# SNAQ: A Fast, Space-Efficient, and Practical Surface Code Architecture for Spin Qubits

Jason D. Chadwick, Willers Yang, Joshua Viszlai, and Frederic T. Chong
University of Chicago
Chicago, IL, USA

***Abstract*—Spin qubits in silicon quantum dot arrays are a promising quantum computation platform for long-term scalability due to their small qubit footprint and compatibility with advanced semiconductor manufacturing. However, spin qubit devices face a key architectural bottleneck: the large physical footprint of readout components relative to qubits prevents a dense layout where all qubits can be measured simultaneously, complicating the implementation of quantum error correction. This challenge is offset by the platform's unique rapid shuttling capability. In this work, we explore the design constraints and capabilities of spin qubits in silicon and propose the SNAQ (Shuttling-capable Narrow Array of spin Qubits) surface code architecture, which relaxes the 1:1 readout-to-qubit assumption by leveraging spin shuttling to time-multiplex ancilla qubit initialization and readout. Our analysis shows that, given sufficiently-high (experimentally demonstrated) qubit coherence times, this tradeoff achieves an orders-of-magnitude reduction in chip area per logical qubit. Additionally, using a denser grid of physical qubits, SNAQ enables fast transversal logic for short-distance logical operations, achieving 2.0-14.6× improvement in local logical clock speed while still supporting global operations via lattice surgery. This translates to a 60-64% reduction in spacetime cost of 15-to-1 magic state distillation, a key fault-tolerant subroutine. Our work demonstrates the architectural consequences of this new tradeoff, pinpoints critical hardware metrics, and provides a compelling path toward high-performance fault-tolerant computation on near-term-manufacturable spin qubit arrays.**

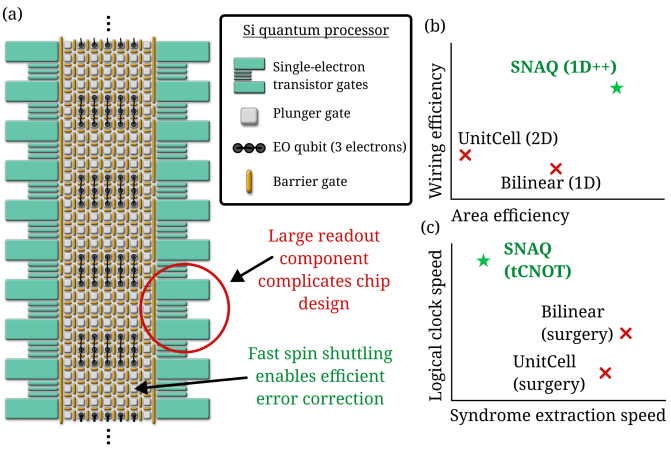


Fig. 1. (a) Depiction of a possible implementation of a narrow 7-dot-wide array of quantum dots, analogous to similar devices fabricated by Intel and HRL [27]–[29]. Electrons (black) are held in place at the plunger gates (light gray) and can interact with their neighbors (or shuttle to adjacent plungers) by manipulating the voltages on the barrier gates (orange). Single-electron transistors (green) allow for qubit readout on the two sides of the array. (b) SNAQ achieves similar performance to baselines while improving chip area efficiency and wiring efficiency. (c) SNAQ avoids the downside of a longer syndrome extraction cycle by enabling transversal CNOTs (tCNOTs), achieving significantly faster logical clock speeds compared to lattice-surgery-based approaches.

## I. Introduction

Scaling quantum computers is an immense architectural challenge, and it is not yet known which hardware modality (transmons, neutral atoms, trapped ions, etc.) will ultimately be able to scale up to the million-qubit processors we need to enable powerful applications [1]–[5]. Spin qubits in silicon quantum dot arrays are a promising candidate for large-scale quantum computing, primarily due to their small footprint (on the order of $100 \times 100$ nm per qubit) and their compatibility with existing advanced semiconductor fabrication techniques [6]–[12]. However, the very scalability that makes spin qubits attractive creates a fundamental architectural challenge. The physical footprint of readout components, such as a single-electron transistor (SET) or single-electron box (SEB), requires large reservoirs and ohmic contacts that are many times the size of a quantum dot [13]–[18]. This makes it difficult to build a dense array of dots while dedicating a unique readout component to each qubit, instead incentivizing asymmetric fixed-width arrays as shown in Figure 1(a). On the other hand, spin qubits also offer a unique and compelling capability: extremely fast spin shuttling between quantum dots [19], [20], which can enable novel shuttling-based hardware layouts that can overcome the readout bottleneck [21]–[26]. However, shuttling is not free. Qubits accumulate errors proportional to the distance shuttled, making it crucial to minimize shuttling distance when possible.

Previous spin qubit architecture proposals have addressed these challenges by considering sparser arrays, creating large interior spaces to fit the needed components, but at the cost of a large shuttling overhead required to enable the dense 2D qubit connectivity needed for error correction. In this work, instead of designing an entirely new hardware layout, we instead treat the fixed-width array as the fundamental building block, posing the central question: can we efficiently run the surface code in this constrained geometry by trading parallel measurement for a denser qubit array?

We propose a hardware-aware surface code architecture named SNAQ (Shuttling-capable Narrow Array of spin Qubits) that leverages fast spin shuttling to overcome the limited read-

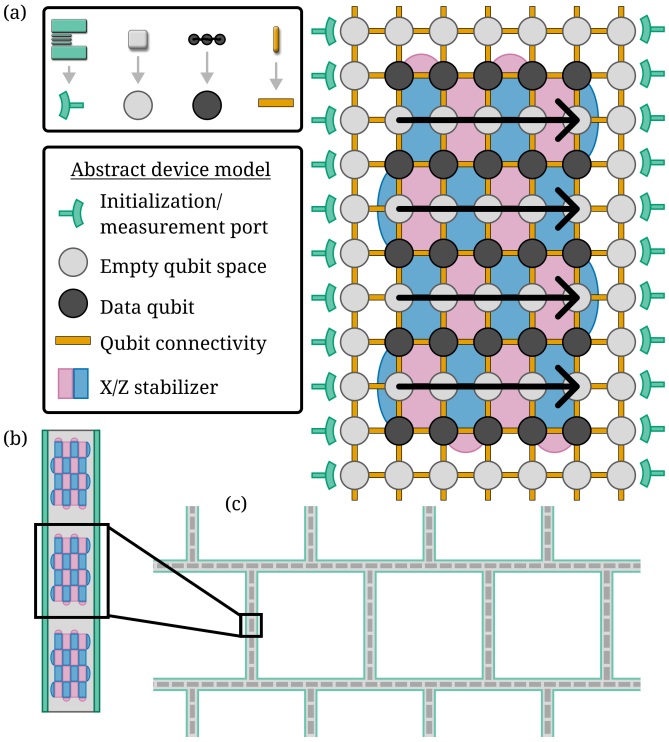


Fig. 2. Proposed SNAQ (Shuttling-capable Narrow Array of spin Qubits) architecture. (a) For ease of visualization, we translate from the specific chip components in Figure 1 to the abstract components we will use in the rest of the paper. Depicted is a layout of a distance-5 surface code patch on a 7-dot-wide array. Initialization and readout components (green) are only available on the edges of the array. Data qubits (dark gray) are interleaved with channels of empty dots (light gray) used to shuttle surface code ancilla qubits from edge to edge in the direction of the black arrows. Surface code X and Z stabilizers (blue and pink) are shown behind the array, supported on the data qubits. (b) A logical 1 × N layout of surface codes in a narrow array. (c) Possible design of a fully-scalable architecture consisting of loops of SNAQ channels, with large spaces in between to allow for integration of control wires and readout.

out capabilities in silicon quantum dot hardware. We show that for sufficiently low idle error rates demonstrated by existing spin qubit hardware, the initialization and measurement of physical qubits can be serialized with manageable impact on the logical performance of the surface code. This creates a significantly denser qubit array, which reduces the required amount of spin shuttling and enables the transversal CNOT for short-range logical multiqubit gates, delivering logical clock speed improvements and potentially improved parallelism.

This work makes the following main contributions:

- We identify and characterize the qubit-to-readout area mismatch as a key design challenge for scalable spin qubit architectures.
- We propose SNAQ, a novel surface code architecture that leverages fast spin shuttling to time-multiplex syndrome extraction and overcome this bottleneck in near-term-manufacturable arrays.
- We analyze the physical parameters for this new design with detailed circuit-level simulations, pinpointing the readout density $\rho$ and the qubit idle error rate $p_{\text{id}}$ as the most critical hardware metrics for its viability. Although the idle error restricts the lowest achievable logical error rate, our method shows competitive performance against baselines for algorithmically-relevant error rates.
- We demonstrate through simulation that SNAQ enables low-latency transversal CNOTs (tCNOTs) for local multiqubit gates, creating a two-tiered latency hierarchy between fast-local tCNOTs and slow-global lattice surgery.
- We quantify the benefits of the fast-local latency tier, demonstrating up to 2-14× speedup per operation and a 60-64% spacetime cost reduction for the 15-to-1 magic state distillation benchmark.

## II. Background

In this section, we explain the components and terminology of silicon quantum dot devices and provide some high-level background on quantum error correction and the surface code. We urge the reader to refer to Ref. [30] for a more thorough background on spin qubits and to Refs. [31], [32] for more information on the surface code.

### A. Spin qubits in silicon

Semiconductor spin qubits generally refer to nanoscale qubits encoded in the spin states of electrons or holes in a semiconductor substrate. In this work, we will specifically consider electrons held in quantum dots in silicon, which is a platform that has seen considerable recent progress [28], [30], [33]–[36]. Here we will describe the general terminology and architectural implications of the silicon spin qubit platform, leaving a specific discussion of state-of-the-art performance metrics to Section III. From an architectural perspective, these devices present a compelling substrate for large-scale fault-tolerant quantum computing: the qubit footprint is comparably small (dot pitches on the order of tens of nanometers), fabrication is compatible with CMOS semiconductor processes, and spin shuttling is a novel and powerful tool.

A silicon quantum dot array is defined by electrostatic gates, as depicted in Figure 1. In this work, we identify three important categories of these gates: “plunger” gates, each of which is tuned to create a quantum dot that can hold one electron; “barrier” gates, which control inter-dot tunnel coupling, used for shuttling and for electron-electron interaction via the exchange interaction; and the gates used to form a single-electron transistor (SET), which define large electron reservoirs. The simplest way to use a quantum dot array as a quantum processor is to treat individual electrons as qubits, where the two logical states are the spin states $|\uparrow\rangle$ and $|\downarrow\rangle$; alternative encodings can exploit up to four electrons per qubit to trade control complexity, magnetic gradient requirements, and coupling speed. Generally, increasing the number of electrons that encode each qubit simplifies the control requirements. In all qubit encodings, logical operations between qubits are mediated by the exchange interaction: by tuning the inter-dot tunnel barrier and detuning, one can dynamically adjust the exchange coupling and thereby implement controlled-phase,

CNOT-equivalent, or SWAP-like operations. The exchange-only (EO) qubit encoding uses three electrons to define each qubit, and is a promising choice for scalability because it relaxes the requirement for localized magnetic field control and instead can be operated by only tuning DC voltages on barriers and plungers [37], [38]. It may even be beneficial to use multiple different encodings in one architecture [39].

The architectural implications of silicon quantum processors diverge from more familiar transmon, neutral atom, or trapped-ion technologies in several key ways. First, readout of spin qubits typically uses spin-to-charge conversion, which tends to require a significantly larger component area than the qubit [13]–[18]. Designing the layout of a silicon spin processor is thus similar to the "pitch matching" exercise studied extensively in classical architecture, where physical components of different sizes must be aligned in a layout [40]–[44]. Because the qubit-qubit interactions are fundamentally constrained to be short-distance, the large readout size often results in architectures in which readout sensors are sparsely distributed or placed on the edges, and the internal qubit array must rely on qubit shuttling for readout access. Second, the capability of "shuttling" (moving an electron spin coherently between quantum dots) has a different cost model than qubit movement in neutral atoms or trapped ions. In spin qubits, the shuttling is very fast (a few nanoseconds per dot-to-dot transfer), but shuttling errors accumulate more rapidly with distance than in atomic systems.

Silicon spin qubit technology is a highly promising candidate for truly scalable quantum processors. At the same time, its distinct challenges and opportunities call for novel architectural ideas to make the best use of this hardware and enable efficient fault-tolerant computation.

### B. Quantum error correction and the surface code

Quantum error correction (QEC) is the process by which many physical qubits are used to encode a smaller number of logical qubits with much higher lifetimes and operating fidelities. A typical QEC code is encoded into a large number of *data qubits* and operates by repeatedly measuring a set of stabilizers of the code, which are joint observables on multiple data qubits. Typically, each stabilizer is measured using an *ancilla qubit* that interacts with each of the relevant data qubits before being measured. Measurement of all stabilizers of the code produces a syndrome that indicates the difference between the expected stabilizer parities and the observed values. This syndrome must then be decoded by a classical algorithm to determine which physical qubits experienced errors. We will refer to this stabilizer measurement process as syndrome extraction (SE). Repeatedly performing SE rounds and decoding allows individual physical qubit errors to be corrected, stabilizing the logical qubit and extending its lifetime.

Important metrics of a QEC code in the context of this work are the code distance $d$, which determines the error robustness and size of a code, and the threshold $p^*$, which is the physical error rate below which the code can be scaled up to increase logical lifetimes. When the physical error rate $p$ is below $p^*$, the logical error rate $p_L$ scales approximately as

$$p_L \approx A\left(\frac{p}{p^*}\right)^{(d+1)/2}. \tag{1}$$

Among the many QEC codes proposed, the rotated surface code stands out for its easy-to-build planar connectivity requirements, relatively high threshold, ease of decoding [31], [32], and well-understood logical operations [45]–[47]. One logical qubit in a distance-$d$ surface code is encoded on a square $d \times d$ grid of physical qubits. Each stabilizer within this grid is either a joint X or Z operator supported on four neighboring qubits (or two on a boundary). Including edge stabilizers, there are $d^2 - 1$ stabilizers that must be measured in each SE round. Multiqubit logical operations in the surface code can be performed in two primary ways, through lattice surgery or transversal CNOTs, both of which are supported on SNAQ.

## III. Hardware requirements

SNAQ is designed to be compatible with current or near-future silicon spin qubit technology. This section reviews state-of-the-art spin qubit capabilities and lists the concrete device-level capabilities we assume for SNAQ and the experimental evidence that these capabilities are achievable with current or near-future silicon spin qubit technology.

### A. Array geometry and readout placement

A SNAQ device consists of a narrow and dense two-dimensional grid of quantum dots with readout components (SETs or SEBs) on two edges of the array, such as the 7-dot-wide example shown in Figure 1(a). This design maximizes compatibility with existing semiconductor fabrication techniques [28], [29], [48], and is a directly scaled-up version of existing devices such as Intel's 4x27-dot array [27] or HRL's 3x3-dot array [29]. The choice of a fixed width means that the array can be constructed using a fixed number of interconnect layers, which are required to control the interior dots in the array, regardless of the size along the second axis of the array. A key parameter that affects the performance of SNAQ is the *readout density*, $\rho$, which is the density of readout devices along the sides of the array relative to the rows of qubits. We will discuss this parameter in more detail in Section IV-A. Readout density can be increased by building more complex routing on the array edges or by using a qubit encoding that occupies more dots, such as the three-electron exchange-only qubits depicted in Figure 1.

### B. Spin shuttling

SNAQ relies on coherent electron transport to move ancilla qubits between readout/initialization regions and interior interaction sites. There are two distinct methods for shuttling an electron between dots. "Bucket-brigade" shuttling is performed by a series of discrete single-dot jumps, which has the advantage of only requiring discrete control pulses on plungers and barriers, and has achieved error rates of less than 0.3% per

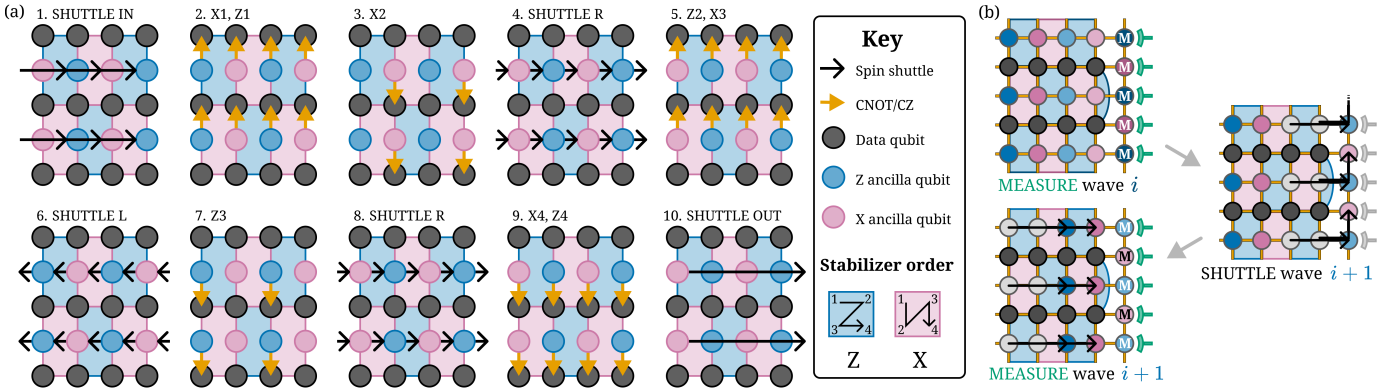


Fig. 3. (a) Schedule of shuttle and CNOT/CZ operations that implement the surface code X and Z checks in the required order. Each ancilla qubit (colored circle) is responsible for interacting with four data qubits in a rectangle, and it must interact in a specific order to avoid hook errors [31], as shown in the lower right. (b) Serialized measurement of a group of ancilla qubits for readout density $\rho = 1$. Ancilla qubits in group $i$ are measured first, followed by qubits in group $i + 1$.

hop [49], [50]. "Conveyor-mode" shuttling involves operating the metal gates along the path with smoothly oscillating voltages to engineer a continuously-moving potential wave that carries the electron along the path. This second approach is more complex to engineer, but can reach faster shuttling rates and has been achieved with per-hop fidelities of 99.99% in experiment [20]. At a nominal dot pitch of 100 nm, these results correspond to per-dot shuttling latencies around 2 ns. We find that these experimentally-achieved speeds and fidelities are already sufficient to enable effective quantum error correction, and we expect further improvement in the future.

### C. Coherence

Data qubits in SNAQ are idle for longer intervals than a standard surface code due to serialized ancilla initialization and measurement. The idling coherence time of the data qubits is therefore a crucial parameter that will significantly affect the performance of a SNAQ device. Quantum-dot electron-spin coherence times vary significantly between device types and choice of qubit encoding. State-of-the-art coherence times range from around 1 ms to 0.56 s for one and two-electron qubits [51]–[54]. Three-electron EO qubits, whose appeal has grown recently, have been stabilized along one axis for up to 720 µs [55]. Coherence is expected to improve further as device uniformity and material purity improve [30]. We find that a coherence time of at least 200 µs is required to enable competitive error correction on SNAQ, with the specific number depending on the code distance, readout element density, and the strength of other error sources.

### D. Physical qubit gate fidelities

SNAQ does not impose any unique constraints on the fidelity of single- and two-qubit gates compared to surface code proposals on other hardware modalities, simply requiring that the gate fidelity is below the threshold of the surface code (typically around $10^{-2}$). We set a realistic gate error target of $10^{-3}$, which has been nearly reached or exceeded in many experimental spin qubit demonstrations [33]–[36], [56], [57].

### E. Initialization and readout

Spin qubit readout via spin-dependent tunneling [13], [58] offers a promising path towards fast high-fidelity readout. Early experiments have demonstrated readout fidelity above 99% in 6 µs with a single-electron box (SEB) [17], 1.6 µs using a single-electron transistor (SET) [15], and 990 ns using a dot-charge sensor (DCS) [59]. Theoretical analyses predict that a readout fidelity of 99.97% and duration as short as 100 ns could be achieved with the SEB [17], and readout above 99% fidelity could be performed in only 36 ns with fully optimized SET device parameters [60]. In this work, we assume that initialization and readout each take 500 ns and have fidelities comparable to physical quantum gates.

### F. Practical summarized requirements for near-term SNAQ

The following target parameters capture the regime in which serialized readout with shuttling yields competitive logical performance, as we will show in simulation in the following sections:

- **Array geometry:** Dense fixed-width rectangular array of nearest-neighbor-connected dots with readout elements confined to edges [27]–[29]. Readout density $\rho$ at least 1, ideally closer to 2.
- **Shuttling:** Per-hop error $p_{\text{sh}} \leq 10^{-4}$ at 2 ns latency [20].
- **Idling:** Coherence time exceeding 200 µs (effective idle-error-per-µs $p_{\text{id}} \lesssim 5 \times 10^{-3}$) under active stabilization such as dynamical decoupling [51]–[55].
- **Physical logic gates:** CNOT error $p_{\text{g}} \gtrsim 99.9\%$ [33]–[36], [56], [57].
- **Initialization and measurement:** Initialization and readout fidelity equal to $p_{\text{g}}$ with 500 ns latency each [60].

To provide a clear characterization of the architecture, we restrict our evaluations to the most impactful regions of the

design space, prioritizing sensitivity studies of the variables that drive the primary tradeoffs.

## IV. The SNAQ architecture

The layout of our proposed architecture is shown in Figure 2(a), where a $(d+2)$-qubit wide array is used to support a surface code of distance $d$.

### A. Hardware layout

SNAQ consists of a dense fixed-width array of quantum dots with readout components on both sides, as shown in Figure 1(a). Although previous work has proposed building long shuttling channels to route around these large readout zones [24], [61], in this work we explore the more near-term friendly approach of placing readout elements along two outer edges of a fixed-width array [27], keeping the interior free for uniform dot placement and nearest-neighbor exchange coupling. A fixed array width $w$ limits the largest surface code that we can embed in the array using the mapping of Figure 2(a) to $d_{\max} = w - 2$. A two-column layout to support multiqubit interactions via lattice surgery would support a maximum distance of $d_{\max} = \lceil \frac{w-3}{2} \rceil$.

The readout density $\rho$, the ratio of readout sensors to array rows, is a key hardware parameter. A density of 1 means that each row has a readout sensor, as in the device shown in Figure 2(a). The readout density directly determines the amount of serialization required to initialize or measure some fixed number of spin qubits, so it has a significant effect on the speed and error rate of the surface code. Achieving $\rho \gg 1$ in a manufacturable device is difficult, but a small constant-factor increase (such as $\rho = 2$) may be attainable with careful routing.

### B. Scheduling a surface code on SNAQ

Figure 3(a) shows the specific schedule of shuttles and CNOTs that perform a stabilizer measurement cycle of the surface code. The ancilla qubits are first initialized on the left side of the array and are shuttled into place as they become ready. The ancilla qubits and data qubits then perform several layers of CNOTs. Finally, the ancilla qubits are shuttled to the right side of the array, where they can be measured. Note that the depiction of shuttle operations here is slightly simplified; in reality, buffer space would be needed in between adjacent ancilla qubits for them to both be shuttled at the same time. This means that the group shuttle operation cannot be fully parallelized.

In many quantum error correction codes, including the rotated surface code, the specific schedule of CNOT gates is crucial, as some orderings of these gates can allow hook errors, where a single error on an ancilla qubit can propagate to multiple data qubit errors during a round of syndrome extraction [31]. This would reduce the effective code distance, so care must be taken to choose a gate ordering that avoids hook errors. In our schedule, we implement the ordering shown in the lower right of Figure 3(a), where the number

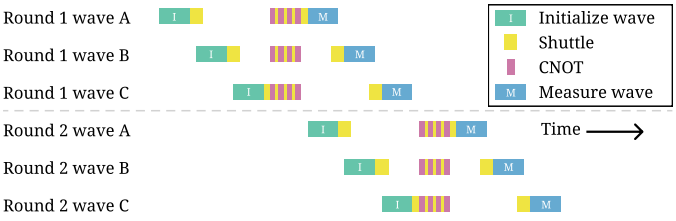


Fig. 4. Example pipeline schedule showing initialization, data entanglement, and measurement for two successive rounds of ancilla qubits, split into three waves each. The measurement of the first round of ancillae can be done simultaneously with the initialization of the second round. Any blank space in the schedule indicates idling, during which dynamical decoupling techniques could be used to improve coherence.

in each corner of a stabilizer indicates the order of the gates from the perspective of the Z or X ancilla qubit.

The primary bottleneck of the SNAQ surface code is the initialization and measurement of $d^2 - 1$ ancilla qubits in each round of syndrome extraction. In the proposed syndrome extraction schedule, these ancilla qubits are all initialized on the left side of the array, shuttled into place in the interior, and then eventually measured and removed on the right side of the array. Because a given surface code patch will only have $O(d)$ available initialization/measurement devices on one of its edges for any constant density $\rho$, a sufficiently-large code distance will require multiple serial waves of ancilla qubit initialization/measurement, as depicted in Figure 3(b). The number of ancilla waves is described by

$$n_w = \left\lceil \frac{d^2 - 1}{2\rho(d+1)} \right\rceil, \tag{2}$$

where $d^2 - 1$ is the total number of ancilla qubits in a distance-$d$ surface code and $2\rho(d+1)$ is the number of readout ports on one side of a SNAQ surface code. This serialization is the primary novel source of error that the SNAQ surface code must mitigate. Logical initialization and measurement of a surface code require serialized readout and measurement of the $d^2$ data qubits as well, which can be done in a similar method to the ancilla qubits (and can be done in roughly half the number of waves by using both edges of the array at the same time).

Importantly, we note that steps 1 and 10 of Figure 3(a), which are by far the longest due to serialization, can be pipelined: while layers of ancillae are being moved to the readout locations on the right side of the array, fresh waves of ancillae can be initialized at the same time to fill in the empty spots in the array, as shown in Figure 4. Our circuit-level simulations account for this pipelining.

### C. Multiqubit logical gates

Two main methods have been proposed to entangle logical surface code qubits. Hardware that is restricted to nearest-neighbor connectivity can perform lattice surgery [45]–[47] (LS), which involves the merging and splitting of surface code patches to perform multiqubit Pauli measurements. In addition, hardware with long-range connections can perform

| Name | Near-term | Low-$d$ perf. | Ultrahigh-$d$ perf. | Logical operations | Logical clock cycle time ($d = 11$) | Logical parallelism | Refs. |
|---|---|---|---|---|---|---|---|
| UnitCell | × | × | ✓ | LS | 28.6 µs | 2D | [21], [24], [61] |
| Bilinear | ✓✓ | ✓ | × | LS | 20.8 µs | 1D | [25] |
| **SNAQ** | ✓ | ✓✓ | × | LS + tCNOT | 5.3 µs (local) - 55.6 µs (global) | 1D++ | This work |

TABLE I
COMPARISON TO PRIOR FAULT-TOLERANT SILICON ARCHITECTURE PROPOSALS

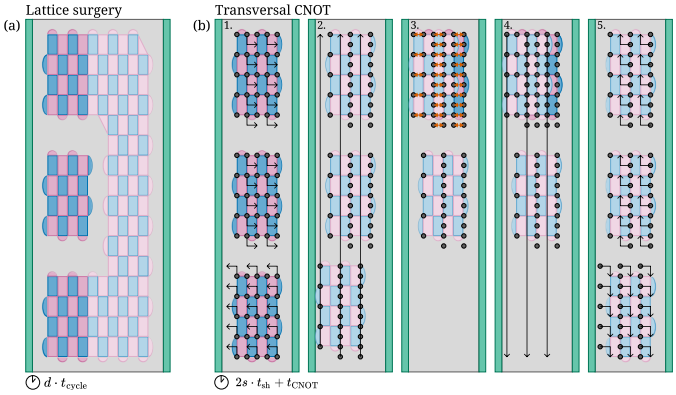


Fig. 5. Two possible implementations of multiqubit operations in the SNAQ architecture. (a) Lattice surgery can connect distant logical qubits via a channel of "routing patches" (lighter shaded stabilizers), which requires a wider dot array to support a second column of surface codes. (b) Spin shuttling enables transversal CNOT operations for sufficiently-close tiles. Between stabilizer measurement rounds, physical qubits can shuttle past intermediate tiles to reach the target tile, then perform the transversal operation and shuttle back.

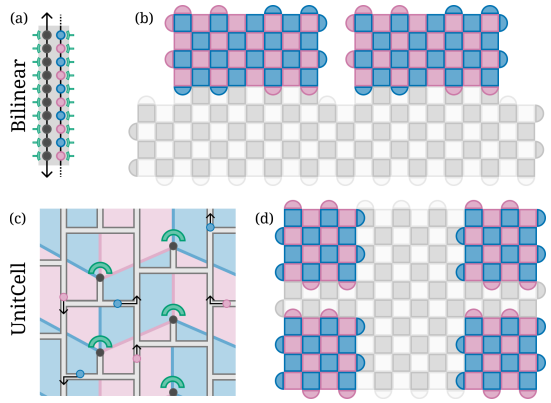


Fig. 6. Two prior spin qubit architecture proposals. (a) The Bilinear architecture, based on Refs. [25], [26], consists of two rails of qubits, one of which can shuttle back and forth. Data and ancilla qubits of a surface code are placed column-by-column into the linear array. (b) The Bilinear architecture supports a linear logical-level arrangement, where each (wide) surface code patch is connected to others by a shared lattice surgery routing bus (gray). (c) The UnitCell architecture, similar to Refs. [21], [24], [61], consists of large unit cells, each of which contains a readout port. A data qubit can be placed in each unit cell and the ancilla qubits must shuttle long distances between unit cells. (d) The two-dimensional physical qubit array of the UnitCell architecture naturally produces a two-dimensional logical array, with lattice surgery routing space between surface code patches.

a transversal CNOT (tCNOT), where each data qubit of one patch interacts with its corresponding data qubit in the second patch. Using specialized decoding techniques [62], [63], transversal CNOTs only require $\Theta(1)$ SE rounds after each operation, saving significant temporal costs compared to lattice surgery (the actual number can be less than 1 SE round per tCNOT [62]). Additionally, while parallel shuttling of qubits is possible, parallel lattice surgery through the same shared routing space is not, so transversal CNOTs may have compilation benefits for certain workloads. Both lattice surgery and transversal operations are supported in the SNAQ architecture, and are depicted in Figure 5. Importantly, a two-column logical layout is required to support lattice surgery, while a processor that only uses tCNOTs can be operated in a single-column layout, which may be significantly easier to manufacture for the same target code distance. The tradeoff to using tCNOTs is that they natively only work well for relatively small separation distances, before shuttling errors accumulate. We explore this and suggest possible methods to extend the tCNOT's range in Section V-E.

### D. Prior proposals

We compare SNAQ to two baseline spin qubit architectures based on prior proposals, which we call UnitCell and Bilinear, described in Figure 6. These architectures both support a 1:1 ratio of readout ports to qubits, but with very different means. The Bilinear architecture restricts the physical qubits to a 2×N physical layout, as proposed in Ref. [25]. This design makes fabrication relatively straightforward and cost-effective, but performance depends more heavily on shuttling (as not all data-ancilla pairs can be placed near each other) and yields a logical layout with only one shared lattice surgery routing bus, limiting achievable logical parallelism.

On the other hand, the approach of the UnitCell architecture is to create large unit cells that each contain a readout component and have enough space for all readout and qubit wiring, as proposed in Refs. [21], [24], [61]. Each unit cell is surrounded by shuttling channels, allowing qubits to be moved between nearby unit cells. We assume a relatively optimistic unit cell size of $5 \times 5$ µm, corresponding to 50 dot-to-dot shuttling distances per side. The surface code can be implemented by assigning each data qubit to its own unit cell and shuttling ancilla qubits to each data qubit cell. The benefit of this approach is that the amount of required shuttling does not change with the code distance, so this architecture is the only one of the three to exhibit a true threshold and to enable arbitrarily-high code distance. The downside is a very large constant amount of shuttling, which can be thought of as effectively adding to the baseline physical error rate, thereby shifting the code threshold. This will reduce UnitCell's relative performance for smaller code distances. The resulting logical layout is also two-dimensional, as shown in Figure 6d.

SNAQ's layout lies somewhat in between these two baselines with its fixed-width design, though fundamentally differs from both in purposefully deviating from the 1:1 readout-to-qubit ratio of the baselines. This allows SNAQ to have a much higher qubit density than either baseline, which is critical to enable its transversal logic capabilities.

We provide a high-level, qualitative comparison of these architectural tradeoffs in Table I. This summary contrasts the UnitCell and Bilinear baselines with SNAQ, highlighting our proposal's focus on near-term manufacturability, strong low-distance performance, and efficient logical operations. Unlike the baselines, which are limited to lattice surgery (LS), SNAQ's dense design supports fast transversal CNOTs (tCNOTs), enabling a more flexible, high-parallelism logical layout. The "1D++" logical connectivity of SNAQ refers to the potential parallelism of transversal operations compared to lattice surgery.

## V. Performance Evaluations

We develop a custom simulation framework that can accurately model the pipelined schedule of shuttled waves of ancilla qubits, allowing us to precisely quantify both the clock speed of the code and the logical performance. We simulate the performance of our proposed architecture using Stim [64] and decode with PyMatching [65]. For a given code distance $d$, we assume a fixed array width of $d+2$. For all experiments, we perform both X and Z-basis experiments and combine the two error rates to obtain an overall error rate.

### A. Noise model

Our noise model is inspired by exchange-only spin qubits [33]; while this limits the direct application of our results for other spin encodings, our general takeaways should apply for a wide range of possible implementations. We use a noise model with three error parameters $p_{\text{g}}$, $p_{\text{sh}}$, and $p_{\text{idle}}$. The parameter $p_{\text{g}}$ sets the strength of gate and readout errors. Each CNOT gate causes a two-qubit depolarizing error with probability $p_{\text{g}}$, each readout encounters a bit flip with probability $p_{\text{g}}$, and each Hadamard gate causes a one-qubit depolarizing error with probability $p_{\text{g}}/10$. The parameter $p_{\text{sh}}$ is the error-per-shuttle such that shuttling a qubit over a distance of $m$ dots incurs a depolarizing error with probability $m \cdot p_{\text{sh}}$. Finally, the idle error $p_{\text{idle}}$ is defined as the chance that a qubit experiences a depolarizing idle error in a 1 µs idle interval (in our noise model, this is equivalent to the time for the five layers of CNOTs in one SE round), so a coherence timescale of $T$ would correspond to $p_{\text{id}} = e^{-(1\ \mu\text{s})/T}$.

We set the durations of physical operations assuming an exchange pulse duration of 10 ns, which means that a CNOT takes approximately 200 ns (the exact CNOT latency depends on dot-level connectivity [66], [67], but here we leave it fixed), a single-qubit Hadamard gate takes 30 ns, and the shuttling latency is 2 ns per dot. We assume that initialization and readout can each be done in 500 ns. Although our numerical simulation results depend on these specific choices, the important underlying ratio is the idle coherence timescale relative to the rate of syndrome extraction rounds. Our results can therefore be interpreted for different hardware latency assumptions by rescaling $p_{\text{id}}$ accordingly.

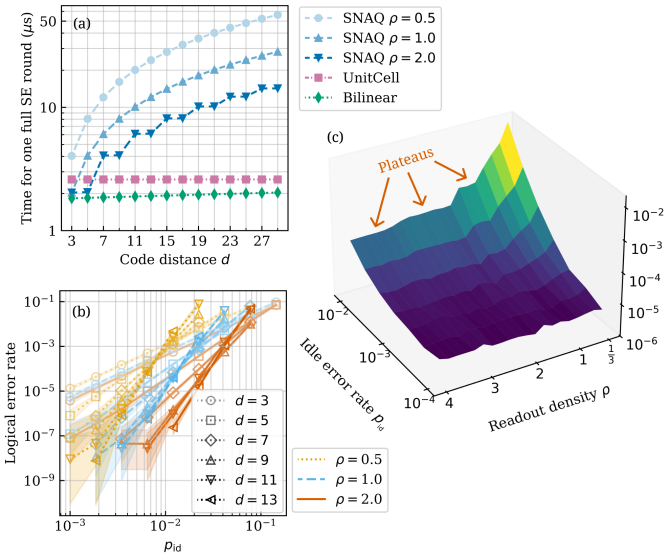


Fig. 7. Exploring the effect of readout serialization in SNAQ. (a) Time to complete one (non-pipelined) SE round in SNAQ compared to the baselines, for various values of $\rho$. (b) Logical error rate as a function of $p_{\text{id}}$ for various code distances and densities. Doubling $\rho$ can give an order of magnitude improvement in logical error rate. (c) Logical error rate of a $d = 7$ surface code on SNAQ, showing "plateaus" where ranges of density settings lead to the same number of ancilla waves, yielding the same performance.

In some qubit encodings, the shuttling and idling errors can be strongly biased towards dephasing noise [20], which would incentivize asymmetric surface codes with $d_z \neq d_x$. A benefit of SNAQ is that the readout serialization would be determined by the smaller of $d_x$ or $d_z$. Other options to consider include bias-tailored variants of the surface code [68]–[71].

### B. Impact of readout serialization

The readout density $\rho$ is a critical architectural parameter that directly affects logical performance by changing the duration of each syndrome extraction (SE) round, as shown in Figure 7(a). For $\rho = 2$, we see that the SE duration remains below 10 µs up to $d = 21$. Pipelining (Figure 4) will reduce SNAQ's effective SE duration by half. Compared to the two baseline architectures, SNAQ's SE duration at the same code distance is still significantly longer even with a relatively high $\rho$; however, we will show later that SNAQ can still achieve a faster logical clock cycle due to its ability to execute transversal gates.

To quantify the impact of readout serialization on SNAQ's logical performance, we simulate logical performance at different readout densities $\rho$ and different idle error rates $p_{\text{id}}$, with $p_{\text{g}} = p_{\text{sh}} = 0$. Due to the increase in serialization for larger distances, we do not expect to observe a clear threshold (a point where the lines of the same group would all cross). As shown in Figure 7(b), doubling the density can improve

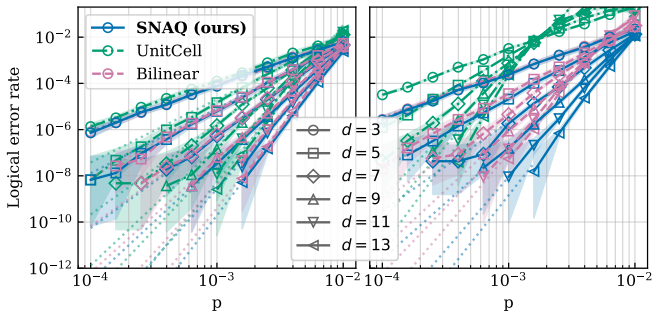


Fig. 8. Logical error rate comparison to baseline architectures with SNAQ readout density $\rho = 2$. *Left:* Performance under noise model $(p_g, p_{id}, p_{sh}) = (p, p/10, p/100)$, where idling errors are stronger than shuttling errors. *Right:* Performance under noise model $(p_g, p_{id}, p_{sh}) = (p, p/100, p/10)$ where shuttling errors dominate idling errors.

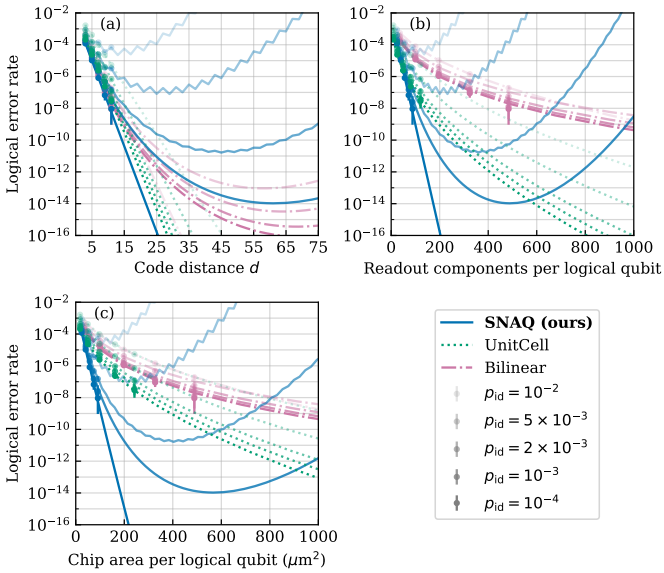


Fig. 9. Projecting logical performance to higher code distances under fixed $(p_g, p_{sh}) = (10^{-3}, 10^{-5})$ for various $p_{id}$ with SNAQ $\rho = 2$. (a) SNAQ and Bilinear give diminishing returns as code distance increases due to an increase in shuttling and idling errors with larger codeblock size, while UnitCell exhibits consistent error suppression as distance is increased. (b) Logical error rate as a function of readout count, which may be an important driver of packaging cost, showing that SNAQ outperforms Bilinear and is competitive with UnitCell. (c) Logical error rate as a function of total component area per logical qubit. For sufficiently-low $p_{id}$, SNAQ achieves better logical performance at significantly reduced total chip area compared to both baselines.

the logical performance at the same distance by more than an order of magnitude, underscoring the importance of $\rho$ for the SNAQ architecture. Figure 7(c) visualizes this relationship for fixed $d = 7$, showing performance "plateaus" where ranges of $\rho$ yield the same number of ancilla waves.

### C. Comparison to other architectures

Next, we compare the logical performance of SNAQ with that of Bilinear and UnitCell. Figures 8(a-b) show the results of circuit-level simulations under two different noise models, one dominated by the idle error and one by the shuttling error. When idling errors are stronger, SNAQ is competitive with baseline methods, and outperforms baselines when shuttling errors are stronger.

Although we cannot directly simulate code performance at larger distances with current Monte Carlo methods, as the number of required shots becomes impractical, we can instead fit the lower-distance logical error rates to the expected scaling behavior of SNAQ. With $p_g$, $p_{sh}$, and $p_{id}$ fixed, we can model the logical error rate of SNAQ as a modification of Eq. 1:

$$p_L(d) = A\left(\frac{p_g}{p_g^*} + \frac{p_{sh}^{eff}}{p_{sh}^*} + \frac{p_{id}^{eff}}{p_{id}^*}\right)^{(d+1)/2} \tag{3}$$

where $p_{sh}^{eff}$ is the cumulative shuttling error experienced by the physical qubits, $p_{id}^{eff}$ is the cumulative idling error experienced by the qubits. Here we have made the first-order assumption of a linear threshold surface [72] characterized by three thresholds $p_g^*$, $p_{sh}^*$, and $p_{id}^*$. To fit circuit-level simulation data to this model, we use the form

$$p_L(d) = A\big(\alpha + \beta d + \gamma n_w(\rho, d)\big)^{(d+1)/2}, \tag{4}$$

where $A$, $\alpha$, $\beta$, and $\gamma$ are fitting parameters, and $n_w(\rho, d)$ is given by Eq. 2. This model accounts for the three ways in which physical errors in SNAQ scale with increasing code distance: gate errors remain constant, shuttling errors increase linearly in $d$, and idle errors increase with the amount of ancilla serialization $n_w$. We have derived this model from the structure of the SE circuit, and we find that it fits well to the available simulation data, but we caution that it is an approximation and circuit-level simulation still remains the most reliable performance predictor. To fit the Bilinear baseline, we fix $\gamma = 0$, and for the UnitCell baseline, we fix $\beta = \gamma = 0$.

In Figure 9(a), we simulate the performance of SNAQ and baselines for distances up to $d = 11$ under $(\rho, p_g, p_{sh}) = (2, 10^{-3}, 10^{-5})$ and various values of $p_{id}$, fitting each set of results to Eq. 4 and extrapolating to large distances. As expected, SNAQ is highly sensitive to $p_{id}$, but we see that it can still reach utility-scale error rates below $10^{-6}$ with $p_{id} = 5 \times 10^{-3}$, roughly corresponding to a coherence time of 200 μs, which is well within achievable ranges for spin qubits. This figure shows that at a fixed code distance, SNAQ's logical error rate is far more sensitive to $p_{id}$ than the baselines. We also can clearly see that $p_{id}$ defines a clear error floor for SNAQ, limiting the achievable logical error rate. To reach extremely low error rates below $10^{-12}$ for long-term resource-intensive applications, $p_{id} = 10^{-3}$ or $10^{-4}$ will be required. This translates to a spin qubit coherence time near the single-ms range. Although this has already been demonstrated in small devices [51]–[54], it remains to be seen whether the same performance can be achieved at scale.

However, comparing these architectures at a fixed code distance is misleading because it obscures the starkly different physical resource costs of the architectures. In SNAQ, where

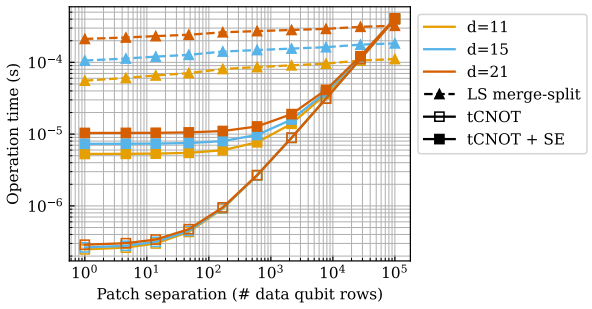


Fig. 10. Duration of lattice surgery merge-split and tCNOT operations in SNAQ as a function of patch separation, assuming a per-shuttle latency of 2 ns and a density of $\rho = 1$. For separation distances under 200 dots, tCNOT takes less than a microsecond (without the SE round), and is competitive with lattice surgery out to separation distances near 100,000.

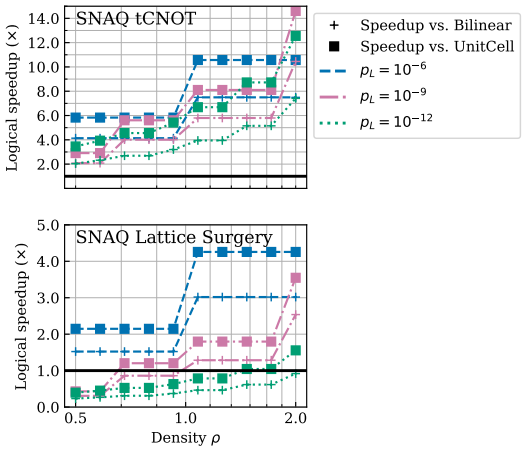


Fig. 11. Comparing SNAQ logical operations to baselines under $(p_g, p_{sh}, p_{id}) = (10^{-3}, 10^{-5}, 10^{-4})$ for various $\rho$. SNAQ's tCNOT outperforms baseline lattice surgery methods by over 2× even for low $\rho$ and can provide up to 14.6× improvement in the range studied here. Lattice surgery on SNAQ is fast for higher $p_L$ and $\rho$, but becomes less efficient for lower $p_L$ due to the significant increase in required code distance (making each SE round take much longer).

I/O costs and qubit count are decoupled, the number of qubits may no longer be the primary cost driver. We therefore turn to more relevant architectural metrics by studying two fabrication-limited resources: the total number of readout components and the resulting chip area. Figure 9(b) shows the same data but with code distance converted to the number of readout components required per logical qubit, revealing SNAQ's much more efficient use of readout components and wiring. Finally, in Figure 9(c), we convert code distance to chip area by assuming that each qubit (plunger and adjacent barrier gates) takes up a $100 \times 100$ nm $= 0.01$ μm$^2$ space and we use an order-of-magnitude estimate of the readout footprint as 1 μm$^2$ based on Ref. [18] and chip images from e.g. Refs. [13]–[17]. These estimates account for the footprint of the components themselves, but do not include interconnect routing, which will be implementation-specific. For UnitCell, we therefore only consider the area taken up by the readout component and shuttling channels (not the large interior spaces reserved for wiring). This analysis reveals that, for achievable $p_{id}$, SNAQ produces logical qubits that are orders of magnitude more area-efficient than the baselines, already providing a clear advantage at $p_{id} = 5 \times 10^{-3}$ that becomes much stronger with lower $p_{id}$.

### D. Logical clock speed

Figure 10 shows the durations of the tCNOT and LS merge-split operations over varying separation for three surface code distances $d \in (11, 15, 21)$. For the tCNOT, we show both the time to complete the tCNOT itself and the time to complete a tCNOT and half a (non-pipelined) SE round, assuming that two tCNOTs are performed for every SE round [5], [62]. For separation distances up to $10^4$, we observe that the tCNOT+SE is significantly faster than the LS operation by up to 10×, depending on the code distance. The LS operation slightly increases in duration for larger separation distances due to the need to preserve performance (by adding SE rounds) against an increasing number of possible temporal error chains [73].

We can compare these temporal costs to those of lattice surgery on the two baseline architectures. Distance-$d$ lattice surgery consists of $d$ SE rounds, so at $d = 11$ the minimum total LS merge+split time is 20.8 μs for Bilinear and 28.6 μs for UnitCell. Although these durations are significantly shorter than SNAQ's $\rho = 1$ minimum LS merge+split time of 55.6 μs, they are approximately twice as long as SNAQ's $d = 11$ (tCNOT + half SE) duration of 5.3 μs. At the same code distance, SNAQ therefore provides a logical clock cycle speedup of 3.9-5.4× with $\rho = 1$ compared to prior work. SNAQ maintains a latency advantage over separation distances of over 7000 physical qubits. The relative latency advantage is similar across a wide range of code distances, with the $d = 5$ tCNOT providing a 4.1-5.8× speedup with a 2.2 μs latency, and $d = 27$ providing a 4.1-5.2× speedup with a latency of 13.4 μs.

To more thoroughly quantify SNAQ's potential speedup and its sensitivity to the key hardware parameter $\rho$, we simulate logical performance for various values of $\rho$. For this analysis, we use fixed physical error rates of $(p_g, p_{sh}, p_{id}) = (10^{-3}, 10^{-5}, 10^{-4})$, simulate distances up to 11, and fit the model in Eq. 4 to the data. We apply the same process to the two baselines to provide a direct comparison. As in Section V-C, these fits allow us to determine the code distance at which each architecture can achieve a given logical error rate. We can then convert these code distances to logical clock speeds. Figure 11 shows the comparison of the logical operation speeds in SNAQ with those of UnitCell and Bilinear. We find that SNAQ's tCNOT is significantly faster than either of the baselines across the entire studied range. For a near-term target $p_L$ of $10^{-6}$, we find a 4.1-10.6× speedup across the density range. For lower $p_L$, the improvement is smaller for low $\rho$, but still remains above 2×. Because the interaction distance of the tCNOT is limited, as we will investigate in the following subsection, we also compare the lattice surgery

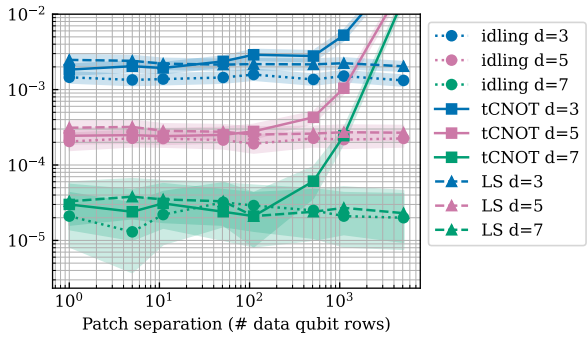


Fig. 12. Error rates of transversal CNOT (tCNOT) and lattice surgery merge-split (LS) as a function of patch separation. Patch separation is defined as the number of dot rows between the two patches.

time between SNAQ and the baselines. Additionally, we find that SNAQ's lattice surgery still outperforms the baselines for a target $p_L$ of $10^{-6}$, but is less competitive for lower $p_L$ and lower $\rho$.

### E. Transversal CNOT fidelity

However, the separation distances at which a standard tCNOT will outperform lattice surgery are limited: in SNAQ, the *physical* shuttling and idling errors that occur during a transversal CNOT will increase linearly with the distance between the patches, meaning that the *logical* error rate will increase exponentially. We thus expect transversal operations to perform better for sufficiently-close patches, while lattice surgery will be preferred for sufficiently long-distance communication. To investigate this tradeoff, we perform circuit-level simulations of tCNOT and LS operations in the SNAQ architecture using parameters $(p_{\mathrm{g}}, p_{\mathrm{id}}, p_{\mathrm{sh}}, \rho) = (10^{-3}, 10^{-4}, 10^{-5}, 1.0)$. We prepare two surface code patches in the logical $|0\rangle$ state and either perform a tCNOT or an LS merge-split between them, calculating the resulting error rate (the chance that either logical qubit experiences a bit-flip for the tCNOT, and the chance that we correctly measure the ZZ observable for the LS operation). We decode the lattice surgery experiment with PyMatching and use the BP+OSD decoder [74], [75] to decode the transversal CNOT. The results are shown in Figure 12, where we see that tCNOT and LS both exhibit performance near memory idle error rates for short separation distances, but for sufficiently large distances the performance of tCNOT begins to degrade. The distance at which this occurs depends on the relative strength of the per-dot shuttling error compared to other error sources: performance suffers once the separation distance $s$ is large enough that $s \cdot p_{\mathrm{sh}}$ is the dominant source of error. In the case of our simulations, $p_{\mathrm{g}} = 10^{-3}$ is the dominant error source for an idling surface code, so with $p_{\mathrm{sh}} = 10^{-5}$, the limit will be around $s \approx 100$, which matches what we observe in the simulation results.

We envision several potential ways to extend the tCNOT to longer ranges, all of which involve a space tradeoff in the SNAQ array. First, the shuttled surface code could stop along the way to perform multiple SE rounds during a long-distance tCNOT, which would avoid accumulating shuttle errors [76]. Second, extra spaces could be used to prepare Bell pairs to use for gate teleportation or entanglement swapping [77]. Finally, the most aggressive method is to designate every other surface code as a dedicated logical ancilla for GHZ state preparation, allowing for constant-depth CNOTs [78]–[80]. Although these methods are compelling, an evaluation of the tradeoffs involved is beyond the scope of this work.

### F. Architectural implications

The speed and error rate studies discussed above suggest that both tCNOTs and lattice surgery are valid options for logical operations on the SNAQ architecture, with tCNOTs providing a significant speed advantage for sufficiently-short separation distance and LS operations enabling longer-distance communication without degrading fidelity. However, the LS operations are approximately $10\times$ slower that the tCNOT (depending on patch separation and code distance),

A SNAQ device of width $w$ can support a surface code distance of up to $d = w - 2$ in a logical $1 \times N$ configuration if logical operations are restricted to only tCNOTs. This processor would have some maximum allowed interaction distance within which the tCNOT's performance is acceptable, which depends on the relative strengths of shuttling errors to other sources of error. Compiling programs to such an architecture is reminiscent of the challenges faced in nearest-neighbor-connectivity noisy intermediate-scale quantum (NISQ) processors, where long-range interactions necessitate multi-hop, higher-cost operations, except in this case the extra cost is in time instead of added error. In these NISQ processors, locality-aware mapping and routing techniques were developed to minimize the amount of long-distance communication needed [81]; SNAQ may benefit from the resurgence of similar techniques to make use of fast tCNOTs whenever possible.

On the other hand, a SNAQ device of width $w$ could also support a surface code distance of up to $d = \lfloor \frac{w-3}{2} \rfloor$ in a logical $2 \times N$ layout, which could then support both transversal CNOTs and lattice surgery if one logical channel is left free to use as routing space. Such a device with two available modes of communication is reminiscent of the classic latency hierarchy in classical computer architecture, where care must be taken to avoid unnecessary use of main memory accesses or network connections. Such a processor presents a compelling challenge for future compilation research.

## VI. Resource estimates for magic state distillation

The numerics in the prior section showed that, using transversal CNOTs, the SNAQ architecture can perform individual logical operations over twice as fast as competing architectures for sufficiently-close logical qubits. As an example of compiling a logical circuit to the SNAQ architecture, we consider 15-to-1 T state distillation, a crucial task for fault-tolerant quantum computation which serves as a useful microbenchmark. 15-to-1 distillation is derived from the

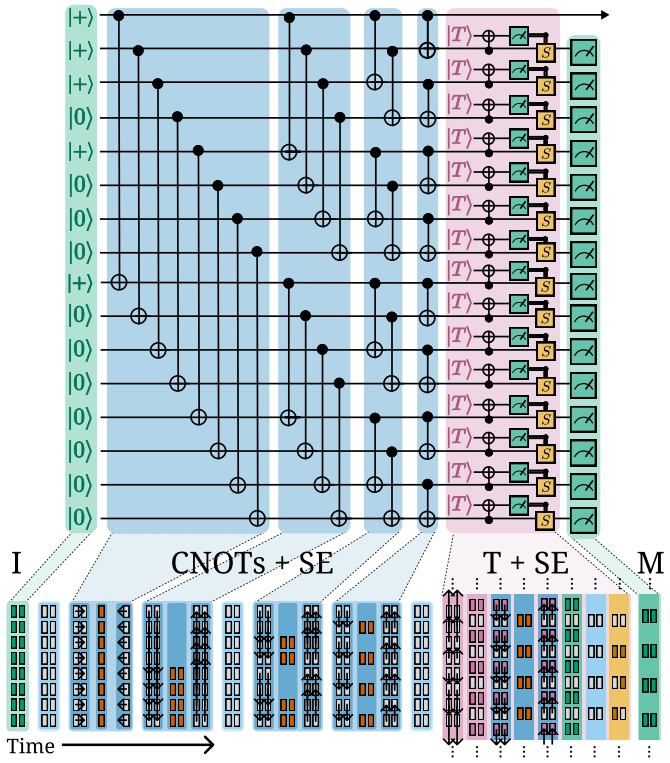


Fig. 13. *Top:* Logical circuit implementing 15-to-1 distillation. CNOTs are shaded in groups to show those that can be performed in parallel using transversal CNOTs. *Bottom:* A $2 \times 8$ SNAQ layout executing the circuit. The schedule involves parallel tCNOTs and T gate injection, interleaved with four SE rounds.

[[15,1,3]] Hastings-Haah code [82] and involves preparing 15 noisy T states and "distilling" them to produce one higher-fidelity T state. There are a variety of ways to compile the distillation circuit onto the surface code, trading space and time [46], [47], [83], [84].

Neither of the two prior spin qubit architecture proposals are compatible with parallelizable transversal CNOTs, so lattice surgery is needed instead. For the UnitCell architecture, we can use the 15-patch two-dimensional layout from Figure 11 of Ref. [83] that requires $6d$ SE rounds. For the Bilinear architecture, using a virtual $2\times$N layout of logical surface code qubits as depicted in Figure 3 of Ref. [25], we can implement the same circuit, but must restrict to a $2\times5$ layout. For SNAQ, we choose to use the circuit shown in Figure 13, which uses 16 logical qubits. The benefit of this encoding circuit is that the transversal CNOT implementation is highly parallelizable, requiring only four layers if the shuttles are properly scheduled. We assume readout density $\rho = 1$, T state prep time comparable to surface code initialization time, and no decoding delay for conditional S corrections. Our construction assumes that four SE rounds are performed during the distillation, as shown in Figure 13. Each SE round here takes twice as long due to the two-column logical layout. We assume the implementation of a fold-transversal S gate [85], [86], which we set to have a latency equal to $2d \cdot t_{\text{shuttle}} + t_{\text{CNOT}}$.

| | $d = 7$ | | $d = 15$ | |
|---|---|---|---|---|
| | Time (µs) | Vol. (qubit·s) | Time (µs) | Vol. (qubit·s) |
| Bilinear | 156.2 | 0.152 | 346.3 | 1.555 |
| UnitCell | 109.2 | 0.159 | 234.0 | 1.576 |
| **SNAQ** | **58.9** | **0.060** | **134.1** | **0.561** |

TABLE II
15-TO-1 DISTILLATION COST COMPARISON

The results are shown in Table II, where we see that SNAQ achieves a 60-63% volume reduction compared to the baselines at $d = 7$ and 64% volume reduction at $d = 15$. These cost reductions, while significant, are smaller than the improvements in the logical clock speed. We attribute this to two aspects of this particular example: (1) the compiled distillation circuit is very shallow, with only five layers of tCNOTs, so placing an SE round after every two tCNOTs yields an average SE count per tCNOT of 0.8 instead of the 0.5 we would aim for in a deep circuit, and (2) the double-column logical layout, which was chosen to reduce shuttling distance, makes each SE round twice as slow.

Reducing the spatial code distance in one direction as proposed in Ref. [83] could allow a lower number of SNAQ ancilla waves, further reducing its runtime. The specific code distances chosen would depend on the error rate of the noisy Ts, which could be close to the physical gate error rate if fast injection is used [87], or far lower if cultivation techniques are used first [88]. Modeling cultivation on the SNAQ architecture is beyond the scope of this work but is an important future step to reduce the cost of preparing high-quality T states.

## VII. Conclusion

In this work, we proposed the SNAQ surface code, a novel surface code implementation tailored for spin qubits. To overcome the readout component size problem, we introduced the idea of serialized ancilla qubit initialization and readout, showing that this is an effective way to enable error correction on a near-term-manufacturable silicon quantum dot array. SNAQ is more space-efficient than prior proposals while providing a substantial logical clock speedup through transversal logic by leveraging rapid spin shuttling and a dense dot array. The choice between fast, short-range tCNOTs and slow, long-range LS operations creates a compelling latency hierarchy that deserves future exploration.

The shuttling capabilities of silicon quantum dots naturally leads to the consideration of more complex quantum low-density parity check (qLDPC) codes, which provide better encoding rates than the surface code but require nonplanar qubit connectivity. Spin shuttling may enable the implementation of such codes by connecting distant physical qubits. However, these codes would impose stronger constraints on the viable range of $\rho$ and $p_{\text{id}}$ in a SNAQ implementation. We leave this exploration as important future work.

SNAQ demonstrates that the often-assumed 1:1 readout-to-qubit ratio is neither necessary nor optimal on a shuttling-equipped architecture such as silicon spin qubits. Enabled by

the unique capability of fast spin shuttling, SNAQ is orders-of-magnitude more area-efficient and up to 2-14× faster than proposals that mimic 2D grid layouts. Our work highlights qubit coherence and readout density as the most critical device metrics to enable this new architecture, providing a compelling, practical path toward fault-tolerant spin qubit processors.

## ACKNOWLEDGEMENTS

We are grateful to Prithvi Prabhu, Fahd Mohiyaddin, and Nathaniel Bishop at Intel for helpful discussions on spin qubit manufacturing constraints and performance metrics.

This work is funded in part by the STAQ project under award NSF Phy-232580; in part by the US Department of Energy Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing Program; and in part by the NSF Quantum Leap Challenge Institute for Hybrid Quantum Architectures and Networks (NSF Award 2016136), in part by the NSF National Virtual Quantum Laboratory program, in part based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, and in part by the Army Research Office under Grant Number W911NF-23-1-0077. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

FTC is the Chief Scientist for Quantum Software at Infleqtion and an advisor to Quantum Circuits, Inc.

## REFERENCES

[1] Isaac H. Kim, Ye-Hua Liu, Sam Pallister, William Pol, Sam Roberts, and Eunseok Lee. Fault-tolerant resource estimate for quantum chemical simulations: Case study on Li-ion battery electrolyte molecules. *Physical Review Research*, 4(2):023019, April 2022. Publisher: American Physical Society.

[2] Michael E. Beverland, Prakash Murali, Matthias Troyer, Krysta M. Svore, Torsten Hoefler, Vadym Kliuchnikov, Guang Hao Low, Mathias Soeken, Aarthi Sundaram, and Alexander Vaschillo. Assessing requirements to scale to practical quantum advantage, November 2022. arXiv:2211.07629 [quant-ph].

[3] Tyler Leblond, Christopher Dean, George Watkins, and Ryan Bennink. Realistic Cost to Execute Practical Quantum Circuits using Direct Clifford+T Lattice Surgery Compilation. *ACM Transactions on Quantum Computing*, 5(4):1–28, December 2024.

[4] Craig Gidney. How to factor 2048 bit RSA integers with less than a million noisy qubits, May 2025. arXiv:2505.15917 [quant-ph].

[5] Hengyun Zhou, Casey Duckering, Chen Zhao, Dolev Bluvstein, Madelyn Cain, Aleksander Kubica, Sheng-Tao Wang, and Mikhail D. Lukin. Resource Analysis of Low-Overhead Transversal Architectures for Reconfigurable Atom Arrays. In *Proceedings of the 52nd Annual International Symposium on Computer Architecture*, pages 1432–1448, Tokyo Japan, June 2025. ACM.

[6] 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(1):13575, November 2016. Publisher: Nature Publishing Group.

[7] 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(1):1766, December 2017. Publisher: Nature Publishing Group.

[8] 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(7):eaar3960, July 2018. Publisher: American Association for the Advancement of Science.

[9] Xiao Xue, Bishnu Patra, Jeroen P. G. van Dijk, Nodar Samkharadze, Sushil Subramanian, Andrea Corna, Brian Paquelet Wuetz, Charles Jeon, Farhana Sheikh, Esdras Juarez-Hernandez, Brando Perez Esparza, Huzaifa Rampurawala, Brent Carlton, Surej Ravikumar, Carlos Nieva, Sungwon Kim, Hyung-Jin Lee, Amir Sammak, Giordano Scappucci, Menno Veldhorst, Fabio Sebastiano, Masoud Babaie, Stefano Pellerano, Edoardo Charbon, and Lieven M. K. Vandersypen. CMOS-based cryogenic control of silicon quantum circuits. *Nature*, 593(7858):205–210, May 2021. Publisher: Nature Publishing Group.

[10] Stephan G. J. Philips, Mateusz T. Mądzik, Sergey V. Amitonov, Sander L. de Snoo, Maximilian Russ, Nima Kalhor, Christian Volk, William I. L. Lawrie, Delphine Brousse, Larysa Tryputen, Brian Paquelet Wuetz, Amir Sammak, Menno Veldhorst, Giordano Scappucci, and Lieven M. K. Vandersypen. Universal control of a six-qubit quantum processor in silicon. *Nature*, 609(7929):919–924, September 2022. Publisher: Nature Publishing Group.

[11] Samuel Neyens, Otto K. Zietz, Thomas F. Watson, Florian Luthi, Aditi Nethwewala, Hubert C. George, Eric Henry, Mohammad Islam, Andrew J. Wagner, Felix Borjans, Elliot J. Connors, J. Corrigan, Matthew J. Curry, Daniel Keith, Roza Kotlyar, Lester F. Lampert, Mateusz T. Mądzik, Kent Millard, Fahd A. Mohiyaddin, Stefano Pellerano, Ravi Pillarisetty, Mick Ramsey, Rostyslav Savytskyy, Simon Schaal, Guoji Zheng, Joshua Ziegler, Nathaniel C. Bishop, Stephanie Bojarski, Jeanette Roberts, and James S. Clarke. Probing single electrons across 300-mm spin qubit wafers. *Nature*, 629(8010):80–85, May 2024. Publisher: Nature Publishing Group.

[12] Paul Steinacker, Nard Dumoulin Stuyck, Wee Han Lim, Tuomo Tanttu, MengKe Feng, Andreas Nick, Santiago Serrano, Marco Candido, Jesus D. Cifuentes, Fay E. Hudson, Kok Wai Chan, Stefan Kubicek, Julien Jussot, Yann Canvel, Sofie Beyne, Yosuke Shimura, Roger Loo, Clement Godfrin, Bart Raes, Sylvain Baudot, Danny Wan, Arne Laucht, Chih Hwan Yang, Andre Saraiva, Christopher C. Escott, Kristiaan De Greve, and Andrew S. Dzurak. A 300 mm foundry silicon spin qubit unit cell exceeding 99% fidelity in all operations, October 2024. arXiv:2410.15590.

[13] Andrea Morello, Jarryd J. Pla, Floris A. Zwanenburg, Kok W. Chan, Kuan Y. Tan, Hans Huebl, Mikko Möttönen, Christopher D. Nugroho, Changyi Yang, Jessica A. van Donkelaar, Andrew D. C. Alves, David N. Jamieson, Christopher C. Escott, Lloyd C. L. Hollenberg, Robert G. Clark, and Andrew S. Dzurak. Single-shot readout of an electron spin in silicon. *Nature*, 467(7316):687–691, October 2010. Publisher: Nature Publishing Group.

[14] M. J. Curry, M. Rudolph, T. D. England, A. M. Mounce, R. M. Jock, C. Bureau-Oxton, P. Harvey-Collard, P. A. Sharma, J. M. Anderson, D. M. Campbell, J. R. Wendt, D. R. Ward, S. M. Carr, M. P. Lilly, and M. S. Carroll. Single-Shot Readout Performance of Two Heterojunction-Bipolar-Transistor Amplification Circuits at Millikelvin Temperatures. *Scientific Reports*, 9(1):16976, November 2019. Publisher: Nature Publishing Group.

[15] Elliot J. Connors, JJ Nelson, and John M. Nichol. Rapid High-Fidelity Spin-State Readout in $\mathrm{Si}$/$\mathrm{Si}$-$\mathrm{Ge}$ Quantum Dots via rf Reflectometry. *Physical Review Applied*, 13(2):024019, February 2020. Publisher: American Physical Society.

[16] N. W. Hendrickx, W. I. L. Lawrie, L. Petit, A. Sammak, G. Scappucci, and M. Veldhorst. A single-hole spin qubit. *Nature Communications*, 11(1):3478, July 2020. Publisher: Nature Publishing Group.

[17] G. A. Oakes, V. N. Ciriano-Tejel, D. F. Wise, M. A. Fogarty, T. Lundberg, C. Lainé, S. Schaal, F. Martins, D. J. Ibberson, L. Hutin, B. Bertrand, N. Stelmashenko, J. W. A. Robinson, L. Ibberson, A. Hashim, I. Siddiqi, A. Lee, M. Vinet, C. G. Smith, J. J. L. Morton, and M. F. Gonzalez-Zalba. Fast High-Fidelity Single-Shot Readout of Spins in Silicon Using a Single-Electron Box. *Physical Review X*, 13(1):011023, February 2023. Publisher: American Physical Society.

[18] Quentin Schmidt, Baptiste Jadot, Brian Martinez, Thomas Houriez, Adrien Morel, Tristan Meunier, Gaël Pillonnet, Gérard Billiot, Aloysius Jansen, Xavier Jehl, Yvain Thonnart, and Franck Badets. Compact frequency multiplexed readout of silicon quantum dots in monolithic FDSOI 28nm technology. In *2024 IEEE European Solid-State Electron-*

*ics Research Conference (ESSERC)*, pages 157–160, September 2024. ISSN: 2643-1319.

[19] Inga Seidler, Tom Struck, Ran Xue, Niels Focke, Stefan Trellenkamp, Hendrik Bluhm, and Lars R. Schreiber. Conveyor-mode single-electron shuttling in Si/SiGe for a scalable quantum computing architecture. *npj Quantum Information*, 8(1):100, August 2022. Publisher: Nature Publishing Group.

[20] Maxim De Smet, Yuta Matsumoto, Anne-Marije J. Zwerver, Larysa Tryputen, Sander L. de Snoo, Sergey V. Amitonov, Amir Sammak, Nodar Samkharadze, Önder Gül, Rick N. M. Wasserman, Maximilian Rimbach-Russ, Giordano Scappucci, and Lieven M. K. Vandersypen. High-fidelity single-spin shuttling in silicon, June 2024. arXiv:2406.07267 [cond-mat].

[21] Jelmer M. Boter, Juan P. Dehollain, Jeroen P.G. Van Dijk, Yuanxing Xu, Toivo Hensgens, Richard Versluis, Henricus W.L. Naus, James S. Clarke, Menno Veldhorst, Fabio Sebastiano, and Lieven M.K. Vandersypen. Spiderweb Array: A Sparse Spin-Qubit Array. *Physical Review Applied*, 18(2):024053, August 2022.

[22] Zhenyu Cai, Adam Siegel, and Simon Benjamin. Looped Pipelines Enabling Effective 3D Qubit Lattices in a Strictly 2D Device. *PRX Quantum*, 4(2):020345, June 2023. Publisher: American Physical Society.

[23] Armands Strikis and Lucas Berent. Quantum Low-Density Parity-Check Codes for Modular Architectures. *PRX Quantum*, 4(2):020321, May 2023. Publisher: American Physical Society.

[24] Matthias Künne, Alexander Willmes, Max Oberländer, Christian Gorjaew, Julian D. Teske, Harsh Bhardwaj, Max Beer, Eugen Kammerloher, René Otten, Inga Seidler, Ran Xue, Lars R. Schreiber, and Hendrik Bluhm. The SpinBus architecture for scaling spin qubits with electron shuttling. *Nature Communications*, 15(1):4977, June 2024. Publisher: Nature Publishing Group.

[25] Adam Siegel, Armands Strikis, and Michael Fogarty. Towards Early Fault Tolerance on a $2\ifmmode\times\else\texttimes\fi{}N$ Array of Qubits Equipped with Shuttling. *PRX Quantum*, 5(4):040328, November 2024. Publisher: American Physical Society.

[26] Adam Siegel, Zhenyu Cai, Hamza Jnane, Balint Koczor, Shaun Pexton, Armands Strikis, and Simon Benjamin. Snakes on a Plane: mobile, low dimensional logical qubits on a 2D surface, January 2025. arXiv:2501.02120 [quant-ph] version: 1.

[27] Eric Henry. 2D Silicon Spin Qubits. In *APS March Meeting 2025*, MAR-C19:8. American Physical Society, March 2025.

[28] Hubert C. George, Mateusz T. Mądzik, Eric M. Henry, Andrew J. Wagner, Mohammad M. Islam, Felix Borjans, Elliot J. Connors, J. Corrigan, Matthew Curry, Michael K. Harper, Daniel Keith, Lester Lampert, Florian Luthi, Fahd A. Mohiyaddin, Sandra Murcia, Rohit Nair, Rambert Nahm, Aditi Nethwewala, Samuel Neyens, Bishnu Patra, Roy D. Raharjo, Carly Rogan, Rostyslav Savytskyy, Thomas F. Watson, Josh Ziegler, Otto K. Zietz, Stefano Pellerano, Ravi Pillarisetty, Nathaniel C. Bishop, Stephanie A. Bojarski, Jeanette Roberts, and James S. Clarke. 12-Spin-Qubit Arrays Fabricated on a 300 mm Semiconductor Manufacturing Line. *Nano Letters*, 25(2):793–799, January 2025. Publisher: American Chemical Society.

[29] Sieu D. Ha, Edwin Acuna, Kate Raach, Zachery T. Bloom, Teresa L. Brecht, James M. Chappell, Maxwell D. Choi, Justin E. Christensen, Ian T. Counts, Dominic Daprano, J. P. Dodson, Kevin Eng, David J. Fialkow, Christina A. C. Garcia, Wonill Ha, Thomas R. B. Harris, nathan holman, Isaac Khalaf, Justine W. Matten, Christi A. Peterson, Clifford E. Plesha, Matthew J. Ruiz, Aaron Smith, Bryan J. Thomas, Samuel J. Whiteley, Thaddeus D. Ladd, Michael P. Jura, Matthew T. Rakher, and Matthew G. Borselli. Two-dimensional Si spin qubit arrays with multilevel interconnects, February 2025. arXiv:2502.08861 [quant-ph].

[30] Guido Burkard, Thaddeus D. Ladd, Andrew Pan, John M. Nichol, and Jason R. Petta. Semiconductor spin qubits. *Reviews of Modern Physics*, 95(2):025003, June 2023. Publisher: American Physical Society.

[31] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, September 2002.

[32] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, September 2012.

[33] Aaron J. Weinstein, Matthew D. Reed, Aaron M. Jones, Reed W. Andrews, David Barnes, Jacob Z. Blumoff, Larken E. Euliss, Kevin Eng, Bryan H. Fong, Sieu D. Ha, Daniel R. Hulbert, Clayton A. C. Jackson, Michael Jura, Tyler E. Keating, Joseph Kerckhoff, Andrey A. Kiselev, Justine Matten, Golam Sabbir, Aaron Smith, Jeffrey Wright, Matthew T. Rakher, Thaddeus D. Ladd, and Matthew G. Borselli. Universal logic with encoded spin qubits in silicon. *Nature*, 615(7954):817–822, March 2023. Publisher: Nature Publishing Group.

[34] Yi-Hsien Wu, Leon C. Camenzind, Patrick Bütler, Ik Kyeong Jin, Akito Noiri, Kenta Takeda, Takashi Nakajima, Takashi Kobayashi, Giordano Scappucci, Hsi-Sheng Goan, and Seigo Tarucha. Simultaneous High-Fidelity Single-Qubit Gates in a Spin Qubit Array, July 2025. arXiv:2507.11918 [quant-ph].

[35] Joseph D. Broz, Jesse C. Hoke, Edwin Acuna, and Jason R. Petta. Demonstration of an always-on exchange-only spin qubit, August 2025. arXiv:2508.01033 [quant-ph].

[36] Paul Steinacker, Nard Dumoulin Stuyck, Wee Han Lim, Tuomo Tanttu, MengKe Feng, Santiago Serrano, Andreas Nickl, Marco Candido, Jesus D. Cifuentes, Ensar Vahapoglu, Samuel K. Bartee, Fay E. Hudson, Kok Wai Chan, Stefan Kubicek, Julien Jussot, Yann Canvel, Sofie Beyne, Yosuke Shimura, Roger Loo, Clement Godfrin, Bart Raes, Sylvain Baudot, Danny Wan, Arne Laucht, Chih Hwan Yang, Andre Saraiva, Christopher C. Escott, Kristiaan De Greve, and Andrew S. Dzurak. Industry-compatible silicon spin-qubit unit cells exceeding 99% fidelity. *Nature*, 646(8083):81–87, October 2025. Publisher: Nature Publishing Group.

[37] D. Bacon, J. Kempe, D. A. Lidar, and K. B. Whaley. Universal Fault-Tolerant Quantum Computation on Decoherence-Free Subspaces. *Physical Review Letters*, 85(8):1758–1761, August 2000. Publisher: American Physical Society.

[38] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley. Universal quantum computation with the exchange interaction. *Nature*, 408(6810):339–342, November 2000. Publisher: Nature Publishing Group.

[39] Mauricio Gutiérrez, Juan S. Rojas-Arias, David Obando, and Chien-Yuan Chang. Comparison of spin-qubit architectures for quantum error-correcting codes, June 2025. arXiv:2506.17190 [quant-ph].

[40] Daniel P. Siewiorek, C. Gordon. Bell, and Allen. Newell. Computer structures : principles and examples / [edited by] daniel P. Siewiorek, C. Gordon bell, and allen newell. In *Computer structures : principles and examples*, McGraw-Hill computer science series. McGraw-Hill Book Company, Principles and examples, 1982. tex.lccn: 80027926.

[41] Mark Oskin, Frederic T. Chong, Isaac L. Chuang, and John Kubiatowicz. Building quantum wires: the long and the short of it. *SIGARCH Comput. Archit. News*, 31(2):374–387, May 2003.

[42] Neil H. E. Weste and David Money Harris. *CMOS VLSI design: a circuits and systems perspective*. Addison Wesley, Boston, 4th ed edition, 2011.

[43] Hossein SeyyedAghaei, Mahmood Naderan-Tahan, Magnus Jahre, and Lieven Eeckhout. Memory-Centric MCM-GPU Architecture. *IEEE Computer Architecture Letters*, 24(1):101–104, January 2025.

[44] Aditya Sridharan Iyer, Daehyun Kim, Saibal Mukhopadhyay, and Sung Kyu Lim. Multi-Tier 3D SRAM Module Design: Targeting Bit-Line and Word-Line Folding. In *Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design*, number 40, pages 1–9. Association for Computing Machinery, New York, NY, USA, April 2025.

[45] Dominic Horsman, Austin G. Fowler, Simon Devitt, and Rodney Van Meter. Surface code quantum computing by lattice surgery. *New Journal of Physics*, 14(12):123011, December 2012. arXiv:1111.4022 [quant-ph].

[46] Austin G. Fowler and Craig Gidney. Low overhead quantum computation using lattice surgery, August 2019. arXiv:1808.06709 [quant-ph].

[47] Daniel Litinski. A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery. *Quantum*, 3:128, March 2019. arXiv:1808.02892 [cond-mat, physics:quant-ph].

[48] A. Elsayed, M. M. K. Shehata, C. Godfrin, S. Kubicek, S. Massar, Y. Canvel, J. Jussot, G. Simion, M. Mongillo, D. Wan, B. Govoreanu, I. P. Radu, R. Li, P. Van Dorpe, and K. De Greve. Low charge noise quantum dots with industrial CMOS manufacturing. *npj Quantum Information*, 10(1):70, July 2024.

[49] A.M.J. Zwerver, S.V. Amitonov, S.L. De Snoo, M.T. Mądzik, M. Rimbach-Russ, A. Sammak, G. Scappucci, and L.M.K. Vandersypen. Shuttling an Electron Spin through a Silicon Quantum Dot Array. *PRX Quantum*, 4(3):030303, July 2023.

[50] Akito Noiri, Kenta Takeda, Takashi Nakajima, Takashi Kobayashi, Amir Sammak, Giordano Scappucci, and Seigo Tarucha. A shuttling-based

two-qubit logic gate for linking distant silicon quantum processors. *Nature Communications*, 13(1):5740, September 2022. Publisher: Nature Publishing Group.

[51] 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(12):986–991, December 2014.

[52] 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(12):981–985, December 2014.

[53] Filip K. Malinowski, Frederico Martins, Peter D. Nissen, Edwin Barnes, Łukasz Cywiński, Mark S. Rudner, Saeed Fallahi, Geoffrey C. Gardner, Michael J. Manfra, Charles M. Marcus, and Ferdinand Kuemmeth. Notch filtering the nuclear environment of a spin qubit. *Nature Nanotechnology*, 12(1):16–20, January 2017.

[54] Arne Laucht, Rachpon Kalra, Stephanie Simmons, Juan P. Dehollain, Juha T. Muhonen, Fahd A. Mohiyaddin, Solomon Freer, Fay E. Hudson, Kohei M. Itoh, David N. Jamieson, Jeffrey C. McCallum, Andrew S. Dzurak, and A. Morello. A dressed spin qubit in silicon. *Nature Nanotechnology*, 12(1):61–66, January 2017.

[55] Bo Sun, Teresa Brecht, Bryan H. Fong, Moonmoon Akmal, Jacob Z. Blumoff, Tyler A. Cain, Faustin W. Carter, Dylan H. Finestone, Micha N. Fireman, Wonill Ha, Anthony T. Hatke, Ryan M. Hickey, Clayton A. C. Jackson, Ian Jenkins, Aaron M. Jones, Andrew Pan, Daniel R. Ward, Aaron J. Weinstein, Samuel J. Whiteley, Parker Williams, Matthew G. Borselli, Matthew T. Rakher, and Thaddeus D. Ladd. Full-Permutation Dynamical Decoupling in Triple-Quantum-Dot Spin Qubits. *PRX Quantum*, 5(2):020356, June 2024. Publisher: American Physical Society.

[56] K. Takeda, A. Noiri, J. Yoneda, T. Nakajima, and S. Tarucha. Resonantly Driven Singlet-Triplet Spin Qubit in Silicon. *Physical Review Letters*, 124(11):117701, March 2020. Publisher: American Physical Society.

[57] Adam R. Mills, Charles R. Guinn, Michael J. Gullans, Anthony J. Sigillito, Mayer M. Feldman, Erik Nielsen, and Jason R. Petta. Two-qubit silicon quantum processor with operation fidelity exceeding 99%. *Science Advances*, 8(14):eabn5130, April 2022. Publisher: American Association for the Advancement of Science.

[58] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven. Single-shot read-out of an individual electron spin in a quantum dot. *Nature*, 430(6998):431–435, July 2004. Publisher: Nature Publishing Group.

[59] Jacob Z. Blumoff, Andrew S. Pan, Tyler E. Keating, Reed W. Andrews, David W. Barnes, Teresa L. Brecht, Edward T. Croke, Larken E. Euliss, Jacob A. Fast, Clayton A.C. Jackson, Aaron M. Jones, Joseph Kerckhoff, Robert K. Lanza, Kate Raach, Bryan J. Thomas, Roland Velunta, Aaron J. Weinstein, Thaddeus D. Ladd, Kevin Eng, Matthew G. Borselli, Andrew T. Hunter, and Matthew T. Rakher. Fast and High-Fidelity State Preparation and Measurement in Triple-Quantum-Dot Spin Qubits. *PRX Quantum*, 3(1):010352, March 2022.

[60] D. Keith, M. G. House, M. B. Donnelly, T. F. Watson, B. Weber, and M. Y. Simmons. Single-Shot Spin Readout in Semiconductors Near the Shot-Noise Sensitivity Limit. *Physical Review X*, 9(4):041003, October 2019. Publisher: American Physical Society.

[61] Rubén M. Otxoa, Josu Etxezarreta Martinez, Paul Schnabl, Normann Mertig, Charles Smith, and Frederico Martins. SpinHex: A low-crosstalk, spin-qubit architecture based on multi-electron couplers, April 2025. arXiv:2504.03149 [quant-ph].

[62] Madelyn Cain, Chen Zhao, Hengyun Zhou, Nadine Meister, J. Pablo Bonilla Ataides, Arthur Jaffe, Dolev Bluvstein, and Mikhail D. Lukin. Correlated decoding of logical algorithms with transversal gates, March 2024. arXiv:2403.03272 [cond-mat, physics:quant-ph].

[63] Hengyun Zhou, Chen Zhao, Madelyn Cain, Dolev Bluvstein, Casey Duckering, Hong-Ye Hu, Sheng-Tao Wang, Aleksander Kubica, and Mikhail D. Lukin. Algorithmic Fault Tolerance for Fast Quantum Computing, June 2024. arXiv:2406.17653 [quant-ph].

[64] Craig Gidney. Stim: a fast stabilizer circuit simulator. *Quantum*, 5:497, July 2021. arXiv:2103.02202 [quant-ph].

[65] Oscar Higgott. PyMatching: A Python package for decoding quantum codes with minimum-weight perfect matching, July 2021. arXiv:2105.13082 [quant-ph].

[66] F. Setiawan, Hoi-Yin Hui, J. P. Kestner, Xin Wang, and S. Das Sarma. Robust two-qubit gates for exchange-coupled qubits. *Physical Review B*, 89(8):085314, February 2014.

[67] Jason D. Chadwick, Gian Giacomo Guerreschi, Florian Luthi, Mateusz T. Mądzik, Fahd A. Mohiyaddin, Prithviraj Prabhu, Albert T. Schmitz, Andrew Litteken, Shavindra Premaratne, Nathaniel C. Bishop, Anne Y. Matsuura, and James S. Clarke. Short two-qubit pulse sequences for exchange-only spin qubits in two-dimensional layouts. *Physical Review A*, 111(5):052616, May 2025. Publisher: American Physical Society.

[68] Xiao-Gang Wen. Quantum Orders in an Exact Soluble Model. *Physical Review Letters*, 90(1):016803, January 2003. Publisher: American Physical Society.

[69] Alexey A. Kovalev, Ilya Dumer, and Leonid P. Pryadko. Design of additive quantum codes via the code-word-stabilized framework. *Physical Review A*, 84(6):062319, December 2011. Publisher: American Physical Society.

[70] Barbara M. Terhal, Fabian Hassler, and David P. DiVincenzo. From Majorana fermions to topological order. *Physical Review Letters*, 108(26):260504, June 2012. Publisher: American Physical Society.

[71] J. Pablo Bonilla Ataides, David K. Tuckett, Stephen D. Bartlett, Steven T. Flammia, and Benjamin J. Brown. The XZZX surface code. *Nature Communications*, 12(1):2172, April 2021. Publisher: Nature Publishing Group.

[72] Bence Hetényi and James R. Wootton. Tailoring quantum error correction to spin qubits. *Physical Review A*, 109(3):032433, March 2024.

[73] Bálint Domokos, Áron Márton, and János K. Asbóth. Characterization of errors in a CNOT between surface code patches, May 2024. arXiv:2405.05337 [quant-ph].

[74] Joschka Roffe, David R. White, Simon Burton, and Earl Campbell. Decoding across the quantum low-density parity-check code landscape. *Physical Review Research*, 2(4):043423, December 2020. Publisher: American Physical Society.

[75] Joschka Roffe. LDPC: Python tools for low density parity check codes, 2022.

[76] Sergey Bravyi, Andrew W. Cross, Jay M. Gambetta, Dmitri Maslov, Patrick Rall, and Theodore J. Yoder. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627(8005):778–782, March 2024. Publisher: Nature Publishing Group.

[77] Austin G. Fowler, David S. Wang, Charles D. Hill, Thaddeus D. Ladd, Rodney Van Meter, and Lloyd C. L. Hollenberg. Surface Code Quantum Communication. *Physical Review Letters*, 104(18):180503, May 2010. Publisher: American Physical Society.

[78] Willers Yang and Patrick Rall. Harnessing the Power of Long-Range Entanglement for Clifford Circuit Synthesis. *IEEE Transactions on Quantum Engineering*, 5:1–10, 2024.

[79] Elisa Bäumer and Stefan Woerner. Measurement-based long-range entangling gates in constant depth. *Physical Review Research*, 7(2):023120, May 2025.

[80] Rich Rines, Benjamin Hall, Mariesa H. Teo, Joshua Viszlai, Daniel C. Cole, David Mason, Cameron Barker, Matt J. Bedalov, Matt Blakely, Tobias Bothwell, Caitlin Carnahan, Frederic T. Chong, Samuel Y. Eubanks, Brian Fields, Matthew Gillette, Palash Goiporia, Pranav Gokhale, Garrett T. Hickman, Marin Iliev, Eric B. Jones, Ryan A. Jones, Kevin W. Kuper, Stephanie Lee, Martin T. Lichtman, Kevin Loeffler, Nate Mackintosh, Farhad Majdeteimouri, Peter T. Mitchell, Thomas W. Noel, Ely Novakoski, Victory Omole, David Owusu-Antwi, Alexander G. Radnaev, Anthony Reiter, Mark Saffman, Bharath Thotakura, Teague Tomesh, and Ilya Vinogradov. Demonstration of a Logical Architecture Uniting Motion and In-Place Entanglement: Shor's Algorithm, Constant-Depth CNOT Ladder, and Many-Hypercube Code, September 2025. arXiv:2509.13247 [quant-ph].

[81] Gushu Li, Yufei Ding, and Yuan Xie. Tackling the Qubit Mapping Problem for NISQ-Era Quantum Devices. In *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS '19, pages 1001–1014, New York, NY, USA, April 2019. Association for Computing Machinery.

[82] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, February 2005. Publisher: American Physical Society.

[83] Daniel Litinski. Magic State Distillation: Not as Costly as You Think. *Quantum*, 3:205, December 2019. arXiv:1905.06903 [quant-ph].

[84] Nicholas Fazio, Mark Webster, and Zhenyu Cai. Low-overhead Magic State Circuits with Transversal CNOTs, October 2025. arXiv:2501.10291 [quant-ph].
[85] Jonathan E. Moussa. Transversal Clifford gates on folded surface codes. *Physical Review A*, 94(4):042316, October 2016.
[86] Zi-Han Chen, Ming-Cheng Chen, Chao-Yang Lu, and Jian-Wei Pan. Transversal Logical Clifford gates on rotated surface codes with reconfigurable neutral atom arrays, December 2024. arXiv:2412.01391 [quant-ph] version: 1.
[87] Ying Li. A magic state's fidelity can be superior to the operations that created it. *New Journal of Physics*, 17(2):023037, February 2015.
[88] Craig Gidney, Noah Shutty, and Cody Jones. Magic state cultivation: growing T states as cheap as CNOT gates, September 2024. arXiv:2409.17595 [quant-ph].