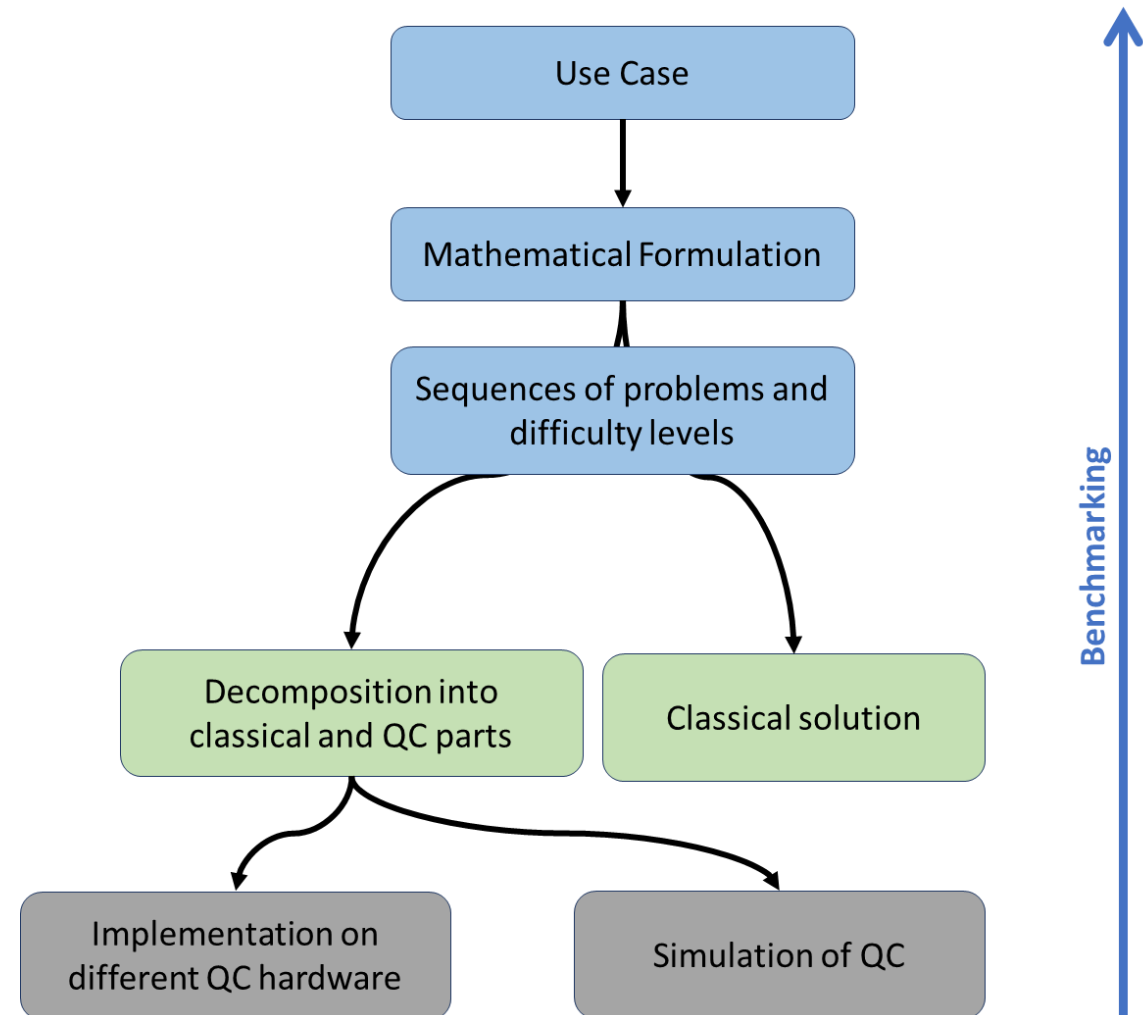


Application-driven benchmarking of quantum computers

Potential analysis from the application perspective

Application-driven benchmarking of QC:

- Identification of use cases suited for a **practical** quantum advantage.
- Consider mathematical **decomposition** and **sequence scaling** of problem
- Which **QC-hardware** technology (in which size, in which quality)?
- Considering the **full path** to the solution of a problem (e.g. including the data transfer to quantum computers, and post-processing of a quantum solution and integration into an industrial workflow)



```
plugins: ["quark_plugin_tsp", "quark_plugin_dwave"]

first_layer:
  &first_layer [
    "tsp_graph_provider": { nodes: 4, seed: 42 },
    "tsp_graph_provider": { nodes: 5, seed: 42 },
    "tsp_graph_provider": { nodes: 6, seed: 42 },
  ]

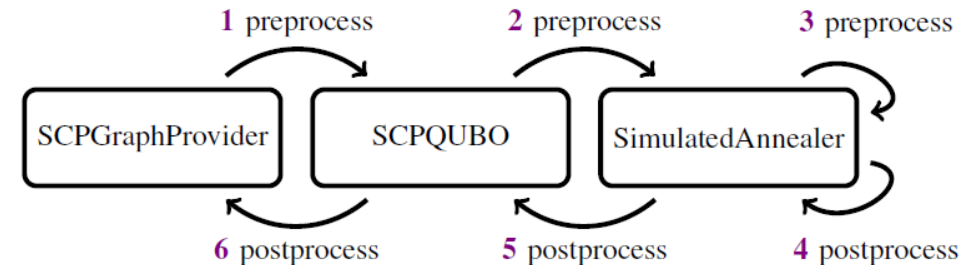
second_layer: &second_layer "tsp_qubo_mapping_dnx"

third_layer: &third_layer [
  "simulated_annealer_dwave": {num_reads: 1},
  "simulated_annealer_dwave": {num_reads: 1000},
]

pipeline1: &pipeline1 [
  *first_layer,
  *second_layer,
  *third_layer,
]

pipeline2: &pipeline2 [
  *first_layer,
  "classical_tsp_solver",
]

pipelines: [*pipeline1, *pipeline2]
```



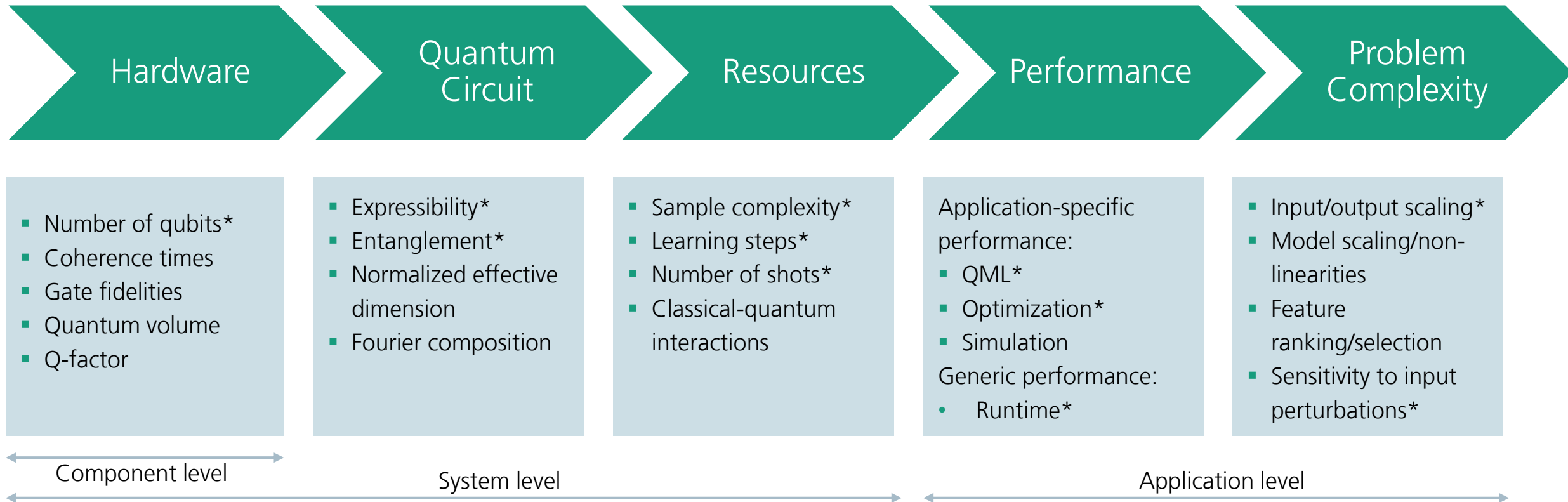
Key features

- Modular structure via **plugins**, application-agnostic core
- Simple **pipeline configuration** in single file
- Customization of **pre-/post-processing**
- Integrated **metrics**
- Diverse **hardware** interfaces
- **Datasets** handled by individual user

Benchmarking along the solution pipeline

Metrics

*Implemented in Quark
for selected applications



BenchQC

Use cases

Category	Model	Use case
	Classification	Defect detection
QML	Generative models	Generative design for constructional elements
		Data augmentation for quality inspection
	QRL	6G Beam management
Optimization	VQE, QAOA, Quantum Annealing	Multipath-connectivity across aerial vehicles
		Sensor placement in water networks
		Software verification strategies
		LiDAR sensor configuration
		Assembly line scheduling and packet sequencing
		Financial asset optimization
Simulation	Hamiltonian simulation	Dynamics of electrons in materials
		Low-temperature state preparation of materials
		Nuclear magnetic resonance simulation

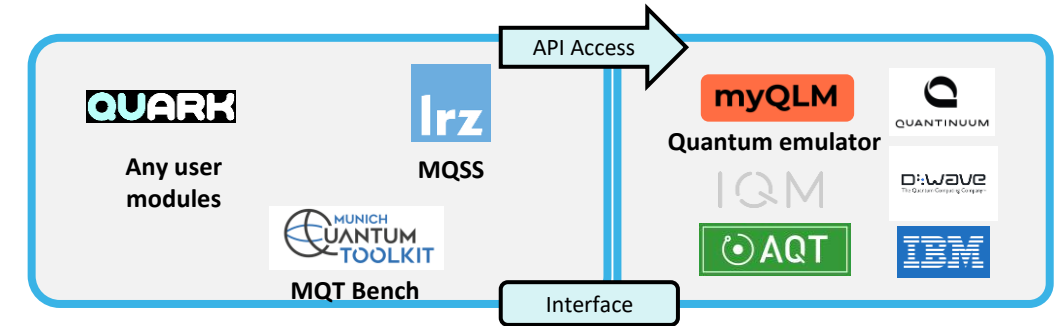


Road crack dataset: Binary classification problem, image size 224x224

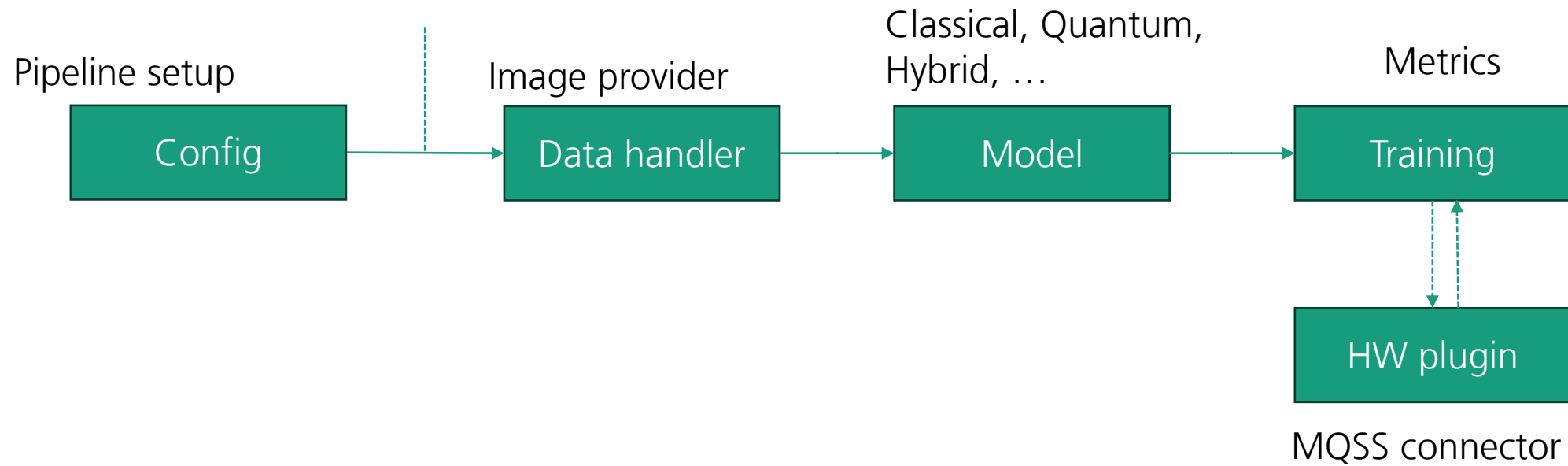
<https://arxiv.org/abs/2504.11204>

Quantum-Classical Image Classification (I)

Quark pipeline

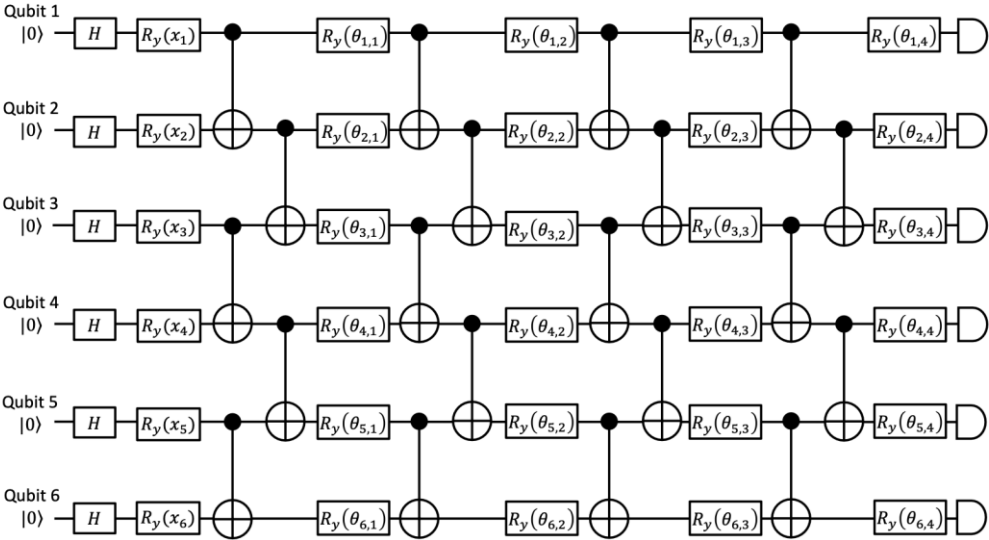
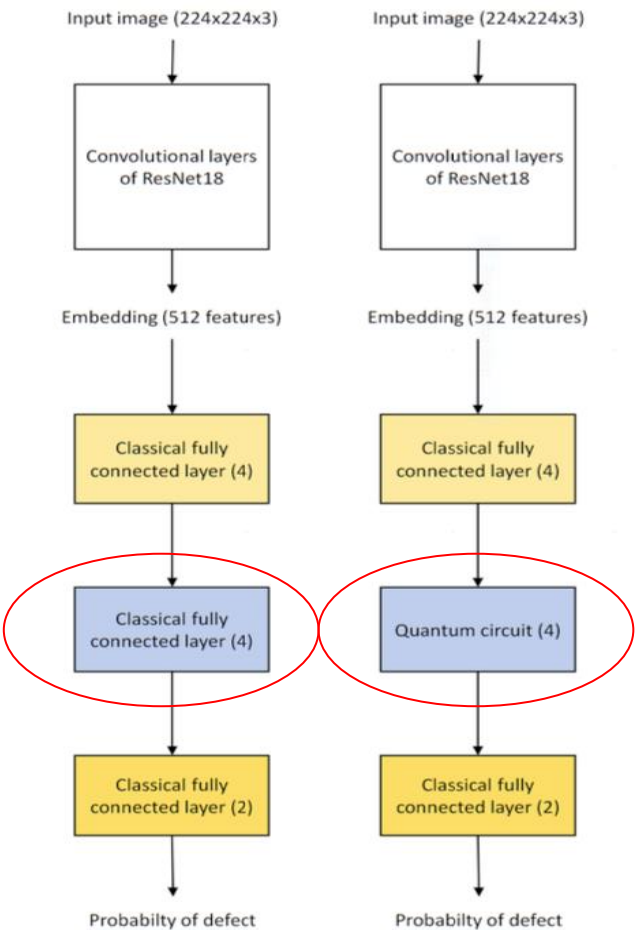


(classical, hybrid,
use case parameters, scaling)



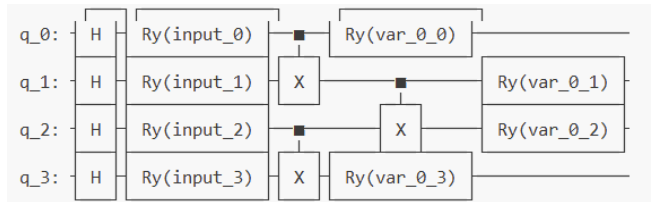
Quantum-Classical Image Classification (II)

System design



Angle encoding, 6 qubits, strongly entangled Ansatz

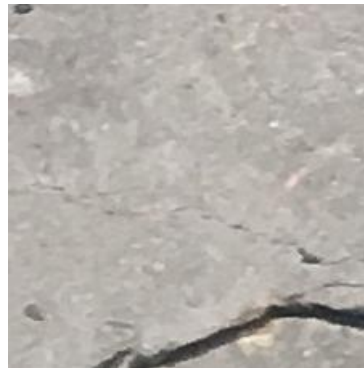
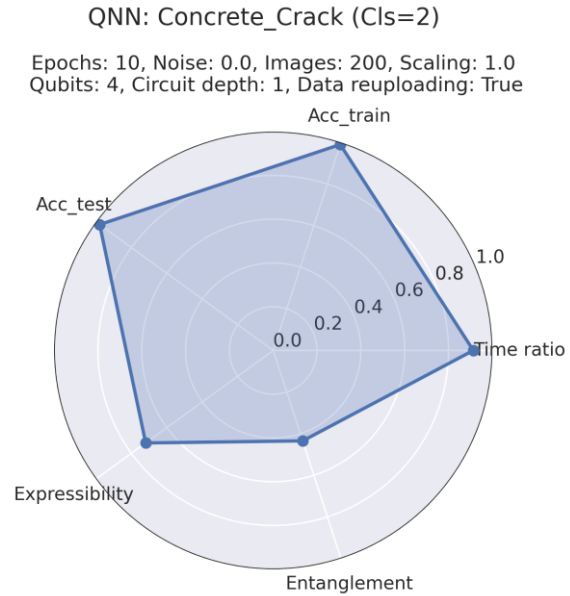
Noise variation study



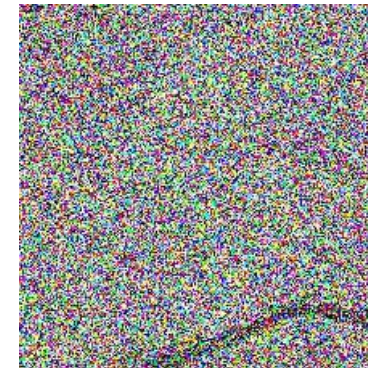
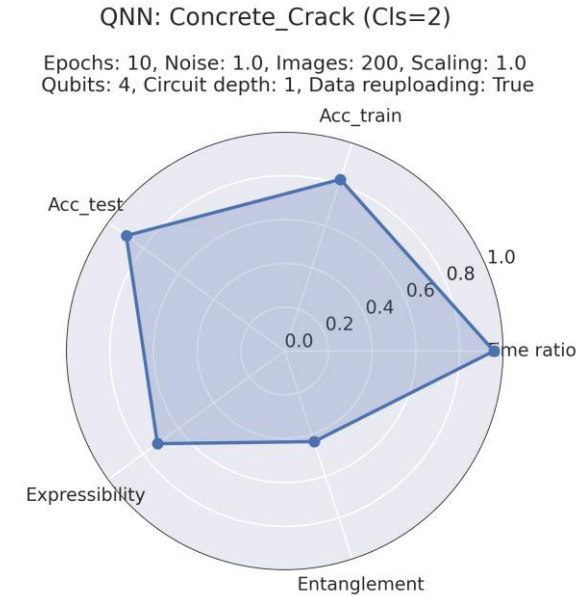
Observations

- **Addition of noise** is an effective measure to gradually increase problem complexity
- **Fixed:**
 - Circuit design
 - Number of training epochs

Noise = 0



Noise = 1



Noise = 2

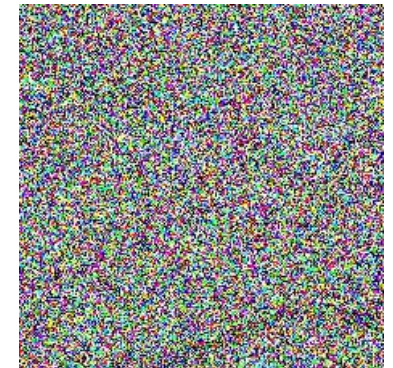
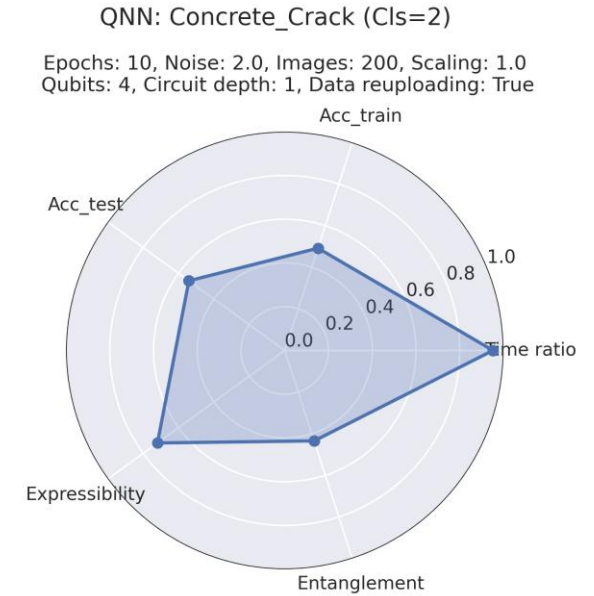
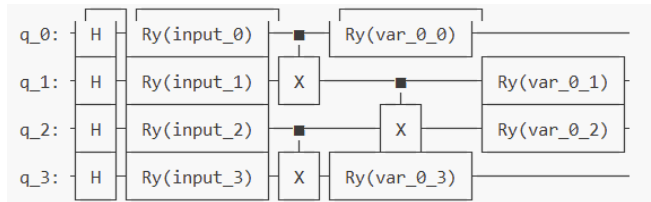


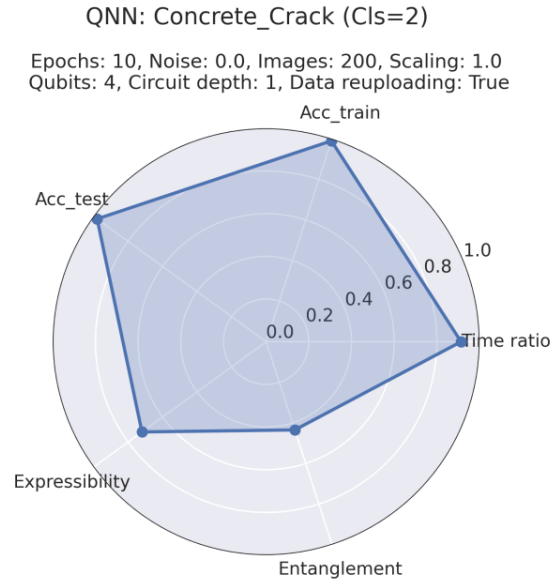
Image rescaling study



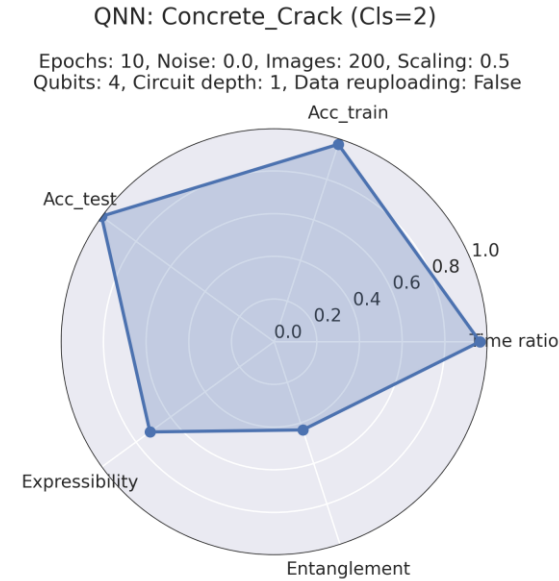
Observations

- Image **rescaling** can be tolerated to some extent, but eventually accuracy collapses
- Fixed:**
 - Circuit design
 - Number of training epochs

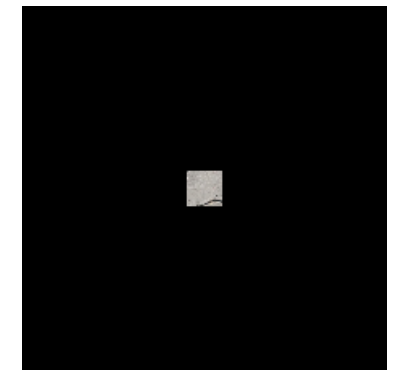
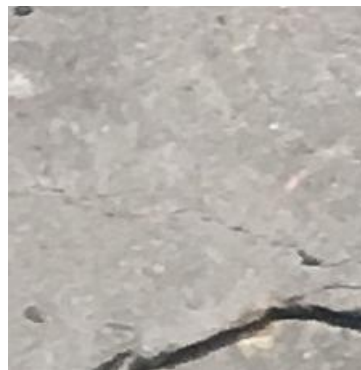
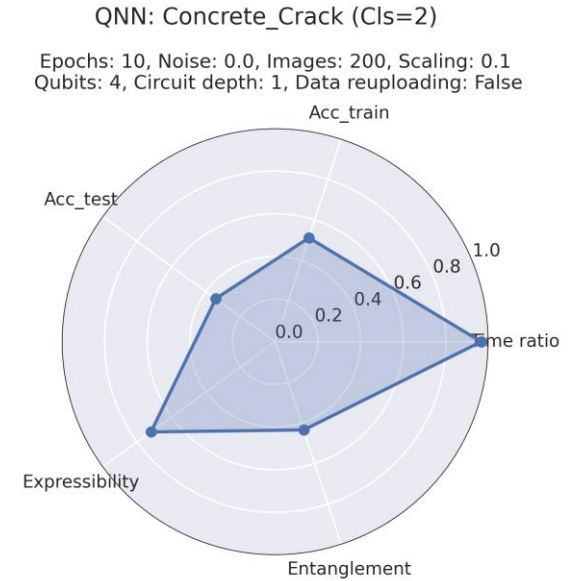
Scale_factor = 1.0



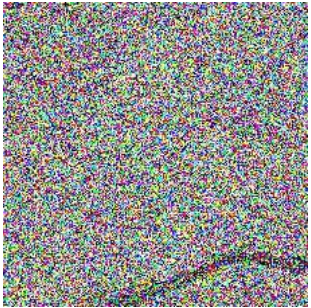
Scale_factor = 0.5



Scale_factor = 0.1



Circuit depth study (with data reuploading)



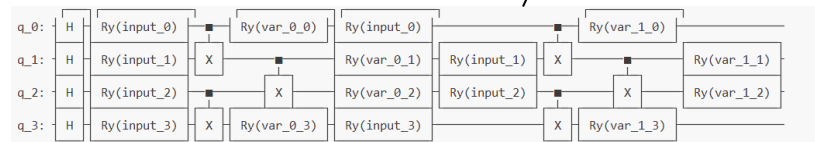
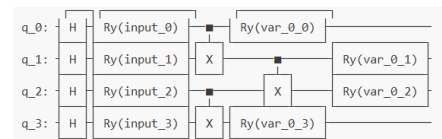
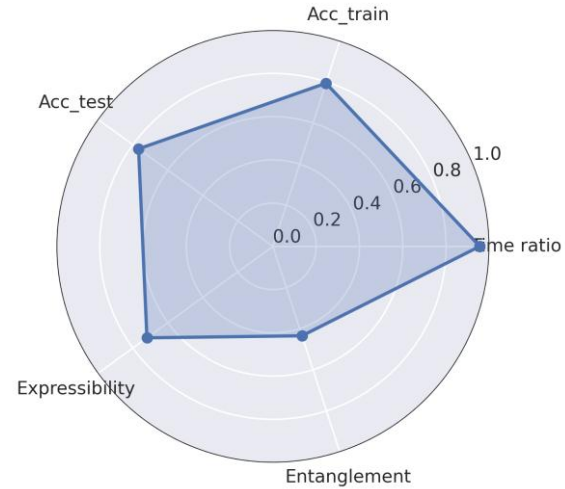
Observations

- Deeper circuits **increase expressibility and entanglement**
- Accuracy **not necessarily** much better?
- Fixed: 10 Epochs of training, average of 3 runs

$d = 1$

QNN: Concrete_Crack (Cls=2)

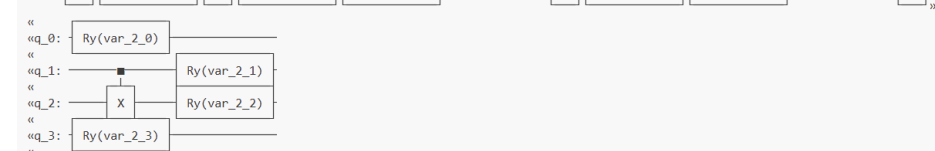
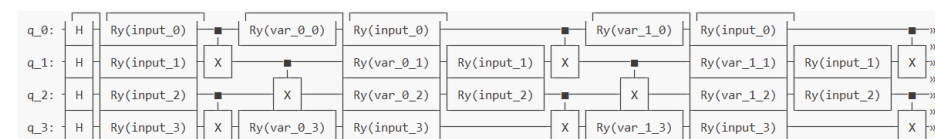
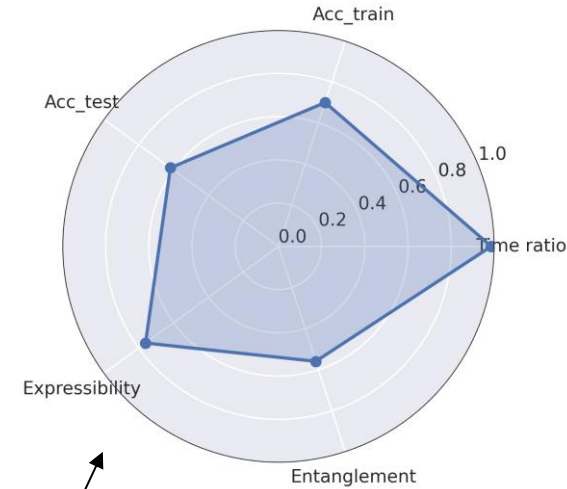
Epochs: 10, Noise: 1.0, Images: 200, Scaling: 1.0
Qubits: 4, Circuit depth: 1, Data reuploading: False



$d = 2$

QNN: Concrete_Crack (Cls=2)

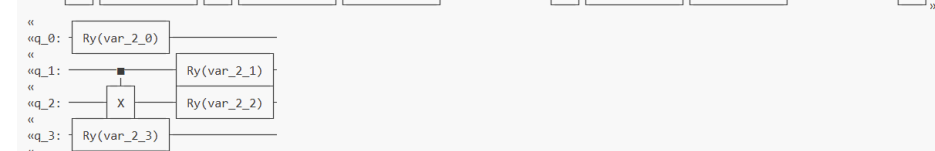
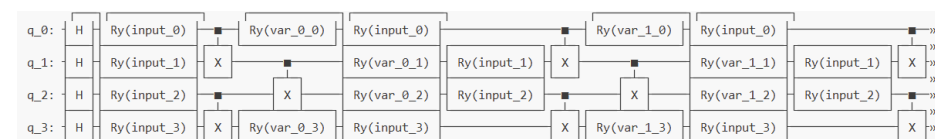
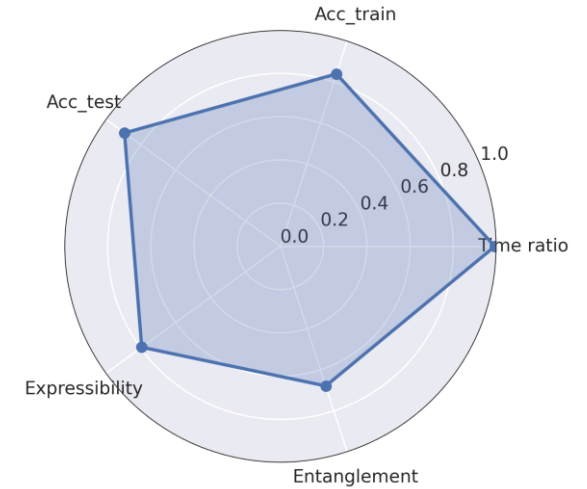
Epochs: 10, Noise: 1.0, Images: 200, Scaling: 1.0
Qubits: 4, Circuit depth: 2, Data reuploading: True



$d = 3$

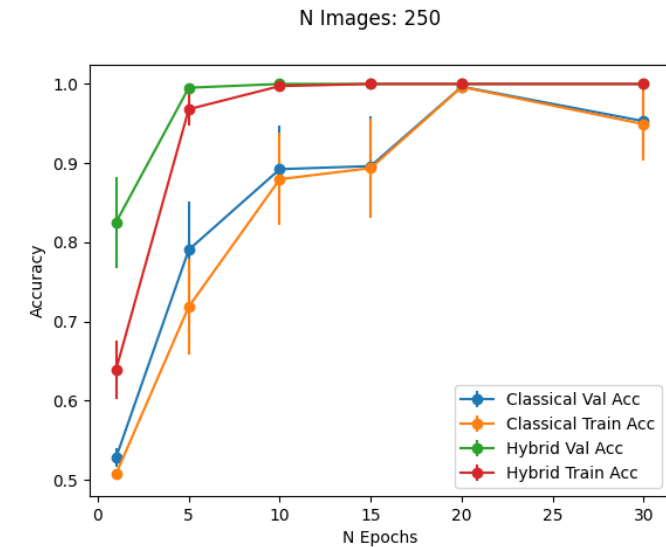
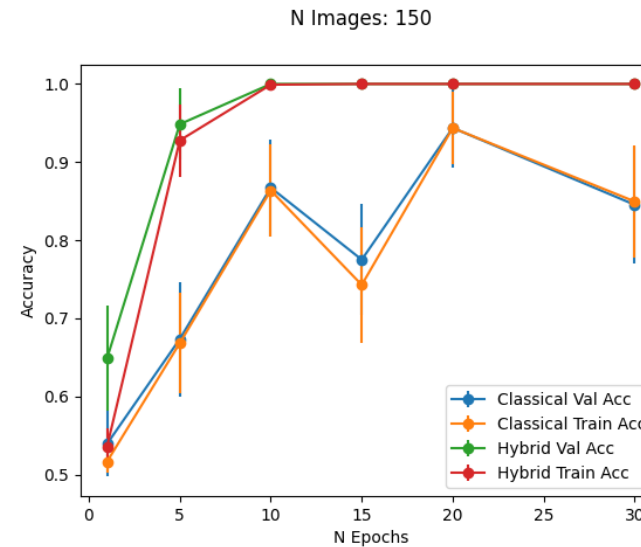
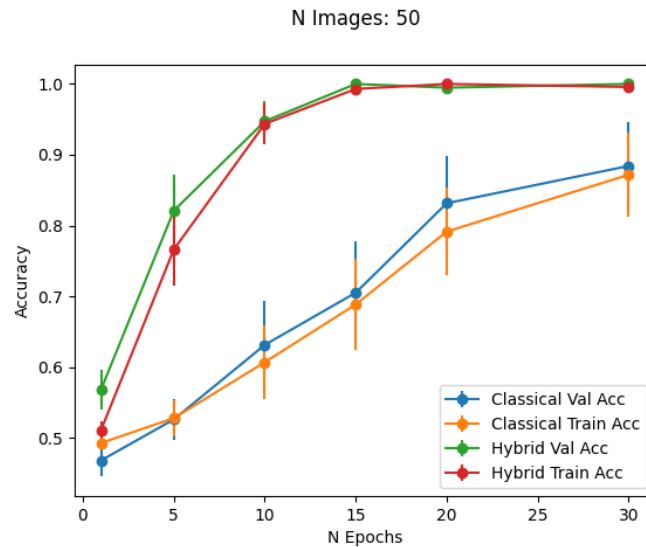
QNN: Concrete_Crack (Cls=2)

Epochs: 10, Noise: 1.0, Images: 200, Scaling: 1.0
Qubits: 4, Circuit depth: 3, Data reuploading: True



Dataset size study

- Road crack dataset
- 4 qubits
- Strongly entangled Ansatz

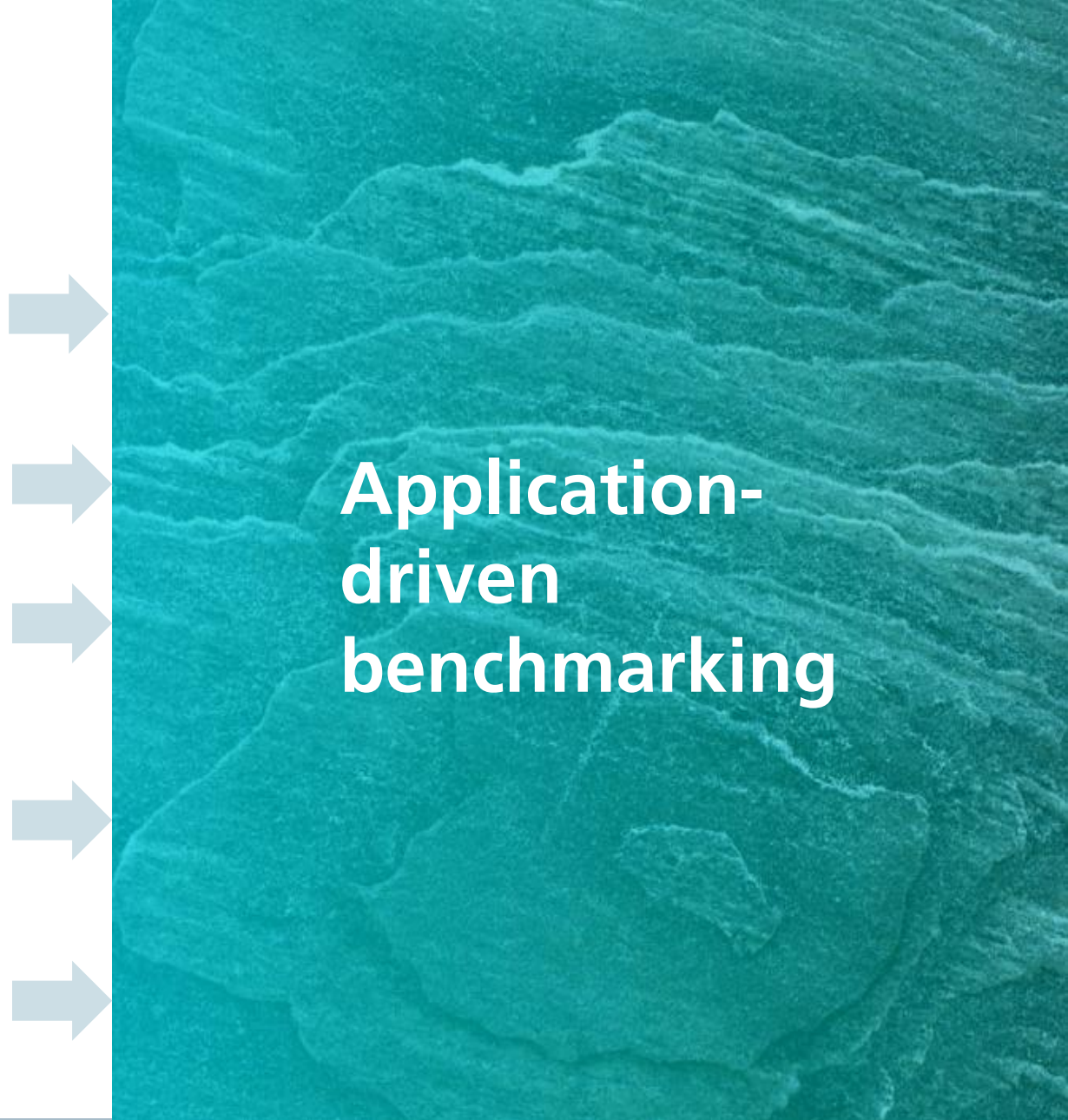


Benchmark potential benefit of hybrid-quantum NN compared to classical:

- Learns a given performance with **less training steps**, if **data is scarce**
- Gap closes with **more epochs or more samples**
- **Quantum circuit metrics** (expressibility, entanglement) is still **low**

Summary

- To benefit from quantum computing in real applications, a **practical quantum advantage** is required
- This advantage will not only depend on quantum hardware properties, but on **the full hardware and software stack**
- We need the **right tools** to do holistic benchmarking (e.g. QUARK)
- In the field of QML, a practical quantum advantage is still unclear, yet promising trends emerge. Benchmarking can guide the way to **best combinations of algorithms, metrics, hardware, etc.**
- Results can help establish a route towards **standards**



Application-
driven
benchmarking