The *spider web* array–a sparse spin qubit array

Jelmer M. Boter, ^{1,2} Juan P. Dehollain, ^{1,2,3} Jeroen P. G. van Dijk, ^{1,2} Toivo Hensgens, ^{1,2} Richard Versluis, ^{1,4} James S. Clarke, ⁵ Menno Veldhorst, ^{1,2} Fabio Sebastiano, ¹ and Lieven M. K. Vandersypen ^{1,2,5,*}

¹ QuTech, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
² Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
³ School of Mathematical and Physical Sciences, University of Technology Sydney, Ultimo NSW 2007, Australia
⁴ Netherlands Organisation for Applied Scientific Research (TNO),

P.O. Box 155, 2600 AD Delft, The Netherlands

⁵ Components Research, Intel Corporation, 2501 NE Century Blvd, Hillsboro, OR 97124, USA (Dated: 2020-01-10)

One of the main bottlenecks in the pursuit of a large-scale quantum computer is the large number of control signals needed to operate qubit systems, which become unmanageable at the off-chip level as the system size scales up. Here, we discuss a quantum-dot, spin-qubit architecture that integrates on-chip control electronics, allowing for a significant reduction in the number of connections at the chip boundary. By arranging the qubits in a two-dimensional array with $\sim 12~\mu m$ separation, we create space to implement a locally integrated sample and hold circuit. This allows to offset the inhomogeneities in the potential landscape across the array and to share the majority of the control signals for qubit operations across the array. This work presents a novel and complementary approach to previously proposed architectures, focusing on a feasible approach to integrating quantum and classical hardware, and significantly closing the gap towards a fully CMOS compatible quantum computer implementation.

Semiconductor quantum dots [1], particularly in silicon [2], are attractive hosts for spin qubits in large-scale quantum computation applications, because of their assumed compatibility with conventional CMOS integration processes. The last several years have seen significant progress in spin qubit research that resulted in the demonstration of long coherence times [3], high-fidelity single- [3–5] and two-qubit gates [6, 7], quantum algorithms [8], quantum non-demolition measurements [9, 10], strong spin-photon coupling [11–13] and long distance spin-spin coupling [14].

GENERAL INTRO:

- Challenges of scaling up
- Previously proposed architectures with pros and cons
- Include reference to Alexeev preprint

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend con-

sequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Here, we present the spider-web array, a sparse twodimensional spin qubit array with single qubit nodes separated by gate-based shuttling channels. Classical control electronics is integrated locally with the quantum hardware in order to minimize the need for off-chip interconnects, resulting in a scalable Rent's exponent [15]. We first describe in Section the gubit array structure and components that allow for the implementation of the basic single- and two-qubit operations. In Sec. we describe the integration of on-chip classical electronics at both the local qubit level and outside the qubit plane, aimed at exponentially reducing the number of control lines between qubit plane and the chip boundaries, while providing local DC biasing to mitigate inhomogeneity and effectively operate the gubits in the array with a surface code implementation [16]. Section describes in detail the implementation of the surface code cycle as well as how logical



operations are performed via lattice surgery. Finally, we discuss how the locally integrated electronics reduces the number of connections at the quantum plane boundary and the corresponding footprint in Sections and .

ARRAY DESIGN

The overarching concept of the spider-web array is presented in Figure 2. A large two-dimensional square-lattice of spin-qubits constitutes the the quantum plane. The spin-qubits consist of single electrons confined in electrostatically defined quantum dots in silicon. To implement the surface code [17], we concatenate half of the qubits in the array as data qubits with the other half as ancilla qubits, and allow single-qubit operations as well as nearest-neighbour two-qubit operations. In contrast to most other spin-qubit architectures, this array is sparse, with a qubit pitch $d \sim 12~\mu m$. This has two major implications:

- 1. It requires a non-conventional qubit operation protocol, which we implement by shuttling the qubits to interaction regions—where single- and two-qubit operations are performed along with readout.
- 2. It facilitates the local integration of classical control electronics, consisting here of sample-and-hold circuits that provide independent DC biasing of each gate electrode, which effectively offsets inhomogeneities in the potential landscape across the array. Consequently, the array can be considered fully homogeneous allowing the majority of qubit control signals to be shared across the entire array, which significantly reduces the number of control lines at the quantum plane boundary.

The quantum computing architecture we propose here consists of a two-dimensional array of electron-spin qubits. Linear arrays of gate electrodes are arranged to form a square lattice of electrostatically-defined, nearestneighbour-coupled semiconductor quantum dots. While generally each quantum dot hosts a qubit, the design we present here uses a sparse qubit array with a qubit pitch of $\sim 12 \, \mu m$. Electron shuttling channels connect the vertices of the array to transport electrons to and from interactions regions halfway the shuttling arms. The sparseness of the array opens up spaces locally and an open spider web structure is created. This facilitates the local integration of sample-and-hold circuits and thereby independent DC biasing of gate electrodes enables offsetting inhomogeneities in the potential landscape across the array. Consequently, the array can be considered fully homogeneous henceforth and the majority of control signals for qubits operations can be shared across the entire array, which significantly reduces the number of control lines at the quantum plane boundary.

Figure 1 gives a schematic overview of the proposed spider web architecture. A unit cell (Fig. 1(a)), the smallest operational unit, contains four qubits separated by the qubit pitch d, that is mainly determined by footprint constraints, as will be discussed later. The unit cell has an area of $(2d)^2$ and a perimeter of 8d. A module (Fig. 1(b)) is formed from $N \times N$ unit cells and has an area and a perimeter of $(2dN)^2$ and 8dN, respectively. Within a module, DC biasing and readout takes place sequentially, while all operations that are part of the surface code cycle occur simultaneously in all unit cells in the module. To scale the number of qubits beyond the size of a module, the quantum plane (Fig. 1(c)) consist of $M \times M$ modules, covering an area of $(2dNM)^2$ and having a perimeter of 8dNM, for a total of $4M^2N^2$ qubits. Control of the different modules in the quantum plane is fully parallel for DC biasing, the surface code cycle as well as readout. What determines N (the module size) and M will be discussed later. The area outside of the quantum plane can be used to reduce the off-chip wire count even further.

Figure 2(a) shows the unit cell, containing four qubits along with all the elements required to operate them. Qubits are separated by the qubit pitch d and remain at the vertices of the lattice (Fig. 2(b)) while idle, where they are surrounded by barrier gates (in red) that control both the confinement potential that keeps the electrons at the vertex, as well as the tunnelling in and out of the shuttling channels. The qubits are shuttled to the operations regions between the vertices (Figs. 2(c,d)) in order to perform single- and two-qubit operations, as well as readout and initialization. Initialization of the array is done by loading electrons from the reservoirs in the operations region shown in Fig. 2(c) and shuttling them to the vertices. Single-qubit gates are performed in the same operations region via electric dipole spin resonance (EDSR) is a transverse magnetic field gradient provided by a pair of micromagnets [18]. Two-qubit gates are performed in the operations regions shown in Figs. 2(c,d) and rely on the exchange interaction between electron spins to perform a \sqrt{SWAP} operation [1, 19, 20]. Finally, qubit readout is performed in the operations region shown in Fig. 2(c), where a readout quantum dot connected to source/drain ohmic contacts is used for charge sensing and spin readout is achieved via spin-to-charge conversion based on Pauli spin blockade [2]. Two-qubit gates have to be performed between all qubit pairs and therefore a two-qubit gate can be performed in all operations regions. Qubits are able to share an operations region to perform single-qubit gates and only ancilla qubits have to be read out, so not all operations regions require these capabilities. Therefore, as shown in Fig. 2(a), a unit cell contains four qubit idle regions (Fig. 2(b)), two regions to perform single- and two-qubit gates as well as initialization and readout (Fig. 2(c)), and six regions to only perform two-qubit gates (Fig. 2(d)).

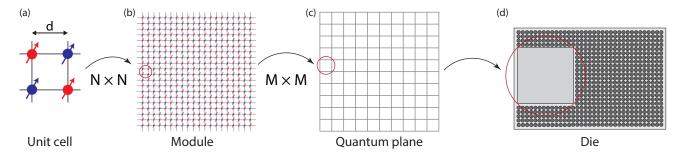


FIG. 1. Overview of the *spider web* architecture with a schematic breakdown of its components, as described in the main text. (a) A unit cell contains four qubits that are separated by the qubit pitch d. Qubits are colour coded to distinguish data qubits (blue) and ancilla qubits (red), as defined in the surface code. (b) A module constitutes of N×N unit cells, and in turn M×M modules together form (c) the quantum plane. Generally, the quantum plane occupies only part of (d) the die and the remaining area can be used to further reduce the off-chip wire count.

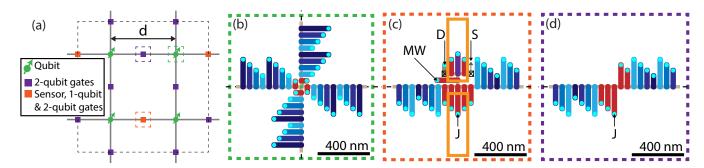


FIG. 2. (a) Schematic (not to scale) of a unit cell containing four spin qubits (green), operations regions (purple/orange), connected via shuttling channels (grey lines). (b) Qubit idling region. Four barrier gate electrodes (red) define the confinement potential and allow qubits into the shuttling channels (defined by the blue gate electrodes). Cyan circles represent vias to higher interconnect layers. (c) Qubit operations region including control gates (red), sensing dot plunger (purple), source (S)/drain (D) ohmic contacts (squares) and micromagnets (orange rectangles). The gate electrodes labelled MW and J are used for single- and two-qubit operations (see main text), respectively. (d) Two-qubit operation only regions.

ELECTRONICS

In this section we discuss the (local) electronics that is needed for DC biasing and for supplying the control signals for qubit operations, and to perform qubit readout.

DC biasing

To mitigate inhomogeneity in the qubit lattice, we design a sample-and-hold scheme to apply a local DC bias to the gate electrodes. Based on the different gates functionalities, we define two bias voltage resolutions. For gates acting as barriers to shuttling channels only a resolution sufficient to maintain an electron in a quantum dot is required and therefore we can afford a coarse resolution of 1 mV. A resolution of 1 μ V is required for all other plunger and barrier gates [21]. To achieve the coarse resolution a minimum hold capacitance of ~ 0.16 fF is required (limited by the electron charge $e/\Delta V$), while ~ 14 pF (limited by thermal noise $k_B T/V^2$, assuming power dissipation from the local electronics raises the operating

temperature to 1 K) is required for the fine resolution. In the qubit idling region (Fig. 2(b)) the four barrier gates only require coarse resolution DC biasing. In the qubit operations region shown in Fig. 2(c), a coarse resolution is sufficient for the two outer gates in the shuttling channel, while the seven other gates require a fine resolution. In the two-qubit operations region (Fig. 2(d)), the three inner gates require a fine resolution, while for the two outer gates a coarse resolution is sufficient. These requirements are summarized in Table I. A total of 64 control gates per unit cell require DC biasing, out of which 32 require fine resolution, while for the other 32 a coarse resolution is sufficient. The need for DC biasing of the shuttling gates is eliminated by making the travelling wave potential large enough to overcome potential landscape inhomogeneities.

The DC biasing scheme is schematically depicted in Fig. 3. The open area in between the qubits and shuttling channels created by the sparseness of the *spider web* array allows for local integration of demultiplexers and capacitors that together form sample-and-hold circuits. Local demultiplexers (four per unit cell) each distribute

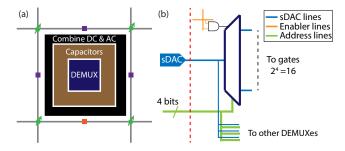


FIG. 3. (a) Schematic of a unit cell containing locally integrated classical electronics. Demultiplexers in combination with capacitors form sample-and-hold-circuits to provide DC bias voltages. Additionally, circuits that combine the DC bias with AC control signals are required. (b) Input/output schematic of the demultiplexers. A demultiplexer (dark blue), once enabled via crossbar addressing (orange), ports the DC voltages coming from the sDAC (light blue) to the output selected by the 4-bit address bus (green). The dashed red line indicates the quantum plane boundary. The colour coding in (a) and the legend in (b) represent the same components in both panels as well as in Fig.4.

DC voltages generated remotely (i.e. outside the quantum plane) to 16 local capacitors connected to the gate electrodes, by selecting one of their sixteen outputs based on a 4-bit address.

Figure 4 shows the DC biasing scheme on the unit cell, module and quantum plane levels. All demultiplexers within a module share the same input DC biasing signal (Fig. 4(b)), and all demultiplexers in the quantum plane share the same address bus (Fig. 4(c)). The demultiplexers in a module are enabled sequentially by crossbar addressing and in turn sequentially update each gate. This way, all modules are updated in parallel and therefore one module refresh cycle is required to refresh the entire qubit array. The number of gates that can be updated sequentially will be limited by the required gate voltage refresh rate, that will be set by the current leakage of the hold circuit and the time required to update each gate. This, in turn, will set the module size (i.e. the number of unit cells N^2 per module).

	# per unit cell	Fine $(1 \mu V)$	Coarse (1 mV)	Pulsed gates
Qubit idling region	$4\times$	-	4	4
All-operations region	$2\times$	7	2	6
Two-qubit operations	6×	3	2	5
region				
Total per unit cell		32	32	58

TABLE I. Summary of the gates per unit cell requiring DC biasing and their resolution, and the number of pulsed signals.

Signals for qubit operations

To perform single- and two-qubit operations, as well as readout and initialization, the qubits are shuttled to the operations regions in Figs. 2(c,d), where the operations are performed by applying pulsed signals to the appropriate gates.

Linear arrays of gate electrodes (blue gates in Fig. 2) define shuttling channels and a travelling wave potential can trap and shuttle an electron along these channels. Four phase-shifted sinusoidal signals that are applied to four consecutive gates (different shades of blue in Fig. 2) and reused every fourth gate, are used to generate the travelling wave potential. Electrons can be forced to tunnel into a shuttling channel by using a barrier gate (Fig. 2(b)). The shuttling signals are always on and the phase shifts control the direction of shuttling. The traveling wave potential is made large enough to overcome the inhomogeneities in the potential landscape, eliminating the need to apply DC biasing on the shuttling gates. Electrons will travel the distance corresponding to four gate electrode in one period of the shuttling signal, so a 1 GHZ shuttling signal applied to the gates makes the electrons travel the distance corresponding to four gates per nanosecond. Assuming a gate pitch of 50 nm, the electrons will be shuttled with a velocity of 200 nm/ns and a distance of 10 µm takes 50 ns.

For single-qubit operations, a microwave pulse is applied to the control gate labelled MW in Fig. 2(c) in order to perform EDSR in the magnetic field gradient provided by a pair of micromagnets. A pulsed signal to the gates labelled J in Figs. 2(c,d), that activates an exchange interaction between electrons underneath the adjacent gates, is used to perform a two-qubit operation. Figure 5 schematically shows how AC signals for qubit operations are combined with DC biasing by employing a complementary switching circuit (see φ_{AC} and $\overline{\varphi_{AC}}$).

All red gates in Fig. 2(b,d), and the five bottom gates and the gate labelled MW in Fig. 2(c) require pulsed signals. This totals to 58 gates per unit cell requiring a pulsed signal, as summarised in the rightmost column of Table I.

Readout

Qubit readout is performed at the operations region shown in Fig. 2(c). Charge sensing is done with a readout quantum dot connected to source/drain ohmic contacts and spin-to-charge conversion based on Pauli spin blockade is employed for spin readout [2]. Additionally, the array is initialised by shuttling electrons that are provided by the ohmics in this operations region to the unit cell vertices.

Readout is performed sequentially across the unit cells of each module, while the modules are read out in paral-

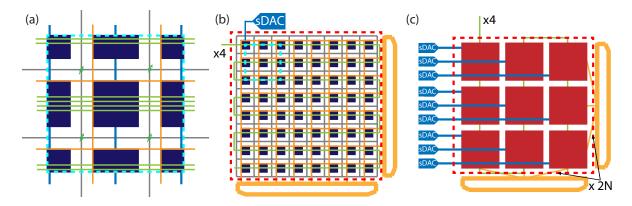


FIG. 4. Schematics of (e) a unit cell and (d) a module. Demultiplexers are sequentially enabled by crossbar addressing controlled by multiplexers (orange blocks). (f) Schematic of the array of modules completing the quantum plane. Dashed red lines in (d) and (f) denote the quantum plane boundary.

lel. This is done by connecting the drain contacts of all sensor dots in a module to a single line at the quantum plane boundary and consecutively pulsing their plungers to bring them to the low-impedance, electrostatically sensitive regime, while all other sensor dots in the module are in Coulomb blockade (i.e. in the high-impedance regime). A global readout demultiplexer is used for the sequential control of the sensor plungers in a module. This demultiplexer can be shared between all modules across the entire array.

Since qubit readout is the most time consuming operation and is performed sequentially within a module, it will most likely limit the surface code cycle rate as well as the ultimate size of the array. Instead of using the same module size for both DC biasing and readout, we define readout modules that consist of N_{RO}^2 unit cells (with $N_{RO} \leq N$). A total of M_{RO}^2 readout modules make

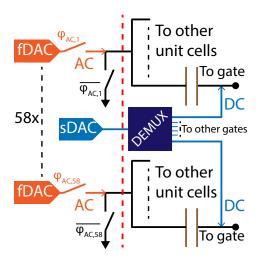


FIG. 5. Circuit schematic to combine AC and DC signals as described in the main text. fDAC (sDAC) are voltage sources for pulsed signals (DC biasing). The dashed red line indicates the quantum plane boundary.

up the full quantum plane (note: $NM = N_{RO}M_{RO}$).

Additionally, an extra level of parallel readout can be implemented by means of for example amplitude or frequency modulation. Amplitude multiplexing can be achieved by tuning the DC bias of the sensor plungers that are read out at the same time to voltages that result in distinct amplitudes for the sensor response. It is possible to use frequency multiplexing by using resonant circuits with different resonance frequencies. Without going into details of the multiplexing strategy, we assume in the following four sensors to be read out simultaneously. This does reduces the number of required sequential readouts and also slightly the number of lines. The latter, however, does not make a significant difference.

SURFACE CODE OPERATION

The proposed array sustains the surface code by using a cyclic sequence of pulsed signals within a unit cell, with the same sequence performed in parallel across all unit cells in the entire array. This is facilitated by the local DC biases, which allow to use the same control signals in all unit cells. A total of 58 remote pulsed voltage sources are used to generate the required cyclic pulsed signals for all quantum control gates (i.e., one source per pulsed gate in a unit cell).

Figure 6 depicts the surface code cycle. Figure 6(a) shows the circuit diagram for a single surface code cycle and is the 4-qubit analogue of the circuit shown in Ref. [22]. For the present architecture only four qubits are required per unit cell, because the exchange coupling is not always present as is the case in superconducting qubits. Steps 1–10 consist of either single- or two-qubit operations for which one or two of the qubits are shuttled from their idle position to an interactions region to undergo either a single- or a two-qubit gate. At the end of the cycle, the ancilla qubits are measured. As an example, Figs. 6(b,c) depict the qubit movement in steps 1

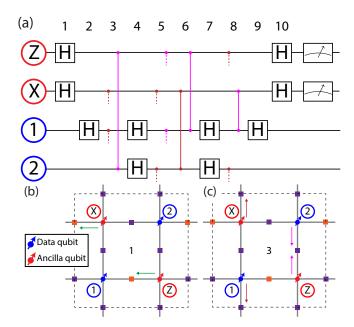


FIG. 6. Surface code operation of the spider web array. (a) Circuit diagram showing the surface code cycle for the four-qubit unit cell. H-blocks represent a Hadamard gate, solid lines represent a intra-unit cell two-qubit gates and dashed lines represent inter-unit cell two-qubit gates. (b,c) Schematics showing the qubit movement in step 1 (b) and step 3 (c) of the surface code cycle showing an example of single- and two-qubit gates respectively.

and 3, corresponding to single- and two-qubit gates, respectively. Assuming both single- and two-qubit gates to take 100 ns to perform, a 1 GHz shuttling signal and 240 gates per lattice arm (see Sec.) each step of the surface code cycle takes 160 ns.

Logic gates in the surface code with lattice surgery are achieved by creating defects in the lattice. We implement these defects by preventing shuttling of a subset of data qubits via locally integrated switches. The signals used to control these switches are arranged in a crossbar fashion across the entire quantum plane. In practice, we propose to use x crossbars over the entire array, in order to allow for x defects to be simultaneously created and manipulated. By activating a certain switch, the corresponding data qubit is not shuttled over to the interactions region. At the same time, the ancilla qubit with which the data qubit was supposed to have a two-qubit interaction is still shuttled over, but since the data qubit is not present, the two-qubit gate does not take place.

LINE SCALING

The routing scheme for quantum control signals that we have described is summarized in Tab. II. The outlined DC biasing scheme allows for significant sharing of control lines, which results in a very efficient scaling of ratio of the number of interconnects at the quantum plane boundary to the number of lines at the unit cell level. The total number of gate electrodes at the unit cell level scales with number of qubits $(4M^2N^2)$ and we will now discuss the scaling of the number of connections at the quantum plane boundary allowed by the operation schemes described before. The line scaling at different levels of the spin qubit array for the control schemes described before is summarized in Table III. A Rent's exponent as low as p=0.5 is obtained at the quantum plane boundary.

Independent DC biasing of the sparse array by means of sample-and-hold circuits is provided with $O(M^2 + N)$ lines at the quantum plane boundary. Concretely, at the unit cell level four digital address lines and four enabler line (two horizontal and two vertical) select a specific output of one demultiplexer, which is set to the correct voltage via a single connection to a voltage source. This makes a total of nines line required for DC biasing, which all have to enter and leave the unit cell. At the module level only the enabler lines scale with the number of unit cells as 4N, that enter and leave the module. Meanwhile, the four digital address lines and the connection to the voltage source is shared between all unit cells in a module. The digital address lines are common with other modules, so also have to leave the module. Every module has its own voltage source, so at the quantum plane level M^2 connections to voltage sources are needed, while both the digital address lines and the enabler lines are shared between all modules.

Shuttling of electrons across the array only requires four signals that can be fully shared over the entire array. At the unit cell and module level these connection have to enter and leave, so they count double.

A constant number of 58 lines (see Tab. I) at the quantum plane boundary (irrespective of the number of qubits) suffices to sustain the surface code by delivering all pulsed and microwave control signals across the entire chip. These lines are counted double at the unit cell and module level, as they enter and leave.

Switches are used to deactivate data qubits in order to implement logical operations in the surface code. The signals to control those switches are arranged in a crossbar fashion over the entire quantum plane with as few as

Shuttling gates	Source \rightarrow gate
(blue)	
Pulsed gates	DC: source \rightarrow local demultiplexer \rightarrow gate
(red)	AC: source \rightarrow gate
Sensing dot	DC: source \rightarrow local demultiplexer \rightarrow gate
plunger (purple)	AC: source \rightarrow global demultiplexer \rightarrow gate
Drain contacts	$Measurement device \leftarrow ohmic$

TABLE II. Summary of signal routing for the four different type of control lines in the array design.

	Unit cell	Module	Quantum plane
DC biasing	18	8N+9	$M^2 + 4N + 4$
Shuttling	8	8	4
Pulsed signals	116	116	58
$\&~\mathrm{MW}$			
Logical	8x	8Nx	4NMx
operations			
Readout	4	$4\log_2 N_{RO} - 1$	M_{RO}^2+
			$2\log_2 N_{RO} - 1$
	146+8x	8N(1+x)+	$M^2 + M_{RO}^2$
Total		$4\log_2 N_{RO} + 132$	+4N(1+Mx)
			$+2\log_2 N_{RO} + 65$

TABLE III. Line scaling at different levels of the array. At the unit cell and module level most lines are counted double, since they have to enter and leave through the perimeter.

O(xMN) lines. A total of x defects can be created and manipulated simultaneously by x parallel crossbars over the entire array. For every crossbar, two horizontal and two vertical lines have to enter and leave a unit cell and this scales with N and NM at the module and quantum plane level, respectively. The lines enter and leave at the unit cell and module level, so are counted double.

Decoding is used to address the sensor plungers in a readout module for the sequential readout scheme and it thereby obtains a line scaling at the quantum plane boundary as $O(M_{RO}^2 + \log(N_{RO}))$. A unit cell contains two readout SETs that both require a connection to their plunger and share a single ohmic line that has to enter and leave the unit cell. At the module level a single ohmic line is required to enter, and $2\log_2 N_{RO} - 1$ lines enter and leave the module to address the readout demultiplexers, assuming four plungers to be connected to a single line for multiplexed readout. To read out the full quantum plane, all M_{RO}^2 modules require their own ohmic line, while the address lines for the readout demultiplexer are shared.

The wire count at different levels of the array does depend on the total number of qubits, but also on the exact choice of the module size N and the corresponding value for M (for a given total number of qubits), as well as the number of parallel crossbars for lattice surgery. Assuming a total of 2^{20} ($\approx 10^6$) qubits to make a concrete assessment of the number of lines, a module size of N = $2^6 = 64$ requires $M = 2^3 = 8$ modules. For the sake of this example we assume x = 32 parallel crossbars. The number of connections for this choice of N, M and xis summarized in Tab. IV. With a qubit pitch $d=12~\mu\mathrm{m}$ (see next section) and these numbers for N and M, a unit cell, module and the quantum plane have a perimeter of $8d \approx 96 \mu m$, $8dN \approx 12 mm$ and $8dNM \approx 49 mm$, respectively. The line density is highest at the unit cell level where $O(10^2 - 10^3)$ wires pass through the unit cell

	Unit cell	Module	Quantum plane
DC biasing	18	8N+9	$M^2 + 4N + 4$
Shuttling	8	8	4
Pulsed signals	116	116	58
$\&~\mathrm{MW}$			
Logical	8x	8Nx	4NMx
operations			
Readout	4	$4\log_2 N_{RO} - 1$	M_{RO}^2+
			$2\log_2 N_{RO} - 1$
	146+8x	8N(1+x)+	$M^2 + M_{RO}^2$
Total		$\left 4\log_2 N_{RO} + 132\right $	+4N(1+Mx)
			$\left +2\log_2 N_{RO} + 65 \right $

TABLE IV. TO BE UPDATED - Number of connections at different levels of the array and for different purposes for N = 64, M = 8 and x = 32.

perimeter, using multiple interconnect layers.

FOOTPRINT

Next, we make an assessment of the footprint of the classical control electronics we have discussed before. Since this electronics has to be integrated locally, the corresponding footprint set the qubit pitch d. The most significant contribution to the footprint comes from the capacitors that are required for the sample-and-hold scheme. As summarized in Tab. I, a total of 32 gate electrodes per unit cell require a fine voltage resolution and for another 32 gate electrodes a coarse voltage resolution is sufficient. The total capacitance per unit cell therefore is \sim 450 pF. State-of-the-art deep-trench capacitor technology achieves $\sim 1 \text{ pF/}\mu\text{m}^2$ [23], so we estimate a total capacitor footprint of $\sim 450 \ \mu m^2$. In order to estimate the footprint for the demultiplexers used for DC biasing and readout, we extrapolate to 28-nm technology from a modelled demultiplexer circuit in 40-nm technology. We obtain an estimate of the total footprint of the DC biasing and readout demultiplexers of $\sim 60 \ \mu m^2$ per unit cell. The total footprint of the classical control electronics per unit cell therefore is $\sim 510 \ \mu m^2$. This allows to set the qubit pitch to d≥12 μm and results in a unit cell area of $4d^2 = 576 \ \mu m^2$. With a 50-nm gate pitch, each of the lattice arms would comprise 240 gate electrodes.

For a concrete assessment of the footprint of the array design discussed here, let us again consider a total of $(2MN)^2 = 2^{20} \ (\approx 10^6)$ qubits. A total area of $(2dNM)^2 \approx 151 \ \mathrm{mm^2}$ is covered by the quantum plane. The remaining area on, for example, a typical 22 mm \times 33 mm (726 mm²) die is $\sim 575 \ \mathrm{mm^2}$, and can be used to implement classical control circuits, i.e. among others the pulsed voltage sources we have described. In addition, additional levels of multiplexing can be employed

to bring the off-chip wire count, typically being the real bottleneck for Rent's rule, to well below the wire count at the quantum plane boundary.

We now consider the footprint requirements of the control electronics that need to be locally integrated in the quantum plane, which sets the qubit pitch d. The bulk of the footprint will be taken up by the capacitors required for the sample-and-hold scheme. Coarse resolution is required for 32 gates and another 32 gates require fine resolution, which comprise a total capacitance per unit cell of ~ 450 pF. Assuming ~ 1 pF/ μ m² (using state-of-theart deep-trench capacitor technology [23]), we estimate a total capacitor footprint of $\sim 450 \ \mu \text{m}^2$. In addition, we modelled a demultiplexer circuit using 40-nm technology, extrapolated to 28-nm technology and obtained an estimate of the total footprint of the DC biasing and readout demultiplexers of $\sim 60 \ \mu \text{m}^2$ per unit cell. This adds to a total footprint per unit cell of $\sim 510 \ \mu \text{m}^2$, which allows to set the qubit pitch to d \geq 12 μ m. Assuming a 50 nm pitch between gate electrodes, this would require linear arrays of 240 gate electrodes per lattice arm. A unit cell has an area of $4d^2 \approx 576 \,\mu\text{m}^2$.

In order to make a concrete assessment of the feasibility of the proposed design and scheme let us assume a total of $(2MN)^2 = 2^{20}$ qubits. The total area covered by the quantum plane is $(2dNM)^2 \sim 151 \text{ mm}^2$. The remaining area on a typical 22 mm \times 33 mm (726 mm²) die is $\sim 575 \text{ mm}^2$, and can be used to implement classical control circuits and to bring the wire count going off-chip, typically the real bottleneck for Rent's rule, to well below the wire count at the quantum plane boundary by means of additional levels of multiplexing.

CAPACITANCE AND HEAT LOAD?

- * To whom correspondence should be addressed: l.m.k.vandersypen@tudelft.nl
- D. Loss and D. P. DiVincenzo, Physical Review A 57, 120 (1998).
- [2] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Reviews of Modern Physics 85, 961 (2013).
- [3] 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, Nature Nanotechnology 9, 981 (2014).
- [4] E. Kawakami, T. Jullien, P. Scarlino, D. R. Ward, D. E. Savage, M. G. Lagally, V. V. Dobrovitski, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, PNAS 113, 11738 (2016).

- [5] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, Nature Nanotechnology 13, 102 (2018).
- [6] X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and L. M. K. Vandersypen, Physical Review X 9, 021011 (2019).
- [7] 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, Nature 569, 532 (2019).
- [8] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nature 555, 633 (2018).
- [9] J. Yoneda, K. Takeda, A. Noiri, T. Nakajima, S. Li, J. Kamioka, and T. Kodera, .
- [10] X. Xue, B. D'Anjou, T. F. Watson, D. R. Ward, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, W. A. Coish, and L. M. K. Vandersypen,
- [11] N. Samkharadze, G. Zheng, N. Kalhor, D. Brousse, A. Sammak, U. C. Mendes, A. Blais, G. Scappucci, and L. M. K. Vandersypen, Science 359, 1123 (2018).
- [12] X. Mi, M. Benito, S. Putz, D. M. Zajac, J. M. Taylor, G. Burkard, and J. R. Petta, Nature 555, 599 (2018).
- [13] A. J. Landig, J. V. Koski, P. Scarlino, U. C. Mendes, A. Blais, C. Reichl, W. Wegscheider, A. Wallraff, K. Ensslin, and T. Ihn, Nature 560, 179 (2018).
- [14] F. Borjans, X. G. Croot, X. Mi, M. J. Gullans, and J. R. Petta, Nature, 1 (2019).
- [15] D. P. Franke, J. S. Clarke, L. M. K. Vandersypen, and M. Veldhorst, Microprocessors and Microsystems 67, 1 (2019).
- [16] C. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, New Journal of Physics 14, 123011 (2012).
- [17] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. 43, 4452 (2002).
- [18] M. Pioro-Ladrière, Y. Tokura, T. Obata, T. Kubo, and S. Tarucha, Applied Physics Letters 90, 024105 (2007).
- [19] D. P. DiVincenzo, Fortschritte der Physik 48, 771 (2000).
- [20] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, 10th ed. (Cambridge University Press, Cambridge, 2010).
- [21] 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, npj Quantum Information 3, 34 (2017).
- [22] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, L. Di-Carlo, and A. Bruno, Physical Review Applied 8, 034021 (2017).
- [23] J. M. Park, Y. S. Hwang, S.-W. Kim, S. Y. Han, J. S. Park, J. Kim, J. W. Seo, B. S. Kim, S. H. Shin, C. H. Cho, S. W. Nam, H. S. Hong, K. P. Lee, G. Y. Jin, and E. S. Jung, in 2015 IEEE International Electron Devices Meeting (IEDM) (IEEE, New York, 2015) pp. 1–26.