Gmail

Samudra Dasgupta <samudra.dasgupta@gmail.com>

IEEE Transactions on Quantum Engineering: TQE-21-06-QC-0068.R1

2 messages

IEEE Transactions on Quantum Engineering <onbehalfof@manuscriptcentral.com> Wed, Dec 15, 2021 at 12:22 AM

Reply-To: jbardin@engin.umass.edu
To: samudra.dasgupta@gmail.com, dasguptas@ornl.gov

15-Dec-2021

Dear Mr. Dasgupta:

"Stability of Noisy Quantum Computing Devices" (Manuscript ID TQE-21-06-QC-0068.R1), which you submitted to IEEE Transactions on Quantum Engineering, has been reviewed. The comments from the reviewers appear below and in separate files which may be read by going into the "Manuscripts with Decisions" queue in your Author Center and clicking the "view decision letter" link for this paper.

Unfortunately, on the basis of these reviews, we cannot accept your paper for publication in IEEE Transactions on Quantum Engineering.

While we regret not being able to accept your manuscript, we hope that our prompt decision will enable you to pursue other ways to disseminate your results in a timely manner.

Thank you for sending your manuscript to us for consideration. We look forward to further contributions from you in the future.

Sincerely,

Joseph Bardin
Editor, IEEE Transactions on Quantum Engineering

Manuscript Title: Stability of Noisy Quantum Computing Devices
ID: TQE-21-06-QC-0068.R1

Reviewers' comments to author, if any (please also check your Author Center for other files):
Reviewer: 2

Comments to the Author
(There are no comments. Please check to see if comments were included as a file attachment with this e-mail or as an attachment in your Author Center.)

Reviewer: 1

Comments to the Author
I agree that both computational accuracy and stability are important and not equivalent, but I am somewhat confused by the authors' definition of reproducibility in quantum computing. For example, if a qubit has an average zero and one state readout fidelity of exactly 50% for one month and then the processor manager improves the calibration and the readout fidelity goes to 99% for a month (which would produce a Hellinger distance of 1) I might think that you ought to be able to reproduce everything you were able to do with 50% readout fidelity with 99% readout fidelity, unless of course you were particularly interested in measuring readout fidelity. In this sense I think the symmetric nature of the Helligner metric is somewhat uninformative. I might define reproducibility as some quantum algorithm or circuit having a non-zero fidelity at some circuit depth with some number of qubits, but perhaps if the author's elaborate on what reproducibility means then the Hellinger distance may make more sense.

I have some concern about how the Hellingering distance is calculated, but perhaps the analysis method just requires a further explanation so I'll walk through my understanding. If I look at the top plot in figure 1a. The red 'Mean' trace is the source characterization dataset, one data point per day corresponding to the readout fidelity averaged over zero and one state preparation. Then (this is where I get a little confused and borrowed some information from the description of Figure 3 to fill in the blanks), to compute the Hellinger distance you use the first 30 days of data from may 2019 (which are not actually plotted in the top plot–this would be good to add), and compute the Hellinger distance between the May data and a moving 30 day window so the left most point in the blue Hellinger distance trace is the overlap of the reference May-19 data and the first point the moving window which covers June-19. I am not sure I have understood everything to this point correctly, but presuming I have, I am imagining that you used the discrete computation of the Bhattachary coefficient in equation 8 where the x you are summing over is the readout fidelity. Since the readout fidelity is a continuous variable, I am imagining that this computation is very dependent on some bin width. When I look at the top plot in figure 1a it looks to me like the readout fidelity is both quite good, and quite stable from about August-2019 to April-2020 and yet the Hellinger distance varies from near 0.5 to 0.0 in the same time period. I imagine this is because the reference dataset from May-2019 just didnt have any data point with good readout fidelity, but it is pretty difficult to understand why the Hellinger distance is a helpful metric when it varies half of its full scale during a time when the processor performance was largely stable.

If the above description of your analysis is correct then my imprecise suggestion of plotting the variance was slightly misinterpreted and I do not think the Bernoulli variance is informative. What I intended to suggest was to plot the standard deviation (not variance, my mistake) of the 30 data points in the moving window average. The standard deviation of performance metrics seems like a naive way to inspect stability and I was hoping the comparison to the Hellinger distance might highlight some useful comparison.

I wanted to follow up regarding the details of the source dataset from the Yorktown processor. If I understand correctly the data set you are using is the calibration record results for this processor reported by the processor manager. If we take the case of readout fidelity, I could imagine that for a superconducting qubit measured with dispersive readout the readout fidelity is dependent on some choice of state descrimination of an analog signal. When calibrating readout it could be the case that one would prepare the zero and one states and measure the resulting signal and then choose the state descrimination that maximizes the readout fidelity. If that is the reported readout fidelity (where state descrimination was optimized knowing which state was prepared) that would be potentially different than what users would see if there was some time for analog performance to drift between when the state descrimination was chosen and when readout fidelity was subsequently measured later without updating the calibration. This is why I made the comment in my first review that depending on how this number is reported it may or may not be particularly representative of what users of this processor would observe throughout the day. Give your dataset on the toronto device includes your own readout characterizations throughout the day, comparing those results to that day's calibration report might be in interesting line of comparison.

Regarding duty cycle–I agree with your definition for a single qubit, but how do you deal with a multi-qubit algorithm where gates may need to be scheduled synchronously and the effective gate length becomes as long as the slowest gate because shorter gates get padded with an idle? It seems like perhaps the more useful metric is just to use the gate fidelity as some fixed length across the device? E.g. if the longest gate is 441 ns on a given day, then it seems like the ideal metric to track might be the gate fidelity for a 441 ns gate on each target (where shorter gates idle for the remainder of the time between their gate length and 441 ns).

Is the temporal analysis in Figures 2, 3, and 4 all the same and as described in the appendix on Stability Analysis? In particular there seem to be some differences in the description of the analysis (1 month vs 3 month moving window and the reference is either fixed as the month of may 2019, or some moving/trailing window?). It was not obvious to me why the analysis of the different metrics (readout fidelity, gate fidelity, and duty cycle) should follow a different procedure. If the procedures are intentionally different, please include an explanation as to why.

Page 10, line 47: "first set"--maybe use "Set 0" since "Set 1" is referred to in the next sentence (or make Set 1, set 2 in the next sentence).

Samudra Dasgupta <samudra.dasgupta@gmail.com> Wed, Dec 15, 2021 at 12:27 AM
To: "Humble, Travis S." <humblets@ornl.gov>

Dear Travis,

Please see the TQE review result below (not good news unfortunately),

Thanks,

Samudra
[Quoted text hidden]