Comments on An IEEE Framework for Metrics and Benchmarks of Quantum Computing

John Gamble, Christopher Granade, Nathan Wiebe
Microsoft Research, Redmond, WA, USA

The current draft of the document lays out terminology that serves as a framework for discussion about quantum computing technology. Four main terminologies are introduced to track the development of quantum technologies: metrics, benchmarks, use cases and technology layers.

While detailed examples are provided for use cases and technology layers, no examples of metrics or benchmarks are given. Clarifying examples of metrics and benchmarks for multiple platforms would be useful as clarifying examples. For example, the average gate fidelity or diamond norm of a gate, or the process matrix of an error channel may be metrics of interest to a use case. Further, techniques such as randomized benchmarking or various flavors of quantum process and gate set tomography would be examples of benchmarking methods.

There are, of course, subtle assumptions behind benchmarking methods that need to be made in order for a benchmarking method to accurately estimate a metric. Mentioning the existence of such caveats is critical as assumptions such as Markovianity, locality, or gate-independence of noise underlie the majority of proposed benchmarking methods. Similarly, the estimation protocol used to infer a metric from a benchmark is critical to understanding the information about a quantum device provided by an estimate. We thus suggest redefining the word “metric” to be a “quantity that characterizes a feature or behavior in a quantum computing device or system, and that may be inferred from the results of a benchmark.”

The term “computational model” may be inappropriate for analog simulation. This is because there is no clear set of operations nor cost function associated with it. For analog quantum computing it is even more challenging to define a concrete computational model and so we recommend changing the notation used in Table 1. An alternative term such as “quantum information processing platform”, or simply “quantum information processor”, that does not have a formal meaning to computer scientists would be useful.

An additional issue raised by the taxonomy proposed as Table 1 is that many benchmarks are inapplicable to adiabatic quantum devices and to analog quantum simulation devices. Benchmarks for quantum devices that are not capable of executing quantum programs (colloquially, “circuits”) tend to be focused on individual instances of use cases, rather than on the broader use case itself. As a result, it is more difficult to provide a consistent taxonomy for metrics which allow predicting interesting behavior of non-universal devices.

Regarding Table 2, quantum annealers and other technologies do not in general yield the global minimum for an objective function. Since finding global optima for spin glasses is known to be NP-hard, this would imply that these technologies can solve NP-hard problems efficiently, which is broadly considered to be unlikely. Also, fault tolerant quantum computers typically are not capable of self-correction. Hence, it should be clarified that a classical decoder is often needed to detect and correct errors, necessitating the involvement of many technology layers and a rich control stack.

On Table 3, we find the term “physical circuit” confusing. Used in the context of controlling networks of physical devices, the “physical controller” might be clearer.