

# Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking

J T Muhonen<sup>1</sup>, A Laucht<sup>1</sup>, S Simmons<sup>1</sup>, J P Dehollain<sup>1</sup>, R Kalra<sup>1</sup>,  
F E Hudson<sup>1</sup>, S Freer<sup>1</sup>, K M Itoh<sup>2</sup>, D N Jamieson<sup>3</sup>, J C McCallum<sup>3</sup>,  
A S Dzurak<sup>1</sup> and A Morello<sup>1</sup>

<sup>1</sup> Centre for Quantum Computation and Communication Technology, School of Electrical Engineering and Telecommunications, UNSW Australia, Sydney, NSW 2052, Australia

<sup>2</sup> School of Fundamental Science and Technology, Keio University, 3-14-1 Hiyoshi, 223-8522, Japan

<sup>3</sup> Centre for Quantum Computation and Communication Technology, School of Physics, University of Melbourne, Melbourne, Victoria 3010, Australia

E-mail: [juha.muhonen@unsw.edu.au](mailto:juha.muhonen@unsw.edu.au) and [a.morello@unsw.edu.au](mailto:a.morello@unsw.edu.au)

Received 7 August 2014, revised 3 October 2014

Accepted for publication 10 November 2014

Published 18 March 2015



CrossMark

## Abstract

Building upon the demonstration of coherent control and single-shot readout of the electron and nuclear spins of individual <sup>31</sup>P atoms in silicon, we present here a systematic experimental estimate of quantum gate fidelities using randomized benchmarking of 1-qubit gates in the Clifford group. We apply this analysis to the electron and the ionized <sup>31</sup>P nucleus of a single P donor in isotopically purified <sup>28</sup>Si. We find average gate fidelities of 99.95% for the electron and 99.99% for the nuclear spin. These values are above certain error correction thresholds and demonstrate the potential of donor-based quantum computing in silicon. By studying the influence of the shape and power of the control pulses, we find evidence that the present limitation to the gate fidelity is mostly related to the external hardware and not the intrinsic behaviour of the qubit.

Keywords: quantum information, silicon, single dopant

(Some figures may appear in colour only in the online journal)

## 1. Introduction

The discovery of quantum error correction is among the most important landmarks in quantum information science [1–4]. Together with the steady improvement in the coherence and fidelity of various physical quantum bit (qubit) platforms, it represents one of the main motivations for the investment in quantum information technologies. In general, uncontrolled interactions between the qubit and its environment can cause loss of coherence and gate errors. Quantum error correcting codes, however, guarantee that the errors can be recovered, provided that the average error rate is below a certain fault-tolerance threshold [5]. The threshold is strongly dependent on the quantum computer architecture: for instance, a bilinear nearest-neighbour array requires errors below  $10^{-6}$  to achieve

fault-tolerance [6]. More recent ideas have shown the tantalizing prospect of fault-tolerance error correction with error rates that can be as high as  $10^{-2}$  [7, 8]. This level of gate fidelity has become accessible to the most advanced qubit platforms, such as superconducting qubits [9] and trapped ions [10].

Quantifying gate errors is, however, not trivial. The most commonly used protocol, quantum process tomography (QPT) [11], is based upon preparing different input states (chosen to form a complete basis) and processing them identically many times. The output states are then characterized with quantum state tomography to extract all the components of the process matrix and quantify the process errors. QPT can in principle be applied to any process and state space, but it is not scalable to large number of qubits since the number of measurements

required to completely map a multi-qubit process increases exponentially with the number of qubits involved.

In addition and more importantly for the single qubit system discussed here, quantum process tomography is also sensitive to errors in the state preparation and measurement (SPAM errors) and cannot distinguish between these and pure control errors [12]. Therefore, with QPT it is only possible to characterize gate fidelities if the gate errors are larger than preparation and readout errors. This is an unsatisfying situation, since fault-tolerant quantum computation normally poses much more stringent requirements on gate fidelities than on initialization and readout.

Randomized benchmarking [13, 14] has gained attention in recent years as a scalable solution for determining gate fidelities and has been experimentally demonstrated with e.g. atomic ions [15–17], nuclear magnetic resonance [18] and superconducting qubits [9, 12]. The goal of randomized benchmarking protocols is to extract the *average gate fidelity*, defined as the average fidelity of the output state over pure input states. The fidelity of the output state  $\mathcal{E}(\rho)$  as compared to the ideal output state  $U(\rho)$  is defined as [3]

$$\mathcal{F}_U = \left( \text{tr} \sqrt{\sqrt{\mathcal{E}(\rho)} U(\rho) \sqrt{\mathcal{E}(\rho)}} \right)^2, \quad (1)$$

where  $U$  is the ideal gate operation,  $\mathcal{E}$  the actual operation and  $\rho$  is the density matrix describing the input state.

Here we adopt a benchmarking method based upon the construction and application of random sequences of gates that belong to the Clifford group. It has been shown [19–21] that the average fidelity of a Clifford gate can be extracted from the fidelity of the final state of the qubit, averaged over several random sequences of Clifford gates, having started from a fixed input state. This provides a scalable benchmarking method with well-defined conditions of applicability and which does not depend on SPAM errors [19–21]. The Clifford group does not provide a universal set of quantum gates, but a universal set can be constructed from them by the addition of only one gate or, alternatively, by using ancilla qubits and their measurements [3]. Also, the Clifford gates play an important role in the error correction schemes based on stabilizer codes [22].

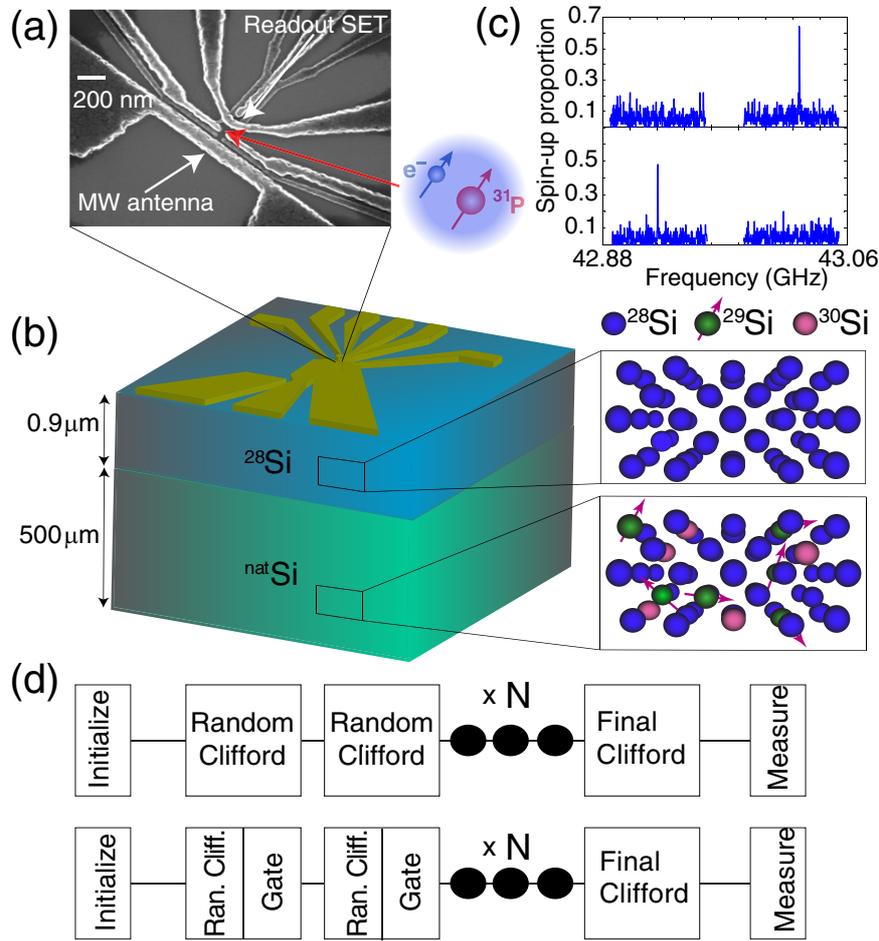
Spin qubits based upon the electron or nuclear spin of phosphorus atoms in isotopically purified  $^{28}\text{Si}$  [23] are known to have extraordinary coherence times [24, 25] and recent experiments have shown that such record coherence can be retained also at the single-atom level in functional nanoelectronic devices [26]. Here, we present a thorough investigation of the gate fidelities of the single-atom spin qubit device described in [26]. We focus on the gate fidelity of qubits represented by the electron ( $e^-$ ) and ionized nuclear spin ( $^{31}\text{P}^+$ ) with randomized benchmarking protocols using Clifford-group gates. We show that all 1-qubit gate fidelities are consistently above 99.8% for the electron (with a deduced average of 99.95% from a long sequence of Clifford gates) and above 99.9% for the nucleus (with an average of 99.99%). These results, combined with the previously reported record coherence times, show that individual dopants in  $^{28}\text{Si}$  are one of the best physical realizations of quantum bits.

## 2. Single-atom spin qubit device

Our qubit system is a single substitutional  $^{31}\text{P}$  donor in silicon, fabricated and operated as described in detail in [26–30]. In particular, the system used in the present experiment is the same as Device B in [26]. Both the donor-bound electron ( $e^-$ ) and the nucleus ( $^{31}\text{P}$ ) possess a spin 1/2 and each encode a single qubit. The qubit logic states are the simple spin up/down eigenstates, which we denote with  $|\uparrow\rangle, |\downarrow\rangle$  for the electron and  $|\uparrow\rangle, |\downarrow\rangle$  for the nucleus. The electron and the nucleus are coupled by the contact hyperfine interaction  $A$ . In the presence of a large magnetic field  $B_0$ , the resulting eigenstates are, to a very good approximation, the separable tensor products of the electro-nuclear basis states ( $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$ ).

Arbitrary quantum states are encoded on the  $e^-$  qubit by applying pulses of oscillating magnetic field  $B_1$  at the frequencies corresponding to the electron spin resonance (ESR),  $\nu_{e1,2} \approx \gamma_e B_0 \mp A/2$  [29], where  $\gamma_e = 27.97 \text{ GHz T}^{-1}$  and  $A = 96.9 \text{ MHz}$  in this specific device [26]. In the randomized benchmarking experiments discussed below, we operated the  $e^-$  qubit at  $\nu_{e2} \approx \gamma_e B_0 + A/2$  while the nuclear spin was in the  $|\uparrow\rangle$  state. The nuclear spin qubit was operated in the ionized charge state of the donor ( $\text{D}^+$ ), where the nuclear magnetic resonance (NMR) frequency is simply  $\nu_{n+} = \gamma_n B_0$ , where  $\gamma_n = 17.23 \text{ MHz T}^{-1}$  is nuclear gyromagnetic ratio [30]. The experiments were performed in high magnetic fields ( $B_0 \approx 1.5 \text{ T}$  applied along the [1 1 0] crystal axis of Si) and low temperatures (electron temperature  $T_{\text{el}} \approx 100 \text{ mK}$ ). Microwave control fields  $B_1$  are produced by a broadband on-chip microwave antenna [31] terminating  $\sim 100 \text{ nm}$  away from the donor qubit. At microwave frequencies, the cable and the cold attenuators connecting the source to the on-chip antenna provide  $\sim 40 \text{ dB}$  attenuation.

The device structure is shown in figure 1(a). The substrate is a  $0.9 \mu\text{m}$  thick epilayer of isotopically purified  $^{28}\text{Si}$ , grown on top of a  $500 \mu\text{m}$  thick  $^{\text{nat}}\text{Si}$  wafer (figure 1(b)). The  $^{28}\text{Si}$  epilayer contains 800 ppm residual  $^{29}\text{Si}$  isotopes. A stack of aluminium gates above the  $\text{SiO}_2$  is used to induce a single-electron transistor (SET) underneath the oxide. The single-donor qubit is selected out of a small number of ion-implanted  $^{31}\text{P}$  atoms [32], placed  $\approx 10 \text{ nm}$  below the Si/SiO<sub>2</sub> interface and underneath an additional stack of control gates. The single-shot electron spin readout is based on its spin-dependent tunneling [28, 33] into the nearby SET island. The readout process also leaves the electron spin initialized in the  $|\downarrow\rangle$  state. Since this only occurs when the donor is brought in resonance with the Fermi level of the detector SET, we can tune the electrostatic gates fabricated above the donor implant area to ensure that, at any time, only a single donor is able to undergo electron tunneling events. The successful isolation of a single  $^{31}\text{P}$  donor is unequivocally proven by ESR experiments where only one of the ESR frequencies is active (figure 1(c)) at any time. The ESR frequency sporadically jumps between  $\nu_{e1}$  and  $\nu_{e2}$ , signalling the flip of the nuclear spin state [30]. The nuclear spin readout is obtained by applying a  $\pi$ -pulse at frequency  $\nu_{e2}$  to an electron initially  $|\downarrow\rangle$ . If the electron is successfully flipped to the  $|\uparrow\rangle$  state, the nuclear spin state is declared  $|\uparrow\rangle$ .



**Figure 1.** Device structure and schematic of the benchmarking protocols. (a) Scanning electron micrograph of a device similar to the one used in the experiment, highlighting the position of the P donor, the microwave (MW) antenna and the SET for spin readout. (b) Schematic of the Si substrate, consisting of an isotopically purified  $^{28}\text{Si}$  epilayer (with a residual  $^{29}\text{Si}$  concentration of 800 ppm) on top of a natural Si wafer. (c) Electron spin resonance measurements. Each of the two resonance lines correspond to one of the possible nuclear spin orientations. Since we observe a single  $^{31}\text{P}$  donor, any data trace shows only one resonance. (d) Schematic of the randomized benchmarking protocol for measuring the average Clifford gate fidelity (top) and the fidelity of individual interleaved gates (bottom).

### 3. Randomized benchmarking protocol for average Clifford gate fidelity

The randomized benchmarking protocol we adopt is based upon the one studied theoretically in [19–21] and demonstrated experimentally with e.g. superconducting qubits and silicon quantum dots in [9, 34]. It is based on measuring the probability of preserving an initial quantum state while applying a variable number of Clifford gate operations. Within the validity of certain assumptions [20, 21], this probability is predicted to decay exponentially with the number of gate operations and the decay constant is related to the average Clifford gate fidelity.

Specifically, for a sequence of  $N$  gate operations, we choose randomly (with uniform probability) the  $N$  gates from the 24 Clifford gates, where each Clifford gate is composed of  $\pi$  and  $\pi/2$  rotations around different axes (table 1). We then add a final Clifford gate to the sequence, chosen to ensure that—if all the previous operations were ideal—the final qubit state will be an eigenstate of the observable  $\sigma_z$  accessible to our measurement. Since the qubit readout is single-shot, we

have to repeat  $r$  times the same gate sequence to extract the probability of recovering the expected state. For the  $e^-$  qubit we used  $r = 200$ . To average over the whole gate space, the process described above is repeated  $K$  times (at constant  $N$ ) for different random gate sets. It has been shown [21] that the measured fidelity converges to the correct average gate fidelity for  $K \gg 10$ . As a reasonable compromise between convergence and measurement time, we have chosen the value  $K = 15$ . Finally, the sequence is repeated for  $n$  different values of  $N$ , to extract the decay of the process fidelity upon increasing  $N$ . A rigorous analysis of the confidence intervals for randomized benchmarking as a function of  $K$  and  $N$  is given in [35].

In order to respect the assumptions upon which the theoretical framework of Clifford-group randomized benchmarking is based, one should normally choose the final Clifford gate such that the target output state coincides with the input state ( $|\downarrow\rangle$  in the case of the  $e^-$  qubit) [19]. Our measurement procedure, however, benefits considerably from choosing the target output state randomly, since this allows us to exactly cancel out certain measurement errors. We then need

**Table 1.** The Clifford group gates written as the physical gates applied. The gate numbering is arbitrary, we follow the grouping from [21]. Physical gate X denotes a  $\pi$ -pulse around the X-axis (i.e. unitary operation  $\exp(-i\pi\sigma_x/2)$ ), Y/2 a  $\pi/2$ -pulse around the Y-axis ( $\exp(-i\pi\sigma_y/4)$ ), etc. The average number of physical gates per Clifford gate is 1.875.

Clifford gate	Physical gates
1	I
2	Y/2, X/2
3	-X/2, -Y/2
4	X
5	-Y/2, -X/2
6	X/2, -Y/2
7	Y
8	-Y/2, X/2
9	X/2, Y/2
10	X, Y
11	Y/2, -X/2
12	-X/2, Y/2
13	Y/2, X
14	-X/2
15	X/2, -Y/2, -X/2
16	-Y/2
17	X/2
18	X/2, Y/2, X/2
19	-Y/2, X
20	X/2, Y
21	X/2, -Y/2, X/2
22	Y/2
23	-X/2, Y
24	X/2, Y/2, -X/2

less free parameters in our fits, which considerably improves their accuracy. Below we show that results obtained while taking a random final state do not differ quantitatively from analysing just the gate sets where the output state is  $|\downarrow\rangle$ . This is understandable, since the errors only accumulate significantly after  $\sim 100$  gates and the effect of one last gate that breaks the symmetry of the sequence is not expected to be important.

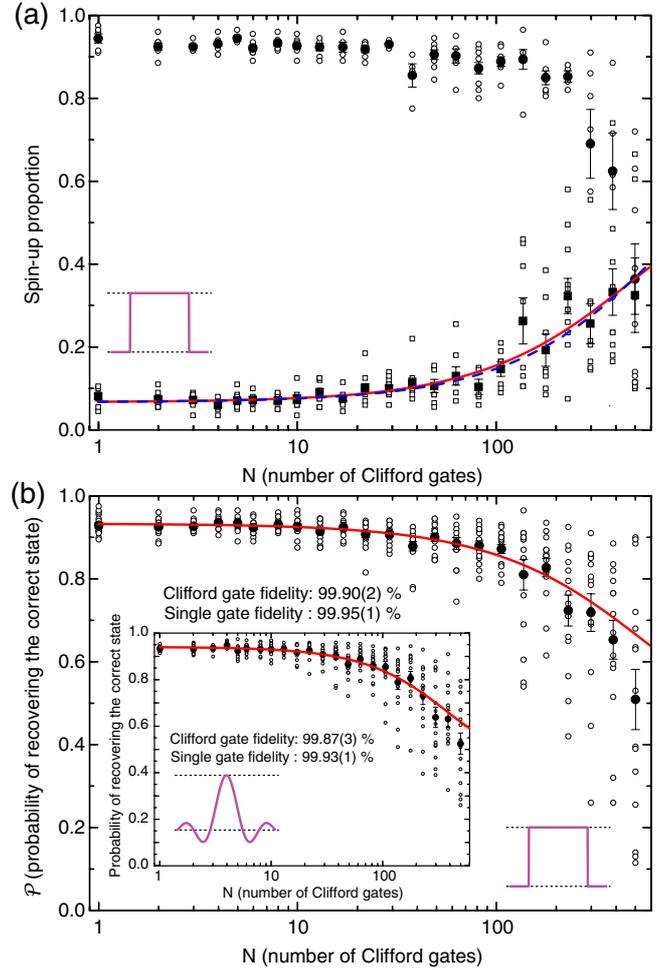
Physically, the electron spin readout process [28] yields the probability  $P_\uparrow$  of the qubit being  $|\uparrow\rangle$  at the end of the sequence. Even with a perfect sequence designed to output  $|\uparrow\rangle$ , the measured  $P_\uparrow$  is  $< 1$  due to SPAM errors. Similarly, a perfect sequence designed to output  $|\downarrow\rangle$  will yield  $P_\uparrow > 0$ . In figure 2(a) we plot  $P_\uparrow$  as a function of  $N$  separately for sequences designed to output either  $|\uparrow\rangle$  (circles) or  $|\downarrow\rangle$  (squares). We define  $P_\uparrow^{|\downarrow\rangle}$  ( $P_\uparrow^{|\uparrow\rangle}$ ) to mean the measured  $P_\uparrow$  when the target state was  $|\downarrow\rangle$  ( $|\uparrow\rangle$ ). We can fit the  $|\downarrow\rangle$  data points with the function:

$$P_\uparrow^{|\downarrow\rangle}(N) = P_0^{|\downarrow\rangle} p_c^N + P_\infty, \quad (2)$$

where  $p_c$  is related to the average Clifford gate fidelity through [19]

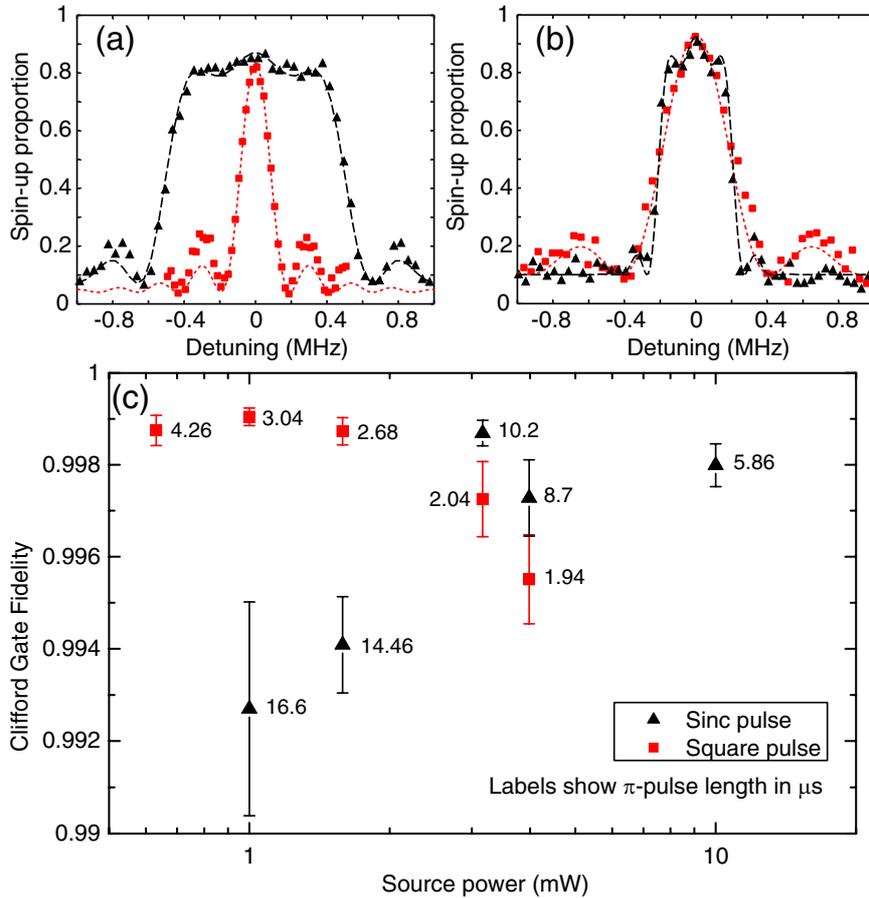
$$\mathcal{F}_c = (1 + p_c)/2, \quad (3)$$

$P_0^{|\downarrow\rangle} < 0$  is related to ‘dark counts’ arising from SPAM errors and  $P_\infty$  represents the probability of measuring  $|\uparrow\rangle$  on a completely random state. If only the sequences designed to yield  $|\downarrow\rangle$  had been measured, we should fit the data leaving  $P_0^{|\downarrow\rangle}$ ,  $p_c$  and  $P_\infty$  as free fitting parameters (dashed line in figure 2(a)). The uncertainty on  $P_\infty$  is rather large, since it



**Figure 2.** (a) Randomized benchmarking measurements on the electron spin qubit, using square pulses at 1 mW source power ( $\sim 100$  nW at the device). Data has been separated to show the case where the target output state is  $|\uparrow\rangle$  (circles) or  $|\downarrow\rangle$  (squares). Open symbols show the electron spin-up probability  $P_\uparrow$  from each of the individual sequences of Clifford gates. Solid symbols show the mean and the standard error of mean for each sequence. The lines are fits of the  $|\downarrow\rangle$  data to equation (2), with  $P_\infty$  either as a free (dashed line) or fixed (solid line) fitting parameter. See text for details. (b) Main panel: same data as in (a), but combining the data sets for both target output states to obtain the overall probability  $\mathcal{P}$  to recover the correct output state. The solid line is a fit to equation (4) with  $P_\infty = 0.5$  as a fixed parameter. Inset: similar data obtained with sinc pulses at 3.16 mW source power. The gate fidelity error margins are calculated from the 95% confidence interval for parameter  $p_c$  from the non-linear least-squares fit of  $\mathcal{P}(N)$ . Bootstrapping the residuals of the fit gave similar results for the confidence intervals. All fits have been weighted with the inverse of the unbiased sample variance at each  $N$ .

depends on the highly scattered data points at large  $N$ . This affects the overall confidence interval for  $p_c$  and, therefore,  $\mathcal{F}_c = 99.92(16)\%$ . However, having measured also sequences whose target output state is  $|\uparrow\rangle$ , we can extract a better estimate for  $P_\infty$  since to a good approximation it is simply given by  $(P_\uparrow^{|\uparrow\rangle}(N=1) + P_\uparrow^{|\downarrow\rangle}(N=1))/2$ . If we re-fit the data using a fixed value for  $P_\infty$  as obtained above, we find a nearly identical value of  $\mathcal{F}_c = 99.90(2)\%$ , but now with a much improved confidence interval. This confidence interval, however, does not account for statistical errors in determining  $P_\infty$ .



**Figure 3.** (a, b) Measured (symbols) and simulated (lines) spin-up proportion after applying a  $\pi$ -pulse to a spin-down  $e^-$  state as a function of pulse detuning from the resonance frequency. Triangles and dashed lines refer to sinc pulses; squares and dotted lines refer to square pulses. In (a), the microwave source power is adjusted to make the  $\pi$ -pulse length 4.5  $\mu$ s for both pulses, whereas in (b) the peak source power is kept identical for both pulses, with  $\pi$  length of 2.08  $\mu$ s (11.06  $\mu$ s) for a square (sinc) pulse. (c) Average Clifford gate fidelities measured as a function of source power. Labels indicate the  $\pi$ -pulse length in  $\mu$ s and triangles (squares) correspond to sinc (square) pulses. All fits have been done similarly as in figure 2(b).

We can go one step further and fit the combined data sets  $P_{\uparrow}^{|\uparrow\rangle}$  and  $1 - P_{\uparrow}^{|\downarrow\rangle}$  (figure 2(b)) with the function

$$\mathcal{P}(N) = \mathcal{P}_0 p_c^N + \mathcal{P}_{\infty}, \quad (4)$$

where  $\mathcal{P}$  represents the probability of retrieving the correct output state and the value of  $\mathcal{P}_{\infty} = [P_{\infty} + (1 - P_{\infty})]/2 = 0.5$  is fixed by construction. This fit yields an average Clifford gate fidelity  $\mathcal{F}_c = 99.90(2)\%$  for the electron spin, using square pulses. In the remainder of this paper, we will use combined  $P_{\uparrow}^{|\uparrow\rangle}$  and  $1 - P_{\uparrow}^{|\downarrow\rangle}$  data sets to extract  $\mathcal{F}_c$ . The ability to fix  $\mathcal{P}_{\infty} = 0.5$  greatly reduces the uncertainty in the fit to the data, but importantly does not cause measurable deviation of the value of  $\mathcal{F}_c$  from that extracted using only the sequences where the target state coincides with the initial state. We confirmed this conclusion on multiple datasets.

#### 4. Dependence of $e^-$ gate fidelities on pulse power and shape

Gate errors in spin qubit systems can arise from two broad classes of physical phenomena: uncertainty in the instantaneous resonance frequency of the qubit (noise in  $\sigma_z$ ) or

imperfections in the control field power and/or duration (noise in  $\sigma_x$ ). The errors arising from noise in  $\sigma_z$  can be reduced to some extent by minimizing the duration of the control pulses (provided enough microwave power is available). This maximizes the excitation spectrum in the frequency domain and ensures that the qubit frequency is well within the pulse spectrum. In practice, however, very short pulses can become imperfect due to the finite time resolution and bandwidth of the hardware used to gate the microwave power. This can then translate into an inaccurate calibration of the pulse length, i.e. noise in  $\sigma_x$ . One way to mitigate this problem is to use shaped pulses.

We have applied the randomized benchmarking protocol for the electron spin qubit with two different pulse shapes: (i) simple square pulses and (ii) sinc-function ( $\sin(x)/x$ ) shaped pulses. We chose a sinc function truncated at  $\pm 3\pi$  (sinc-3). For a similar length  $\pi$ -pulse, the sinc pulse gives an excitation spectrum that is roughly 5 times wider than that of a square pulse (figure 3(a)), but correspondingly requires a higher peak microwave power. At the same peak microwave power, the sinc pulse produces an excitation profile that is similar to the square pulse (figure 3(b)), but has a 5 times longer total duration. This can help relaxing the requirements on the time resolution of the pulse generator.

We also investigated the gate fidelities at different microwave powers, as explained in detail below. For the optimal power, the results are shown in the main panel of figure 2(b) for the square pulse and in its inset for the sinc pulse. From this data we extract an average Clifford gate fidelity of 99.90(2)% (99.87(3)% for sinc). Since a Clifford gate consists on average of 1.875 single ( $\pi$  or  $\pi/2$ ) gates (see table 1), these values can be converted to average single gate fidelities as  $1 - (1 - \mathcal{F}_c)/1.875$ . This produces 99.95(1)% for the square pulse and 99.93(1)% for the sinc pulse.

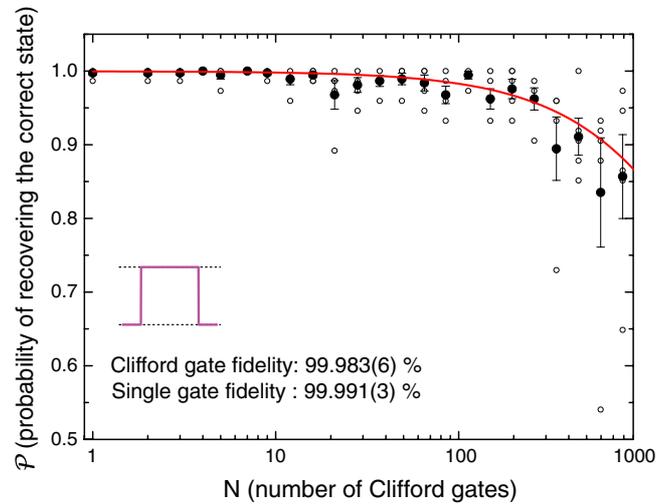
In figure 3(c) we present results from the randomized benchmarking protocol as a function of the source power for the two different pulse shapes. We have also indicated the time it takes to perform a  $\pi$ -rotation with the corresponding power and pulse shape. From this data we can see that (i) at low powers the square pulse outperforms the sinc and (ii) at higher powers the square pulse fidelity is markedly reduced. These effects are probably both related to the pulse length: at low powers the sinc pulse duration becomes exceedingly long, so that over the long sequence of Clifford gates some low-frequency noise begins to affect the outcome. At high powers, the square pulses become short compared to the time resolution (20 ns) of the baseband generator used to drive the digital in-phase / in-quadrature (I/Q) modulator in our vector microwave source. Below 2  $\mu$ s pulse length, the 20 ns resolution of the hardware already corresponds to a 1% precision, which starts to be the limiting factor. Conversely, we do not see a clear improving trend while increasing the microwave power. This seems to indicate that the gate fidelities are not limited by  $\sigma_z$  noise. This is consistent with the very narrow resonance linewidth ( $\approx 1.8$  kHz) measured in this sample [26], which is indeed always much smaller than the spectral width of the control pulses.

### 5. Interleaved single-gate fidelities

In addition to the average fidelity of Clifford gates, a similar sequence has been proposed [36] and applied [9, 34] to extract the average fidelity of a specific individual physical gate operation. This procedure is called interleaved randomized benchmarking and is based on interleaving the gate under study within a random Clifford sequence [36] (see figure 1(c)). The random sequence of Clifford gates has the effect of randomizing the input states for the interleaved gate. Since an extra gate will normally introduce additional errors, the average fidelity of the interleaved gate can be extracted by comparing the interleaved benchmarking decay to the reference Clifford decay (figure 2). The fitting function is the same as for the reference decay (equation (4), replacing  $p_c$  with  $p_{\text{gate}}$ ) and the interleaved gate fidelity is then extracted as

$$\mathcal{F}_{\text{gate}} = (1 + p_{\text{gate}}/p_c)/2. \quad (5)$$

The results of the interleaved randomized benchmarking for the electron spin qubit are shown in table 2. We note that the average of the individual gate fidelities is close, although slightly lower than the average single gate fidelity deduced from the reference decay. Also notable is the similarity of



**Figure 4.** Randomized benchmarking for the nuclear spin qubit with gate operations performed using square pulses. The duration of a  $\pi$ -pulse is 150  $\mu$ s. The open circles show results from individual measurements and solid circles the mean of the  $K = 5$  different random sequences. Error bars, fit and the error margins as in figure 2(b).

the gate fidelities obtained with square and sinc pulse shapes, even though they are measured at different source powers and different pulse lengths. The  $\pi$ -pulses appear to give higher fidelities than the  $\pi/2$ -pulses. This could be due to the finite time resolution of the pulse generator, which has more effect on the shorter  $\pi/2$ -pulses, or it could indicate some refocussing effect of the  $\pi$ -pulses, which compensate for some slight qubit dephasing during the long sequences of gates.

### 6. Nuclear spin gate fidelities

As evidenced in earlier experiments [25, 26, 30], the nuclear spin of the  $^{31}\text{P}$  donor constitutes an excellent qubit, particularly when the donor is in the ionized state. We have performed randomized benchmarking of the nuclear spin qubit, focusing on the  $^{31}\text{P}^+$  ionized state. After the gate operations, a high-fidelity, quantum non-demolition readout of the  $^{31}\text{P}$  nuclear state [30] can be accomplished by loading a spin-down electron and applying a  $\pi$ -pulse on the electron spin conditional on the nuclear spin state. The electron spin is then read out [28]. In the present experiments, the electron readout is repeated 50 times to obtain the nuclear state with high fidelity. For the randomized benchmarking we use  $r = 75$  repetition of each sequence of  $N$  Clifford gates. At every  $N$  we use  $K = 5$  different, randomly chosen sequences. This low value of  $K$  is sub-optimal from the point of view of converging to the average Clifford gate fidelity, but was necessary to avoid having exceedingly long experiments. These measurements already took 7 h with  $K = 5$ , partly due to the time needed to load the random sequences of up to  $N = 1000$  Clifford gates onto the baseband generator. Here we used square pulses, with 150  $\mu$ s duration for a  $\pi$ -pulse. This means that, even for the highest  $N$ , the total duration of the sequence remains well below the dephasing time of the qubit,  $T_{2n0}^* = 600$  ms [26].

**Table 2.** Results of the interleaved gate benchmarking for the electron spin qubit. Notation is the same as in table 1. The source power used in these experiments was 0.63 mW for square pulses and 3.16 mW with sinc pulses. The error margins are calculated by summing squarely the relative 95% confidence limits for parameters  $p_{\text{gate}}$  and  $p_c$ . All fits have been done similarly as in figure 2(b).

Electron spin		
Interleaved Gate	Square pulse (%)	Sinc pulse (%)
X	99.93(4)	99.97(4)
Y	100.00(5)	99.95(4)
X/2	99.93(5)	99.87(6)
Y/2	99.85(8)	99.81(6)
−X/2	99.94(4)	99.92(5)
−Y/2	99.90(5)	99.81(6)
Average	99.93	99.89

**Table 3.** Results of the interleaved gate benchmarkings for the nuclear spin qubit. All notation identical to table 2.

Nuclear spin	
Interleaved gate	Square pulse (%)
X	99.94(4)
Y	99.98(1)
X/2	99.97(2)
Y/2	99.99(1)
−X/2	99.88(2)
−Y/2	99.93(5)
Average	99.95

The average Clifford gate fidelity and the individual interleaved gate fidelities are shown in figure 4 and table 3, respectively. We found a  $^{31}\text{P}^+$  gate fidelity of order 99.99%, matching even some of the stricter error correction thresholds. The individual interleaved gate fidelities are somewhat worse, but still of order 99.95%.

## 7. Conclusions

We have quantified the quantum gate fidelities of  $^{31}\text{P}$  electron and nuclear spin qubits in isotopically enriched  $^{28}\text{Si}$  and showed that they exceed some fault-tolerance thresholds [7, 8]. An analysis of  $e^-$  gate fidelities as a function of microwave power suggests that further improvements could be obtained with higher-resolution baseband generators and other engineering solutions. The behaviour of the gate errors as a function of pulse power and shape suggests that the 1-qubit gate fidelities are not affected by device-intrinsic phenomena such as fluctuations of the qubit energy splitting. A complete assessment of the viability of donor-based spin qubits for large-scale quantum computing in the solid state will require the realization and the benchmarking of 2-qubit logic gates. A recent proposal predicts fault-tolerant 2-qubit gate fidelities with logic operations based on weak exchange interactions [37] and ESR pulses, whereas very strong exchange has been measured experimentally [38] in a donor pair. While awaiting the experimental demonstration of a universal set of 1- and 2-qubit logic gates, the present results already show great promise for the realization of donor-spin-based quantum computers in silicon.

## Acknowledgments

We thank S Bartlett and S Flammia for illuminating discussions. This research was funded by the Australian Research Council Centre of Excellence for Quantum Computation and Communication Technology (project number CE11E0001027) and the US Army Research Office (W911NF-13-1-0024). We acknowledge support from the Australian National Fabrication Facility and from the laboratory of Prof Robert Elliman at the Australian National University for the ion implantation facilities. The work at Keio has been supported in part by FIRST, the Core-to-Core Program by JSPS and the Grant-in-Aid for Scientific Research and Project for Developing Innovation Systems by MEXT.

## References

- [1] Shor P W 1995 Scheme for reducing decoherence in quantum computer memory *Phys. Rev. A* **52** R2493
- [2] Steane A 1996 Multiple-particle interference and quantum error correction *Proc. R. Soc. Lond.* **452** 2551–77
- [3] Nielsen M and Chuang I 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
- [4] Raussendorf R 2012 Key ideas in quantum error correction *Phil. Trans. R. Soc.* **370** 4541–65
- [5] Knill E, Laflamme R and Zurek W H 1998 Resilient quantum computation *Science* **279** 342–5
- [6] Stephens A M, Fowler A G and Hollenberg L C L 2008 Universal fault tolerant quantum computation on bilinear nearest neighbor arrays *Quantum Inform. Comput.* **8** 330–44
- [7] Knill E 2005 Quantum computing with realistically noisy devices *Nature* **434** 39–44
- [8] Fowler A G, Mariantoni M, Martinis J M and Cleland A N 2012 Surface codes: towards practical large-scale quantum computation *Phys. Rev. A* **86** 032324
- [9] Barends R *et al* 2014 Superconducting quantum circuits at the surface code threshold for fault tolerance *Nature* **508** 500–3
- [10] Benhelm J, Kirchmair G, Roos C F and Blatt R 2008 Towards fault-tolerant quantum computing with trapped ions *Nat. Phys.* **4** 463–6
- [11] Chuang I L and Nielsen M A 1997 Prescription for experimental determination of the dynamics of a quantum black box *J. Mod. Opt.* **44** 2455–67
- [12] Chow J M, Gambetta J M, Tornberg L, Jens Koch, Bishop L S, Houck A A, Johnson B R, Frunzio L, Girvin S M and Schoelkopf R J 2009 Randomized benchmarking and process tomography for gate errors in a solid-state qubit *Phys. Rev. Lett.* **102** 090502
- [13] Emerson J, Alicki R and Życzkowski K 2005 Scalable noise estimation with random unitary operators *J. Opt. B: Quantum Semiclass. Opt.* **7** S347
- [14] Knill E, Leibfried D, Reichle R, Britton J, Blakestad R B, Jost J D, Langer C, Ozeri R, Seidelin S and Wineland D J 2008 Randomized benchmarking of quantum gates *Phys. Rev. A* **77** 012307
- [15] Olmschenk S, Chicireanu R, Nelson K D and Porto J V 2010 Randomized benchmarking of atomic qubits in an optical lattice *New J. Phys.* **12** 113007
- [16] Brown K R, Wilson A C, Colombe Y, Ospelkaus C, Meier A M, Knill E, Leibfried D and Wineland D J 2011 Single-qubit-gate error below  $10^{-4}$  in a trapped ion *Phys. Rev. A* **84** 030303
- [17] Harty T P, Allcock D T C, Ballance C J, Guidoni L, Janacek H A, Linke N M, Stacey D N and Lucas D M 2014

- High-fidelity preparation, gates, memory, and readout of a trapped-ion quantum bit *Phys. Rev. Lett.* **113** 220501
- [18] Ryan C A, Laforest M and Laflamme R 2009 Randomized benchmarking of single- and multi-qubit control in liquid-state nmr quantum information processing *New J. Phys.* **11** 013034
- [19] Magesan E, Gambetta J M and Emerson J 2011 Scalable and robust randomized benchmarking of quantum processes *Phys. Rev. Lett.* **106** 180504
- [20] Magesan E, Gambetta J M and Emerson J 2012 Characterizing quantum gates via randomized benchmarking *Phys. Rev. A* **85** 042311
- [21] Epstein J M, Cross A W, Magesan E and Gambetta J M 2014 Investigating the limits of randomized benchmarking protocols *Phys. Rev. A* **89** 062321
- [22] Gottesman D 1998 The heisenberg representation of quantum computers (arXiv:quant-ph/9807006)
- [23] Ager J W, Beeman J W, Hansen W L, Haller E E, Sharp I D, Liao C, Yang A, Thewalt M L W and Riemann H 2005 High-purity, isotopically enriched bulk silicon *J. Electrochem. Soc.* **152** G448–51
- [24] Tyryshkin A M *et al* 2012 Electron spin coherence exceeding seconds in high-purity silicon *Nat. Mater.* **11** 143–7
- [25] Saeedi K, Simmons S, Salvail J Z, Dluhy P, Riemann H, Abrosimov N V, Becker P, Pohl H-J, Morton J J L and Thewalt M L W 2013 Room-temperature quantum bit storage exceeding 39 min using ionized donors in silicon-28 *Science* **342** 830–3
- [26] Muhonen J T, Dehollain J P, Laucht A, Hudson F E, Sekiguchi T, Itoh K M, Jamieson D N, McCallum J C, Dzurak A S and Morello A 2014 Storing quantum information for 30 s in a nanoelectronic device *Nat. Nanotechnol.* **9** 986–91
- [27] Morello A, Escott C C, Huebl H, Willems van Beveren L H, Hollenberg L C L, Jamieson D N, Dzurak A S and Clark R G 2009 Architecture for high-sensitivity single-shot readout and control of the electron spin of individual donors in silicon *Phys. Rev. B* **80** 081307
- [28] Morello A *et al* 2010 Single-shot readout of an electron spin in silicon *Nature* **467** 687–91
- [29] Pla J J, Tan K Y, Dehollain J P, Lim W H, Morton J J L, Jamieson D N, Dzurak A S and Morello A 2012 A single-atom electron spin qubit in silicon *Nature* **489** 541–5
- [30] Pla J J, Tan K Y, Dehollain J P, Lim W H, Morton J J L, Zwanenburg F A, Jamieson D N, Dzurak A S and Morello A 2013 High-fidelity readout and control of a nuclear spin qubit in silicon *Nature* **496** 334–8
- [31] Dehollain J P, Pla J J, Siew E, Tan K Y, Dzurak A S and Morello A 2013 Nanoscale broadband transmission lines for spin qubit control *Nanotechnology* **24** 015202
- [32] Jamieson D N *et al* 2005 Controlled shallow single-ion implantation in silicon using an active substrate for sub-20 keV ions *Appl. Phys. Lett.* **86** 202101
- [33] Elzerman J M, Hanson R, Willems Van Beveren L H, Witkamp B, Vandersypen L M K and Kouwenhoven L P 2004 Single-shot read-out of an individual electron spin in a quantum dot *Nature* **430** 431–5
- [34] Veldhorst M *et al* 2014 An addressable quantum dot qubit with fault-tolerant control fidelity *Nat. Nanotechnol.* **9** 981–5
- [35] Wallman J J and Flammia S T 2014 Randomized benchmarking with confidence *New J. Phys.* **16** 103032
- [36] Magesan E *et al* 2012 Efficient measurement of quantum gate error by interleaved randomized benchmarking *Phys. Rev. Lett.* **109** 080505
- [37] Kalra R, Laucht A, Hill C D and Morello A 2014 Robust two-qubit gates for donors in silicon controlled by hyperfine interactions *Phys. Rev. X* **4** 021044
- [38] Dehollain J P, Muhonen J T, Tan K Y, Saraiva A, Jamieson D N, Dzurak A S and Morello A 2014 Single-shot readout and relaxation of singlet and triplet states in exchange-coupled  $^{31}\text{P}$  electron spins in silicon *Phys. Rev. Lett.* **112** 236801