

Supplementary Materials for

A crossbar network for silicon quantum dot qubits

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, Menno Veldhorst*

*Corresponding author. Email: m.veldhorst@tudelft.nl

Published 6 July 2018, *Sci. Adv.* **4**, ear3960 (2018)
DOI: 10.1126/sciadv.aar3960

This PDF file includes:

- Section S1. Tolerance to quantum dot inhomogeneity
- Section S2. Column-by-column alternating static magnetic field
- Section S3. Inhomogeneity of the ESR stripline
- Section S4. GRAPE pulse for spin rotation
- Section S5. Shuttling fidelity
- Section S6. Pauli spin blockade spin-to-charge conversion with ancillary qubit in the spin-up state
- Section S7. Shuttling bus for a 2D array module
- Fig. S1. Impact of misalignment and errors in gate and dot dimensions.
- Fig. S2. Overall resonance frequency error as a function of fabrication error.
- Fig. S3. Stripline schematic and simulation results.
- Fig. S4. GRAPE pulse optimization for high-fidelity single-qubit gates.
- Fig. S5. Charge shuttling process.
- Fig. S6. Scheme for Pauli spin blockade spin-to-charge conversion with ancillary qubit in the spin-up state.
- Fig. S7. Connecting qubit modules.
- References (41–44)

Section S1. Tolerance to quantum dot inhomogeneity

In this section, we discuss the required homogeneity for the shared gate control. Firstly, we estimate the upper bound of the inter-dot tunnel coupling when the tunnel barriers are set to the off-state. Finite tunnel coupling in the off-state can result in unwanted shuttling of electrons. These shuttle processes need to be error corrected in quantum algorithms. Here, we consider the surface code operation (2), to estimate the error correction cycle. Importantly, undesired shuttle events could occur at any coordinate of the quantum module. Consequently, we need to take into account the idle qubits in each step and consider the complete cycle time of the qubit module. We target for a shuttle-error rate below 0.1% in a complete error-correction cycle. Most errors are expected during the readout of the measurement qubits, due to speed and pulsing requirements. Row-controlled Pauli spin blockade spin to charge conversion can be performed at a frequency of 3 MHz, such that spin-to-charge conversion on a 50×50 qubit module size can be within 10 μ s. In order to achieve a 0.1% error rate, electrons should have a shuttle rate at least 10^3 smaller than the cycle time, and we require $t_{0,\text{off}} < 10$ Hz in the off-state.

Now we discuss the tunneling rate range when the barriers are set to the on-state. Since the desired qubit shuttling rate is at least 1 GHz, we require a minimum coupling of $t_{0,\text{on}} > 10$ GHz. The upper bound in the tunnel rate follows from the requirement that a larger tunneling rate needs a larger detuning to prevent charge state mixing between different quantum dots (see Section 5 for details). While in principle larger voltage pulses may enable this, larger detuning could lead to overhead in operations and bring down undesired higher orbital levels. Therefore, we require the tunneling rate variation to be within one order of magnitude with an upper bound for the tunneling rate $t_{0,\text{on}} < 100$ GHz.

Next to the tunnel coupling, the uniformity requirement on the quantum dot chemical potential $\Delta\mu$ are also crucial. When $\Delta\mu < E_C$, where E_C denotes the charging energy, all the quantum dots under the same QL can be controlled to have the same charge state. Importantly, this qubit occupation configuration is the ground state even when the inter-dot tunnel couplings are set to on. We envision this to be beneficial in order to correct shuttle errors. Although the qubit configuration in Fig. 1B can also be achieved with $\Delta\mu > E_C$, by sequentially shuttling the qubits to the desired sites, we point out that lower uniformity demands higher detuning for compensation. This could slow down the gate pulsing speed and hence the overall operation rate, and impractical voltage pulses may be required.

The regime $\Delta\mu < E_C$ is also important for operations like parallel Pauli spin blockade spin to charge conversion, as shown in the main text, Fig. 5C. When variations in μ are large, electrons such as the one labelled 2 in the figure may shuttle to an adjacent column. While such errors could be corrected by another phase-controlled shuttle, we envision this to be impractical and providing a significant overhead. First, it will significantly slow down pulses to ensure adiabaticity. Second, it will require large voltage pulses in order to overcome the variations. Third, after a Pauli spin blockade step, a shuttle step has to be implemented purely for pulsing back electrons to their targeted positions. Instead, when $\Delta\mu < E_C$, these requirements are avoided altogether, since the spread in chemical potential energies is smaller than the charging energy that separates different electron occupations, such that electrons will remain in their targeted positions.

In conclusion, we require the chemical potential variations $\Delta\mu < E_C$, and the tunnel coupling $t_{\text{off}} < 10$ Hz and $10 \text{ GHz} < t_{\text{on}} < 100 \text{ GHz}$.

Section S2. Column-by-column alternating static magnetic field

In this section, we estimate the column-by-column alternating magnetic field generated by the constant current through the CL . The CL have width $w_{CL} = 30$ nm and height $h_{CL} = 60$ nm. Since the effective superconducting penetration depth λ_{eff} can be much larger than these dimensions, we assume a uniform current density through the CL . The resulting magnetic field is calculated along the row direction and in the plane 20 nm below the bottom of the CL grids, where the quantum dot qubits are located. For generality, we take the CL s infinite in length and calculate the field strength via the Biot-Savart Law. We note that this assumption will hold for the large arrays assumed here, although the edges will require corrections. The rectangle shape of CL is approximated by dividing it into square node points that are uniformly spaced in the rectangle with equal currents (30×60 nodes). A current density of $j_{CL} \sim 4 \times 10^{10}$ A/m² in the gate lines could generate the targeted $\delta\nu_{CL} = 10$ MHz frequency difference between columns. Though this current density is significant, it is below the superconducting critical current density of for example NbN (41). In total, we calculate 40 CL s and convert the field to resonance frequency using $g=2$, and plot the middle region as shown in Fig. 2D of the main text.

We further estimate the local resonance frequency variation, $\Delta\nu_0$, due to imperfect device fabrication or inhomogeneity in quantum dot position and size. In fig. S1A-D, we show the influence of a deviation of a certain geometry in CL gate on $\Delta\nu_0$, by comparing it to ideal CL gates. The left panels shows the color-coded relative resonance frequency error with respect to the designed value along the row direction (x-axis) for different fabrication errors (y-axis). Although we find that $\Delta\nu_0$ can be significantly, the maximum is in between the quantum dots and the amplitude is strongly reduced at the center of the qubit location. To estimate the influence on the quantum dot position, we calculate the average resonance frequency error, $\Delta\nu_{0,ave}$, for different dot sizes, as shown in the right panels. By comparing the various results, we can see that with the same absolute fabrication error, the offset of the CL gate height has the strongest effect, especially for the misalignment on the bottom side (fig. S1C). We envision therefore that it will be crucial to choose an integration scheme that minimizes the roughness under the gate area. The influence of quantum dot geometry on the field is relatively weak, as shown in fig. S1E-S1F, since the out-of-plane field is not sensitive to the dot position as shown in Fig. 2D of the main text.

In fig. S2, we plot the sum of all the errors presented in fig. S1 with a dot size of 20 nm. For a 1 nm error in fabrication or dot geometry homogeneity, the maximum change in magnetic field $\delta\nu_{fab} = 100$ kHz (Note that this number is not sensitive to the quantum dot size. The limiting factor in $\Delta\nu_0$ is CL height, which shows weak dependence on the quantum dot size).

We estimate the magnetic field at the dot sites by taking linear averages for different dot sizes. The real electron wave function will have a different distribution, and will distribute more in the middle than the edge. Since we find that the center is most insensitive to fabrication errors, the results shown here can be taken as the upper bound on requirements. This also becomes clear from fig. S1, where we see that larger dot sizes will generally contribute to higher magnetic field deviations. In addition, if the self-correlation length of the geometry error is smaller than the effective span of the electron wave function, the overall resonance frequency variation will be even smaller. Therefore, we estimate that for 1 nm root-mean-square (rms) variation in the gate geometry, the qubit to qubit resonance frequency variation is in the range of $\delta\nu_{fab} = 100$ kHz. While these numbers are certainly challenging, industrial fabrication has pushed uniformity to the limit, such that alignments with nm resolution are possible (16).

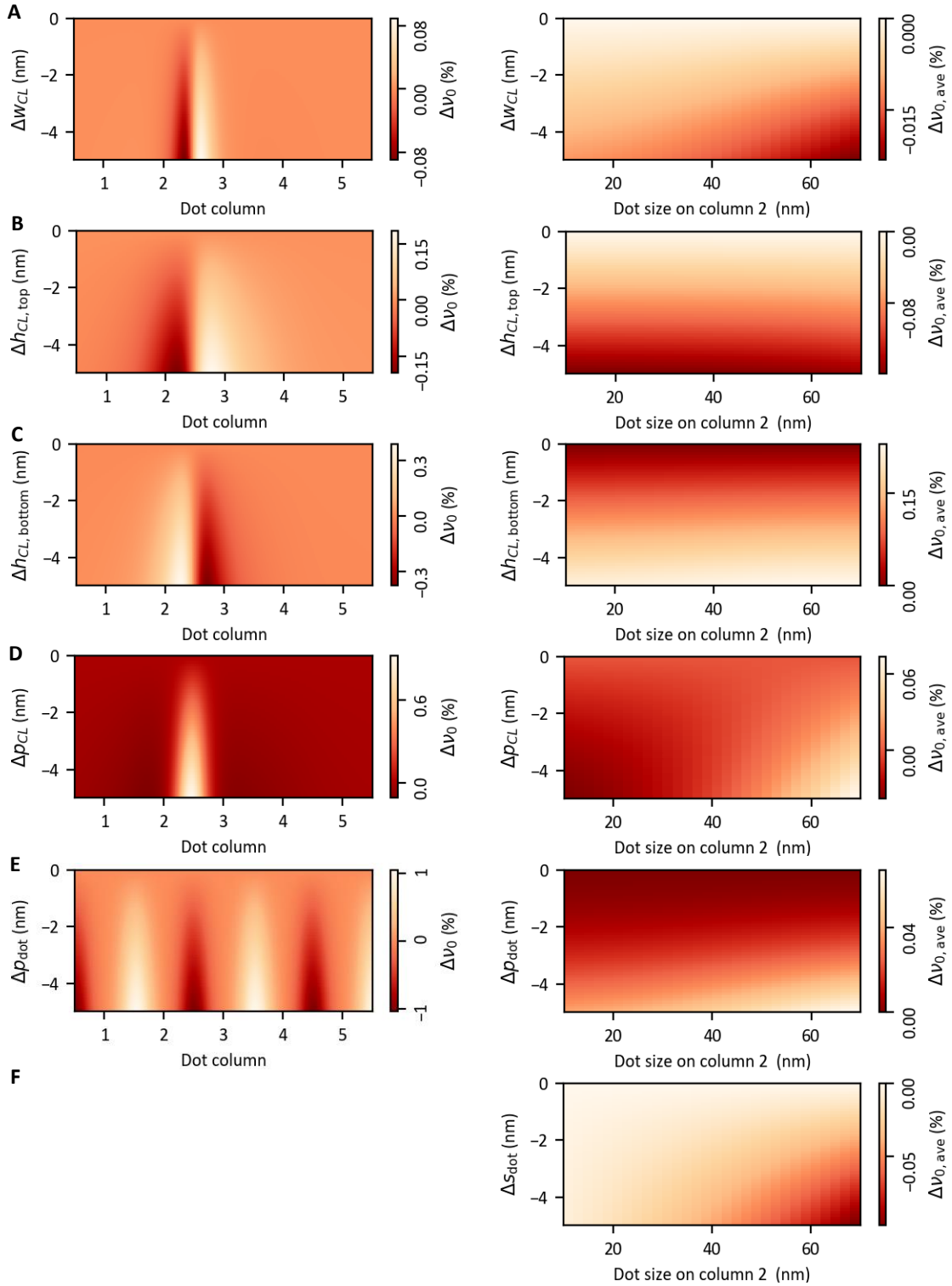


Fig. S1. Impact of misalignment and errors in gate and dot dimensions. Error in ΔV_0 for errors in (A) the CL gate width Δw_{CL} , (B) the CL gate top location $\Delta h_{CL,top}$, (C) the CL gate bottom location $\Delta h_{CL,bottom}$, (D) the CL gate lateral location Δp_{CL} , (E) the dot location Δp_{dot} , and (F) the dot size Δs_{dot} . The left panels show the errors normalized with respect to the targeted perpendicular magnetic field. The right panels show the averaged errors, $\Delta V_{0,ave}$, taking into account the finite size of the quantum dots.

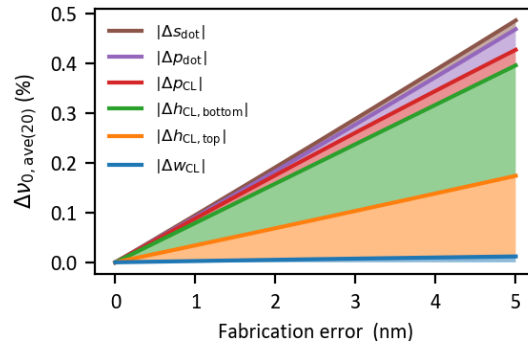


Fig. S2. Overall resonance frequency error as a function of fabrication error. The ν_0 error is calculated based on a dot size of 20 nm. Different resonance frequency errors are stacked on top of each other to show the worst case scenario assuming the same deviation in gate or dot geometry. Controlling the vertical dimension is most critical.

Section S3. Inhomogeneity of the ESR stripline

Simultaneous qubit control requires the amplitude of the spin-resonant magnetic field to be highly homogenous, such that all resonant qubits respond with the same Rabi frequency. In order to estimate the homogeneity and optimize the design of our stripline, we turn to the Microwave Studio simulation package from Computer Simulation Technology (CST-MWS) (20). With this 3D simulator of high-frequency devices we can create a 3D model of our stripline structure, define ports for excitations and solve Maxwell's equations over a finite-element mesh of our model.

Figure S3A shows a schematic of the stripline model we have designed and simulated. A qubit module includes a pair of narrow striplines placed above the qubit plane. Current flowing through the striplines generates a magnetic field that wraps around the striplines. Therefore, the qubit module experiences the in-plane component of this field. A stripline pair is chosen because we can obtain the same homogeneity as for the case of a single but wider stripline, while significantly less current is required. The stripline pair is furthermore simple in design. The model consists of a lossless silicon substrate with superconducting striplines on the surface. CST-MWS models these lines with a frequency dependent surface impedance and equal penetration depth λ over all frequencies. The striplines fan out to a short-circuited coplanar waveguide structure, similar to those described in reference (7,8). We used the frequency domain solver and analyzed our results at $\nu_0 = 1$ GHz.

Using the parametric optimization function built into the CST-MWS simulator, we run a sweep of simulations to optimize for field homogeneity. Here, we vary the stripline width ($w_{\text{stripline}}$), the pitch between the striplines ($d_{\text{stripline}}$), and the separation between the striplines and the qubit module ($h_{\text{stripline}}$). Figure S3B shows the plots of the homogeneity along one axis of the qubit module for parameter combinations we tested. We find RF field inhomogeneity across the 2D array $\delta v_{\text{stripline}} < 2\%$, for $w_{\text{stripline}} = 2 \mu\text{m}$, $d_{\text{stripline}} = 6 \mu\text{m}$, $h_{\text{stripline}} = 4 \mu\text{m}$.

To achieve homogenous fields the current distribution through the striplines has to be taken into account. For superconducting striplines, this is to a large extent determined by the superconducting penetration depth λ . In thin films with thickness d , the effective penetration depth is given by $\lambda_{\text{eff}} = \lambda_{\text{bulk}} \coth(d/\lambda_{\text{bulk}})$ (42). As a result, λ_{eff} in thin films can reach several micrometers, for example when using NbN with λ_{bulk} close to $0.5 \mu\text{m}$ (41). We have analyzed a range of superconductor penetration depths (with λ ranging from $0.5 \mu\text{m}$ to $5 \mu\text{m}$) and found only minor variations, demonstrating the robustness of our design and enabling to use a range of superconducting materials and film-thicknesses for the stripline.

Additionally, we can extract the current density along our striplines by integrating the magnetic field along a cross-section of the stripline, which is readily available as part of the simulation results.

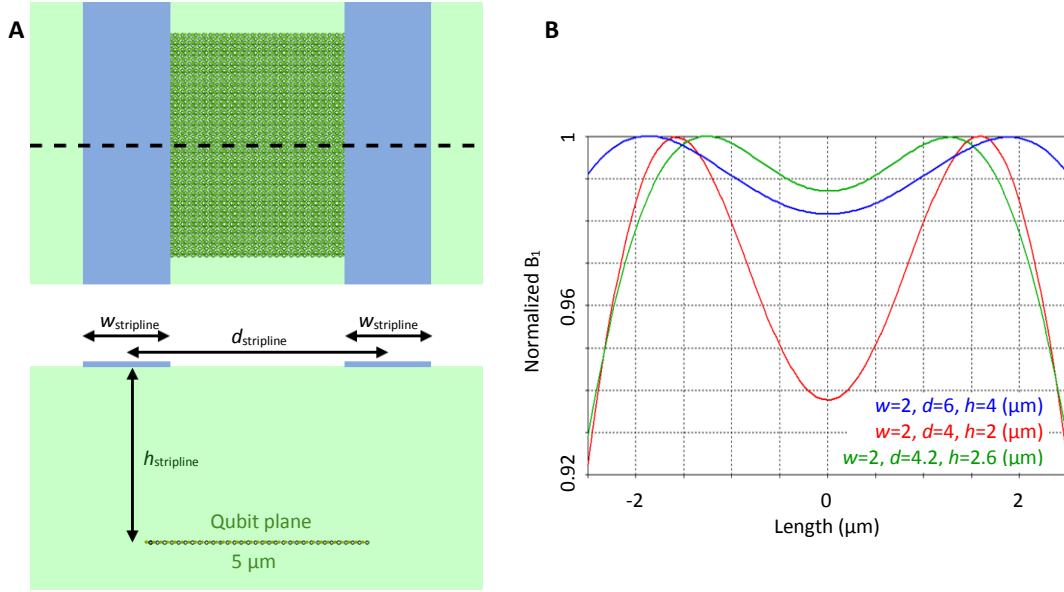


Fig. S3. Stripline schematic and simulation results. (A) Top image shows a top view of the stripline design with the superconducting metal strips in blue, the silicon substrate in green and the qubits represented as small circles. Bottom image is a cross-section along the dashed line, showing the qubits directly under the stripline pair. The relevant dimensions for the design are labelled. (B) Field homogeneity for different stripline design dimensions. We configured the optimization algorithm from CST to test different design dimensions to maximize the homogeneity. The figure shows some of the results obtained, with the optimal result found in blue.

Section S4. Grape pulse for spin rotation

A crucial point in the single qubit manipulation across the 2D array is that the applied ESR pulse can address the qubits with the larger (smaller) ν_0 without effecting the qubit with the smaller (larger) ν_0 . At the same time, however, it needs to tolerate variations in the static field and in the ESR field. As discussed in the main text and Supplementary Materials Section 3, the ESR field inhomogeneities can be engineered to be $\delta\nu_{Stripline} < 2\%$. The variation in the qubit resonance frequency are estimated to be $\delta\nu_0 \sim 150$ KHz for a frequency difference of 10 MHz between the columns. In fig. S4A and S4B we show the gate fidelity of the targeted qubits ($\nu_0 = 105$ MHz) and idle qubits ($\nu_0 = 95$ MHz) as a function of variations in ν_0 and ν_1 when naively applying a 1 MHz square ESR pulse. The target gate for the 105 MHz qubit is a $\pi/2$ rotation, while for the 95 MHz qubit it is the identity operator. The resonant qubit is rotated with $> 99.9\%$ fidelity with tolerances for detuning around 100 kHz in ν_0 and 3% in ν_1 . However, the off-resonant qubit does not achieve the targeted fidelity even for null detuning.

For these reasons, we have made use of numerical techniques to identify a composite pulse that can meet our requirements. The scheme we adopted is the Gradient Ascent Pulse Engineering (GRAPE) (22). In the algorithm, the time evolution of the system is split in small timeslots in which the amplitude of the pulse is assumed to be constant. For each timeslot, the amplitude is then optimized using standard multi-variable optimization methods in order to maximize the overlap between the actual gate and the target gate. Since the goal is to obtain a selective pulse tolerant to detuning in resonance frequency, we evaluate the simulation result as an average of the fidelities of four qubits: two qubits with frequencies 105 ± 0.1 MHz targeted to be on resonance and two qubits with resonant frequencies 95 ± 0.1 MHz targeted to be off resonance. In the simulation, it is important to limit the number of qubits, since increasing the search space can also increase the number of local maxima with insufficient high fidelity such that the algorithm is incapable of solving the problem. We note that because of the low Rabi frequency compared to the Larmor frequency, we do not take the rotating wave approximation.

In fig. S4 C (D) we show the average gate fidelity for a $\pi/2$ (null) rotation using the optimized GRAPE pulse for the respective qubits. Fidelities beyond 99.9% can be achieved up to 300 kHz in ν_0 and over 3% in ν_1 . The gate can be executed in the same time as a 2 MHz Rabi pulse, i.e. in 250 ns. This comes at the cost of a slightly larger rms amplitude (1.1 MHz compared to 0.7 MHz) with a maximum peak of ~ 3 MHz.

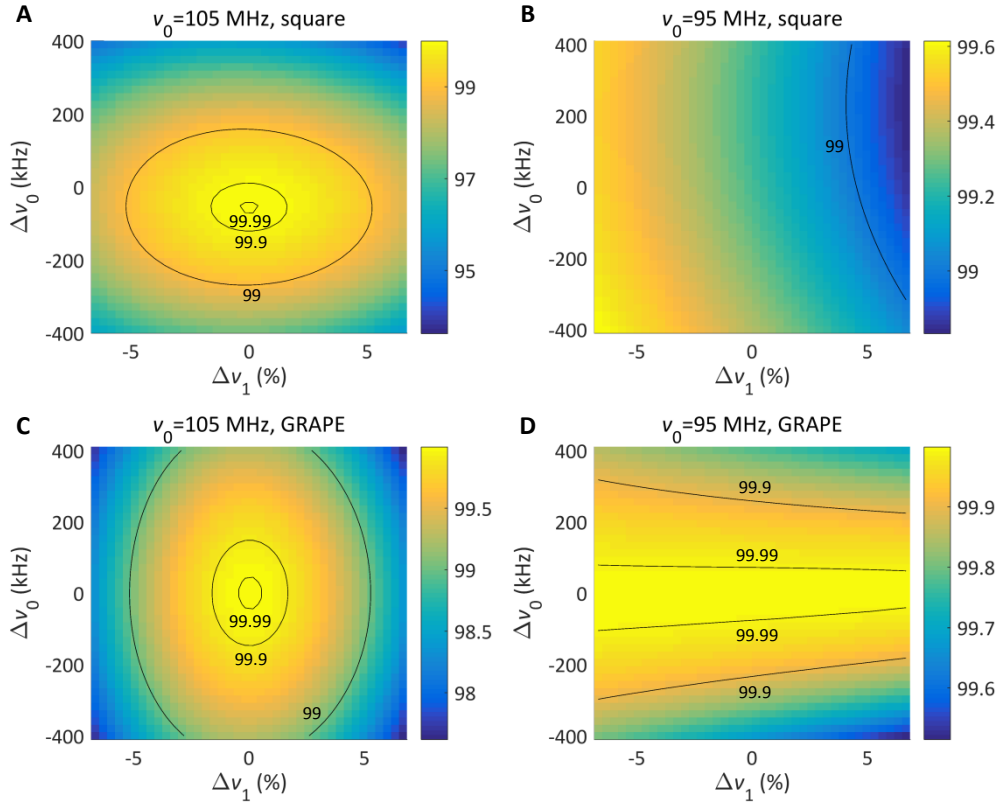


Fig. S4. GRAPE pulse optimization for high fidelity single qubit gates. Gate fidelity for a square pulse (A) and (B) and for an optimized GRAPE pulse (C) and (D), targeting in both cases a $\pi/2$ -rotation and identity-gate on the qubits with higher and lower resonant frequency, respectively. The pulse is shown as a function of variations in resonance frequency v_0 and ESR field v_1 to test the robustness of the pulse. For the optimized GRAPE pulse we can find that the fidelity can be readily above 99.9 %.

Section S5. Shuttling fidelity

In this section, we estimate the shuttling fidelity considering various inter-dot energy detunings, tunnel couplings and shuttling speeds. Firstly, we calculate the required detuning to isolate the qubits. When the tunnel barrier is set to on by CL or RL , the tunnel coupling t_0 mixes different charge states also around idle qubits. Here, we consider the lower bound, which is the situation where the tunneling barrier between an idle qubit and an empty dot is turned on diabatically. Then the charge state, say, $(1,0)$, will process around the eigenstate. To maintain the minimum $(1,0)$ fraction higher than F , the inter-dot energy detuning between the idle qubit and the empty dot needs to be larger than $\varepsilon_F = t_0|2Fh - 1|/\sqrt{Fh - Fh^2}$, with the $(1,0)$ fraction in the eigenstate $Fh = \sqrt{(F+1)/2}$. We find consequently for the case $t_0 = 100$ GHz and charge fraction $F = 99.9\%$ that the minimum required detuning is $\varepsilon_F = 26$ meV.

To separate different charge state during the shuttling process, we also consider the effect of $\Delta\mu$. As shown in fig. S5A, the clearance for each state is $\Delta\mu + \varepsilon_F$. When shuttling a single qubit as shown in Fig. 3B and C in the main text, their neighboring qubits controlled by $QL3$ should not be affected. fig. S5A shows a scheme that with high QL pulsing amplitude of $2(\Delta\mu + \varepsilon_F)$, the $QL3$ qubit remains with a negative detuning. fig. S5B shows an alternative scheme where the pulsing amplitude can be halved by applying a pulse on $QL3$ to compensate the change in detuning. The choice between fewer control signals or faster operation together with smaller pulsing amplitude can be made based on the physical qubit properties and control circuitry specifications.

We now discuss a linear and adiabatic pulsing scheme. We describe the non-adiabaticity by the Landau-Zener formula, $P_n = \exp(-4\pi^2 t_0^2 \Delta t / \Delta\varepsilon)$, where Δt is the shuttling time and $\Delta\varepsilon$ is the total detuning sweep range. To implement parallel linear shuttling scheme, we need to account for the largest pulsing amplitude with the smallest tunneling rate in order to tolerate quantum dot variations. The tunneling rate of $t_{0,max} = 100$ GHz requires $\varepsilon_{F,max} = 26$ meV (although higher t_0 allows faster shuttling, the required ε_F could be impractical to achieve). Combining this with the smallest tunneling rate $t_{0,min} = 10$ GHz and $P_n = 10^{-3}$, we find the line-by-line parallel shuttling rate $f_{parallel} \sim 45$ MHz.

In order to pulse faster, we have developed a new protocol. The results are shown in fig. S5C and S5D. In this protocol we first reduce the detuning while the tunnel coupling is off, then we turn on the coupling adiabatically (0.2 ns), apply the linear adiabatic detuning pulse (0.6ns), and finally we turn of the coupling adiabatically (0.2 ns) and set the detuning back to the idle value; corresponding to 1 GHz shuttling rate. When the tunnel barrier is set to off, there is no hard boundary on the detuning pulsing speed. The tunnel barrier can be adiabatically turned on with a timescale below nano-seconds, since the detuning energy is generally much larger than the tunnel coupling. As marked by the black contour line in fig. S5C, higher than 99.9% shuttling fidelity can be achieved requiring a minimal tunneling rate $t_{0,min} > 10$ GHz, and it does not pose an upper bound to $t_{0,max}$. Charge noise and disorder could introduce errors and complicate shuttling. Overcoming disorder in Si MOS could be particularly challenging and Si/SiGe heterostructures may prove superior in this area. We envision that semiconductor qubits fabricated in industrial facilities are particularly suitable to tackle these challenges and enable high shuttling fidelities. Furthermore, we note that a high shuttling fidelity can be achieved for a wide set of parameter values, while further optimization is possible (43).

For global shuttling (Fig. 5E) we need to take into account the chemical potentials variations between different dots, and $\Delta\varepsilon > 2\Delta\mu$, such that the pulsing amplitude is larger than the variation. We can use this requirement together with the result shown in fig. S5C to find the associated tunnel coupling for a 1 GHz shuttling rate. Considering $\Delta\mu = 2$ meV, we find $t_{0,min} > 20$ GHz.

Now we consider the global RF-dispersive based charge readout using the frequency multiplexing scheme as shown in Fig. 5D. Here, we focus on the quantum capacitance C_q , which affect the signal strength, and the full width at half maximum (FWHM), which affect the simultaneous measurement range. In the low temperature high t_0 limit, $\text{FWHM} \sim 3t_0$; and at zero detuning, $C_{q0} \sim (e\alpha')^2/4t_0$, where α' is the lever-arm different between two dots (44). For $\Delta\mu \sim 2$ meV and $t_{0,min} = 10$ GHz, we need ~ 16 measurement cycles to cover $\Delta\mu$ with FWHM. For $t_{0,min} = 10$ GHz, the minimal quantum capacitance $C_{q0}/2 \sim 19$ aF. For $t_{0,max} = 100$ GHz, $C_{q0} \sim 3.8$ aF, but the FWHM is much wider and it can be integrated over several measurement cycles.

In the last part of this section, we discuss the charge pulsing for Pauli spin blockade spin-to-charge conversion (results shown in fig. S5E and S5F) and assume long spin lifetimes. The first step involves turning on t_0 . This step is limited in speed due to the small direct coupling between the $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ components. Next, we apply a linear detuning pulse to shift the lower

energy eigenstate, e.g. the $|\uparrow\downarrow\rangle$ -like state, into $S(0,2)$. Now, because of the small direct coupling, the pulsing speed is mainly limited by t_0 rather than the Zeeman energy difference. Finally, we turn off t_0 . This step can also be fast as there is no other states close to $S(0,2)$ at positive detuning. A high-fidelity 3 MHz Pauli spin blockade spin-to-charge conversion rate is achieved by turning on t_0 in 200 ns, followed by linear sweeping the detuning in 110 ns, and turning off t_0 in 20 ns. Figure S5E shows the operation fidelity as a function of t_0 and ε , where the black contour line denote the region where the fidelity is beyond 99.9%.

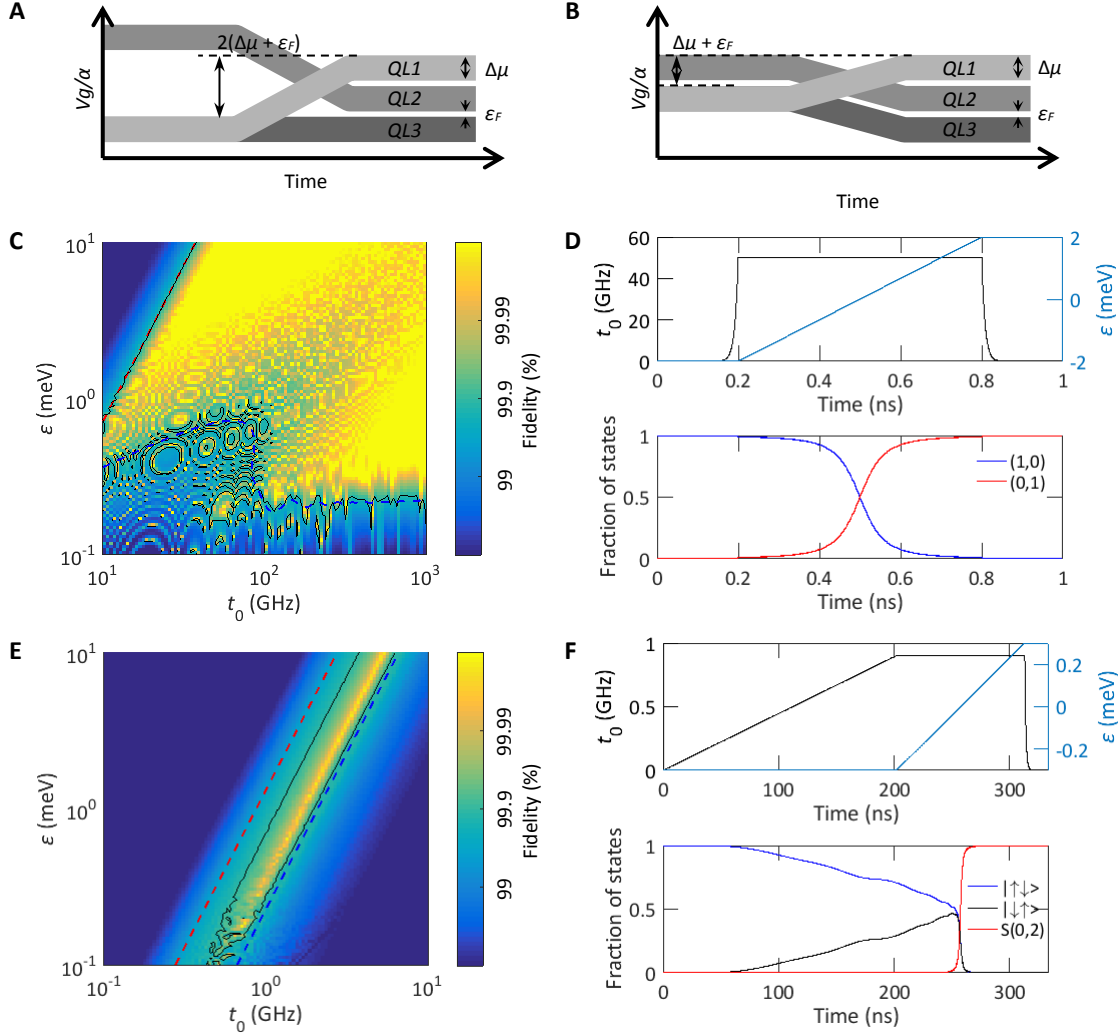


Fig. S5. Charge shuttling process. (A) and (B) Qubit shuttling scheme for the same operation as Fig. 3B or C in the main text. (A) $QL1$ and $QL2$ on top of the shuttling sites are pulsed with a larger amplitude such that the qubit under $QL3$ is not affected. (B) $QL3$ is also pulsed to compensate the reduced detuning, and the pulsing amplitude is reduced. (C) The fidelity of a three-step shuttling with an operation speed of 1 GHz. The first step is turning on the inter-dot tunnel coupling from 1 Hz to t_0 in 0.2 ns. The second step is a linear sweep of the detuning from $-\varepsilon$ to ε in 0.6 ns. The last step involves turning off the tunnel coupling in 0.2 ns. Here, we consider a fast, linear control voltage on the barrier gate and approximate the tunneling rate change by an exponential scale. (D) An example shuttling process with $t_0 = 50$ GHz and $\varepsilon = 2$ meV. (E) The fidelity of a three-step Pauli spin blockade spin-to-charge conversion with an operation speed of 3 MHz. The first step is turning on the inter-dot tunnel coupling linearly from 1 Hz to t_0 in ~ 20 ns. The second step is linearly sweep the detuning from $-\varepsilon$ to ε in ~ 110 ns. The last step is turning off the tunnel coupling back to 1 Hz in ~ 20 ns. (F) An example Pauli spin blockade process with $t_0 = 1$ GHz and $\varepsilon = 0.3$ meV, where $|\uparrow\downarrow\rangle$ has a lower energy. In (C) and (E), black contour lines denote the 99.9% fidelity-threshold, red dashed lines correspond to a non-adiabatic probability of 10^{-3} , and the blue dashed contour lines denote the 99.9% fidelity-threshold during tunnel coupling control.

Section S6. Pauli spin blockade spin to charge conversion with ancillary qubit in the spin up state

In this section, we explain how to implement Pauli spin blockade readout with the ancillary qubit in the spin up state. This complements the protocol with the ancillary qubit in the spin down state, as described in the main text and visualized in Fig. 4B. In this protocol, the spin up ancillary qubit is located in the column with the smaller magnetic field (red column) at the starting of the Pauli spin blockade process. Consequently, the qubit for readout is pulsed to the ancillary qubit site, as shown in fig. S6A denoted by step (i). We note that an alternative sequence is possible as well, obtained by reversing the pulsing direction. As shown in fig. S6B, if the target qubit is in the spin up state, it will remain in the (1,1) charge state. In contrast, if the state is spin down, it will move to the $S(0,2)$ state. Consequently, the charge occupation can be readout with the gate-based dispersive readout as described in the main text.

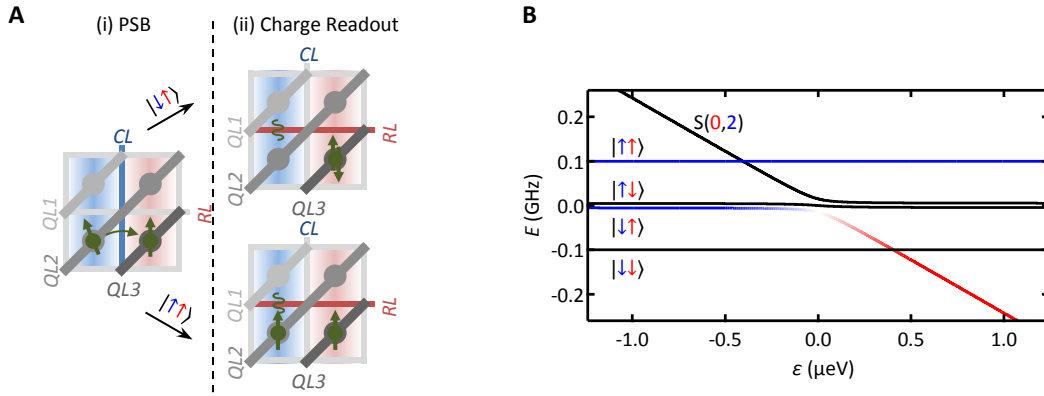


Fig. S6. Scheme for Pauli spin blockade spin to charge conversion with ancillary qubit in the spin up state. (A) Schematic for (i) the Pauli spin blockade (PSB) and for (ii) the charge readout process. **(B)** Double dot energy diagram. The ancillary qubit with smaller Zeeman energy (on the red column) is tuned to the up state. The readout state are the spin states of the qubit with the larger Zeeman energy. By increasing adiabatically the detuning beyond the anticrossing location, the individual spin states are projected to a singlet (denoted by the blue to red line) or triplet state (blue line). The resulting difference in charge state can be consequently measured as described in the main text.

Section S7. Shuttling bus for 2D array module

In this section, we discuss briefly a direction how to shuttle qubits between different qubit modules as a means towards truly large-scale quantum computation. The architecture is shown in fig. S7. With this approach it is possible to shuttle individual or complete arrays of qubits using the shuttling gates. In this implementation, the shuttling gates have the same geometry as the *CLs* and *RLs*. Figure S7(i) denotes the starting position of the qubits for a possible shuttle sequence. In fig. S7(ii), we lower the tunnel barriers in the forward direction of the qubits with label 1. Next, we raise the tunnel barriers in the backward direction, and hence they are shuttled forward. In fig. S7(iii), we repeat the same process as step (ii), but now for the qubits with label 2. After these steps, the qubit array is shuttled forward by one dot site as shown in fig. S7(iv). In this scheme, the shuttling gates are grouped in four, where the gates in each group perform similar operations. Consequently, all the gates belonging to the same group can be connected together to further reduce the number of wires interfacing to external electronics.

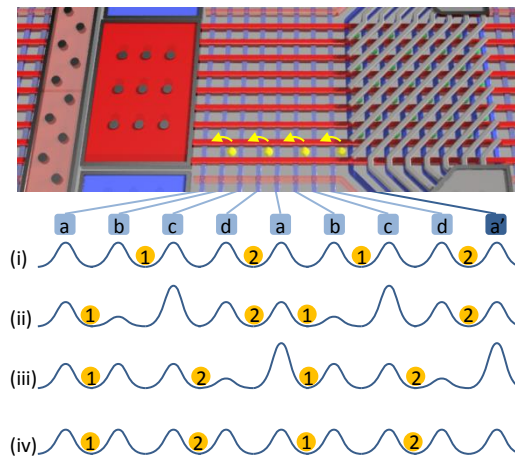


Fig. S7. Connecting qubit modules. The shuttling highways need fewer gate lines as compared to the qubit module. This limits functionality but does allow to shuttle qubits between different qubit modules. In addition, the gate lines can be grouped, such that space becomes available for interconnects or local electronics. The bottom section of this figure schematically shows a particular shuttle scheme, where a column of qubits is shuttled by one dot site by advancing from steps (i) to (iv).