

# Review of “Randomized benchmarking with gate-dependent noise”

*Chris Ferrie* @csferrie, csferrie.com

This is an open peer review of arxiv:1703.09835v2, Randomized benchmarking with gate-dependent noise by Joel J Wallman.

**TO JOURNAL EDITORS:** You are free to use the contents of the repository as you see fit.

**DISCLAIMER:** A referee necessarily spends *far* less time than the authors on the paper. As such, this report surely contains errors. I would love for you to point them out so I can correct them. Otherwise, this report is provided without warranty of any kind, including its suitability as a resource for making decisions affecting the career of the authors of the paper.

**CONFLICTS OF INTEREST:** I know the author of this paper personally. Though we have not collaborated on an existing paper, it is quite possible that we may be co-authors on a near-future paper. I was also funded, at least in part, by the same grant that this research was funded by. First, I currently have no competing or common financial interests with the author. Second, on a personal note, I will not be holding back any punches for the author. I trust that they takes this in stride. [For non-academics reading this, this situation is quite common and colleagues often (quite critically) review each others work for publication (albeit anonymously).]

## Tweetable summary of my report

Great and timely result with solid theoretical backing. Would like more implementation details, incl. data and code. #openscience?

## Summary of claims and results

To me, the work serves two purposes: (1) a thinly veiled criticism of Ref. [28], and (2) a mathematical justification for the empirical success of randomized benchmarking in the presence of gate-dependent errors, which is always the relevant regime in practice. It would pain me too much to have to disentangle the two, so I will focus on the former for the most part.

First, let me comment on the state-of-play prior to this work. Randomized benchmarking is the go-to tool for experimental groups to validate the performance of their qubits. It’s importance cannot be understated in this regard. However, the existing theoretical analysis consisted only of useful results in the case of gate-independent (aka “nice”) noise models. Such models are a fantasy in real experiments. Nevertheless, every time the protocol was performed, the

results looked exactly as predicted by the obviously incorrect model. This paper shows the curve often depicted in randomized benchmarking experiments is in its necessary functional form: an exponential decay. Moreover, the author identifies the source of the potential paradox in about the simplest way that could be imagined by showing the effect of gate-dependence is a small perturbation to the model. This is an important and welcome result.

### **Correctness**

I trust the results are correct. The proofs of the theorems are easy enough to follow. The “numerics” section leaves much to be desired but ultimately is only a swift sanity check of the results. Trust in those results cannot extend beyond trust in the author as it currently stands.

The only broad criticism of the technical results is the now possible added confusion over what the author refers to as “noise between the gates.” As the author notes as an open problem, this new definition of noise is not necessarily physical. I suspect much debate over this notion will follow in the community. It is a shame that this will probably lead to more confusion until resolved or expounded upon.

### **Accessibility**

The most beautiful thing about this paper is Sec. I. If you only understand randomized benchmarking at a superficial level and can only afford the brain cycles to keep it that way, Sec. I tells you everything you need to know. The theorems are clearly stated and proofs well-written.

Much of the prose is difficult to follow and feels as if it is written solely for someone who eats, breathes and sleeps randomized benchmarking. There is just too much detail left out—many of the numbers, for example  $0.01r$ , are stated without explaining their practical consequence. (Is that type of accuracy uncommon, difficult, or impossible in this example?) Moreover, as mentioned above, there is a constant thread of vague comments and statements about Ref. [28]. Context is important, but this is unnecessary and obfuscates the author’s result.

### **Reproducibility**

I was able to follow the calculations and fill in any implicit steps. On the numerics section... where is the code?! I spent far too many hours trying to decode what the author is talking about here. Am I bitter? Absolutely. I have nothing to show for several days of wasted effort in trying to figure this out. I’m sure the results are sound and if I had more time it could probably

be reproduced, but it would definitely take some guess work and that is not acceptable.

### **Final verdict**

This is a great result. The paper, however, suffers from two flaws. The first is less crucial but greatly hinders the accessibility of the result.

Ref. [28] is mentioned 17 times! 17! That’s more than once per page. This must be a record. To actually understand a good chunk of what is said in this paper, a detailed reading of Ref. [28] is a prerequisite. I should hope this was not the intention of the author. I’ve not the time to read Ref. [28] in the detail required. The situation would be rectified by reviewing the pertinent points of Ref. [28] with either explicit quotes or equations up front. That, or move all mention of Ref. [28] to a single section “relation to other work” because, ultimately, the current paper stands on its own.

The second major flaw is crucial but easily remedied. The numerics section does not give enough detail to reproduce the work. This is easily rectified by supplying the code which produced them.

### **Detailed criticism, questions, and comments**

#### **More important**

1. The first time Ref. [28] is mentioned is the sentence, “This shows that the looseness of the bounds for the particular example of Ref. [28] is in fact generic.” Wait, what? Where did that come from? Why is Ref. [28] even relevant here? We need to be given the example and summary Ref. [28], or this shouldn’t be mentioned here.
2. In the first paragraph of Sec. IV, the author refers to the “higher order terms in the analysis of Ref. [27].” This is completely meaningless to probably even the authors of Ref. [27] since that was a 20-page paper with series expansions all over the place and written 5 years ago. If Ref. [27] is going to be referred to in this way, the relevant parts should be restated. If that is not feasible, minimally, explicit equation numbers of the reference need to cited.
3. I don’t know what “ $k$ th order terms in  $\triangle_G$ ” means. In fact, “terms” are mentioned several times. Terms of what?
4. What are  $h_1$  and  $h_2$  in Eq. (15)? Somehow I’m guessing they appear in the 0th and 1st order analysis of Ref. [27], but how?
5. “...rigorously hold to within 0.01.” Within 0.01 *of what?*

6. In the statement of Theorem 4, it states that  $t$  is the loss of trace and refers to Eq. (22), which only has  $p$ . Something is missing.
7. In Eq. (38), what is  $H$ ?
8. In the description of the numerics, what is a “Haar-random unitary of a fixed infidelity”? How is this measure defined, or how are such unitaries generated?
9. I must confess I have no clue how to calculate  $p(\mathcal{RL})$  from Theorem 2. Obviously I am missing something or this needs to be made more explicit. Perhaps a worked example would help.
10. Why are we setting  $B = 0$ ? How does this come about? The SPAM-free value is  $1/d$  is it not?

#### **Less important**

11. The definition of average is troublesome. It is defined with the standard notation of a generic expectation value but only allows uniform averages. Is every average in randomized benchmarking an arithmetic mean?
12. Nitpicking here, but the blackboard font is used for both expected values and groups, and the latter is not defined in the list of notation. If the brackets  $[\cdot]$  in the expectation value was used consistently, this wouldn’t be so bad.
13. In Eq. (11),  $\mathcal{C}_{11}$  is introduced to be equivalent to  $t$  but then never used.
14. “...gate dependent fluctuations, which may be as large as 0.1...” Fluctuation in what? The fidelity? The bound? The estimated error rate?