Fraunhofer IKS

Fraunhofer Institute for Cognitive Systems IKS

Systematic benchmarking of quantum computers: status and recommendations

PD Dr. habil. Jeanette Miriam Lorenz

25.06.2025

Emerging technology:
Quantum computing

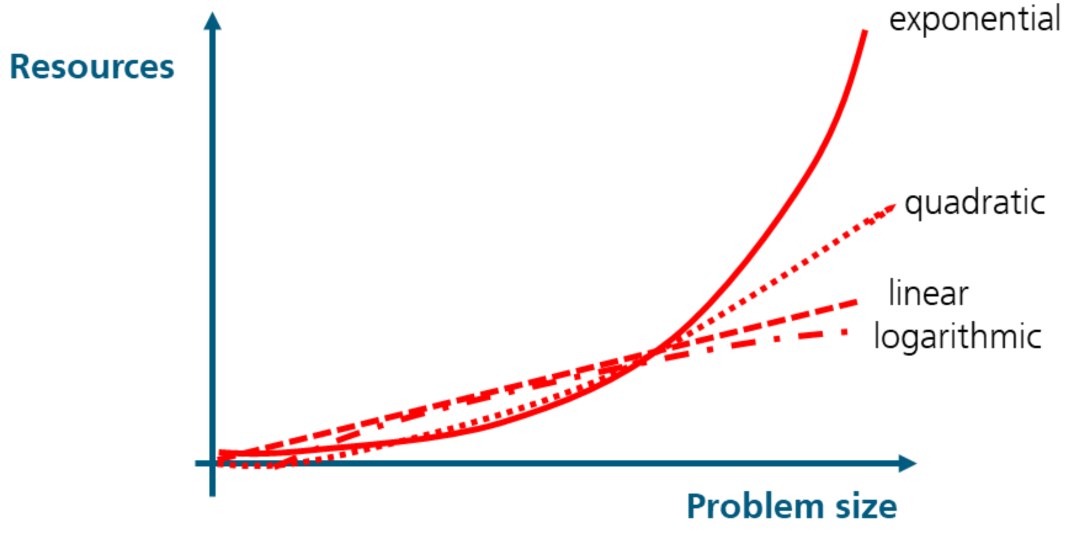
Quantum computing (QC) promises to be useful in many application areas and therefore industries:

- **Simulation** of quantum mechanical systems
  - (Development of new drugs, chemical sector with battery development,…)
- **Solving linear systems of equations** (fluid dynamics, e.g., in aerospace)
- **Optimization** problems (Logistics, production, pharma,…)
- **Quantum machine learning** (Computer vision, mobility,…)

But which combination of quantum computing hardware + software + algorithms can achieve this?

 © Fraunhofer IKS

A word on complexity and what we need to achieve



In QC we are interested in achieving at least a polynominal speedup with respect to classical algorithms, probably even super-polynomial required [1]

→ Complexity Theory, comparisons quantum & classical algorithms typically in ‚Big-O'-notation

Note: simply measuring benefits in terms of computational complexity might be too short-sighted: accuracy, energy efficiency…

Also tendency to go to more ‚empirical' benchmarks (like in ML)

[1] Torsten Hoefler, Thomas Haener, Matthias Troyer, Disentangling Hype from Practicality: On Realistically Achieving Quantum Advantage, arXiv:2307.00523 [quant-ph]

Which application categories are promising?

[T. Hoefler, T. Haener, M. Troyer, Disentangling Hype from Practicality: On Realistically Achieving Quantum Advantage, arXiv:2307.00523 [quant-ph]]

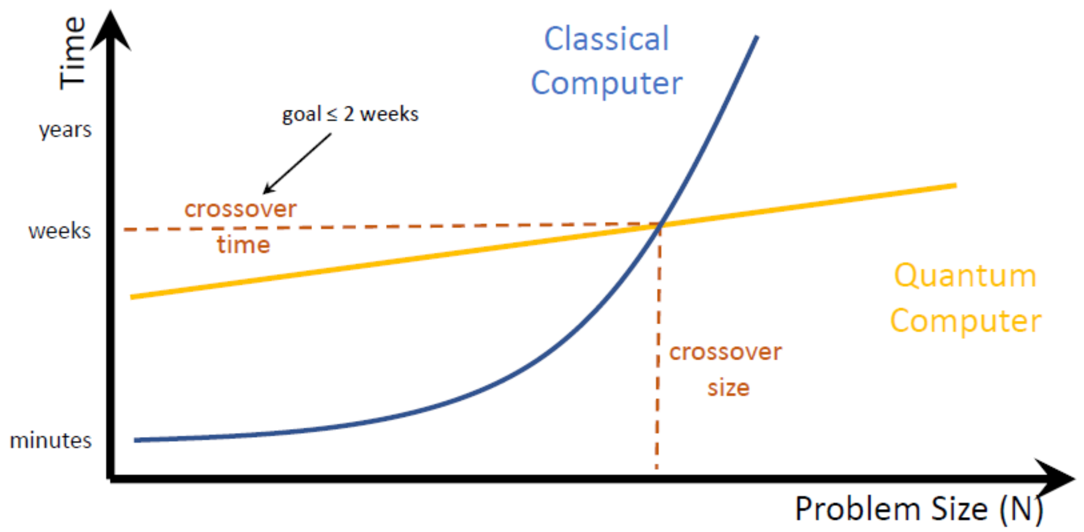


Fig. 1. **Quantum speedup**: The time needed to solve certain problems with quantum algorithms increases more slowly than that of any known classical algorithm as the problem size N increases. To be practical, however, we need more than an asymptotic speedup: the crossover time where quantum advantage gets realized needs to be reasonably short and the crossover problem size not too large. (For illustration, the time axis is scaled such that the quantum algorithm is a straight line.)

To achieve a **practical quantum advantage** it is required that a quantum computer is also faster in real time compared to a classical computer – this means not just assymptotically.

Possibly for every application/algorithm specific cross-over point.

Fraunhofer IKS

Motivation for benchmarks

Quantum computing itself already proposed in the 1980s:

- R.P. Feynman
- D. Deutsch
- …

Key fault-tolerant algorithms proposed in the 1990's:

- Shor's algorithm
- Grover's algorithm

Despite small laboratory experiments of realizing quantum computers in the 1990s, interest in quantum computing only raised significantely in 2019 with the Google Sycamore experiment

- Since then various claims of quantum advantage, quantum utility,…
- But no practical quantum advantage achieved yet.
- One difficulty is the slowness of quantum hardware versus speed-ups in computional complexity.

➔ Benchmarking of all quantum hardware, software & algorthms and indeed solving use cases required.

➔ Different levels of benchmarks:

- **Component-level:** benchmarks of hardware implementation of individual components
- **System-level:** benchmarks of hardware performance of entire implementations
- **Application-level:** comping from concrete use cases

+ also finer level benchmarks

 © Fraunhofer IKS Public information

Fraunhofer IKS

What is a benchmark?

A **(quantum) metric** is a quantitative measure to assess the performance of a (quantum) device, specified by a clear protocol. The verb **(quantum) benchmarking** details the act of performing a (quantum) benchmark. A **(quantum) benchmark** is a set of results of implementing one or multiple quantum metric(s) on a certain set of (quantum) devices aimed at characterizing the accuracy of the implementation.

 © Fraunhofer IKS

Public information

Fraunhofer IKS

Objectives and SOTA of benchmarking

- **Target groups:** General public & specialists; HPC centers to guide procurement of devices; users; investors

- **Objectives:**
  - Comparisons of different hardware platforms: **developments towards large-scale quantum computing, accuracy, time of computations**
  - Progress towards fault-tolerant quantum computers; requires **scalable** benchmarks
  - Different benchmarks serve different target groups

- **SOTA**:
  - First benchmarks centered around component-level benchmarks, but these are **difficult to scale**
  - Overall system performance rather depends on operations executed **most often**, or being particularly bad.
  - For fault-tolerant quantum computers, **overhead by quantum error correction** will be important
  - Combining individual computional building blocks on quantum computers demonstrated that the total system behaves differently (e.g. through cross-talk) -> **holistic benchmarks**
  - More recently, **HPC-level benchmarks & application-level benchmarks**

 © Fraunhofer IKS Public information

Benchmarking initiatives
Application-driven (examples)

Europe		
BACQ (France)	Benchmarks Applicatifs des Calculateurs Quantiques	THALES, LNE, CEA, CNRS, EVIDEN & TERATEC
Bench-QC (Germany)	Application-driven benchmarking of quantum computers	Fraunhofer IIS & IKS, BMW, ML Reply, Optware, Quantinuum
QPack (Netherlands)	Benchmark for quantum computing	TU Delft
CUCO (Spain)	Computacion Cuantica En Industrias Estrategicas	GMV
EuroQCS-Poland	Application Performance Benchmarks for Quantum Computers	PSNC
UK		
UK Quantum Computing Metrics		UK Quantum Metrology Institute of PNL
USA		
QED-C Benchmark	Computer programs and software for use in the fields of quantum science and engineering	Quantum Economic Development Consortium
QBI	Quantum Benchmark Initiative	DAPRA
HamLIB	A library of Hamiltonians for benchmarking quantum algorithms and hardware	INTEL and others
MetriQ	Community-driven quantum benchmarks	Unitary Fund
IEEE SA - P7131	WG P7131 “Standard for Quantum Computing Performance Metrics & Performance Benchmarking”	IEEE

 © Fraunhofer IKS

Summarizing and coordinating benchmarking efforts within Europe
The European Quantum Computing Benchmarking Coordination Committee

Different initiatives emerged all over Europe to benchmark the capabilities of quantum computing while reaching higher TRLs

- Important point to understand how different quantum hardware platforms compare to each other – in a systematic and quantitative way.
- Realization that also the interplay of different components within the software stack needs to be considered.

Meeting of the different activities:

- TERATEC seminar close to Paris in May 2023
  → afterwards contribution to the Strategic Research and Industry Agenda for Quantum Technologies (SRIA) with a chapter on benchmarking

- Second TERATEC seminar on benchmarking in Reims in June 2024
  → kick-off of a coordination activity of all European benchmarking efforts:
  European Quantum Computing Benchmarking Coordination Committee
  - Led by Jeanette Lorenz (Fraunhofer IKS, Germany), Frederic Barbaresco (THALES, France) and Ward van der Schoot (TNO, Netherlands)

[https://teratec.eu/Seminaires/TQCI/2023/Seminaire_TQCI-230511.html]

[https://teratec.eu/activites_quantiques/TQCI_240604_programme.html]

 © Fraunhofer IKS Public information

Fraunhofer IKS

Recommencations of the SRIA report

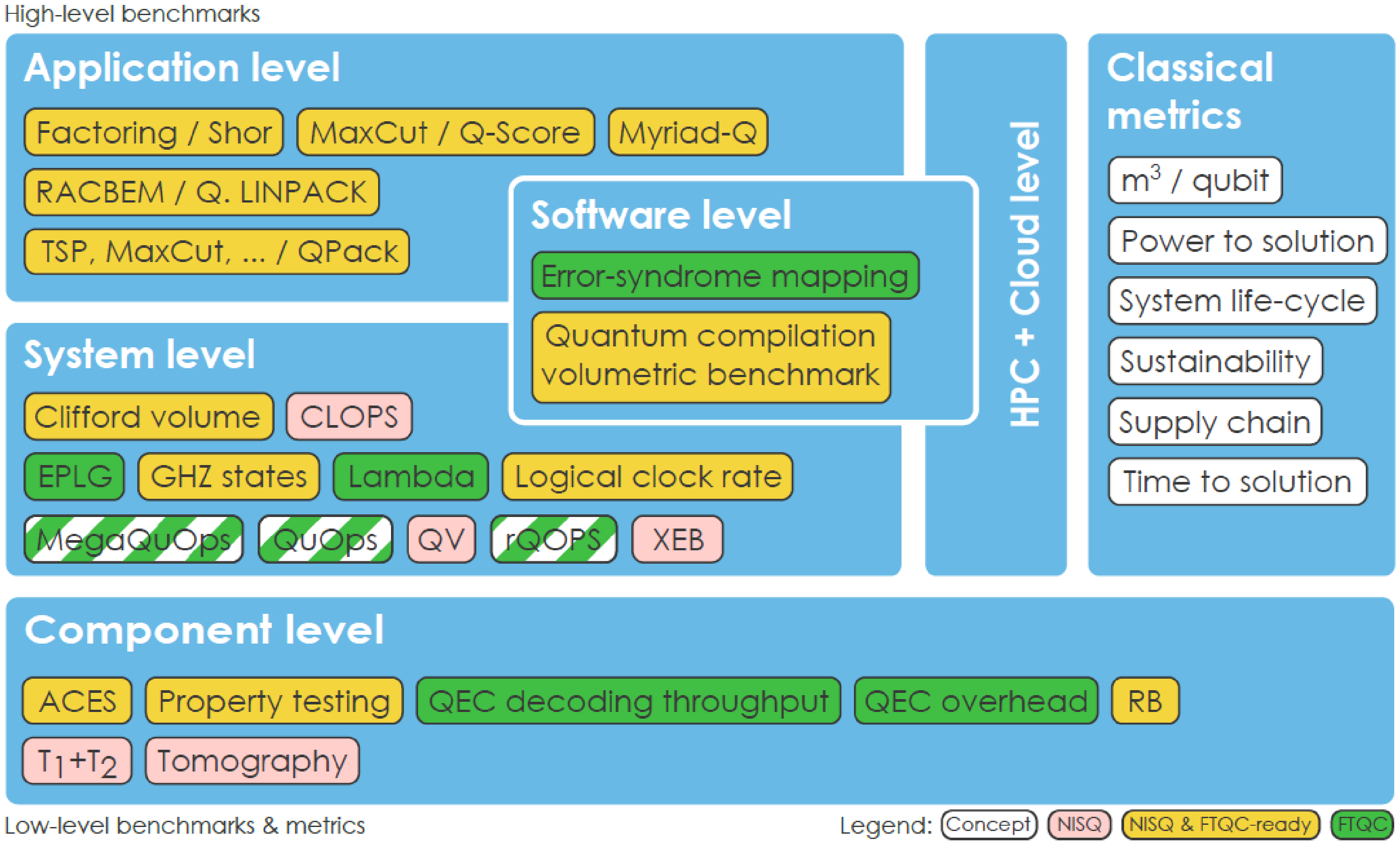
Strategic Research and Industry Agenda (SRIA 2030) for Quantum Technologies in the EU

Recommendations

- To encourage to have a single coordination forum between the different European benchmarking initiatives
- Encourage exchange between standardisation and benchmarking activities
- Define a programme at European level to support R&D effort with cross-disciplinary approach and inclusion of academia and industry for benchmarking.
- Support access to the machines through EuroHPC for benchmark development and testing to enable the development of quantitative and objective benchmarks.

 © Fraunhofer IKS

Different types of benchmarks



 © Fraunhofer IKS Public information

How to construct benchmarks & types

For digital (and universal) quantum computers

Benchmarks need to be reproducible & provide a meaningful comparison across platforms

- Defined in a **SDK-agnostic** and open (e.g., Open-QASM) format
- At least one **reference implementation** using a specific SDK including both quantum circuits + classical post-processing routines
- **Standardized evaluation** code with statistical methods defined + **sample datasets**
- Benchmarks are required to be **transparent**, e.g. error mitigation method clearly outlined
- **Classical resources** in hybrid algorithms need to be clearly defined, delineated and quantified
- **Scalable**

 © Fraunhofer IKS

Component-level benchmarks

Measuring performance of low-level components

T1/T2 times:
- How long a qubit remains stable
- Historically used, but disadvantages
- Unit is time, but misleading, as would need to be compared to gate operation times
- Basis dependent
- Lower bound on performance, but not providing overall predictive statement on performance

Quantum state tomography:
- Recovery of unknown states based on data
- Not scalable

Gate-set tomography:
- Not scalable

Randomized benchmarking:
- Efficient extraction of state preparation and measurement errors + gate operation errors
- Many specialized variants
- Possible to extend to logical qubits/ fault-tolerant QC
- Possible to extract the **average gate fidelity**

Property testing:
- Verifying in a robust and reliable fashion if a certain property is present in an unknown quantum state; e.g., entanglement

Averaged circuit eigenvalue sampling:
- Detect and characterise correlations

Benchmarks of QEC protocols

 © Fraunhofer IKS Public information

System-level benchmarks

To assess performance of entire implementations

Quantum volume:

- Statement if a certain number of qubits + gates provides a certain average success probability
- Problematic, as not scalable

Error per layered gate:

- Measurement how errors propagate through a given architecture

Clifford volume:

- Gates restricted to Clifford gate operations
- Output can therefore be calculated and compared
- Reasonable predictive power about quality of gate operations
- NISQ and FTQC

MegaQuOp:

- Benchmark for large-scale systems

GHZ states:

- Evaluation population & coherence of GHZ states
- Applicable to both physical and logical qubits

Circuit layer operations per second (CLOPS)

Reliable quantum operations per second (rQOPS)

Cross-device verification

Benchmarks related to quantum error mitigation

Quantum operations:

- For FTQC
- Tracks number of operations that a system can reliably execute

 © Fraunhofer IKS

HPC-level & software-level benchmarks

Software-level benchmarks

Consider the entire software stack with all layers and APIs from the end-user to the actual executation on a quantum processor.

Important step: compilation

- Exact compilation of quantum circuits NP-hard in worst case
- **Quantum compilation volumetric benchmark** similar to quantum volume would be sensible: how many qubits and layers of gate operations can be compiled – requires standardized target-computations to be realized

Benchmarks considering the overhead introduced by quantum error correction also desirable

HPC-level benchmarks

Similarities between classical HPC systems and challenges now coming up in QC.

HPC-level benchmarks of increasing importance as quantum computers get integrated into HPC centers. LINPACK-like benchmarks may be relevant.

Cloud-level benchmarks

- Relevant quantities include computational power, storage speed, memory throughput, network latency, instance provisioning time, queue times, cloud front-end overheads

 © Fraunhofer IKS

Application-level benchmarks

Application problem (in mathematical form) + KPIs/metrics + protocol

Benchmarks based on algorithms:

- Need at least superpolynomial speedup to obtain overall better performance than classical algorithms
- **Shor's algorithm** suited as scalable benchmark applicable to both NISQ and FTQC domains
  - One of the European Quantum technology flagship KPIs
- Due to industrial relevance, various benchmarks to assess improvements towards solving industrial optimization problems:
  - Qscore, QPack, QuAS
  - But unclear if QAOA variants can provide any advantage over classical algorithms – refrain from using them

Different more wide benchmarking frameworks and suites:

- QED-C metrics
- SupermarQ
- BaCQ
- QUARK/Bench-QC
- Open QBench

 © Fraunhofer IKS

Recommendations

[https://arxiv.org/abs/2503.04905]

- Benchmarking in quantum computing increases in importance, but benchmarks also need to **be accompanied by standardized evaluation software & sample data. Existing benchmarks and metrics** need to be used.
- The **development and establishment of hardware-agnostic, statistically sound, and numerically efficient evaluation routines** is a significant effort, on top of current activities to establish the benchmarking protocols in an open and standardised fashion → **dedicated projects & funding required**
- Benchmarks need to **be tightly integrated into and aligned with on-going quantum activities** across the entire quantum stack.
- **Integration of benchmarks into standardization** activities need to be covered by entities more suitable for ISO and CEN-CENELEC. Representatives need to be appointed on the national level, and require a significant time of coordination on a EU-wide scale.

Systematic benchmarking of quantum computers: status and recommendations

Jeanette Miriam Lorenz[*,†], Thomas Monz[‡,§], Jens Eisert[¶,‖], Daniel Reitzner[**], Félicien Schopfer[††], Frédéric Barbaresco[‡‡], Krzysztof Kurowski[x], Ward van der Schoot[xi], Thomas Strohm[xii], Jean Senellart[xiii], Cécile M. Perrault[xiv], Martin Knufinke[xv], Ziyad Amodjee[xvi], Mattia Giardini[xvii]

[*]Fraunhofer Institute for Cognitive Systems IKS, 80686 Munich, Germany (email: jeanette.miriam.lorenz@iks.fraunhofer.de)
[†]Ludwig-Maximilians-University Munich, 80539 Munich, Germany
[‡]Universität Innsbruck, Institut für Experimentalphysik, Technikerstrasse 25, 6020 Innsbruck, Austria
[§]Alpine Quantum Technologies GmbH, 6020 Innsbruck, Austria
[¶]Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[‖]Fraunhofer Heinrich Hertz Institute, 10587 Berlin, Germany
[**]CyberSecurity Hub, z. ú., Šumavská 416/15, 602 00 Brno, Czech Republic
[††]Laboratoire national de métrologie et d'essais, 29 avenue Roger Hennequin, 78190 Trappes, France
[‡‡]THALES, 1 Av. Augustin Fresnel, 91120 Palaiseau, France
[x]Poznan Supercomputing and Networking Center IBCH PAS, Noskowskiego 12/14, 61-704 Poznan, Poland
[xi]The Netherlands Organisation (TNO) for applied scientific research, The Hague
[xii]Robert Bosch GmbH, Corporate Research, D-71272 Renningen, Germany
[xiii]Quandela, 7 rue Léonard de Vinci, 91400 Massy, France
[xiv]Alice & Bob, 49 boulevard du general Martial Valin, 75015 Paris, France
[xv]Eviden, science+computing AG, Hagellocher Weg 73, 72070 Tübingen, Germany
[xvi]Laboratoire Charles Fabry, Institut d'Optique Graduate School, CNRS, Universite Paris Saclay, 91127 Palaiseau, France
[xvii]European Quantum Industry Consortium e.V.,Leo-Brandt-Straße, 52428 Jülich

***Abstract*—Architectures for quantum computing can only be scaled up when they are accompanied by suitable benchmarking techniques. The document provides a comprehensive overview of the state and recommendations for systematic benchmarking of quantum computers. Benchmarking is crucial for assessing the performance of quantum computers, including the hardware, software, as well as algorithms and applications. The document highlights key aspects such as component-level, system-level, software-level, HPC-level, and application-level benchmarks. Component-level benchmarks focus on the performance of individual qubits and gates, while system-level benchmarks evaluate the entire quantum processor. Software-level benchmarks consider the compiler's efficiency and error mitigation techniques. HPC-level and cloud benchmarks address integration with classical systems and cloud platforms, respectively. Application-level benchmarks measure performance in real-world use cases. The document also discusses the importance of standardization to ensure reproducibility and comparability of benchmarks, and highlights ongoing efforts in the quantum computing community towards establishing these benchmarks. Recommendations for future steps emphasize the need for developing standardized evaluation routines and integrating benchmarks with broader quantum technology activities.**

***Index Terms*—Quantum computing, quantum algorithms, benchmarking, metrics.**

CONTENTS

LIST OF ABBREVIATIONS

ACES *Averaged circuit eigenvalue sampling* is a component-level benchmarking method that is capable of detecting and characterising correlations.

FTQC *Fault-tolerant quantum computing* relies on careful computational instructions such that errors do not

 © Fraunhofer IKS

Contact

PD Dr. habil. Jeanette Miriam Lorenz
Head of Department

Associate Professor @ LMU Munich
Speaker of the consortium K7-QACI of the Munich Quantum Valley

Tel. +49 89 547088-334
Jeanette.miriam.lorenz@iks.fraunhofer.de

Fraunhofer IKS
Hansastr. 32
80686 München
www.iks.fraunhofer.de

Fraunhofer IKS

Fraunhofer Institute for Cognitive Systems IKS