A CMOS dynamic random access architecture for radio-frequency readout of quantum devices

Simon Schaal,[1,*] Alessandro Rossi,[2] Virginia N. Ciriano-Tejel,[1] Tsung-Yeh Yang,[3] Sylvain Barraud,[4] John J. L. Morton,[1,5] and M. Fernando Gonzalez-Zalba[3,†]

[1] *London Centre for Nanotechnology, University College London, London WC1H 0AH, United Kingdom*
[2] *Cavendish Laboratory, University of Cambridge, J. J. Thomson Ave., Cambridge, CB3 0HE, United Kingdom*
[3] *Hitachi Cambridge Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*
[4] *CEA, LETI, Minatec Campus, F-38054 Grenoble, France*
[5] *Department of Electronic & Electrical Engineering, University College London, London WC1E 7JE, United Kingdom*

(Dated: May 16, 2019)

As quantum processors become more complex, they will require efficient interfaces to deliver signals for control and readout while keeping the number of inputs manageable. Complementary metal-oxide-semiconductor (CMOS) electronics offers established solutions to signal routing and dynamic access, and the use of a CMOS platform for the qubits themselves offers the attractive proposition of integrating classical and quantum devices on-chip. Here we report a CMOS dynamic random access (DRAM) architecture for readout of multiple quantum devices operating at millikelvin temperatures. Our circuit is divided into cells, each containing a control field-effect transistor and a quantum dot (QD) device, formed in the channel of a nanowire transistor. This setup allows selective readout of the QD and charge storage on the QD gate, similar to one-transistor-one-capacitor (1T-1C) DRAM technology. We demonstrate dynamic readout of two cells by interfacing them with a single radio-frequency resonator. Our approach provides a path to reduce the number of input lines per qubit and allow large-scale device arrays to be addressed.

Quantum computers could be used to solve problems that seem intractable with conventional computers [1]. Several different physical implementations of a quantum computer are being developed [2] and state-of-the-art processors are approaching the level of 50 to 100 quantum bits (qubits), a point at which quantum computers are expected to demonstrate capabilities beyond conventional computers for specific tasks [3].

For most physical realisations, quantum processors require cryogenic temperatures to operate, precise low-noise control signals [4] to manipulate the information, and highly sensitive readout techniques to extract the results – all without disturbing the fragile quantum states. In current solid-state quantum processors, signals are generated using general-purpose instruments at room temperature and delivered to the quantum processor at low temperatures. The physical qubits across all platforms are controlled directly with at least one control line per qubit. However, as the size of quantum processors continues to increase, the one-qubit-one-input approach will be unsustainable [5], especially if we consider that a large-scale fault-tolerant quantum computer might ultimately require 10^8 qubits to solve computationally demanding algorithms [6]. Efficiently delivering control and readout signals to increasingly more complex quantum circuits, while reducing the number of inputs per qubit, is a key challenge in developing a large-scale universal quantum computer. Integrated electronics provide a solution to these problems. Some of the challenges that face large-scale quantum computing resemble those that have already been solved for conventional computing. For example, controlling billions of transistors with just a few thousands of input-output connections. Moreover, integrated electronics allows signal generation, data flow management, low-level feedback and high-level operations locally. Therefore, to relax wiring requirements and reduce the latency of solid-state quantum computers, the integration of conventional electronics with quantum devices at cryogenic temperatures could be a promising strategy [7, 8]. However, to apply this approach, understanding the behaviour of integrated circuits at cryogenic temperatures is vital [9].

Digital information processing devices are typically manufactured using silicon as the base material. Coincidentally, electron spins in silicon are amongst the most promising candidates for large-scale quantum computing due to their small footprint (sub 100 nm dimensions) and very long coherence times, particularly in isotopically purified ^{28}Si [10, 11]. Silicon-based spin qubits benefit from a variety of qubit designs and different coupling strategies [12–19] and can be read out dispersively using Pauli spin-blockade [20–22]. To date, operation of one-dimensional arrays has been demonstrated [23], high-fidelity single qubit gates [10, 24–26] and two qubit gates [12, 19] have been achieved and a programmable two-qubit silicon-based processor has been created [18].

Recently, it was shown that CMOS transistors can be used as the basis for spin qubits [27, 28]. Several other silicon-based quantum devices could, in principle, be realised in a manner compatible with industrial CMOS pro-

* simon.schaal.15@ucl.ac.uk
† mg507@cam.ac.uk

cesses, with the potential of large-scale, high-yield fabrication. It seems natural, then, to explore the direct integration of silicon quantum devices and conventional CMOS technology to tackle the challenges in addressing, controlling and reading multi-qubit circuits. Blueprints of such all-silicon systems integrating quantum and classical components have emerged [29–31] and basic demonstrations of direct integration have been reported [32].

In this Article, we report a CMOS dynamic random access architecture for readout of multiple quantum devices. Our design is inspired by the square arrays found in one-transistor-one-capacitor (1T-1C) dynamic random access memory (DRAM) and allows on-demand routing of static and radio-frequency (RF) signals to individual devices. The architecture is composed of individual cells each containing a control field-effect transistor (FET) and a quantum dot (QD) device. In our experiments, the QDs are themselves formed in the channel of a nanowire FET, integrated on the same chip as the control FETs and fabricated using the same CMOS processes. When not addressed, each cell can be used as a node to store charge on the QD device gate that allows trapping single-electrons in the QD device with a time constant approaching 1 s. We demonstrate random access and readout of two individual cells at cryogenic temperatures using capacitive gate-based RF reflectometry [33–35]. We obtain a readout bandwidth of 13 MHz measured from the frequency overlap of two individually addressed cells, and find optimal operation voltage levels for the control transistor. Moreover, we show dynamic readout of the cells and obtain charge stability maps sequentially. Finally, we provide guidelines for scaling the approach by developing an equivalent DC and RF circuit model of the cell, and a 2D architecture with a quadratic reduction in the number of inputs.

CIRCUIT CHARACTERISATION

We show the sequential access circuit in Fig. 1a. It consists of two CMOS single-electron memory cells [32] (cell 1(2) in green(orange)) connected to a lumped-element RF resonator for readout and a single bias line. Each memory cell is made from two transistors which we refer to as Q_i and T_i. Q_i is a 60-nm-wide silicon nanowire transistor with a short gate length (25 and 30 nm for cell 1 and 2 respectively). Such devices are routinely used to trap single-electrons in QDs that form at the top most corners of the nanowire channel when operated in the sub-threshold regime at cryogenic temperatures [36]. Transistor T_i is a wider device with a channel width of 10 μm and gate length of 25 nm and 30 nm for cell 1 and 2 respectively which we refer to as the control FET. The four transistors are manufactured using fully-depleted silicon-on-insulator (FD-SOI) technology following standard CMOS processes. They are located on the same chip and are connected via bond wires (see Methods for details of the fabrication and Fig. 1a for a schematic).

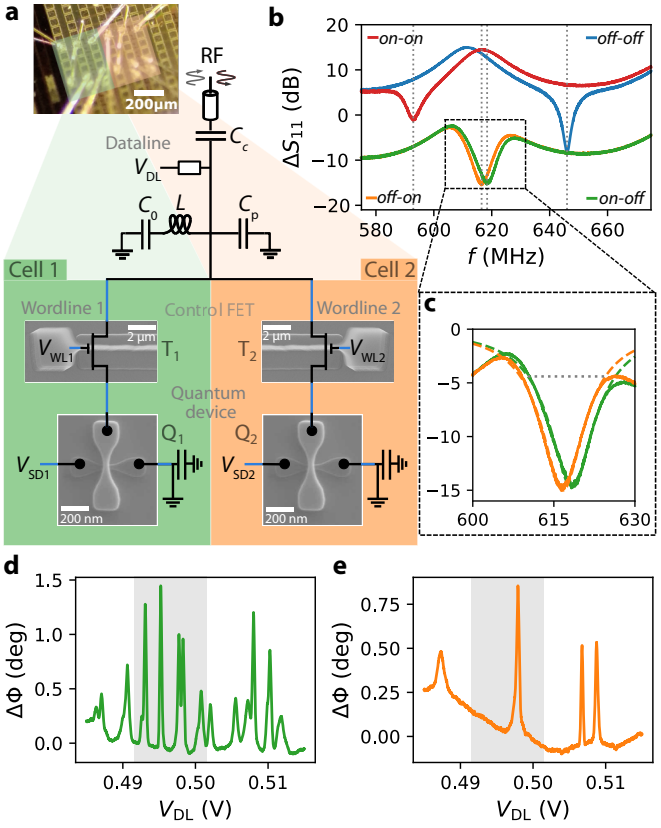

FIG. 1. **Setup and individual device characterisation.** **a**, Sequential access circuit for gate-based RF readout with optical microscope image of parts of the circuit as inset. A single high frequency line and readout resonator is connected to two cells (green and orange) consisting of one control FET, T_i, and quantum device, Q_i, per cell. T_i enables selective readout of Q_i. Electrical connections made via bond wires are represented by blue lines. **b**, Reflection coefficient spectrum of the circuit for different control FET states (T_1 - T_2). Spectra for addressing a single cell have been shifted down by 15 dB for clarity. **c**, Enlarged view of *on-off* and *off-on* state configurations with a spectral overlap and resonance fits indicated (by dashed lines). **d**, Phase response of Q_1 as a function of V_{DL} for $V_{WL1} = 1.2$ V and $V_{WL2} = 0$ V. **e**, Phase response of Q_2 as a function of V_{DL} for $V_{WL1} = 0$ V and $V_{WL2} = 1.2$ V. The regions in grey highlight charge transitions we focus on in further measurements.

We label the primary inputs of the circuit as data and word lines in analogy with memory chips. Each cell has one word line, with voltage V_{WLi}, which connects to the gate of the control FET T_i allowing control over the channel resistance. The data line, with voltage V_{DL}, is shared among the two cells and allows control over the gate voltage on Q_i conditional on the state of T_i. Additionally, a voltage applied to the silicon substrate, V_{BG}, acts as a back-gate. Switching T_i to the *on* state while keeping all the remaining T_j *off* allows for individual addressing of a single quantum device Q_i. Multiple devices can

be addressed sequentially by timing the voltages on T_i accordingly, as we demonstrate further below.

To read the quantum state of the devices, we connect a lumped-element LC resonator in parallel with the memory cells and use RF reflectometry to probe the resonant state of the combined circuit [32]. We couple the RF signal into the data line via the coupling capacitor C_c. The natural frequency of the resonator f_0 is given by $f_0 = 1/2\pi\sqrt{LC_\mathrm{T}}$ where C_T is the total capacitance of the system that includes, in particular, the state-dependent quantum or tunnelling capacitance of any quantum device [37] which is connected to the LC circuit via the control FETs. The whole circuit is operated in a dilution refrigerator with a base temperature of 15 mK.

Next, we show the frequency dependence of the circuit's reflection coefficient S_{11} in Fig. 1b for the four possible states of the two control FETs. A dip in the reflection coefficient occurs when we drive the resonator at its natural frequency of oscillation. This frequency shifts towards lower values (by approximately 28 MHz) for each T_i in the *on* state due to the additional circuit capacitance introduced by the enabled cell. In Supplementary Table 1 we report a comprehensive list of circuit parameters describing the resonance conditions depending on the logic state of each cell. Most importantly, we observe a large spectral overlap of 13 MHz with 3 dB readout bandwidth in the enlarged view in Fig. 1c when addressing one cell at a time. Spectral overlap is vital to dynamical multi-qubit readout as it means that both cells can be read using the same input frequency, while the degree of overlap determines the readout bandwidth of the architecture. In addition to the resonance frequency shift, we observe a reduction in the loaded quality factor Q_L from a value of 96, when both T_1 or T_2 are in the *off* state, to a value of 40, when either T_1 or T_2 are in the *on* state. An *on* state Q_L of 40 is comparable to previous experiments with [32] and without [34] control circuit.

Based on the spectra shown in Fig. 1c, we select a carrier frequency $f_c = 615$ MHz to probe the state of the quantum devices. When using RF reflectometry, changes in the complex impedance of the circuit are probed by driving the circuit close to resonance (using a small signal of -90 dBm) while monitoring the phase and magnitude of the reflected signal (see Methods for details of the circuit). Changes in the capacitance of the quantum device ΔC_G, attributed to tunneling of single electrons, are detected through changes in the reflected phase $\Delta\phi = -2Q_L\Delta C_\mathrm{G}/C_\mathrm{T}$ [35]. In Fig. 1d-e, we observe phase shift peaks as we change V_DL that corresponds to regions of charge instability in Q_i. At these voltages, single electrons cyclically tunnel between the QDs in the channel and the source or drain electron reservoirs in Q_i. For each measurement only one T_i is set to the *on* state while the other is *off*. Next, we discuss measurements focusing on a particular region of this stability diagram (highlighted in grey in Fig. 1d-e) for both quantum devices with the aim to find optimal operation voltage levels for the control transistors.

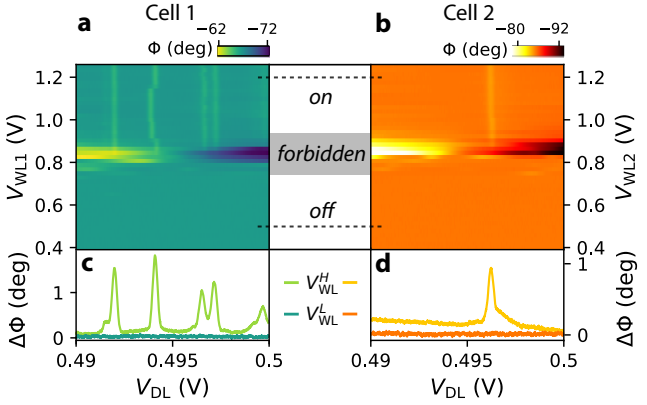

FIG. 2. **Control transistor logic states. a**, Phase response of Q_1 as a function of V_{WL1} and V_DL ($V_{WL2} = 0$ V). **b**, Phase response of Q_2 as a function of V_{WL2} and V_DL ($V_{WL1} = 0$ V). For both cells we observe QD-to-reservoir transitions at large V_WL corresponding to the logical *on* state of the digital transistor. A *forbidden* region of large background signal is found upon approaching the control FET threshold voltage. A region of no signal below threshold corresponds to the *off* state. **c-d**, Line cuts at $V_\mathrm{WL}^\mathrm{L} = 0.5$ V and $V_\mathrm{WL}^\mathrm{H} = 1.2$ V, indicated by dashed lines in (a-b), that highlight the difference between the two digital states for each device.

For a dynamical random-access readout scheme, T_i should fulfil several requirements: In the *on* state, T_i should be sufficiently conductive to allow high-sensitivity gate-based readout of the selected quantum device. In the *off* state, T_i should be sufficiently resistive to block the RF signal towards deselected cells and retain the charge on Q_i's gate for the time operations are being performed in other cells.

As a first step towards dynamically operating the circuit, we identify suitable *on* and *off* state voltages for the control FET gate (i.e. the *high*, $V_{\mathrm{WL}i}^\mathrm{H}$, and *low*, $V_{\mathrm{WL}i}^\mathrm{L}$, signal levels). In Fig. 2a-b, we show the phase of the reflected signal from the resonator as a function of V_DL and $V_{\mathrm{WL}i}$. We can identify three regions: The *on* region for $V_{\mathrm{WL}i} > 0.9$ V, where we observe single electron tunnelling, the *off* region for $V_{\mathrm{WL}i} < 0.7$ V, where we observe no transitions and finally, for $0.7\ \mathrm{V} < V_{\mathrm{WL}i} < 0.9$ V the *forbidden* region. In the latter, T_i is in the depletion regime, where, due to the voltage-dependent gate capacitance of the control FET, the phase varies largely [38]. This region should be avoided when assigning voltage levels. To highlight the different response of the resonator in the digital *on* and *off* states, we show the phase change $\Delta\phi$ as a function of V_DL for cell 1 and 2 in Fig. 2c-d, respectively, at $V_{\mathrm{WL}i}^{\mathrm{L(H)}} = 0.5(1.2)$ V.

We note the close similarity between the operation voltage levels of both T_i for addressing the quantum devices Q_i at millikelvin temperature. In a scaled up architecture, with increasing circuit complexity, reproducible electrical characteristics between cells will be essential.

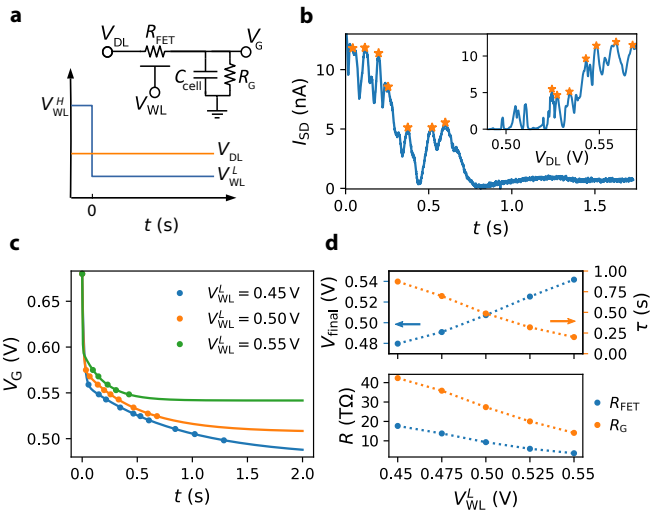

FIG. 3. **Charge retention analysis.** **a**, Equivalent circuit of a single memory cell and pulsing scheme for charge retention analysis. **b**, Source-drain current I_{SD} through the quantum device as a function of time after switching the control transistor to the *off* state $V^L_{WL} = 0.5$ V ($V_{SD} = 2$ mV, $V_{DL} = 0.68$ V). Single electron transitions are observed and peak positions are indicated by stars. The inset shows I_{SD} as a function of V_{DL} where the same transitions are observed and indicated by stars. **c**, Decay of the voltage on the QD gate V_G as a function of time once the control transistor is switched off. Data-points (circles) are obtained from the peak positions (stars) in I_{SD} and solid lines are fits to a double exponential function as described in the text. **d**, Top panel: Quasi-static gate voltage V_{final} and time constant τ as a function of V^L_{WL} obtained from the exponential decay fits (see Eq. 1). Bottom panel: R_{FET} and R_G extracted from τ and V_{final} using Eq. 1. Dashed lines are guides to the eye.

DYNAMIC OPERATION

Random access of a single cell can be achieved by switching the selected T_i *on* while all other T_j are *off*. Since the data line voltage V_{DL} is shared among the cells, the gate voltage on all deselected Q_j floats and decays over time while addressing cell i. Floating gate charge storage is an important feature of dynamic readout and its associated charge retention time sets the maximum time to perform operations on other cells before the information is lost. Charge locking is an established mechanism that is routinely used in dynamic RAM chips and it has recently been used to multiplex the access to GaAs quantum dots [39, 40]. Here, we combine charge locking with gate-based RF readout.

First, we characterise the discharge of one cell in order to determine an appropriate voltage refresh rate. We consider a simplified equivalent circuit model of the memory cell as shown in Fig. 3a. It consist of the FET *off* state channel resistance R_{FET}, the gate leakage resistance R_G, and the cell capacitance C_{cell}. R_G combines the FET and QD gate leakages and C_{cell} is the parallel sum of the QD gate capacitance C_G and the interconnect capacitance C_S, with the latter being dominant in this experiment. The voltage on the QD gate V_G decays over time as

$$V_G(t) = V_{final}\left[1 + \frac{R_{FET}}{R_G}\exp\left(-\frac{t}{\tau}\right)\right] \qquad (1)$$

when the FET is switched to the *off* state. Here $\tau = \frac{C_{cell}R_G R_{FET}}{R_G + R_{FET}}$ is the circuit time constant and $V_{final} = \frac{V_{DL}R_G}{(R_{FET}+R_G)}$ is the equilibrium voltage at the gate of the QD at $t \to \infty$. Since τ and V_{final} depend on R_{FET}, and thus on the operation voltage level V_{WL}, we proceed by investigating their functional dependence to find the optimal voltage operation point that maximises the charge retention time.

We monitor the discharge of the cell in a pulsed experiment by measuring the source-drain current I_{SD} through the QD over time. As shown in Fig. 3a, we keep $V_{DL} = 0.68$ V constant while V_{WL} switches from the *high* level (V^H_{WL}) to a *low* level (V^L_{WL}) at $t = 0$. We set the pulse amplitude to 0.5 V and vary the pulse offset ensuring that the transistor remains *on* in the *high* part of the pulse. We show an exemplary discharge measurement for $V^L_{WL} = 0.5$ V in Fig. 3b, where several single electron transitions (indicated by stars) can be observed in I_{SD} over time. After 1.5 s the current settles to a value determined by V_{final}. We compare the discharge data with a measurement of the same single electron transitions of the device as a function of V_{DL} in quasi-static conditions as shown in the inset of Fig. 3b. By matching peaks in the decay over time to peaks as a function of V_{DL} we reproduce the dynamics of the voltage on the QD gate $V_G(t)$ as shown in Fig. 3c for multiple values of V^L_{WL}. At $t = 0$, we observe an initial fast decay, possibly due to charge-injection and clock-feedthrough [41], followed by a slow decay characterised by Eq. 1. We fit a double exponential to capture the fast and slow dynamics (see Supplementary Equation 1) and extract V_{final} and τ from the slow decay which we show in Fig. 3d as a function of V^L_{WL}. We observe that as V^L_{WL} increases (R_{FET} decreases) V_{final} becomes larger due to the voltage divider characteristic of the cell. In the case of τ, we observe a reduction from 0.9 s to 0.2 s. We note that the time constant could be increased by increasing C_{cell}. The resistance values R_{FET} and R_G extracted from these measurements based on a cell capacitance $C_{cell} = 70$ fF (see Supplementary Table 1) are on the order of $10^{13}\,\Omega$. We can see that R_G increases as V_{WL} decreases which indicates that there is a V_{WL} dependent contribution towards R_G. To summarise, we find that the discharge model fits the data and shows a decrease(increase) in $\tau(V_{final})$ as V_{WL} is increased from 0.45 V to 0.55 V, as expected. Moreover, R_G and R_{FET} decrease by a factor of 3 and 5 respectively as V_{WL} increases.

While initially it may seem beneficial to select a low V^L_{WL} level to maximise τ, one needs to consider that the retention or refresh time is determined by the maximum tolerable gate voltage drop of the cell, δV, which has to

be assessed given a specific qubit implementation. For an optimised circuit with reduced crosstalk and defining the voltage drop ratio $a = \delta V/V_{\rm DL}$ and the resistance ratio $r = (R_{\rm FET} + R_{\rm G})/R_{\rm FET}$, we find that the retention time is given by $t_{\rm r} = R_{\rm G}C_{\rm cell}\ln[(1-ar)^{-1}]/r$ which is a monotonically increasing function of r, given $R_{\rm G}$ varies weakly with r. Then, $t_{\rm r}$ is maximised by operating at large $V_{\rm WL}^{\rm L}$ while remaining in the *off* regime where RF readout of the selected cell is not disturbed. Using circuit simulations we find that as long as $R_{\rm FET} > 10\,{\rm M}\Omega$, the effect of a deselected cell on the readout of a selected cell becomes negligible which is compatible with operating closely below the forbidden region shown in Fig. 2 (see Supplementary Figure 3). Additionally, we note that at $R_{\rm FET} > 10\,{\rm M}\Omega$ the fraction of the RF signal delivered to the deselected cell is $< 1\%$ due to the low-pass filter formed by $R_{\rm FET}$ and $C_{\rm cell}$ (assuming $C_{\rm cell} = 70\,{\rm fF}$ and operation at a few hundreds of MHz). Finally, given a voltage drop ratio of 1%, we estimate a retention time of 20 ms using the parameters extracted from the experiment which is much larger than the coherence time T_2^* of silicon-based electron spin qubits [10].

In this analysis, it is important to note that we keep $V_{\rm DL}$ constant which is approximately what will happen when addressing multiple quantum devices with similar operating voltages. Such operation is a particular feature of our proposal and differs from the 1T-1C DRAM read protocol where $V_{\rm DL}$ is typically set to half the maximum voltage stored in the capacitor. Such voltage level maximises the readout signal and the retention time of both the uncharged and charged memory state of the capacitor [42]. In our proposal, we operate exclusively at the charged state of Q_i. For sequential readout, as demonstrated further below, we select $V_{\rm WL}^{\rm L} = 0.5$ V ($R_{\rm FET} \approx 10^{13}\,\Omega$) to enhance the retention time in the *off* state while preserving good noise margins.

We now turn to demonstrate sequential dynamic readout of quantum devices in two memory cells. We show the pulsing scheme to dynamically read both memory cells in Fig. 4a. In the first half of the cycle, from 0 to 12 ms, we set T_1 and T_2 to the digital *on* and *off* states respectively. Simultaneously, we apply an analogue signal to the common data line $V_{\rm DL}$ (blue trace) that ramps up the gate voltage on the data line (now connected to Q_1). We read the signal dispersively using gate-based readout and detect peaks in the phase due to single-electron transitions between a QD and a reservoir in Q_1. In the second half of the cycle, from 12 to 24 ms, we invert the digital voltages on T_1 and T_2 such that we can now detect the transitions in Q_2 as we ramp down the analogue signal on the data line. The QD-to-reservoir transitions in the phase response are identical to those measured in a static experiment shown in Fig.1d-e. The RF modulation frequency and amplitude is kept constant throughout the measurement. There is a phase offset between the signal detected from Q_1 and Q_2 due to a small difference in reflection coefficient between cells (see Fig. 1c). We therefore show the change in phase $\Delta\phi$ in Fig. 4a (see also Supplementary Figure 2). We obtain a signal-to-noise ratio (SNR) of 10^5 with 100 ms integration time.

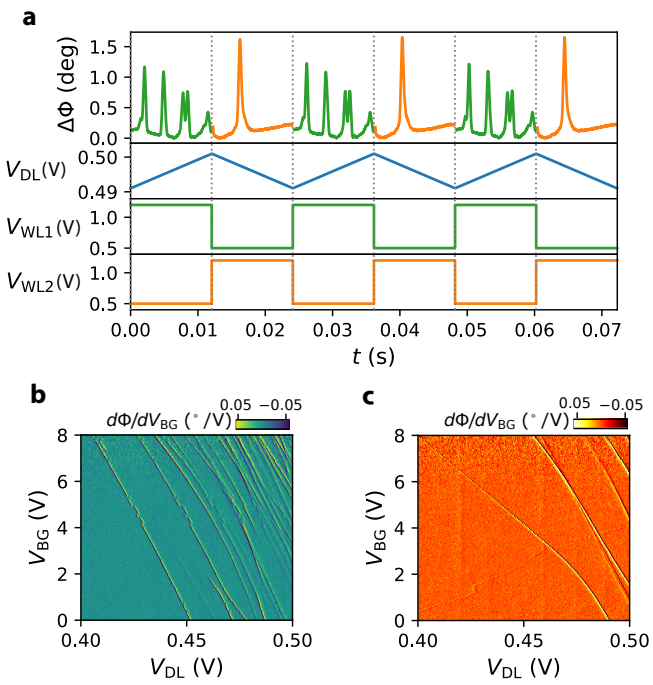

FIG. 4. **Dynamic readout. a**, Pulse scheme for sequential readout and phase response of Q_1 and Q_2. $V_{\rm DL}$ is ramped up and down while $V_{\rm WL1}$ and $V_{\rm WL2}$ are alternating between *high* and *low* states. Pulses are synchronised such that QD reservoir transitions from Q_1 are obtained when $V_{\rm DL}$ is ramped up while Q_2 is measured when $V_{\rm DL}$ is ramped down. **b-c**, Differential phase response obtained sequentially from both cells as a function of data line and back-gate voltages.

Using this interleaved pulsing scheme for sweeping $V_{\rm DL}$ combined with additional stepping of $V_{\rm BG}$ after each cycle, we obtain the charge stability map of both Q_1 and Q_2 sequentially as shown in Fig. 4b-c. The transitions observed in the measurement suggest formation of multiple QDs in both cells (see Supplementary Figure 1 for additional scans), i.e. corner dots [36]. We estimate a maximum power dissipation of $P = C_{\rm FET}f_{\rm op}\Delta V^2 = 25\,{\rm nW/cell}$ when operating at maximum readout bandwidth ($f_{\rm op} = 13$ MHz, $\Delta V = 0.7\,{\rm V}$, $C_{\rm FET} = 4$ fF). However, due to filtering of the lines delivering the control FET signals $V_{\rm WL1,2}$ and data line signal $V_{\rm DL}$, $f_{\rm op}$ was limited to 1 kHz in this demonstration (see Supplementary Figure 2).

INTEGRATED DESIGN AND SCALING UP THE ARCHITECTURE

An integrated design of the readout architecture requires a careful analysis of the relevant circuit parameters and their effect on chip footprint and readout SNR. To simulate performance of the architecture, we put for-

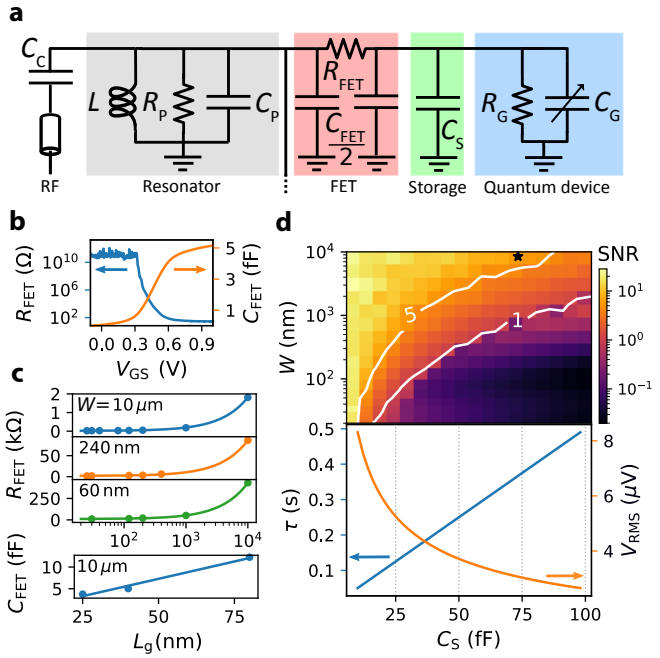

FIG. 5. **Integration.** **a**, Complete circuit model of a single cell composed of the resonator, control FET, storage capacitor and quantum device. The dotted wire indicates the connection to a subsequent cell. **b**, Exemplary resistance and capacitance measurement of a control FET of width $W = 10\,\mu$m and $L_g = 40$ nm at 4 K with turn on at $V_{GS} = 0.5$ V. **c**, Measurements at 4 K (circles) and fit to a model (solid line) of the *on* state capacitance and resistance of FETs with different width W as a function of gate length L_g. **d**, The top panel shows the calculated SNR based on the circuit model with the FET in the *on* state, $L = 400$ nH, $C_p = 480$ fF and R_p=800 kΩ as a function of C_S and W for $L_g = 20$ nm. The black star represents the configuration of the experiment. The bottom panel shows estimations of thermal noise (assuming a temperature of 50 mK) and circuit time constant with the FET in the *off* state.

ward an equivalent circuit of a single cell based on the discharge model in Fig. 3a which we expand to the RF domain (see Fig. 5a). The model consists of a readout resonator (inductance L, capacitance C_p and resistive losses R_p), the control FET (channel resistance R_{FET} and gate capacitance C_{FET}, split equally between source and drain), a charge storage capacitor C_S and the quantum device with state dependent gate capacitance C_G and resistance R_G, respectively [43]. To permit direct integration of classical and quantum devices, the classical control circuit should not exceed critical dimensions of the quantum circuit. In case of a dense array of silicon-based quantum dot qubits, the pitch should be smaller than 100 nm to allow exchange based two qubit gates [44]. Here we give clear guidelines of the circuit values that would enable such integration.

To estimate the dependence of the readout SNR on cell parameters, we first consider the signal to be measured, i.e. the change in quantum device capacitance when electrons tunnel. For quantum dots with a tunnel coupling $\Delta_c = 20\ \mu$eV and lever arm $\alpha = 0.5$, C_G changes by $\Delta C_G = 1$ fF [37] from its geometrical value of ≈ 10 aF, when tunnelling is allowed. Based on the dynamic operation results, we assume R_G is much greater than the impedance of the gate capacitance and can be treated as infinite. Turning to the control FET, we characterise multiple devices of different channel widths W and gate lengths L_g at 4 K to extract the channel resistance and gate capacitance. In Fig. 5b, we show an exemplary measurement of the dependence of these parameters with gate voltage (V_{GS}) for a transistor with $W = 10\ \mu$m and $L_g = 40$ nm and, in Fig. 5c, how the *on* state values depend on device dimensions. From these measurements we generate a model for R_{FET} and C_{FET} as a function of W and L_g (see Supplementary Equations 5&6) and perform circuit simulations assuming a well-matched high-Q RF resonator ($f_c = 310$ MHz and $Q \approx 400$) [35].

The SNR depends on multiple circuit components but here we study its dependence with the parameters that affect the physical dimensions of the cell most significantly: W and C_S. Figure 5d shows the simulated SNR for an integration time of 4 μs (much shorter than the coherence time of electron spins in ^{28}Si, $T_2^* \approx 100\ \mu$s), a noise temperature of 4 K and an optimised applied power at each data point. The SNR decreases as W decreases (*on* state R_{FET} increases) - this can be compensated by decreasing C_S, but only at the cost of reducing the time constant (τ) and increasing the RMS thermal noise voltage (see Fig. 5d). In balancing these various requirements to optimise for C_S, it is also important to consider the capacitor footprint. In DRAM, a storage capacitance of $C_S = 10 - 25$ fF is required to achieve a refresh time in the range of milliseconds. DRAM cells have been continuously scaled down while maintaining a total footprint of $6F^2$ by using trench or stacked capacitors with exotic high-k dielectrics and large capacitor aspect ratios where F is the minimum feature size currently reaching sub 10-nm [45]. We can therefore identify an example set of parameters ($W = 100$ nm, $C_S = 25$ fF) which can fit within an approximate 100 x 100 nm^2 footprint, commensurate with a QD pitch in a dense array, and still obtain SNR > 1 (see Supplementary Figure 6 and Supplementary Table 2). For these parameters thermal noise increases towards 5 μV – comparable to the precision and noise of common low noise voltage sources ($1 - 10\,\mu$V) – and the RC time constant decreases to 0.1 s, which is sufficient for a regular refresh of gate voltages. If longer retention time, better voltage stability (drift and noise) at the same SNR and readout bandwidth is desired, or even higher SNR and readout bandwidth, then the control circuitry requires a larger transistor or capacitor footprint; unless further advance in low *on* state resistance transistors or compact storage capacitors are made. Alternatively, the SNR can be improved by increasing the frequency of operation, the Q_L of the resonator, or using quantum-limited amplifi-

cation [35, 46, 47]. Moreover, requirements on critical dimensions could be relaxed depending on the architectural implementation, which could range from $< 100\,\mathrm{nm}$ for dense arrays using direct exchange to $< 400\,\mathrm{nm}$ when using mediated exchange via an intermediate state [48] or even $1-1000\,\mu\mathrm{m}$ when using sparse qubit arrays and long distance coupling via capacitive couplers, spin shuttles or superconducting resonators [30].

Our strategy for extending this demonstration to a large scale array builds on ideas that have appeared in the literature [29–31] by combining sequential gated readout with frequency multiplexing techniques [49] allowing addressing of an $N \times M$ array (see Supplementary Figure 4). There is a potential for the required inductors to be integrated and CMOS compatible using TiN. The footprint of such inductors can be reduced when operating at higher frequencies, using small critical dimensions and kinetic inductance (see Supplementary Information on control circuit footprint). Additionally, we envision to apply this strategy to a double QD split-gate architecture (see Supplementary Figure 5) where manipulation and readout signals are applied to different gates [50].

CONCLUSION

We have reported a CMOS dynamic random access architecture for radio-frequency readout of QD devices at millikelvin temperatures. We show sequential dispersive readout of individual quantum devices in a two-cell layout, which is an important step in being able to address larger arrays. We find opposing requirements between SNR, charge retention and circuit footprint when analysing scaling towards an integrated design. Our results provide guidelines to find a compromise between the desired measurement bandwidth, voltage drift, noise tolerances and critical dimensions, which can be different for a given qubit implementation. We find exemplary circuit values that can allow the desired level of integration for a particular implementation. Further work towards circuit optimisation could include cross-talk mitigation, adaptation of our circuit model to integrated circuit design, and creation of high quality superconducting inductors with reduced footprint.

METHODS

Fabrication details

All CMOS transistors used in this study were fabricated on SOI substrates with a 145-nm-thick buried oxide and 10-nm-thick silicon layer. The silicon layer is patterned to create the channel by means of optical lithography, followed by a resist trimming process. All transistors share the same gate stack consisting of 1.9 nm HfSiON capped by 5 nm TiN and 50 nm polycrystalline silicon leading to a total equivalent oxide thickness of 1.3 nm. The Si thickness under the HfSiON/TiN gate is 11 nm. After gate etching a SiN layer (10 nm) was deposited and etched to form a first spacer on the sidewalls of the gate. 18-nm-thick Si raised source and drain contacts were selectively grown before the source/drain extension implantation and activation annealing. A second spacer was formed followed by source/drain implantations, activation spike anneal and salicidation (NiPtSi). The wide channel control FETs T_i and nanowire quantum devices Q_i are connected via on-chip aluminium bond wires.

Measurement setup

Measurements were performed at base temperature of a dilution refrigerator (15 mK). Low frequency signals (V_{SD}, V_{DL}, $V_{\mathrm{WL1,2}}$) were delivered through filtered cryogenic loom while a radio-frequency signal for gate-based readout was delivered through an attenuated and filtered coaxial line which connects to a on-PCB bias tee combining the RF modulation with V_{DL}. The resonator is formed from a 82 nH inductor and the sample's parasitic capacitance to ground in parallel with the device. The inductor consists of a surface mount wire-wound ceramic core (EPCOS B82498B series) and the PCB is made from 0.8 mm thick RO4003C with immersion silver finish. The reflected RF signal is amplified at 4 K (LNF-LNC0.6_2A) and room temperature followed by quadrature demodulation (Polyphase Microwave AD0540B) from which the amplitude and phase of the reflected signal is obtained (homodyne detection).

DATA AVAILABILITY

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

[1] Ashley Montanaro, "Quantum algorithms: an overview," npj Quantum Information **2**, 15023 (2016), arXiv:1511.04206.

[2] T D Ladd, F Jelezko, R Laflamme, Y Nakamura, C Monroe, and J L O'Brien, "Quantum computers." Nature **464**, 45–53 (2010), arXiv:1009.2267.

[3] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis, "A blueprint for demonstrating quantum supremacy with superconducting qubits," Science **360**, 195–199 (2018), arXiv:1709.06678.

[4] Jeroen P. G. van Dijk, Erika Kawakami, Raymond N. Schouten, Menno Veldhorst, Lieven M. K. Vandersypen, Masoud Babaie, Edoardo Charbon, and Fabio Sebastiano, "The impact of classical control electronics on qubit fidelity," Preprint at http://arxiv.org/abs/1803.06176 (2018), arXiv:1803.06176.

[5] David P. Franke, James S. Clarke, Lieven M. K. Vandersypen, and Menno Veldhorst, "Rent's rule and extensibility in quantum computing," Preprint at http://arxiv.org/abs/1806.02145 (2018), arXiv:1806.02145.

[6] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, "Surface codes: Towards practical large-scale quantum computation," Physical Review A **86**, 032324 (2012), arXiv:1208.0928.

[7] David J Reilly, "Engineering the quantum-classical interface of solid-state qubits," npj Quantum Information **1**, 15011 (2015).

[8] J. M. Hornibrook, J. I. Colless, I. D. Conway Lamb, S. J. Pauka, H. Lu, A. C. Gossard, J. D. Watson, G. C. Gardner, S. Fallahi, M. J. Manfra, and D. J. Reilly, "Cryogenic Control Architecture for Large-Scale Quantum Computing," Physical Review Applied **3**, 024010 (2015), arXiv:1409.2202.

[9] Harald Homulle, Stefan Visser, Bishnu Patra, Giorgio Ferrari, Enrico Prati, Fabio Sebastiano, and Edoardo Charbon, "A reconfigurable cryogenic platform for the classical control of quantum processors," Review of Scientific Instruments **88** (2017), 10.1063/1.4979611, arXiv:1602.05786.

[10] M Veldhorst, J C C Hwang, C H Yang, a W Leenstra, B de Ronde, J P Dehollain, J T Muhonen, F E Hudson, K M Itoh, A Morello, and a S Dzurak, "An addressable quantum dot qubit with fault-tolerant control-fidelity," Nature Nanotechnology **9**, 981–985 (2014), arXiv:1407.1950v1.

[11] Juha T Muhonen, Juan P Dehollain, Arne Laucht, Fay E Hudson, Rachpon Kalra, Takeharu Sekiguchi, Kohei M Itoh, David N Jamieson, Jeffrey C. McCallum, Andrew S Dzurak, and Andrea Morello, "Storing quantum information for 30 seconds in a nanoelectronic device," Nature Nanotechnology **9**, 986–991 (2014), arXiv:1402.7140.

[12] M Veldhorst, C H Yang, J C C Hwang, W Huang, J P Dehollain, J T Muhonen, S Simmons, A Laucht, F E Hudson, K M Itoh, A Morello, and A S Dzurak, "A Two Qubit Logic Gate in Silicon," Nature **526**, 410–414 (2015), arXiv:1411.5760.

[13] K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, "Isotopically enhanced triple-quantum-dot qubit," Science Adv. **1**, e1500214 (2015).

[14] X Mi, J V Cady, D M Zajac, P W Deelman, and J R Petta, "Strong coupling of a single electron in silicon to a microwave photon," Science **355**, 156–158 (2017), arXiv:1703.03047v1.

[15] Zhan Shi, C. B. Simmons, J. R. Prance, John King Gamble, Teck Seng Koh, Yun-Pil Shim, Xuedong Hu, D. E. Savage, M. G. Lagally, M. A. Eriksson, Mark Friesen, and S. N. Coppersmith, "Fast Hybrid Silicon Double-Quantum-Dot Qubit," Physical Review Letters **108**, 140503 (2012).

[16] Patrick Harvey-Collard, N. Tobias Jacobson, Martin Rudolph, Jason Dominguez, Gregory A. Ten Eyck, Joel R. Wendt, Tammy Pluym, John King Gamble, Michael P. Lilly, Michel Pioro-Ladrière, and Malcolm S. Carroll, "Coherent coupling between a quantum dot and a donor in silicon," Nature Communications **8**, 1029 (2017), arXiv:1512.01606.

[17] E Kawakami, P Scarlino, Daniel R. Ward, F R Braakman, D. E. Savage, M. G. Lagally, Mark Friesen, Susan N. Coppersmith, Mark A. Eriksson, and L M K Vandersypen, "Electrical control of a long-lived spin qubit in a Si/SiGe quantum dot." Nat. Nano. **9**, 666 (2014), arXiv:1404.5402.

[18] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Lagally, Mark Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, "A programmable two-qubit quantum processor in silicon," Nature **555**, 633–637 (2018), arXiv:1708.04214.

[19] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta, "Resonantly driven CNOT gate for electron spins," Science **359**, 439–442 (2018).

[20] A. C. Betz, R. Wacquez, M. Vinet, X. Jehl, A. L. Saraiva, M. Sanquer, A. J. Ferguson, and M. F. Gonzalez-Zalba, "Dispersively Detected Pauli Spin-Blockade in a Silicon Nanowire Field-Effect Transistor," Nano Letters **15**, 4622–4627 (2015), arXiv:1504.02997v2.

[21] Anderson West, Bas Hensen, Alexis Jouan, Tuomo Tanttu, Chih-Hwan Yang, Alessandro Rossi, M. Fernando Gonzalez-Zalba, Fay Hudson, Andrea Morello, David J. Reilly, and Andrew S. Dzurak, "Gate-based single-shot readout of spins in silicon," Nature Nanotechnology **14**, 437–441 (2019), arXiv:1809.01864.

[22] P. Pakkiam, A. V. Timofeev, M. G. House, M. R. Hogg, T. Kobayashi, M. Koch, S. Rogge, and M. Y. Simmons, "Single-Shot Single-Gate rf Spin Readout in Silicon," Physical Review X **8**, 041032 (2018), arXiv:1809.01802.

[23] D. M. Zajac, T. M. Hazard, X Mi, E Nielsen, and J. R. Petta, "Scalable Gate Architecture for a One-Dimensional Array of Semiconductor Spin Qubits," Physical Review Applied **6**, 054013 (2016), arXiv:1607.07025.

[24] Kenta Takeda, Jun Kamioka, Tomohiro Otsuka, Jun Yoneda, Takashi Nakajima, Matthieu R. Delbecq, Shinichi Amaha, Giles Allison, Tetsuo Kodera, Shunri Oda, and Seigo Tarucha, "A fault-tolerant addressable spin qubit in a natural silicon quantum dot," Science Advances **2**, 1–7 (2016), arXiv:1602.07833.

[25] Jun Yoneda, Kenta Takeda, Tomohiro Otsuka, Takashi Nakajima, Matthieu R. Delbecq, Giles Allison, Takumu Honda, Tetsuo Kodera, Shunri Oda, Yusuke Hoshi, Noritaka Usami, Kohei M. Itoh, and Seigo Tarucha, "A quantum-dot spin qubit with coherence limited by charge

noise and fidelity higher than 99.9%," Nature Nanotechnology **13**, 102–106 (2018), arXiv:1708.01454.

[26] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, A. Morello, A. Laucht, and A. S. Dzurak, Preprint at http://arxiv.org/abs/1805.05027 arXiv:1805.05027.

[27] R Maurand, X Jehl, D. Kotekar-Patil, A Corna, H Bohuslavskyi, R. Laviéville, L Hutin, S Barraud, M Vinet, M Sanquer, and S. De Franceschi, "A CMOS silicon spin qubit," Nature Communications **7**, 13575 (2016).

[28] Alessandro Crippa, Romain Maurand, Léo Bourdet, Dharmraj Kotekar-Patil, Anthony Amisse, Xavier Jehl, Marc Sanquer, Romain Laviéville, Heorhii Bohuslavskyi, Louis Hutin, Sylvain Barraud, Maud Vinet, Yann Michel Niquet, and Silvano De Franceschi, "Electrical Spin Driving by g -Matrix Modulation in Spin-Orbit Qubits," Physical Review Letters **120**, 1–5 (2018), arXiv:1710.08690.

[29] M. Veldhorst, H. G. J. Eenink, C. H. Yang, and A. S. Dzurak, "Silicon CMOS architecture for a spin-based quantum computer," Nature Communications **8**, 1766 (2017), arXiv:1609.09700.

[30] L. M. K. Vandersypen, H. Bluhm, J. S. Clarke, A. S. Dzurak, R. Ishihara, A. Morello, D. J. Reilly, L. R. Schreiber, and M. Veldhorst, "Interfacing spin qubits in quantum dots and donorshot, dense, and coherent," npj Quantum Information **3**, 34 (2017), arXiv:1612.05936.

[31] Ruoyu Li, Luca Petit, David P. Franke, Juan Pablo Dehollain, Jonas Helsen, Mark Steudtner, Nicole K. Thomas, Zachary R. Yoscovits, Kanwal J. Singh, Stephanie Wehner, Lieven M. K. Vandersypen, James S. Clarke, and Menno Veldhorst, "A crossbar network for silicon quantum dot qubits," Science Advances **4**, eaar3960 (2018), arXiv:1711.03807.

[32] S Schaal, S Barraud, J. J. L. Morton, and M. F. Gonzalez-Zalba, "Conditional Dispersive Readout of a CMOS Single-Electron Memory Cell," Physical Review Applied **9**, 054016 (2018).

[33] J. I. Colless, A. C. Mahoney, J. M. Hornibrook, A. C. Doherty, H. Lu, A. C. Gossard, and D. J. Reilly, "Dispersive Readout of a Few-Electron Double Quantum Dot with Fast rf Gate Sensors," Physical Review Letters **110**, 046805 (2013), arXiv:1210.4645v1.

[34] M F Gonzalez-Zalba, S Barraud, a J Ferguson, and a C Betz, "Probing the limits of gate-based charge sensing," Nature Communications **6**, 6084 (2015).

[35] Imtiaz Ahmed, James A. Haigh, Simon Schaal, Sylvain Barraud, Yi Zhu, Chang-min Lee, Mario Amado, Jason W. A. Robinson, Alessandro Rossi, John J. L. Morton, and M. Fernando Gonzalez-Zalba, "Radio-Frequency Capacitive Gate-Based Sensing," Physical Review Applied **10**, 014018 (2018), arXiv:1801.09759.

[36] Benoit Voisin, Viet Hung Nguyen, Julien Renard, Xavier Jehl, Sylvain Barraud, François Triozon, Maud Vinet, Ivan Duchemin, Yann Michel Niquet, Silvano De Franceschi, and Marc Sanquer, "Few-electron edge-state quantum dots in a silicon nanowire field-effect transistor," Nano Letters **14**, 2094–2098 (2014).

[37] R. Mizuta, R. M. Otxoa, A. C. Betz, and M. F. Gonzalez-Zalba, "Quantum and tunneling capacitance in charge and spin qubits," Physical Review B **95**, 045414 (2017), arXiv:1604.02884.

[38] A. Rossi, R. Zhao, A. S. Dzurak, and M. F. Gonzalez-Zalba, "Dispersive readout of a silicon quantum dot with an accumulation-mode gate sensor," Applied Physics Letters **110**, 212101 (2017), arXiv:1610.00767.

[39] H. Al-Taie, L. W. Smith, B. Xu, P. See, J. P. Griffiths, H. E. Beere, G. A. C. Jones, D. A. Ritchie, M. J. Kelly, and C G Smith, "Cryogenic on-chip multiplexer for the study of quantum transport in 256 split-gate devices," Applied Physics Letters **102**, 243102 (2013).

[40] R K Puddy, L W Smith, H. Al-Taie, C H Chong, I Farrer, J P Griffiths, D. A. Ritchie, M. J. Kelly, M. Pepper, and C. G. Smith, "Multiplexed charge-locking device for large arrays of quantum devices," Applied Physics Letters **107**, 143501 (2015).

[41] C. Eichenberger and W. Guggenbuhl, "Charge injection of analogue CMOS switches," IEE Proceedings G Circuits, Devices and Systems **138**, 155 (1991).

[42] David Tawei Wang, *Modern DRAM Memory Systems: Performance Analysis And A High Performance, Power-Constrained Dram Scheduling Algorithm*, Ph.D. thesis (2005).

[43] M. Esterli, R. M. Otxoa, and M. F. Gonzalez-Zalba, "Small-signal equivalent circuit for double quantum dots at low-frequencies," Preprint at http://arxiv.org/abs/1812.06056 (2018), arXiv:1812.06056.

[44] B M Maune, M. G. Borselli, B. Huang, T. D. Ladd, P. W. Deelman, K. S. Holabird, A. A. Kiselev, I Alvarado-Rodriguez, R. S. Ross, A. E. Schmitz, M Sokolich, C A Watson, M. F. Gyure, and A. T. Hunter, "Coherent singlet-triplet oscillations in a silicon-based double quantum dot," Nature **481**, 344–347 (2012).

[45] M.A. Siddiqi, *Dynamic RAM: Technology Advancements* (CRC Press, 2017).

[46] J. Stehlik, Y.-Y. Liu, C. M. Quintana, C. Eichler, T. R. Hartke, and J. R. Petta, "Fast Charge Sensing of a Cavity-Coupled Double Quantum Dot Using a Josephson Parametric Amplifier," Physical Review Applied **4**, 014018 (2015), arXiv:1502.01283.

[47] G. Zheng, N. Samkharadze, M. L. Noordam, N. Kalhor, D. Brousse, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, "Rapid high-fidelity gate-based spin read-out in silicon," Preprint at http://arxiv.org/abs/1901.00687 (2019), arXiv:1901.00687.

[48] V Srinivasa, H Xu, and J. M. Taylor, "Tunable Spin-Qubit Coupling Mediated by a Multielectron Quantum Dot," Physical Review Letters **114**, 226803 (2015).

[49] J. M. Hornibrook, J. I. Colless, A. C. Mahoney, X. G. Croot, S. Blanvillain, H. Lu, A. C. Gossard, and D. J. Reilly, "Frequency multiplexing for readout of spin qubits," Applied Physics Letters **104**, 103108 (2014), arXiv:1312.5064.

[50] S De Franceschi, L Hutin, R Maurand, L Bourdet, H Bohuslavskyi, and A Corna, "SOI technology for quantum information processing," 2016 IEEE International Electron Devices Meeting (IEDM) , 13.4.1–13.4.4 (2016), arXiv:1605.07599.

ACKNOWLEDGEMENTS

We would like to thank Sebastian Pauka, Steve A. Lyon and Matthew Schormans for helpful discussions. This research has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 688539 (http://mos-quito.eu) and Seventh Framework Programme (FP7/2007-2013) through Grant Agreement No. 318397 (http://www.tolop.eu.); as well as by the Engineering and Physical Sciences Research Council (EPSRC) through the Centre for Doctoral Training in Delivering Quantum Technologies (EP/L015242/1) and UNDEDD (EP/K025945/1). M.F.G.Z. and A.R. acknowledge support from the Winton Programme for the Physics of Sustainability and Hughes Hall, University of Cambridge.

AUTHOR CONTRIBUTIONS

S.S. and M.F.G.-Z. devised the experiment. S.S., A.R. and M.F.G.-Z. performed the experiments; S.B. fabricated the sample; V.N.C.-T and T.-Y.Y. performed measurements for low-temperature modelling; V.N.C.-T developed and performed simulations towards integration; S.S. did the analysis and prepared the manuscript with contributions from A.R., J.J.L.M. and M.F.G.-Z.

COMPETING INTERESTS

The authors declare no competing interests.