

## A metrological era for quantum computing

*Logic gates at the surface code threshold: Superconducting qubits poised for fault-tolerant quantum computing*

R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. OMalley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis,

Nature **508** 500 503 (2014).

**Recommended and a commentary by David DiVincenzo,  
Forschungszentrum Juelich**

The juggernaut that is superconducting qubits now arrives, according to the announcement of this experiment, at a long-heralded threshold: the theoretically predicted noise threshold for quantum error correction. Errors during operation must be corrected if computation (either classical or quantum) is to be reliable. Redundancy and error correction, such as parity checking, can help when employed effectively. If errors can be dealt with and corrected faster than they occur, then the error rate is “below threshold”, and reliable computation is possible. Years of investigation have led us to the point where the best approach to error correction is the Bravyi-Kitaev surface code [1]. Extensive simulations indicate that the threshold error rate is around 1%.

This means, to be more specific, that successful error correction becomes possible when rates for *all* types of error go below around 1%. These types of errors are: 1) errors during single-qubit gates, 2) errors during two-qubit gates, 3) state preparation errors, and 4) measurement errors. Types 3 and 4 are abbreviated SPAM (State Preparation And Measurement); we will see shortly why they are singled out. The efforts of many experimental groups have been decreasing all these error rates in recent years. In the last few years, SPAM error rates have arrived at the 1% range. One-qubit error rates have already been well *below* 1% for some time: the Barends *et al.* result around 0.08% is typical. The problem has been errors during two-qubit gates: previous work has not succeeding in reducing this error below 4% or so.

Thus, the big breakthrough reported here is a two-qubit error rate for four different 2-qubit gates among five qubits ranging between 0.6 and 1.0%. No dramatic new physics insights have been employed to produce this improvement: gains are achieved by painstaking attention to a comprehensive set of details. This work resembles an advance in metrology, in which systematic characterisation of error is used to minimise every identifiable source of imprecision.

Note that gate errors are defined by a thought experiment in which an ensemble of known quantum states is prepared, subjected to gate operation, and then measured. But the characterisation of error cannot work this way in the present experiments; this direct “tomographic” procedure is unworkable if SPAM errors are significant, and we have now reached the situation where they are very significant compared with the gate error rates to be identified. Thus, gate error characterisation requires something new, something called “randomised benchmarking”. In this scheme the error per gate is estimated by running a protocol that begins with (imperfect) state initialization, followed by randomly-drawn sequences of gates of varying length, and then (imperfect) state measurement. The growth of error with sequence length leads to an estimation of an error per gate averaged over the gate set. By interleaving applications of a specific gate with the randomly-drawn gates, the error rate of that specific gate can be estimated.

There continues to be active theoretical work on how precisely error rates can be inferred using randomized benchmarking. It is agreed that the set from which the random choices of gates should be drawn must be large, but that it is sufficient that they come from the maximal discrete subgroup of the unitary group known as the Clifford group. Theory work now recommends [3] reporting confidence intervals – error bars on the errors, if you will. This construct may strike the reader as slightly absurd – errors on the errors just means more error, no? While this may sound a little bit like the famous “unknown unknown” [2], the theory arguments for the necessity for uncertainty of density operators have become convincing. In the meantime, the poor experimentalists must do the best that they can with the measurements they can presently take. Some of the more conservative recent theory [3] suggests that the Barends *et al* data should be reported as 1% error with a factor of two error bar; tricky error processes, like ones that promote the quantum bits in higher-lying levels, are implausible, but require conservative reporting of the confidence of estimation. This is probably not a situation that will last very long: if the error rates are really close to

the value reported in this experiment, there should soon be successful application of surface code error correction that will conclusively confirm it.

- 
- [1] S. B. Bravyi, A. Yu. Kitaev, “Quantum codes on a lattice with boundary,” arXiv:quant-ph/9811052.
  - [2] D. Rumsfeld, “Known and Unknown: a memoir” (Sentinel-Penguin, 2011).
  - [3] Jeffrey M. Epstein, Andrew W. Cross, Easwar Magesan, and Jay M. Gambetta, “Investigating the limits of randomized benchmarking protocols,” arXiv:1308.2928.