Hardware efficient microarchitectures and quantum error correction codes for large scale quantum processors

A Data Management Plan created using DMPonline.be

Creator: Quinten Eggerickx https://orcid.org/0000-0002-4709-3115

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 11Q4A24N

ID: 206837

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

Quantum computing holds the promise to solve problems which are deemed to be too hard for any classical computer. To make good on this promise, however, quantum computing systems will need to evolve from the small noisy systems that they are today to fully fault-tolerant systems with lots of qubits. An essential ingredient to achieve fault-tolerance are quantum error correcting codes (QECCs). Among all QECCs that have been formulated over the years, quantum low-density parity check codes (qLDPC) have established themselves as the prime candidate for practical implementations, due to their reduced chance of error propagation. Moreover, very recently it was shown that qLDPC codes with optimal scaling properties can be constructed. The aim of this project is therefore to translate these highly theoretical constructions of so-called good qLDPC codes into practically realizable layouts. Furthermore, special attention will be given to the construction and analysis of good qLDPC codes that require a minimal number of qubits. Such codes would be the first ones to be implemented in practice and therefore provide a target for experimentalists.

Last modified: 26-04-2024

Hardware efficient microarchitectures and quantum error correction codes for large scale quantum processors FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Manuscripts	Theoretical findings	Generate new data	Digital	Other	.pdf	<1GB	
Stim data	Quantum error correction code performance simulations using Stim	Generate new data	Digital	Simulation data	.csv .dat	<100GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

/

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

- Manuscripts: No accompanying information is required to keep the the data understandable and usable.
- Stim data: Simulations performed in python with the stim package will be explained in Powerpoint or in the simulation script itself.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

- Manuscripts: No accompanying information.
- Stim data: The simulated circuit, as well as all other simulation parameters are stored as metadata.

3. Data storage & back-up during the research project

Where will the data be stored?

All data will be stored locally on the applicant's computer, as well as on Imec's cloud service provide (Microsoft).
Manuscripts may additionally be stored on overleaf.com and/or arxiv.org when appropriate.

How will the data be backed up?

Backing up of data is carried out by Imec's cloud service provider (Microsoft).

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

Local free storage (360GB) and free cloud storage (1TB) are sufficient for the estimated maximum data size (200GB).

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Data storage and management will be done in accordance to IMEC's strict safety guidelines.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

Costs are covered by Imec.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All the data will be retained for 5 years following the end of the project.

Where will these data be archived (stored and curated for the long-term)?

The data will be left where it's currently saved (local and clouds) as capacity is sufficient.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

Costs are covered by Imec.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

Relevant data will be published in journals or reported in the PhD thesis.

If access is restricted, please specify who will be able to access the data and under what conditions.

Data can be accessed by relevant collaborators or other parties upon request. All data can be be shared with FWO and KU Leuven.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Relevant data will be published on arxiv.org, in journals or reported in the PhD thesis.

When will the data be made available?

The data will be made available upon publication of the research results.

Which data usage licenses are you going to provide? If none, please explain why.

Data usage license will be discussed with Imec/KUL/UGent before any licenses are granted.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

What are the expected costs for data sharing? How will these costs be covered?

No costs are expected.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Quinten Eggerickx will manage data documentation and metadata during the research project.

Who will manage data storage and backup during the research project?

Quinten Eggerickx will manage data storage and backup during the research project.

Who will manage data preservation and sharing?

Quinten Eggerickx will manage data preservation and sharing.

Who will update and implement this DMP?

Quinten Eggerickx will update and implement this DMP.