



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Probing Qubit Memory Errors at the Part-per-Million Level

M. A. Sepiol, A. C. Hughes, J. E. Tarlton, D. P. Nadlinger, T. G. Ballance, C. J. Ballance, T. P. Harty, A. M. Steane, J. F. Goodwin, and D. M. Lucas

Phys. Rev. Lett. **123**, 110503 — Published 13 September 2019

DOI: [10.1103/PhysRevLett.123.110503](https://doi.org/10.1103/PhysRevLett.123.110503)

Probing qubit memory errors at the part-per-million level

M. A. Sepiol, A. C. Hughes, J. E. Tarlton, D. P. Nadlinger, T. G. Ballance,
C. J. Ballance, T. P. Harty, A. M. Steane, J. F. Goodwin,* and D. M. Lucas

Department of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, U.K.

(Dated: 6 August 2019)

Robust qubit memory is essential for quantum computing, both for near-term devices operating without error-correction, and for the long-term goal of a fault-tolerant processor. We directly measure the memory error ϵ_m for a $^{43}\text{Ca}^+$ trapped-ion qubit in the small-error regime and find $\epsilon_m < 10^{-4}$ for storage times $t \lesssim 50$ ms. This exceeds gate or measurement times by 3 orders of magnitude. Using randomized benchmarking, at $t = 1$ ms we measure $\epsilon_m = 1.2(7) \times 10^{-6}$, around ten times smaller than that extrapolated from the T_2^* time, and limited by instability of the atomic clock reference used to benchmark the qubit.

A requirement for any physical realization of a quantum computer is the capability to preserve and to exploit the coherent behaviour of its constituent qubits. During the course of a quantum computation, environmentally-induced noise and imperfections in the control apparatus inevitably lead to a dephasing of qubit superpositions, which introduces memory errors ϵ_m . While contributions to this error stemming from correlated noise can be suppressed using methods such as dynamical decoupling [1–3], quantum error correction (QEC) techniques [4, 5] are necessary to counteract the remaining error. In a typical QEC circuit, ancilla qubits need to be prepared, entangled with logic qubits, and measured, before appropriate feedback is applied to the logic qubits. Thus it is essential that the memory error remains below a correctable level at least for the duration of these operations, and preferably for even longer, to reduce the frequency with which “idle” qubits need to be corrected. Depending on the QEC methods used, and the architecture of the computer, the maximum correctable error can be as high as $\sim 10^{-2}$ [6], but a level of 10^{-4} is often taken as an important target for realistic overheads [7]. The longer ϵ_m remains correctable, the more flexibility there is in both the physical and logical design of the computer, and the greater the scope for increasing the connectivity of qubits (e.g., by moving the qubits around physically [8]).

For qubits with sufficiently long depolarization lifetimes (T_1), the memory error ϵ_m is determined by the relative stability of the qubit and reference oscillator frequencies [9]. The variation of ϵ_m with storage time t depends on the spectrum of the noise processes that affect these frequencies. The benchmark almost universally used to quantify the memory performance is the coherence time T_2^* , the time constant for decay of qubit phase coherence when modelled by exponential decay $\exp(-t/T_2^*)$. The exponential decay model assumes that the spectrum of the frequency noise is white. T_2^* can be obtained by measuring the fringe contrast in Ramsey experiments as a function of the Ramsey delay τ_R . For $\tau_R \sim T_2^*$, the memory error is much larger than qubit state preparation and measurement (SPAM) errors (typically $\epsilon_{\text{SPAM}} \gtrsim 10^{-3}$), and can therefore be measured eas-

ily. In such a method, information regarding the initial stages of decoherence – the regime relevant to quantum computing – has to be inferred by extrapolation, as in this regime $\epsilon_m \lesssim \epsilon_{\text{SPAM}}$ so the large amount of data needed to measure ϵ_m leads to impractically long experiments. Consequently, any non-exponential structure to the coherence decay at short timescales remains undetected, leaving uncertainty about the true impact of memory errors on the computer’s operation.

Previous studies of qubit decoherence have measured T_2^* ranging from minutes to hours in large ensembles of trapped ions or nuclear spins [10–13]. For single physical qubits, a Ramsey $T_2^* \approx 50$ s was measured for a $^{43}\text{Ca}^+$ ion [14], and a coherence time of $T_2 \sim 10$ min was obtained by applying dynamical decoupling pulses to a $^{171}\text{Yb}^+$ ion [15]. Only a single study [16] has attempted to quantify the memory fidelity in the low-error regime of interest for quantum computing: in that work, a technique based on interleaved randomized benchmarking (IRB) was introduced to measure memory errors much smaller than ϵ_{SPAM} , and was used in a superconducting qubit system to show that ϵ_m reached $\sim 10^{-3}$ after a time equal to the typical duration of a single entangling gate. In the present work, we characterize the memory performance of a $^{43}\text{Ca}^+$ trapped-ion hyperfine “atomic clock” qubit (frequency 3.2 GHz) both directly, using Ramsey experiments with short delays and high statistics, and indirectly, using the IRB method.

We use a microfabricated, planar surface-electrode ion trap that incorporates integrated microwave circuitry (resonators, waveguides, and coupling elements), and that is operated at room temperature. Its design, the details of which can be found in [17], allows for single- and two-qubit quantum logic gates to be driven by near-field microwave radiation instead of by lasers, eliminating photon-scattering errors and offering improved prospects for scalability [18–20]. For coherent manipulation of the qubit, we apply microwave pulses to electrodes which lie $75 \mu\text{m}$ below the ion (see §A1). The experimental control and timing of the pulse sequences (with ns precision and ps jitter) is handled via ARTIQ [21].

The main source of memory decoherence in trapped-

ion hyperfine qubits typically stems from fluctuating magnetic fields. An established method for suppressing the effect of these fluctuations is to use a qubit based on an “atomic clock” transition whose frequency is independent of magnetic field to first order [10, 14, 15, 22]. In this work, we use the intermediate-field $^{43}\text{Ca}^+$ clock qubit formed by the $|\downarrow\rangle = 4S_{1/2}^+$ and $|\uparrow\rangle = 4S_{1/2}^{3,+1}$ ground-level hyperfine states [23] at a static magnetic field of $B_0 \approx 146\text{ G}$ [14]. We limit the effect of the qubit’s second-order field dependence by stabilizing the field [24] to within $\Delta B \lesssim 1\text{ mG}$ of the field-independent point B_0 .

We first measure the memory error directly, using conventional Ramsey experiments in which the phase ϕ of the second $\pi/2$ pulse is varied relative to that of the first (fig.1a). We measure the Ramsey fringe contrast by fixing $\phi = \phi_0$ or $\phi = \phi_0 + 180^\circ$, for fixed offset ϕ_0 , rather than scanning ϕ and fitting the resulting fringe with a floated phase offset or scaling factor to compensate for slow phase drifts (as was done in prior work [14, 22]). This ensures that our measurement is sensitive to phase drifts (Z -rotations) of the qubit relative to the microwave local oscillator, as would be the case during a quantum computation. The offset ϕ_0 is calibrated before (but not during) experimental runs to compensate for any residual detuning offset of the microwaves (see §A1).

Since we aim to measure any loss in the fringe contrast (which is ideally 1) at a similar level to our $\epsilon_{\text{SPAM}} \approx 2 \times 10^{-3}$, it is important to monitor drifts in ϵ_{SPAM} , or systematic dependence of ϵ_{SPAM} on the Ramsey delay τ_R . We follow each trial of the Ramsey sequence by two control sequences to measure $\epsilon_{\text{SPAM}} = \frac{1}{2}(\epsilon_\downarrow + \epsilon_\uparrow)$: the first consists of a delay τ_R between qubit $|\uparrow\rangle$ state preparation and measurement, the second contains an additional π pulse to prepare the $|\downarrow\rangle$ state. We find that ϵ_{SPAM} shows negligible systematic variation for $\tau_R < 1\text{ s}$, and remains $\leq 10^{-2}$ for $\tau_R \leq 10\text{ s}$ [26]. Alternating SPAM measurements with Ramsey experiments in this way allows us to capture any drifts in ϵ_{SPAM} with similar statistical uncertainty as for the Ramsey data. We subtract ϵ_{SPAM} from the measured contrast loss for each τ_R .

Results of this experiment, which required a week of continuous data acquisition, are shown in fig.1b. For a direct comparison with the IRB results (below), we can convert the loss in Ramsey fringe contrast to a memory error by scaling it by a factor of $\frac{1}{3}$: this takes into account that the error measured with IRB varies between 0 and $\frac{1}{2}$; and that in the IRB experiments the qubit state spends only $\frac{2}{3}$ of the time near the equator of the Bloch sphere, where it is sensitive to dephasing. The memory error is characterized down to a delay of $\tau_R = 200\text{ ms}$, at which point uncertainty in ϵ_{SPAM} limits our knowledge of ϵ_m to an upper bound only. Collecting more data would reduce this uncertainty, but only with the square root of the acquisition time, making this approach impractical for exploring ϵ_m below the 10^{-4} level; it would also require systematic drifts in ϵ_{SPAM} to be $\ll 10^{-4}$. Fig.1b

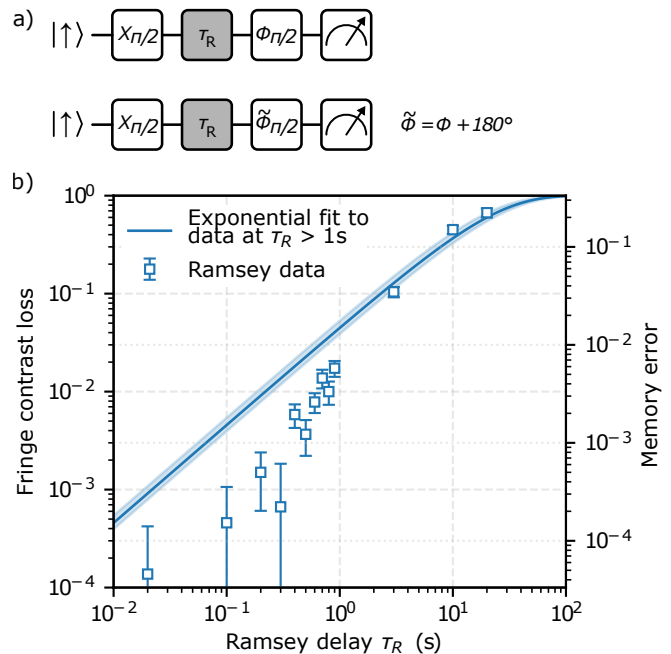


FIG. 1. Two-point Ramsey experiments for measuring the qubit memory error. (a) Sequences used to measure the maximum and minimum of the Ramsey fringe. The phase difference between the final $\pi/2$ pulses remains fixed at $\phi - \phi = 180^\circ$ throughout the experiments. (b) SPAM-corrected loss of Ramsey fringe contrast. The line represents exponential contrast decay (fixed at 0 for $\tau_R = 0$) fitted to the data with $\tau_R > 1\text{ s}$, where the SPAM error is negligible compared with the contrast error; the fit gives $T_2^* = 22(3)\text{ s}$. The right-hand ordinate gives the memory error, averaged over the Bloch sphere, associated with a given Ramsey contrast loss.

also illustrates that assuming an exponential decay based on contrast measurements at long delays, as is customary in T_2^* measurements, would here lead to a significant overestimate of the memory error for shorter delays.

To circumvent the limitation imposed on memory error measurements by the SPAM error, we follow the approach introduced by O’Malley *et al.* [16] which employs the technique of IRB. This method amplifies the memory error relative to the SPAM error by subjecting the qubit to m periods τ of dephasing instead of to a single period, while ensuring that errors add incoherently by interleaving each delay τ with a Clifford gate C_i , sampled randomly from the full single-qubit Clifford group. We call the probability with which a given gate sequence of length m produces the predicted final state the *sequence fidelity*, and for each m we calculate the average sequence fidelity over $k = 50$ distinct random sequences. The average sequence fidelity follows a decay $\frac{1}{2}(ap^m + b)$ with increasing m [27, 29], where $p \in [0, 1]$ is the *depolarizing parameter*, related to the average error rate by $\epsilon = \frac{1}{2}(1 - p)$. Comparing the average sequence fidelity of IRB sequences with interleaved delays to that of “reference” RB sequences without interleaved delays allows us to isolate

the average dephasing error associated with each delay, which we call the memory error $\epsilon_m = \frac{1}{2}(1 - p_{\text{IRB}}/p_{\text{RB}})$. The depolarizing parameter of the reference sequences p_{RB} incorporates Clifford gate errors alone, while that of the interleaved sequences p_{IRB} also includes the memory errors from the delays. The parameter $a = 1 - 2\epsilon_{\text{SPAM}}$ captures SPAM errors, which we ensure are equal in the reference and IRB sequences by adding a delay $m\tau$ after the final gate of the reference sequence. We fix $b = 1$ as $\epsilon_{\text{SPAM}} \ll 1$. In this measurement of ϵ_m , the precision attainable is no longer constrained by SPAM errors, but by the magnitude of the Clifford gate errors, while the accuracy attainable is limited by any systematic changes of the Clifford gate errors when delays are interleaved.

We first characterize the Clifford gate errors *without* extra delays, using the “standard RB” (SRB) method [27], which involves applying random sequences of the form shown in fig.2a. The Cliffords are composed of $\pm X_{\pi/2}$ and $\pm Y_{\pi/2}$ rotations on the Bloch sphere, with an average of $3.50 \pi/2$ rotations per Clifford [28], separated by $12 \mu\text{s}$ to allow the DDS source time to switch between pulse profiles. The duration of each pulse is set to be $\sim 10 \mu\text{s}$ and periodically fine-tuned by optimizing the sequence fidelity for fixed m (typically $m = 2000$). Fig.2b shows the measured sequence fidelity decay, yielding an average error per Clifford gate of $\epsilon_g = 1.7(2) \times 10^{-6}$. A previous measurement of the gate error in this trap using the “NIST RB” method [14, 29] (which used an average of two $\pi/2$ rotations per gate) gave $\epsilon_g = 1.0(3) \times 10^{-6}$. Based on numerical modelling of known experimental imperfections [30], the SRB and NIST RB methods are expected to yield a similar gate error; the 2σ discrepancy may be due to the larger number of physical $\pi/2$ rotations per gate used in the SRB experiment.

Systematic variations of ϵ_g when the delays τ are interleaved can occur because the associated change in microwave duty cycle causes thermally-induced shifts in the microwave power, and these variations would limit the accuracy of the IRB experiment if not prevented. Recalibrating the $\pi/2$ -pulse time for each delay length τ by optimizing the sequence fidelity (as above) is impractical, as the interleaved delays significantly slow down the experiments. Instead, we make the $\pi/2$ -rotations robust to pulse area changes by replacing them with BB1 composite gates [31]. This involves appending a $\{\phi_\pi, (3\phi)_{2\pi}, \phi_\pi\}$ sequence to each $\pi/2$ -pulse (fig.3a), with ϕ appropriately chosen for every different $\pi/2$ -pulse phase. Due to technical limitations of the DDS microwave source, implementing BB1 composite pulses required a Clifford gate decomposition into $+X_{\pi/2}$ and $+Y_{\pi/2}$ gates only, with an average of $3.58 \pi/2$ rotations (i.e. 14.3 microwave pulses) per Clifford. The additional physical pulses necessary for the BB1 gates increase the average Clifford gate error to $\epsilon'_g = 6(1) \times 10^{-6}$; no change in ϵ'_g was measured for a microwave pulse amplitude reduction of 1%, much greater than that observed due to duty cycle effects.

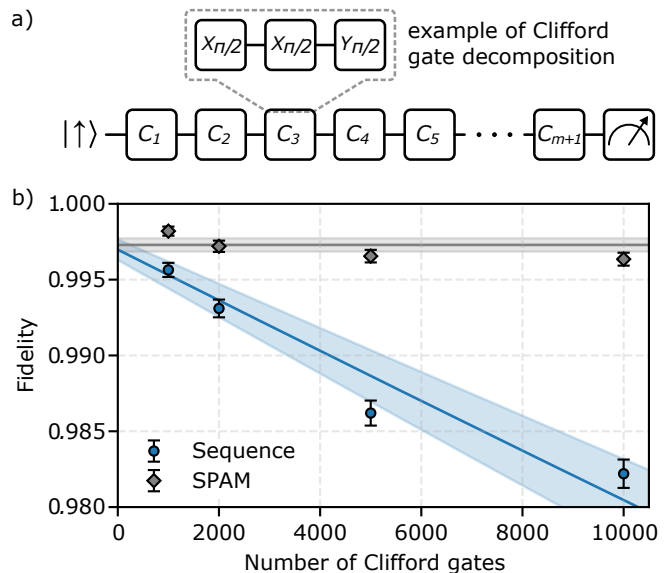


FIG. 2. “Standard” randomized benchmarking of single-qubit gates. (a) Gate diagram of the sequence used, including an example decomposition of a Clifford gate C_i into $\pi/2$ -pulses. The qubit is prepared in $|\uparrow\rangle$ and then subjected to m random gates from the Clifford set, followed by an $(m+1)$ th gate that is chosen to rotate the qubit to one of its two basis states, selected with equal probability. Each RB sequence is alternated with a measurement of ϵ_{SPAM} , for which the delay between state preparation and measurement is equal to the duration of the RB sequence; this allows us to check for any systematic dependence of ϵ_{SPAM} on sequence duration. (b) Measured sequence fidelities (blue circles) and SPAM fidelities (grey diamonds), as a function of sequence length. The fit to a decay $\frac{1}{2}(ap^m + 1)$ (blue line) gives an average Clifford gate error of $\epsilon_g = 1.7(2) \times 10^{-6}$. The fit intercept is consistent with the mean measured SPAM error of $2.7(4) \times 10^{-3}$ (grey line). Error bars on each point are the standard error of the mean over $k = 50$ sequence randomizations. Shaded regions represent the 1σ uncertainties of the fits.

Results of the measurement of the qubit memory error ϵ_m using IRB with BB1 gates are shown in fig.3c (blue circles). During each data point, which requires up to 2 days of continuous acquisition, we measure the detuning between the qubit transition and the microwave source every 4 hours and correct for slow drifts; from these measurements, we estimate that the contribution to ϵ_m due to detuning errors is negligible [26]. Also plotted in fig.3c (grey squares) is the data from the Ramsey experiments (fig.1). The memory error from both methods is consistent, but the superior sensitivity of the IRB approach enables characterization of ϵ_m for delays as short as $\tau = 1 \text{ ms}$, where, at $\epsilon_m = 1.2(7) \times 10^{-6}$, it approaches the noise floor set by the BB1 Clifford gate error. We fit a decoherence model to the IRB data, which assumes white and pink ($1/f$) frequency noise. The memory error for $\tau \lesssim 100 \text{ ms}$ is consistent with that expected from independently measured $1/f^2$ phase noise (white frequency

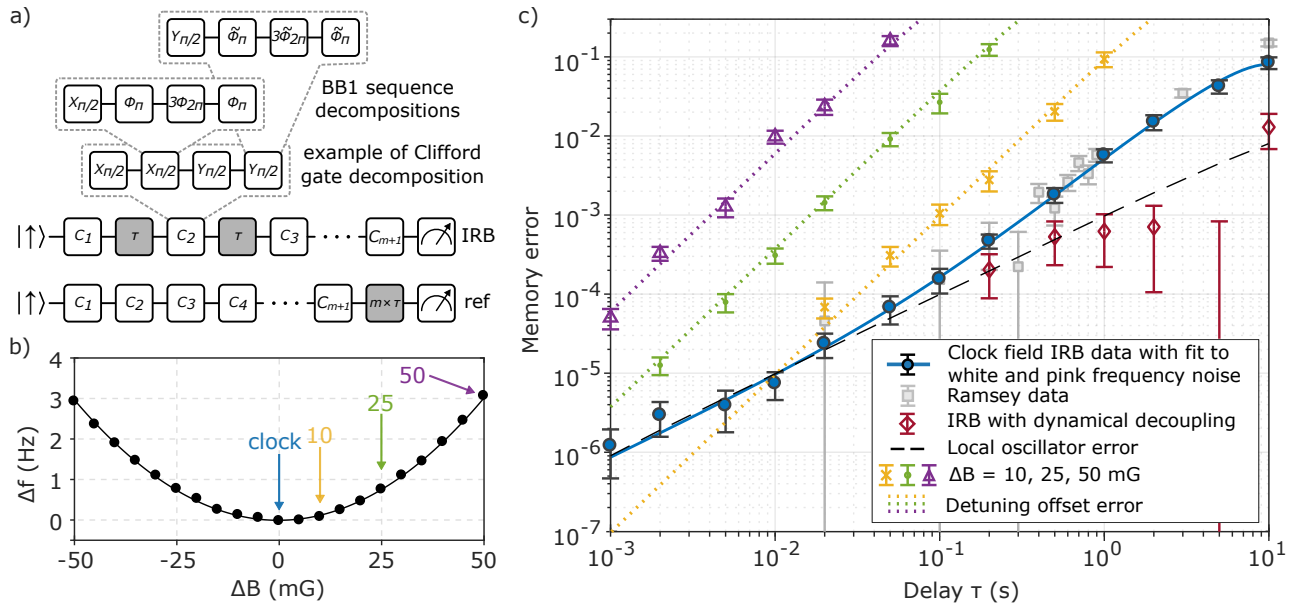


FIG. 3. Single-qubit memory errors ϵ_m measured by interleaved randomized benchmarking. (a) IRB sequence, with delays τ inserted between each Clifford gate C_i . An example decomposition of a Clifford gate C_i into $+X_{\pi/2}$ and $+Y_{\pi/2}$ gates is shown. Also shown are the associated BB1 decompositions: each $X_{\pi/2}$ gate is followed by a sequence of rotations $\{\phi_\pi, (3\phi)_{2\pi}, \phi_\pi\}$; similarly for each $Y_{\pi/2}$ gate, but around an axis $\tilde{\phi}$. For the reference sequence (ref), an additional delay $m\tau$ is inserted after the last Clifford gate to keep the time between the qubit initialization and readout equal to that in the IRB sequence, thus minimizing any systematic differences in ϵ_{SPAM} . (b) Qubit frequency offset Δf versus offset ΔB of the magnetic quantization field, relative to the field-independent “clock” point. (c) Average memory error measured for interleaved delays $1 \text{ ms} < \tau < 10 \text{ s}$. At each τ , we choose m to give a total sequence infidelity of ~ 0.1 , which is much larger than the SPAM error. The IRB data (blue circles) is consistent with the Ramsey results (grey squares, from fig.1). We fit a decoherence model to the IRB data (blue line - see §A3 for details), which contains white noise and $1/f$ noise with a low-frequency cut-off (this cut-off is consistent with the duration of the longest sequences). IRB sequences including dynamical decoupling (red diamonds) show that the error due to $1/f$ noise can be suppressed at long times using this technique. The dashed black line shows the calculated error contribution from the local oscillator (atomic clock) to which the qubit phase is compared: this accounts for the measured memory error in the $\epsilon_m \lesssim 10^{-4}$ regime. Finally the yellow, green and purple data points show the error measured when the magnetic field is deliberately offset from the field-independent point by $\Delta B = (10, 25, 50) \text{ mG}$; note that the local oscillator was left at $\Delta f = 0$, so that the observed error arises mostly from the corresponding qubit detuning $\Delta f \approx (0.12, 0.76, 3.0) \text{ Hz}$ (dotted lines).

noise) on the rubidium atomic clock to which the microwave source is referenced (see §A4). From decoherence measurements of a field-sensitive hyperfine transition, we estimate that contributions to the memory error from magnetic field noise are several orders of magnitude lower than our measured ϵ_m (see §A2). The excess error for $\tau \gtrsim 100 \text{ ms}$ may be due to other slow drifts, for example in the ac Zeeman shift ($\approx -6 \text{ Hz}$) arising from the trap rf fields [32]. We also took IRB data with a simple dynamical decoupling sequence (an X_π pulse inserted every 100 ms during delays $\tau \geq 200 \text{ ms}$), which suppressed the correlated ($1/f$) noise, reducing ϵ_m at longer delays to a level consistent with the atomic clock noise.

An important consideration for a many-qubit processor based on the “quantum CCD” architecture [8] will be inhomogeneity of the static magnetic field across the device, which will lead to departures ΔB of the field from the qubits’ field-independent point B_0 . To simulate the effect of field inhomogeneity, we took further IRB data with ΔB set to several offset values (fig.3b,c).

To obtain a ‘worst case’ error, we did not adjust the microwave detuning for the known offset Δf ; even so, for $\Delta B = 50 \text{ mG}$, we find $\epsilon_m < 10^{-4}$ at $\tau = 1 \text{ ms}$. This error is dominated by the effect of the (known) detuning offset, here $\frac{2}{3} [1 - \cos^2(\pi\tau\Delta f)] \approx 6 \times 10^{-5}$, which could be corrected for by calibrating the field across the processor.

In conclusion, we have characterized the memory errors of a $^{43}\text{Ca}^+$ trapped-ion hyperfine clock qubit to the 10^{-4} level by Ramsey measurement, and to the 10^{-6} level using interleaved randomized benchmarking. The error is consistent with $1/f$ frequency noise at long timescales, contrary to the white noise model often assumed in T_2^* measurements. The memory error remains below the 10^{-4} level relevant to QEC for up to 50 ms, which is around 3 orders of magnitude longer than the time required for entangling gates or qubit measurement [33–36]. At these sub- 10^{-4} error levels, the memory error is consistent with the independently-measured phase noise of the local oscillator, implying that a more stable reference clock would lead to improved performance.

We thank J. Emerson and K. Boone for useful discussions, and M. W. Lucas for the loan of a time-to-amplitude converter. The ion trap was designed by D³.T.C. Allcock. J.E.T. acknowledges funding from the Centre for Doctoral Training on Controlled Quantum Dynamics at Imperial College London. This work was supported by the U.S. Army Research Office (ref. W911NF-14-1-0217) and the U.K. EPSRC “Networked Quantum Information Technology” Hub.

* joseph.goodwin@physics.ox.ac.uk

- [1] G. S. Uhrig, Phys. Rev. Lett. **98**, 100504 (2007).
- [2] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature **458**, 996 (2009).
- [3] D. J. Szwer, S. C. Webster, A. M. Steane, and D. M. Lucas, Journal of Physics B **44**, 025501 (2010).
- [4] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).
- [5] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).
- [6] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).
- [7] A. M. Steane, Quant. Inf. Comp. **7**, 171 (2007).
- [8] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, J. Res. NIST **103** (1998).
- [9] H. Ball, W. D. Oliver, and M. J. Biercuk, npj Quantum Information **2**, 16033 (2016).
- [10] J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, and D. J. Wineland, IEEE Trans. Instr. Meas. **40**, 126 (1991).
- [11] P. T. H. Fisk, M. J. Sellars, M. A. Lawn, C. Coles, A. G. Mann, and D. G. Blair, IEEE Trans. Instr. Meas. **44**, 113 (1995).
- [12] K. Saeedi, S. Simmons, J. Z. Salvail, P. Dluhy, H. Riemann, N. V. Abrosimov, P. Becker, H. J. Pohl, J. J. L. Morton, and M. L. W. Thewalt, Science **342**, 830 (2013).
- [13] M. Zhong, M. P. Hedges, R. L. Ahlefeldt, J. G. Bartholomew, S. E. Beavan, S. M. Wittig, J. J. Longdell, and M. J. Sellars, Nature **517**, 177 (2015).
- [14] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, and D. M. Lucas, Phys. Rev. Lett. **113**, 220501 (2014).
- [15] Y. Wang, M. Um, J. Zhang, S. An, M. Lyu, J.-N. Zhang, L.-M. Duan, D. Yum, and K. Kim, Nature Photonics **11**, 646 (2017).
- [16] P. J. J. O’Malley, J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, Phys. Rev. Applied **3**, 044009 (2015).
- [17] D. T. C. Allcock, T. P. Harty, C. J. Ballance, B. C. Keitch, N. M. Linke, D. N. Stacey, and D. M. Lucas, Applied Physics Letters **102**, 044103 (2013).
- [18] C. Ospelkaus, C. E. Langer, J. M. Amini, K. R. Brown, D. Leibfried, and D. J. Wineland, Phys. Rev. Lett. **101**, 090502 (2008).
- [19] C. Ospelkaus, U. Warring, Y. Colombe, K. R. Brown, J. M. Amini, D. Leibfried, and D. J. Wineland, Nature **476**, 181 (2011).
- [20] T. P. Harty, M. A. Sepiol, D. T. C. Allcock, C. J. Ballance, J. E. Tarlton, and D. M. Lucas, Phys. Rev. Lett. **117**, 140501 (2016).
- [21] S. Bourdeauducq, R. Jördens, P. Zotov, J. Britton, D. Slichter, D. Leibbrandt, D. Allcock, A. Hankin, F. Kermarrec, Y. Sionneau, R. Srinivas, T. R. Tan, and J. Bohnet, “ARTIQ 1.0,” (2016).
- [22] C. Langer, R. Ozeri, J. D. Jost, J. Chiaverini, B. DeMarco, A. Ben-Kish, R. B. Blakestad, J. Britton, D. B. Hume, W. M. Itano, D. Leibfried, R. Reichle, T. Rosenband, T. Schaetz, P. O. Schmidt, and D. J. Wineland, Phys. Rev. Lett. **95**, 060502 (2005).
- [23] Here we use the spectroscopic notation $nL_J^{F,M}$.
- [24] See §A2 for details on the magnetic field stabilization setup, which includes reference [25].
- [25] B. Merkel, K. Thirumalai, J. E. Tarlton, V. M. Schäfer, C. J. Ballance, T. P. Harty, and D. M. Lucas, Review of Scientific Instruments **90**, 044702 (2019).
- [26] J. E. Tarlton, Ph.D. thesis, Imperial College London, UK (2018).
- [27] J. Emerson, R. Alicki, and K. Życzkowski, Journal of Optics B: Quantum and Semiclassical Optics **7**, S347 (2005).
- [28] The minimal Clifford decomposition for $\{\pm X_{\pi/2}, \pm Y_{\pi/2}\}$ physical gates and a ‘virtual’ $Z_{\pi/2}$ leads to an average of 1.41 physical gates per Clifford. The longer decomposition used in this paper was chosen for historical reasons..
- [29] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).
- [30] K. Boone, private communication (2018).
- [31] S. Wimperis, J. Mag. Res., Series A **109**, 221 (1994).
- [32] T. P. Harty, Ph.D. thesis, Oxford University, UK (2013).
- [33] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, Phys. Rev. Lett. **117**, 060504 (2016).
- [34] J. P. Gaebler, T. R. Tan, Y. Lin, Y. Wan, R. Bowler, A. C. Keith, S. Glancy, K. Coakley, E. Knill, D. Leibfried, and D. J. Wineland, Phys. Rev. Lett. **117**, 060505 (2016).
- [35] V. M. Schäfer, C. J. Ballance, K. Thirumalai, L. J. Stephenson, T. G. Ballance, A. M. Steane, and D. M. Lucas, Nature **555**, 75 (2018).
- [36] S. Crain, C. Cahall, G. Vrijsen, E. E. Wollman, M. D. Shaw, V. B. Verma, S. W. Nam, and J. Kim, arXiv:1902.04059 (2019).