Near-Term Spin-Qubit Architecture Design via Multipartite Maximally Entangled States

N. Paraskevopoulos[1,2,†] M. Steinberg[1,2,*,†] B. Undseth[1,3] A. Sarkar[1,2]
L.M.K. Vandersypen[1,3] X. Xue[1,3] and S. Feld[1,2]

[1]*QuTech, Delft University of Technology, Delft, Netherlands*

[2]*Quantum and Computer Engineering Department, Delft University of Technology, Delft, Netherlands*

[3]*Kavli Institute of Nanoscience, Delft University of Technology, Delft, Netherlands*

The design and benchmarking of quantum computer architectures traditionally rely on practical hardware restrictions, such as gate fidelities, control, and cooling. At the theoretical and software levels, numerous approaches have been proposed for benchmarking quantum devices, ranging from, *inter alia*, quantum volume to randomized benchmarking. In this work, we utilize the quantum information-theoretic properties of multipartite maximally entangled quantum states, in addition to their correspondence with quantum error-correction codes, permitting us to quantify the entanglement generated on near-term bilinear spin-qubit architectures. For this aim, we introduce four metrics that ascertain the quality of genuine multipartite quantum entanglement, along with circuit-level fidelity measures. As part of the task of executing a quantum circuit on a device, we devise simulations, which combine expected hardware characteristics of spin-qubit devices with appropriate compilation techniques; we then analyze three different architectural choices of varying lattice sizes for bilinear arrays, under three increasingly realistic noise models. We find that if the use of a compiler is assumed, sparsely connected spin-qubit lattices can approach comparable values of our metrics to those of the most highly connected device architecture. Even more surprisingly, by incorporating crosstalk into our last noise model, we find that, as error rates for crosstalk approach realistic values, the benefits of utilizing a bilinear array with advanced connectivity vanish. Our results highlight the limitations of adding local connectivity to near-term spin-qubit devices, and can be readily adapted to other qubit technologies. The framework developed here can be used for analyzing quantum entanglement on a device before fabrication, informing experimentalists on concomitant realistic expectations.

I. INTRODUCTION

Quantum computers promise to solve certain classes of problems more efficiently than classical computers [1]. While great progress has been made in effectively scaling the physical size of a quantum system, it is not yet clear how to optimally arrange and connect qubits on a lattice, given a particular qubit technology. Indeed, connectivity on solid-state quantum hardware is often determined by the ease of the fabrication process. New manufacturing techniques and qubit technologies are constantly being explored, and the design and benchmarking of a quantum device traditionally takes into account hardware-specific variables such as control electronics, power dissipation, calibration, and crosstalk suppression. However, as a quantum device will ideally be fabricated for the execution of real quantum algorithms involving the generation of highly entangled many-body quantum states, benchmarking and design methods should also factor in how well a quantum device can generate *genuine multipartite entanglement* (i.e., entanglement characterized by quantum correlations between each party in a multiparty state [2,3]).

How such entanglement is generated varies substantially between different flavors of quantum hardware. As a concrete example, superconducting qubits with nearest-neighbor interactions distribute entanglement very differently from trapped ions with all-to-all connectivity. Furthermore, different architectures of a specific quantum

hardware platform may be better or worse at distributing entanglement; neutral atoms [4], trapped ions [5], and semiconductor spins [6,7] can be physically shuttled for dynamic qubit connectivity, enabling more flexibility in certain respects. The ability for these platforms to generate multipartite entanglement therefore depends greatly on the precise architecture, the quality of the primitive operations, and the compilation of the quantum circuit to be executed.

Several methods for benchmarking quantum computer performance currently exist; these range from *randomized benchmarking* [8,9] and various forms of *quantum tomography* [10], to other methods such as calculating the *estimated success probability* (ESP) [11–14] used at the compiler level, *quantum volume* [15] and its derivatives [16,17], as well as through demonstrations of *quantum advantage* [18] and application-level benchmarks [19]. Although there is no shortage of benchmarking techniques, neither is there a current consensus regarding the performance qualification of a quantum device, especially when taking into account different types of qubit technologies, architectural schemes, as well as available elementary gate sets and the topological-graph properties of quantum algorithms [20,21]. In this way, all current benchmarking procedures and protocols exhibit their own advantages and disadvantages; as an example, *gate set tomography* is known to provide self-consistent characterizations of physical gates; unfortunately, most techniques are known to scale exponentially in the number of qubits [22]. Another example, quantum volume, takes into account many essential metrics that factor into processor performance: circuit depth, gate fidelity, qubit count, and qubit connectivity; however, the circuits crafted using this technique are usually not representative of realistic quantum algorithms [15]. This fact has spurred interest in the development of alternative methods for benchmarking at the algorithmic level for existing quantum processors [23–28].

It was first proposed in Ref. [29] that one might be able to utilize *multipartite maximally entangled* (MME) states, in particular, *absolutely maximally entangled* (AME) quantum states, as a suitable and generalizable benchmark for the quality of entanglement generated on a quantum device [29]. The main reason for utilizing AME states lies in the implementation details: the authors describe the preparation of highly entangled states as difficult tasks for today's devices, mainly due to the genuine multipartite nature of the quantum entanglement required among all subsystems [2]. Additionally, utilizing AME states carries several other advantages. Firstly, it is known that AME states and their less highly entangled counterparts, *k-uniform* states, exhibit an exact mathematical duality to *maximal distance separable* (MDS) codes; MDS codes are known to exhibit the highest distance allowed by the *quantum Singleton bound*, and are therefore optimal, in terms of their code-theoretic properties [30–32]. Benchmarking

or designing a quantum device with respect to these many-body states is then directly related to understanding how resilient logical qubits can be fashioned on a device. Secondly, AME states are, by definition, the most highly entangled quantum states possible for a given number of qudits. In this way, using AME and other highly entangled MME states for entanglement characterization on a noisy device provides a benchmark that employs the entire multipartite Hilbert space available in order to generate a state, rather than a subset of qubits. This relationship was illustrated in a recent experiment measuring objects related to *quantum weight enumerators* of AME and *k-uniform* states, and conveyed a direct method for entanglement characterization via a relationship to the *quantum concurrence*, a known entanglement monotone [33].

In the practical implementation of a quantum algorithm, one must also consider not just the difficulty of realizing the quantum circuit generating a particular output state, but just as importantly, the specific compilation requirements arising from the choice of qubit technology, architectural connectivity, and the constraints that follow therein. The challenge of efficient quantum compilation is essential to assess, even at the level of *logical qubits*, since it is understood that concepts from *quantum error correction* (QEC) currently offer the only feasible way by which a quantum computer could be scaled to the large system sizes needed in order to execute useful quantum algorithms [34]. Taking stock of these points, it seems reasonable to incorporate MME- and QEC-motivated design benchmarks, which factor in the benefits (as well as the burden) of quantum compilation, unifying multiple angles of entanglement characterization in a framework.

In this work, we propose a framework consisting of four quantum information-theoretic and compilation-motivated measures for designing and benchmarking quantum devices; our purpose here is not to propose a complete parameterization of device design, but rather to introduce mathematical concepts that can help to inform the design process. The central theme of our choice of measures lies in the fact that they exploit the highly entangled nature of MME states and their relationship to small quantum error-correction codes. Although the methods we introduce can be adapted to any qubit technology, we assess example designs for near-term *spin-qubit* architectures, focusing on those with *bilinear* qubit layout, but with varying connectivities, lattice array sizes, and ultimately, the qubit density (as it relates to the number of empty quantum dots on the device), following recent experimental trends [35–38]. As the spin-qubit community is on the verge of performing its first fault tolerance and QEC demonstrations beyond the repetition code [37,39–46], it is timely and prudent to consider just how much of a role local connectivity could play in the realization of near-term QEC experiments for spin-qubit technology, especially if we compare the relative ease of performing quantum

compilation techniques with the difficulty of fabricating entirely new processors with advanced connectivities.

More explicitly, we consider four main simulations. Firstly, we consider the generation of an AME state of six qubits, which is similar to what was carried out in Ref. [29], and was first discovered in Ref. [47]; we take the additional step of calculating a quantity known as the *Bell operator*, which is widely understood as a metric for gauging the *quantumness* of qubit-qubit correlations among the parties of a genuinely entangled multipartite state. Our second test involves the usage of an MME state related to AME states: the *k-uniform* state. Here, we map the state to the smallest error-detecting surface code (i.e., the $[[4, 1, 2]]$ surface code with three ancilla qubits, as described in Refs. [48,49]), and evaluate the logical qubit's failure rate over many cycles of stabilizer measurements; as it is known that AME states of four qubits do not exist [50], the *k-uniform* state that we have chosen also maximizes the entanglement possible in a four-party Hilbert space. Thirdly, we invoke a modified version of the *estimated success probability* (ESP), a standard quantum compilation measure used to grade the worst-case scenario of circuit performance on a device; our modification takes into account the *decoherence-induced* noise that is expected on near-term spin-qubit devices. Finally, the fourth test employs an *entropic divergence measure* known as the *tripartite mutual information*; this is known to indicate the degree to which quantum entanglement is distributed globally versus locally in a many-body quantum system undergoing unitary evolution and also periodic measurements [51,52].

Our main contributions are as follows. Firstly, we develop a framework for critiquing the design of spin-qubit architectures, utilizing the previously mentioned quantum-information-based objects, together with state-of-the-art spin-qubit compilation techniques [11,53]. Our goal is to evaluate several architectural proposals for bilinear spin-qubit arrays of varying lattice size; we fill these arrays with seven qubits, and then decrease the qubit density of the arrays by increasing the number of quantum dots in the structure. In the semiconductor spin-qubit field, the entire lattice is usually filled with qubits; however, our study shows large benefits to lowering the qubit density, allowing for increased compilation flexibility via *qubit-shuttling* techniques [6,54–56]. The innate flexibility in spin-qubit devices lies in stark contrast to connectivity-constrained devices, such as superconducting qubit technology, for which SWAP gates are needed.

Secondly, we perform a detailed analysis of the aforementioned near-term devices, providing guidance on which of the architectures considered is best suited for small QEC and *fault-tolerance* experiments with bilinear spin-qubit arrays. Our results show that, if we assume the inclusion of hardware-specific quantum compilation techniques, then it is possible for sparsely connected devices to achieve values in our metric that are comparable to those of the most highly connected spin-qubit device in our study. Even further, our simulations with crosstalk indicate that the benefits of adding complex local connectivity are outweighed by increasing crosstalk effects, in essence providing a limitation for architectural connectivity in spin-qubit devices. Our work strongly implies that there are constraints on how much local connectivity can benefit a spin-qubit device at the moment of generating genuine multipartite entanglement, and that appropriate compilation techniques can compensate for the lack of connectivity in an architecture. Our results reveal that efficient quantum compilation techniques can aid in the realization of small quantum error-correction experiments better than advanced connectivity in the actual spin-qubit device.

This article is organized as follows. Section II introduces multipartite maximally entangled states, their relationship to quantum error-correction codes, and the entanglement metrics that we propose in our framework (Secs. II A–II C), as well as our modified version of estimated success probability that we employ in our study (Sec. II D). We additionally cover the basics of quantum compilation (Sec. II E) and some of the details regarding spin-qubit architectures (Sec. II F). Section III showcases the concrete results obtained for each of the simulation experiments; we follow up by analyzing and discussing each of these in Sec. IV. In Sec. V, we provide final comments and summarize potential future directions.

II. BACKGROUND

In what follows, we introduce the four main metrics utilized in this work. Additionally, we provide an introductory section on *quantum compilation* and its role in our investigation, as well as a review of the basics pertaining to spin-qubit architectures.

A. Multipartite maximally entangled states & quantum error correction

Multipartite maximally entangled (MME) states are generalized, highly entangled many-body quantum systems, for which it is understood that certain reductions of these states become *maximally mixed* [2]. Such quantum systems have been shown to exhibit intimate connections with many aspects of quantum information theory such as quantum error correction, as well as quantum communication [57–59] and toy models of quantum gravity theories such as the AdS/CFT correspondence [60–64]. Although formulating a completely exhaustive and general framework for all MME states is very complex, much progress has been made by studying subsets of states. Some examples of these subsets are as follows: *graph states* [65]; *tensor-network states* [66]; and certain classes of quantum error-correction codes, e.g., *maximal distance separable* (MDS) codes [30–32].

In the simplest case, the Bell state $|\Psi_2^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is an example of a maximally entangled quantum state. Performing a partial trace on either of the two subsystems (which we label A) results in a maximally mixed quantum state of the form

$$\mathrm{Tr}_A\big[\,|\Psi_2^+\rangle\,\langle\Psi_2^+|\,\big] \propto \mathbb{I}_2 \,. \qquad (1)$$

One can also define generalized, multipartite quantum systems, in which the potential bipartitions of such a state and the associated reduced density matrix yield a maximally mixed entanglement spectrum [2].

A prominent example of a generalized multipartite maximally entangled state is the *absolutely maximally entangled* (AME) state, which is an n-qudit state $|\psi\rangle$, defined in $\mathcal{H}_q^n := \mathbb{C}_q^{\otimes n}$ (wherein q denotes local dimension), for which the following condition holds:

$$\rho_s = \mathrm{Tr}_{s^c}\big[\,|\psi\rangle\,\langle\psi|\,\big] \propto \mathbb{I} \,, \qquad (2)$$

where the subset $s \subset \{1 \cdots n\}$, $|s| = \lfloor n/2 \rfloor$, and s^c denotes the complementary set of subsystems to s. For the present work, we use the shorthand form AME(n,q) to refer to an n-qudit, q local dimension AME state of interest.

Furthermore, AME states are closely related to quantum error-correction codes. In particular, it was shown in Ref. [31] that an exact mathematical duality exists between a given AME state, AME(n,q), and a *maximal distance separable* (MDS) code with parameters $[\![n,k,d]\!]_q$, where k,d represent the number of encoded logical qubits as well as the distance of the code. In particular, it was shown in Ref. [30] that *code shortening* techniques could be utilized in order to generate entire families of AME states with provably minimal support.

Constructions of particular AME states were provided in Ref. [29], and subsequent translations to the circuit level were given. These constructions mainly relied on mapping some AME(n,q) to a graph state, which are defined as an n-partite pure quantum state composed of n vertices $\mathcal{V} = \{v_1 \cdots v_i \cdots v_n\}$, and edges $\mathcal{E} = \{e_{ij} = \{v_i, v_j\}\}$. Each graph has an associated *adjacency matrix* A, whose entries A_{ij} satisfy $A_{ij} = 1$ if an edge exists; otherwise, the entry's value is zero (self interactions are forbidden in this formalism). As we will only be utilizing the AME(6,2) state in the present work, we define an n-qubit graph state as

$$|G\rangle = \prod_{i<j}^{n} \mathrm{CZ}_{ij}^{A_{ij}}\big[\,|+\rangle\,\big]^{\otimes n} \,. \qquad (3)$$

Note that AME states can be defined using the graph state formalism for any local dimension [32,67,68], with the well-known *cluster states* as a specific case [69].

The AME(6,2) state, given in Refs. [29,47], possesses the following stabilizers:

$$XZIIZZ \,, \qquad (4)$$
$$ZXZZII \,, \qquad (5)$$
$$IZXZIZ \,, \qquad (6)$$
$$IZZXZI \,, \qquad (7)$$
$$ZIIZXZ \,, \qquad (8)$$
$$ZIZIZX \,. \qquad (9)$$

The explicit form of the state is

$$
\begin{aligned}
|\Omega_{6,2}\rangle = \frac{1}{4}\big[\, & |000\rangle\,(|+-+\rangle + |-+-\rangle) \\
& - |001\rangle\,(|+--\rangle - |-++\rangle) \\
& + |010\rangle\,(|++-\rangle - |--+\rangle) \\
& - |011\rangle\,(|+++\rangle + |---\rangle) \\
& - |100\rangle\,(|+++\rangle - |---\rangle) \\
& - |101\rangle\,(|++-\rangle + |--+\rangle) \\
& - |110\rangle\,(|+--\rangle + |-++\rangle) \\
& - |111\rangle\,(|+-+\rangle - |-+-\rangle)\big] \,. \qquad (10)
\end{aligned}
$$

The circuit that generates the AME(6,2) is shown in Fig. 1. Here, the six data qubits in the AME(6,2) circuit are initialized in the $|+\rangle$ state and undergo the commensurate CZ
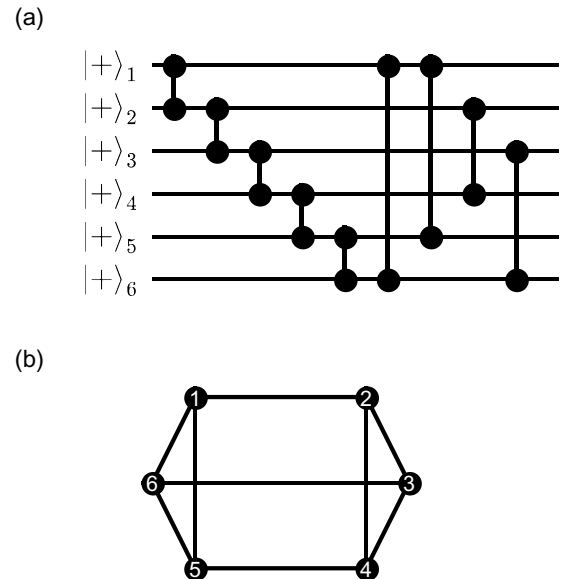
(a)

(b)

FIG. 1.   Generating circuit for the AME(6,2) state. (a) depicts the circuit-level implementation that was utilized for this work, and (b) is a graphical representation of the resulting *graph state* and its qubit-qubit entanglement correlations. Two-qubit gates are shown as *CZ* gates in (a).

gates shown in (a), resulting in entanglement correlations among the six parties, as shown in (b). The circuit and graph state form displayed follow from the definition of AME qubit graph states in Eq. (3).

Using graph states, one can derive the famous *Bell inequalities*, which are known to provide bounds on the correlations allowed by so-called *local-hidden-value* (LHV) theories. Quantum states exhibiting entanglement correlations are known to violate these inequalities, and graph states can be used to directly derive this fact [65,70]. Central to this derivation lies in the calculation of the *Bell operator* $B(G)$, which is defined as

$$B(G) = \sum_{i}^{2^n} S_i \,, \tag{11}$$

where S_i is a *stabilizer* of graph state G. The stabilizer operator is defined as

$$S_i = X_i \bigotimes_{j \in A_{ij}} Z_j \,, \tag{12}$$

which is typical for a graph state.

One can verify via direct computation that the number of stabilizers for a given graph state is exactly 2^n [65]. In this way, the Bell operator provides useful information on the *quantumness* of correlations between neighboring qubits; the Bell operator itself is bounded from above by 2^n, which represents purely quantum correlations, and from below by $2^{(n-1)/2}$, signifying purely classical correlations [29,70]. For the AME(6,2) state, the average expectation value of the Bell operator is found to be within the bounds $4 \leq \langle B \rangle \leq 64$.

It is well known that AME states do not exist for all local dimensions, given a number of qudits; as an example, AME qubit states only exist for $n = 2, 3, 5$, and 6 parties [50,71,72]. As a result, less highly entangled quantum states known as *k-uniform* states that generalize the AME condition were introduced [68]; these states in turn can be defined for any number of parties, for any local dimension. *k*-uniform states are defined as adhering to the same constraint as in Eq. (2), but, in contrast, the requirement $k = |s| \leq \lfloor n/2 \rfloor$ is relaxed. As an example, the famous *Greenberger-Horne-Zeilinger* (GHZ) state constitutes a 1-uniform state, as any reduction to one qubit results in a maximally mixed state [1,73]. In Ref. [74], a generalized formalism for constructing stabilizer quantum error-correction codes from *k*-uniform states was introduced.

B. *k*-uniform quantum states and the surface code

The *surface code* is a family of $[\![L^2 + (L - 1), 1, L]\!]$ stabilizer quantum error-correction codes (with L signifying the length of a square lattice) that originally arose from the celebrated *toric code* [75–77]. Originally arranged on a toroidal lattice, it was discovered thereafter that such a lattice orientation was unnecessary, and planar versions were proposed [78], which eventually became known as the *surface code*. These codes have been proposed and experimentally realized [48,49,79–81] as the current leading candidate for universal, fault-tolerant logical quantum computation [75,76], owing to their resilience against many varieties of noise, the development of *lattice-surgery* techniques [82], ease of experimental implementation via a planar layout [75], and *magic-state distillation* protocols [83], which provide for universal logical gate sets. Many variants of the surface code have been studied and proposed, including the *rotated* [84] and *XZZX* [85] versions, among others [86,87].

The smallest representative of the *rotated surface-code* family, with parameters $[\![d^2, 1, d]\!]$, is the $[\![4, 1, 2]\!]$ *quantum error-detecting* surface code, so named due to the distance of the code. The stabilizers of the code take the form:

$$XXXX\,, \tag{13}$$
$$IZIZ\,, \tag{14}$$
$$ZIZI\,, \tag{15}$$

and the logical operators are $IXIX, IIZZ$. Here we use the convention that $XX = X \otimes X$ for brevity. The code states are defined as

$$|\bar{0}\rangle = \frac{1}{\sqrt{2}} \big[|0000\rangle + |1111\rangle \big]\,, \tag{16}$$

$$|\bar{1}\rangle = \frac{1}{\sqrt{2}} \big[|0101\rangle + |1010\rangle \big]\,. \tag{17}$$

Using Eq. (1), one can easily check that the four-qubit state described above constitutes an example of a *planar 2-uniform quantum state* [88], since only tracing out any two adjacent subsystems leaves the remaining two maximally mixed. Since AME(4,2) states do not exist [50] and are equivalent to 2-uniform states of four parties [68], a planar 2-uniform state represents the maximal amount of genuine quantum entanglement that four qubits can share over an encoding of one logical qubit.

In our simulations, we prepare the encoded logical $|\bar{0}\rangle$ state in the usual manner (that is, with one round of stabilizer measurements), before performing several cycles of stabilizer measurements, as is standard in error-detection experiments [48,49,81]. We check the *logical success rate* and *tripartite mutual information* on a range of one to ten cycles; our reason for this is to evaluate the time dependence of our metrics and the potential for larger-depth quantum-circuit experiments in the selected near-term spin-qubit devices.

C. Entropic measures in monitored quantum circuits

The *von Neumann entropy* (VNE) is a fundamental measure of entanglement in quantum information theory [1], and takes the form

$$\mathcal{S}(\rho) = -\text{Tr}\big[\rho \log \rho\big] , \qquad (18)$$

where all logarithms here are natural bases, and ρ represents a density matrix.

The VNE plays an essential role in assessing the amount of quantum entanglement present in a state. By taking the partial trace of a density matrix, the resulting reduced density matrix's VNE (also known as the *entanglement entropy* of the remaining subsystems) indicates whether nontrivial quantum entanglement is present; if the result is nonzero, then the two bipartitions are in fact entangled. The entanglement entropy is thus bounded as $0 \le \mathcal{S}(\rho_s) \le \log n$, where s represents the set of remaining subsystems after a partial trace operation.

In the study of *monitored quantum circuits* (i.e., quantum circuits that are subject to frequent projective measurements) and *measurement-induced phase transitions* (MIPT), it is common to measure the *tripartite mutual information* [51,52], which is defined as

$$\mathcal{I}_3(A:B:C) = \mathcal{S}_A + \mathcal{S}_B + \mathcal{S}_C + \mathcal{S}_{ABC}$$
$$- \mathcal{S}_{AB} - \mathcal{S}_{BC} - \mathcal{S}_{AC} , \qquad (19)$$

where, for example, $\mathcal{S}_{ABC}$ represents the joint $A \cup B \cup C$ subsystem. If $\mathcal{I}_3 < 0$, then this behavior is generally ascribed to *genuine multipartite entanglement* within the *volume-law phase*, and is commonly considered as a *quantum error-correction (QEC) phase* of a quantum circuit [89–91], due to the proliferation of long-range entanglement present between subsystems (this is apparent since in a system with long-range entanglement, the total entanglement entropy for subsystems AB, BC, AC, and ABC must exceed that of subsystems A, B, and C). $\mathcal{I}_3 = 0$ is known to indicate strictly bipartite correlations among the residual subsystems, and $\mathcal{I}_3 > 0$ can be associated with only classical correlations in the *area-law phase* [92], wherein the individual subsystems begin to behave independently of each other.

One may ask why we do not employ a simpler analysis for estimating multipartite entanglement, such as the *bipartite mutual information* [93,94] or the scaling of entanglement entropy [95]. The goal of the present work is to assess how well quantum entanglement is distributed across a device, given connectivity, size, and shuttling constraints via compilation options. This setup naturally excludes the scaling of entanglement entropy, as such an analysis requires the evaluation of sequentially larger devices and circuits, in order to check if an MIPT is present; that is not the present goal of our work. Moreover, the bipartite mutual information has been used to

confirm the existence of an MIPT via bipartite quantum correlations; in our case, however, we wish to check and affirm how well *genuine multipartite entanglement* (i.e., multipartite entanglement that is shared among all parties, not just in a bipartite manner) exists after performing the circuits in question. As such, our analysis tools are well calibrated for the task at hand.

Other measures of quantum correlations such as *quantum discord*, *entanglement robustness* [96,97], or *concurrence* [33] exist; however, in this work, we opt to understand how well a certain spin-qubit architecture is able to generate quantum entanglement and stay inside of the volume-law (QEC) phase of a quantum circuit's phase diagram. In this sense, both quantum discord and entanglement robustness are not fitting, as the former concerns subtleties of quantum correlations that may not immediately pertain to entanglement, and the latter is related to entanglement properties of mixed states (which we effectively do not treat here). As for the concurrence, this entanglement monotone is related to bipartite entanglement correlations, which does not coincide with the goals of quantifying n-qubit entanglement correlations on a device architecture.

D. Estimated success probability

A standard method for evaluating an architecture's performance in the software regime is known as the *estimated success probability* (ESP) [11–13,98]. Although normally the ESP constitutes a simple multiplication of gate fidelities, we have provided a modification in this work, which can be expressed as

$$\Xi = \left[\prod_k \prod_i F_{i,k} \right] \times e^{-t/T_2} , \qquad (20)$$

where k represents the kth time step in the circuit execution, i the ith gate in the kth time step and F is the fidelity of the corresponding gates. The last term, inspired by Ref. [99] and utilized in Ref. [12], introduces *decoherence-induced errors*, which represent the probability of qubits staying coherent during the execution of the circuit. Here, T_2 and t represent the decoherence time and the circuit-time duration, respectively.

The original ESP model was introduced with only the first term, representing the worst-case scenario to execute a sequence of quantum gates successfully [14,100] on a specific architecture. Unlike other models, ESP does not explicitly account for phase additions or subtractions by gates or noise. Instead, it simplifies the representation by using a single fidelity value that encapsulates the overall operational quality of each gate. This makes ESP the least costly to calculate among the four metrics that we have introduced, as it only scales linearly in computational complexity with the number of gates contained in a circuit. It

does not suffer from the exponential growth of the Hilbert space with the increasing size and depth of a circuit and the accompanying device. Despite the simplicity of this first ESP version, some limited information on the circuit and the underlying device can still be derived indirectly. For instance, one architecture may possess a higher degree of connectivity than another; in this way, the former can achieve a reduced gate overhead when compared with the latter, thus scoring a higher ESP due to the necessity of fewer added gates.

From there, any modification can account for more information about the device's architecture. Taking Eq. (20) as an example, the parallelization capabilities of the device are indirectly reflected in t. More succinctly, when one device enables more gate executions within the same time frame as another, the former will exhibit a lower decoherence error rate than the latter. Another modification could add a term to model *crosstalk*—a common issue in realistic device platforms [101]. In fact, such a version has been successfully used as a figure of merit for spin-qubit architectural explorations in Ref. [12]. For instance, this modified ESP can reveal all kinds of architectural trade-offs. For example, maximizing the parallelization capabilities and/or qubit connectivity can incur a worse ESP result than anticipated due to increased crosstalk from more frequent close proximity interactions and more shared material components.

One should also note that ESP can easily approach zero, especially including the modifications, as it exponentially drops with an increasing number of gates. Although such low values do not have any physical meaning after a certain small number, ESP still remains a reliable way to rank architectures provided *numeric underflows* [102] are avoided. It is important to recognize that ESP is not inherently random, regardless of how small its value becomes. For example, even minor differences between two architectures, such as a single gate variation, will be consistently and reliably captured in their respective ESPs.

Overall, ESP and its modifications do not aim to absolutely predict but rather to establish a relative hierarchy of differentiation between architectures with low computational costs. However, the ESP, in and of itself, differs from the other three metrics, which can be summarized as follows. Firstly, and perhaps most importantly, ESP does not take into account notions of entanglement as in traditional quantum information-theoretic measures [1, 103]. Considering that the goal of our work is to scrutinize near-term spin-qubit architectures from the standpoint of genuine multipartite entanglement, ESP stands as a cheap alternative, but it still holds its value due to its strong correlation with the other metrics, as shown in Sec. III. Finally, and as explained before, no explicit parameters related to the device's architectural properties (such as the parallelization capabilities or connectivity) are inherently factored into the calculation of ESP; instead,

the effects of design features are reflected in the result indirectly.

E. Quantum compilation

Up until this point, we have not described our work as it relates to *quantum compilation*; indeed, this is an important factor when aiming to realize a quantum algorithm in circuit form on a quantum computer. Simply put, quantum compilation constitutes the various adjustments at the level of software to a quantum circuit in order to prepare and execute it on a quantum device [11,104–106]; this includes all of the subprocesses affecting a hardware-agnostic circuit, transforming it into a hardware-compatible version. Each of these procedures aims to maximize the success rate of the quantum circuit by utilizing optimization-based algorithms that take close account of the hardware's constraints [107,108]. Typically, these steps are subdivided into several general stages, which consist of the following, without regard to a particular ordering: (a) *elementary gate-set decomposition*, in which a quantum circuit is translated and simplified as much as possible in the quantum device's native gate set; (b) *scheduling*, wherein the logical time ordering of the circuit is considered, as well as the parallelism of gate operations and the shortening of circuit depth, among other factors; (c) *initial placement* (also known as *qubit assignment* or *initial assignment*), which assigns qubits initially from the circuit to the device; and (d) *qubit routing*, wherein qubits are brought in close proximity (usually adjacent) with the minimal use of a hardware's native communication method in order to facilitate two-qubit gate interactions of the quantum circuit.

One of the central subproblems in quantum compilation involves solving the *quantum circuit mapping problem* (QCMP). The QCMP can be defined as both the initial placement and qubit routing steps of compilation combined, and as explained before, is paramount to the circuit's success rate [109,110], especially in the *noisy intermediate-scale quantum* (NISQ) and *early fault-tolerant* (eFT) eras, where both QPUs and quantum circuits are relatively small, and trade-offs exist regarding the degree to which QEC and fault-tolerance techniques are utilized [111,112]. Lower bounds are known for the QCMP in terms of the number of SWAP gates needed to realize a circuit on a given quantum processor with finite connectivity [108]. However, in this work, instead of scrutinizing the number of SWAP gates added by a compiler, we analyze the number of *shuttles* needed, as it is known that spin-qubit quantum devices can accommodate such hardware-level operations, thereby incurring a smaller resource cost than SWAP gates in superconducting processors over sufficient length scales [113].

Figure 2 describes an example solution to the QCMP. (a) and (c) depict the hardware-agnostic circuit and graph-theoretic representations of the two-qubit gates utilized in
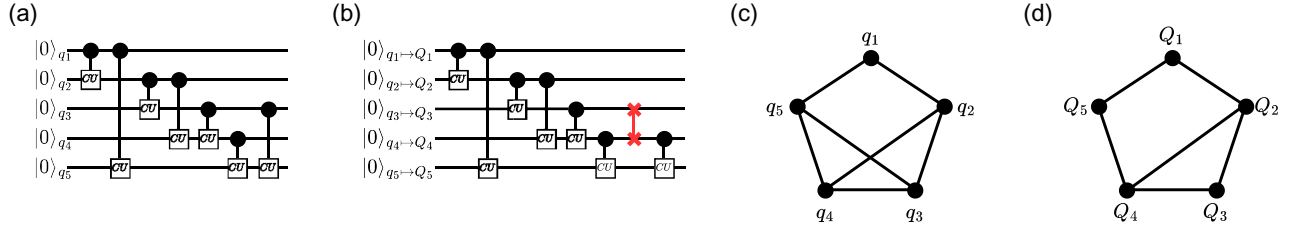
FIG. 2. A small example of *quantum circuit mapping*; gates depicted are general controlled-unitary operations (*CU*). SWAP gates are shown in red. (a) An algorithm in circuit form; the graph-theoretic representation of all two-qubit gates, known as the *interaction graph* (IG), is shown in (c). The task of quantum circuit mapping lies in assigning qubits from a quantum circuit to the physical qubits of a device, and subsequently reorganizing physical qubits such that all two-qubit interactions are performed in an efficient manner. (b) The mapped circuit; as the hypothetical architecture in (d), represented as a *connectivity graph* (CG), cannot accommodate all of the two-qubit gates represented in (c), we therefore must perform a SWAP gate.

an algorithm; this representation is known in the literature as the *interaction graph* (IG) [11]. In (b), a possible solution is presented for realizing the circuit on a quantum computer with connectivity shown as in (d); this graph is known as the *connectivity graph* (CG). We can assign qubits from the algorithm to physical hardware qubits; however, as no connection on (d) exists in order to account for the two-qubit gate CU_{q_3,q_5} required, one solution is to add a SWAP gate between Q_3 and Q_4 such that the final two-qubit interaction can be performed.

In order to evaluate the near-term spin-qubit architectures in our study, we map the quantum circuits presented in Sec. II A onto each architecture through a compilation framework. The dedicated compilation framework for spin-qubit technologies, *SpinQ* [11], was utilized in order to compile circuits for our simulations. As explained in Sec. II F, all circuits considered have been expressed in the hardware's native gates; therefore, the *elementary gate-set decomposition* stage can be skipped, and we can instead concentrate on *initial placement*. As mentioned above, initial placement denotes the process of assigning qubits from an architecture-agnostic circuit to the physical qubits of the device; spin-qubit architectures are no exception to this rule. An exact solution to the QCMP is found when the IG is (subgraph) isomorphic to the CG, inducing no extra gate overhead [108]. In practice, however, there are typically no exact initial placement solutions where all two-qubit gates can be satisfied without qubit movements; this is principally due to the bilinear nature of the considered connectivity graphs, in combination with the number and order of two-qubit gates in the featured circuits. To the best of our knowledge, no specialized initial placement algorithm has been developed for spin-qubit devices that optimizes for the intricacies of *shuttling operations*. In order to compensate, we incorporated the widely used SABRE algorithm [107] into SpinQ [11] for initial placement purposes.

The initial placement algorithm in SABRE works by generating a random qubit placement followed by the utilization of the SWAP-*based heuristic search* routing

algorithm [107] from which the resulting final qubit placement is used as the placement for the reverse circuit. This final placement derived from the reverse circuit can serve as an optimized initial placement for the original circuit. This refined initial mapping is of higher quality, as it incorporates comprehensive information about all gates and qubit interactions in the circuit. Because SABRE's placement quality heavily relies on a SWAP-based routing algorithm, fundamentally, it cannot be optimized for *shuttle-based routing* [53]. This complication signifies a fundamental unpredictability in the initial placement quality, as a higher number of random placement trials does not necessarily imply that better solutions can be found for spin-qubit devices.

After settling on an initial placement, we employ the *beSnake* [53] spin-qubit routing algorithm in order to handle all two-qubit gates in a circuit, in addition to shuttling qubits towards measurement sites for readout. In the architecture CGs that we assess, we assume that shuttling is allowed bidirectionally between any two directly connected nodes and, moreover, that two-qubit gates between any adjacent qubits in the CG can be performed. In this vein, beSnake can be optimized to exploit the imposed shuttling constraints when the movement of one qubit is blocked by others. As outlined in Ref. [53], certain scenarios may arise where shuttling blockages occur. In such cases, specific qubit(s) may obstruct the movement of other qubits along a shortest path. To address these blockages, dedicated mechanisms are employed. These mechanisms have been revisited, ensuring all blockages are resolved exclusively through shuttle operations.

F. Near-term spin-qubit architectures

Spin-qubit technologies possess distinctive physical properties that position them as a promising candidate for scalable quantum computing systems. One of the most notable advantages of semiconductor spin qubits is their extraordinarily small size—up to a thousand times smaller than other qubit technologies [114]. This compactness is

complemented by extensive experience in the semiconductor industry, which supports their development [115–117]. The key component in spin-qubit technologies is the *quantum dot*, which confines an electron or a hole, thus forming a physical qubit [118,119]. Control of a spin qubit is achieved electromagnetically through carefully positioned gate electrodes surrounding the quantum dot. These electrodes facilitate single- and two-qubit operations (in addition to qubit *shuttling*), executed via precise pulse sequences in systems with multiple quantum dots; such systems have been extensively explored in one-dimensional arrays [117,120].

More particularly, spin qubits in gate-defined quantum dots are architecturally interesting as they may be implemented in dense, highly connected 2D quantum dot arrays [121–123] or as sparse, low-connectivity registers connected via coherent long-distance links of multiple qubit modules [113,124–126]. Within the diverse range of architectural designs, various approaches have been developed to support surface-code quantum error correction, each implemented in distinct ways. For the purpose of advancing quantum processor development, it is crucial to evaluate which architectures are most effective based on quantum information-theoretic principles rather than solely on experimental practicality.

While the benefits of semiconductor spin-qubit quantum processors over other technological counterparts are notable, it is known that fabrication challenges emerge in scaling them, particularly in two dimensions [114,118, 127–129]. On the positive side, there have been significant efforts [121,122,125,126,130–132] to tackle these challenges and scale them in higher dimensions. It is forecasted that *bilinear arrays* are especially promising for near-term fabrication and experiments due to their amenity to current technological capabilities, making them more realistic to build [37,38]. Thus, in this work, we consider bilinear arrays of various sizes and connectivities as prospective candidates for near-term spin-qubit devices.

Our goal is to evaluate the merits and drawbacks of these devices by considering chiefly: (a) the expected quantumness in their qubit-qubit entanglement correlations; (b) anticipated logical success rates and ESP resulting from encoding a small error-detection code; and (c) the degree to which genuine multipartite quantum entanglement is generated and maintained on the device. In Fig. 3, we illustrate these bilinear arrays through their CG. In a CG, each node (circle) represents a site of the grid, and each edge connecting the nodes corresponds to the possibility of interaction between the two sites. The CG representing the arrays alongside further hardware characteristics comprises the architectures treated in this work. Additionally, readout can be performed at the bottom-left quantum dot by offsite sensory components. Usually, in experimental devices, it is possible to use multiple measurement sensors. However, we opted to simplify the compilation process by
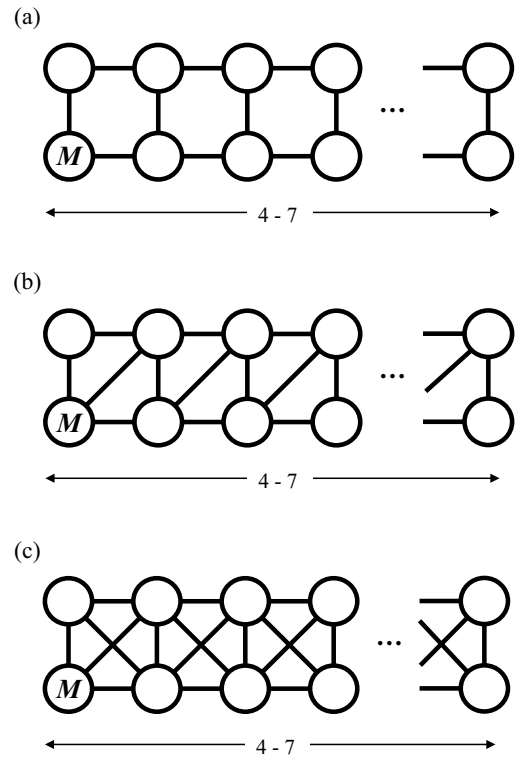

FIG. 3. Connectivity graphs (CG) of possible spin-qubit bilinear-array architectures considered in this work; we label each CG as 1, 2, and 3, in (a),(b),(c), respectively. We analyze the entanglement properties of quantum circuits generating an AME state and a small quantum error-detection code; each circuit utilizes seven qubits in total. In spin-qubit architectures, one typically does fill the entire lattice with qubits, although this is not necessary in general. As such, we test the capabilities of three different lattice connectivities shown in (a), (b),(c) for seven qubits used in quantum circuits. We also increment the size of each bilinear array from 2×4 to 2×7 quantum dots. In all 12 hypothetical architectures, readout can be performed at the bottom-left quantum dot by offsite sensory components.

standardizing the decision of which site should be used for all readouts. Additionally, this approach allowed us to focus exclusively on comparing the sizes and connectivity graphs of the circuits, as introducing an extra sensor could create an unfair advantage for certain configurations, potentially skewing the comparison by introducing size-specific benefits.

In a quintessential semiconductor spin-qubit device, it is optimistic to consider the following quantum device-specific properties at present [114,120,133–139]:

(a) A coherence T_2^* time of 20 µs [116];
(b) Single-qubit [140] and shuttle durations of 100 ns [54,141];
(c) Two-qubit gate times of 150 ns [142,143]; and
(d) Measurement times of 5 µs [144].

In such optimistic models, *thermal relaxation* (i.e., the T_1 time) usually ranges between 100 ms to a few seconds. Regarding gate fidelity, we assume an average of 99.99% for single-qubit gates and nearest-neighbor shuttles, 99.90% for two-qubit gates, and readout [142,143, 145]. Array initialization and qubit resets can also contribute to errors; however, for simplicity in our analysis, we assume that these values are ideal.

To formulate the gate set for our quantum circuits, we ensured that all circuits were expressed using a predefined gate set. Assuming this gate set is not natively supported by the underlying hardware, further decomposition may be required. However, since our focus lies in architectural comparisons rather than in absolute performance, such a tactic will not change the nature of our simulations, since all circuits or produced compilation overhead will experience the same relative change. For these reasons, we selected the gate set $\{H, CX, CY, CZ\}$, which facilitates constructing quantum circuits with a minimal number of gates, highlighting the adaptability and flexibility of our approach to meet various requirements.

It is additionally assumed that no gates can be performed in parallel; there are two main reasons for this. Firstly, it is known that pernicious and notoriously troublesome *crosstalk effects* are known to arise during parallelization in spin-qubit experiments [146]. Secondly, current generation spin-qubit experiments do not incorporate gate-parallelization techniques, although this may change in the future [147]. For these reasons, our first two error models are crosstalk-free. However, as a proof of concept for the utility of our work, we showcase a basic implementation of crosstalk-induced errors in Sec. III D 1, and describe how the current error models can be extended. In accordance with Ref. [101], we define a CPHASE(ζ) gate as

$$\text{CPHASE}(\zeta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\zeta} \end{bmatrix}, \qquad (21)$$

where $\zeta \in [0, 2\pi]$. The actual error model involving crosstalk will be defined more deeply in Sec. III D 1.

As a final note, the routing algorithm beSnake was modified such that when a measurement needs to take place, routing can commence in order to move a qubit toward the measurement zone of a bilinear array. We have enabled the routing optimization feature where a path-selection heuristic evaluates the best shortest path within a 0.05-s time limit, which is plenty for the circuit and device sizes of this study, as tested in Ref. [53].

III. RESULTS

A. Simulation and error model

Results from each of the simulations realized are shown in Figs. 5–8, corresponding to (a) extrapolation of the Bell
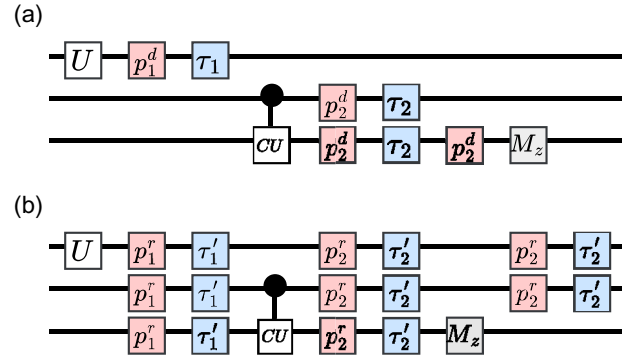

FIG. 4. Both of the error models used in this work. (a) denotes the error model $\mathcal{E}_Q$, while (b) displays the error model $\mathcal{E}_{TN}$.

operator for the AME(6,2) state (Fig. 5); (b) calculation of the logical success rate for the $[\![4, 1, 2]\!]$ error-detecting surface code (Fig. 6); (c) the ESP for circuits pertaining to the error-detecting surface code (Fig. 7); and (d) the results from calculating the tripartite mutual information (Fig. 8). For each data point generated, either 10^4 (Fig. 5), 2×10^4 (Figs. 6 and 7), or 2×10^3 (Fig. 8) Monte Carlo trials were utilized. Furthermore, we used a specific seed number on the nondeterministic aspects of our simulations for reproducibility and consistency in our comparisons.

The error models utilized in our simulations differ depending upon whether we utilized Qiskit [104] or a tensor-network (TN) simulation via the Python package quimb [148]; as such, we name the corresponding error channels applied in our simulations as $\mathcal{E}_Q$ and $\mathcal{E}_{TN}$, respectively. These error models are shown graphically in Fig. 4, and they both take on the following forms:

Error model $\mathcal{E}_Q$ is defined as follows:

(a) After a single-qubit gate U in our model, with probability p_1^d a random Pauli error (that is, an error drawn from $\{X, Y, Z\}$) is applied;

(b) Additionally, after the same single-qubit gate U, with probability $\tau_1 = (1 - e^{-t/T_2})/2$, a Z error is applied, where t depends on the single-qubit gate duration;

(c) After a two-qubit gate CU, we apply with probability $p_2^d = 10 p_1^d$ a random Pauli error (that is, an error drawn from $\{X, Y, Z\}$) on each of the qubits involved;

(d) Moreover, after the same two-qubit gate CU, with probability $\tau_2 = (1 - e^{-t_2/T_2})/2$, a Z error is applied, where t_2 depends on the two-qubit gate duration;

(e) Before each measurement M_z, with probability p_2^d we apply an error drawn from $\{X, Y, Z\}$;

The error model $\mathcal{E}_{TN}$ is defined as follows:

(a) After a single-qubit gate U in this model, with probability p_1^r, a random rotation gate is performed

on each qubit, drawing from the set $\{R_x(\phi), R_y(\phi), R_z(\phi)\}$, where $\phi = \pi/\mu$ and $\mu \in (0, 1]$;

(b) In the same time step and after the same single-qubit gate U, with probability τ_1' a Z-rotation gate $R_z(\psi)$, where $\psi = \pi e^{t/T_2}$, is applied on each qubit;

(c) After a two-qubit gate CU, we apply with probability $p_2^r = 10p_1^r$ a random rotation gate, drawing from the set $\{R_x(\phi), R_y(\phi), R_z(\phi)\}$, where $\phi = \pi/\mu$ and $\mu \in (0, 1]$ for all qubits;

(d) After the same two-qubit gate CU, with probability $\tau_2' = 10\tau_1'$ a Z-rotation gate $R_z(\psi)$, where $\psi = \pi e^{t/T_2}$, is applied on each qubit;

(e) After a measurement M_z is performed, we apply with probability $p_2^r = 10p_1^r$ a random rotation gate, drawing from the set $\{R_x(\phi), R_y(\phi), R_z(\phi)\}$, where $\phi = \pi/\mu$ and $\mu \in (0, 1]$ for all qubits;

(f) Finally, after that same measurement M_z is performed, with probability $\tau_2' = 10\tau_1'$ a Z-rotation gate $R_z(\psi)$, where $\psi = \pi e^{t/T_2}$, is applied on each qubit.

The error models $\mathcal{E}_Q$ and $\mathcal{E}_{TN}$ differ fundamentally in their design and application. With these choices, we aim to demonstrate the diversity of our approach by employing different error models, each offering distinct trade-offs between computational complexity and realism. Further insights are provided in detail in Sec. IV. In this context, the tensor-network simulation can be viewed as a more realistic analog for expected quantum device behavior, wherein errors are excluded after measured qubits, as these states are collapsed and reinitialized with similar *fault-tolerant quantum circuit* models used in the literature [149–152]. Conversely, the $\mathcal{E}_Q$ model combines the standard depolarizing model, implemented with the `depolarizing_error()` function from Qiskit Aer library [104], with decoherence errors cascaded using the thermal relaxation function `thermal_relaxation_error()`.

B. AME state-generation circuit results

In Fig. 5, we have graphed the results obtained from the generating circuit for the AME(6,2) state and the subsequent average measurement of the Bell operator $\langle B \rangle$, alongside the ESP and the shuttle count added to the circuit during the compilation process. Here and in the subsequent section (Sec. III C), the noise model $\mathcal{E}_Q$ is used. The ESP results (expressed as a percentage), the average Bell operator expectation $\langle B \rangle$, and the shuttle count have been graphed together. Subfigures (a)-(c) represent each CG from Fig. 3, with increasing lattice sizes from 2×4 to 2×7 quantum dot sites; as stated earlier, the net result of increasing the number of quantum dots is to reduce the density of qubits relative to unoccupied sites in the device.

In all of the figures, it is clear that the shuttle count closely correlates with the calculated $\langle B \rangle$ and ESP values. For CG$_1$ and CG$_3$, one can surmise via the expectation of $\langle B \rangle$ that qubit-qubit correlations diminish as the size of the lattice is incremented; as a direct result, it is anticipated that the ESP drops commensurately. One exception to this trend can be seen in CG$_2$, wherein the metrics fluctuate on even and odd lattice sizes. One naive reason for this may be due to some extraneous effect that our simulations may not have originally accounted for; however, a more logical reason has to do with the *initial placement* algorithm, SABRE, that was utilized in this work. This point will be analyzed more deeply in Sec. IV. As a final note, we additionally calculated the average (over all lattice sizes) ESP, $\langle B \rangle$, and shuttles required; these data are shown in Table I. The overall trend for all CGs is that, as the connectivity of the device increases, the shuttle count decreases, thereby incrementing the ESP and $\langle B \rangle$ values; nevertheless, we also notice that, as the connectivity increases, the degree to which all three fields change also steadily decreases; we also touch upon this in Sec. IV.
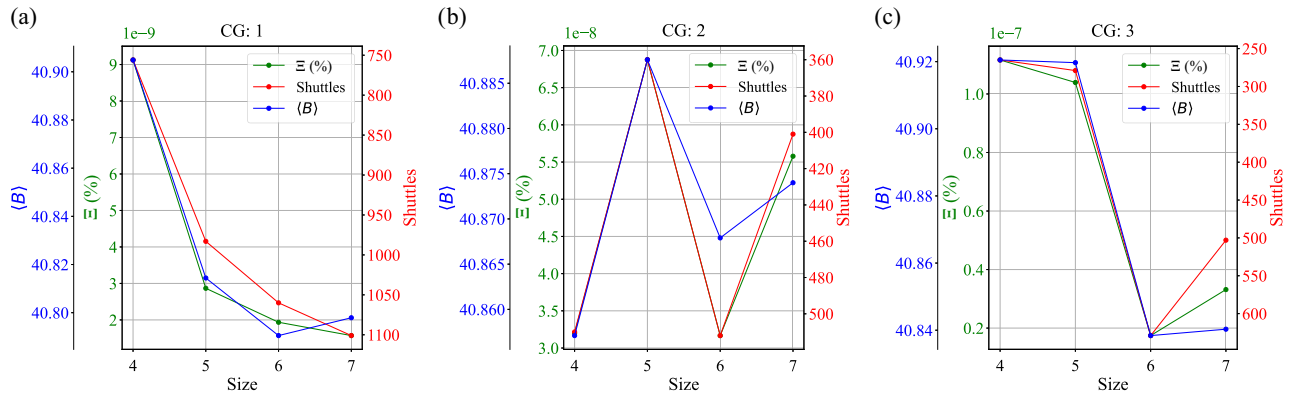

FIG. 5.   Results for the circuit generating the AME(6,2) state. In blue, green, and red, we have measured: the average Bell operator $\langle B \rangle$; the ESP, expressed as a percentage; and the shuttle count added during the compilation process. 10^4 trials were conducted for all three CGs.

TABLE I. Averaged ESP, $\langle B \rangle$, and shuttle operations, as reported in Fig. 5, over all sizes of each CG.

CG	Ξ (10^{-8})	$\langle B \rangle$	Shuttles
CG_1	0.39	40.82	975
CG_2	4.71	40.87	445.75
CG_3	6.65	40.88	419

C. Logical success rate and ESP results

Figure 6 displays the logical success rate p_s obtained for the $[\![4, 1, 2]\!]$ error-detecting surface code, together with the shuttles added. We have reported averages over (a) the number of cycles tested in our simulations as we increase the lattice size, and (b) over all lattice sizes as the number of cycles is incremented. As can be seen in (a), the shuttle count steadily increases for all but CG_2 (for which the shuttle count actually decreases slightly). As a result, the lattice size increases commensurately, showing a drop in the logical qubit's logical success rate. The only exception to this rule is shown for CG_2, which again exhibits an increase in logical success rate as the lattice size is increased. As for (b), averaging over lattice sizes and incrementing the number of stabilizer measurement cycles paints a different picture: initially, after one cycle, it can be seen that the logical success rate and the shuttle count inversely correlate with one another. Indeed, it is also evident that the most highly connected CG, CG_3, exhibits a high logical success rate. However, as we raise the number of measurement cycles, at about the sixth cycle, we see that the logical success rates of all connectivity graphs are roughly indistinguishable from each other, and this trend continues as we continue to add stabilizer measurement cycles; we elaborate more on this in Sec. IV. Finally, we also report on the Pauli logical error content detected in Appendix A.

By and large, we see in (a) and (b) of Fig. 7 that the tendencies of both the ESP and logical success rate agree with each other, although the absolute values of both widely differ. As expected, the ESP exponentially decays with circuit depth, and converges to zero at about six or seven stabilizer measurement cycles. Again, this result is in agreement with our results in Fig. 6 on the logical success rate decay, and are indicative of how, in a more realistic experiment, it is expected that the logical success rate of a logical qubit converges to a specific (low) value. The reason for this convergence is related to the projective (stabilizer) measurements utilized; although errors will build up and propagate throughout a circuit as its depth is increased, projective measurements still will project the logical qubit into a logical eigenstate pertaining to one of the codestates, and in doing so, there will be a finite probability that one of these corresponds to the correct logical qubit codeword.

In Table II, we have calculated the averages of ESP, logical success rate, and shuttles over all cycles for each size and CG. When focusing on the average shuttles, we note a bigger reduction between CG_2 and CG_3 compared to CG_1 and CG_2. These relative differences are also reflected in the ESP and logical success rate in the table. Smaller lattice sizes can offer competitive results when compared to larger ones across all metrics displayed. These observations underscore the fact that improving connectivity or size of the device is not guaranteed to provide an improvement in the entanglement measures that we have chosen to study; we discuss this and other related observations and put them into the larger perspective in Sec. IV.
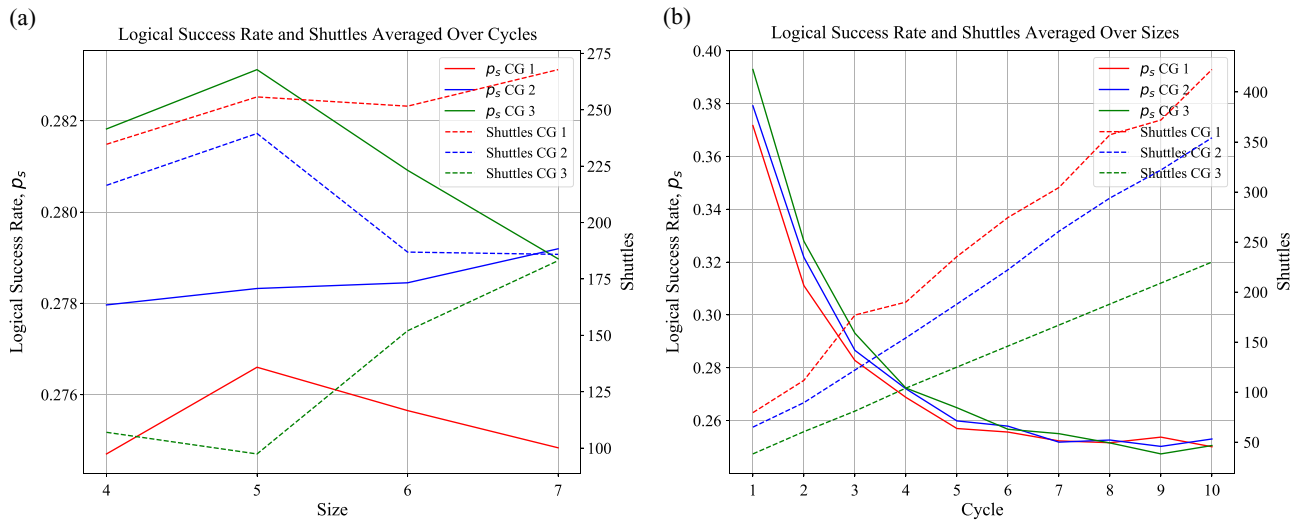

FIG. 6. Finalized logical success rate p_s results for the $[\![4, 1, 2]\!]$ surface error-detecting code, under (a) cycle averaging and (b) averaging over all sizes for a particular CG, and were taken after averaging over 20 000 trials.
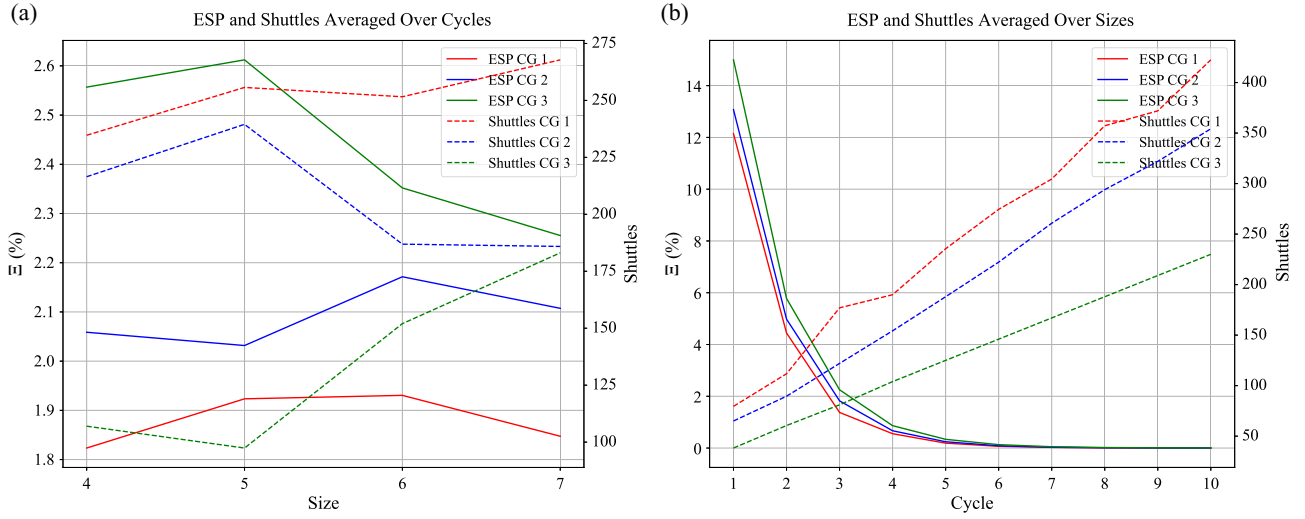
FIG. 7.   Finalized ESP results for the $[\![4, 1, 2]\!]$ surface error-detecting code, under (a) cycle averaging and (b) averaging over all sizes for a particular CG.

D. More invasive noise models

Figure 8 depicts the tripartite mutual information $\mathcal{I}_3$ obtained from tensor-network simulations; these have been derived from the same compiled circuits as previously (Figs. 6 and 7). Here, the noise model $\mathcal{E}_{TN}$ is used. As $\mathcal{I}_3$ is a measure of the global and local distribution of quantum entanglement over subsystems, we note that, unlike previous results, a lower (potentially below zero) $\mathcal{I}_3$ value is desirable, as it implies a global distribution of genuine multipartite entanglement over a circuit (and therefore over the quantum device). Subfigures (a) and (b) again show the averaged $\mathcal{I}_3$ with respect to cycles or lattice sizes, respectively; we then opt to increment lattice sizes and the cycle number, respectively.

In Fig. 8(a), we note a slightly different tendency in the $\mathcal{I}_3$ than in our previous simulations; namely, as the lattice size is increased for CG_2 and CG_3, at first, we observe a *decrease* in the $\mathcal{I}_3$, and subsequently an increase (for CG_1, we notice a trend of increasing $\mathcal{I}_3$). The main reason for this behavior involves the observation that all of the averaged $\mathcal{I}_3$ values are greater than zero; that is, we can surmise

that, averaged over all cycles, the tripartite mutual information for every device size is expected to stay squarely in the *area-law* phase. As such, monitoring changes in the $\mathcal{I}_3$ is important, as this signals the changing nature of multipartite entanglement correlations in the quantum system; however, the overall character of entanglement correlations observed stay within the disentangling (area-law) phase.

Figure 8(b) showcases a similar concept as in Sec. III C; that is, in the first cycle, the average $\mathcal{I}_3$ over all lattice sizes is generally below zero, located inside the *QEC phase*; here, it is presumed that *volume-law* entanglement behavior dominates in the quantum system. However, as we increase the number of cycles in our simulation, we see that already in the second cycle, one can preliminarily note the presence of a *measurement-induced phase transition*, as $\mathcal{I}_3$ crosses over from negative to positive. This information is important, as from here on out, it is known that the individual subsystems of the quantum device will begin to behave in a classically correlated manner; indeed, as we increase the number of cycles, the exact same correlation between shuttle count and $\mathcal{I}_3$ follows the trends of Secs. III B and III C. Additionally, we notice that at around six or seven cycles, the $\mathcal{I}_3$ measured for CG_2 and CG_3 effectively converge, suggesting that, as the number of cycles is increased, the potential gain in connectivity via CG_3 matters less and less.

1. Incorporating crosstalk

As discussed in Sec. II, *crosstalk* refers to a problematic noise source that constrains the degree to which parallelization of noisy operations is possible in most current quantum devices [153]. In many qubit technologies and their accompanying architectures, crosstalk can even arise
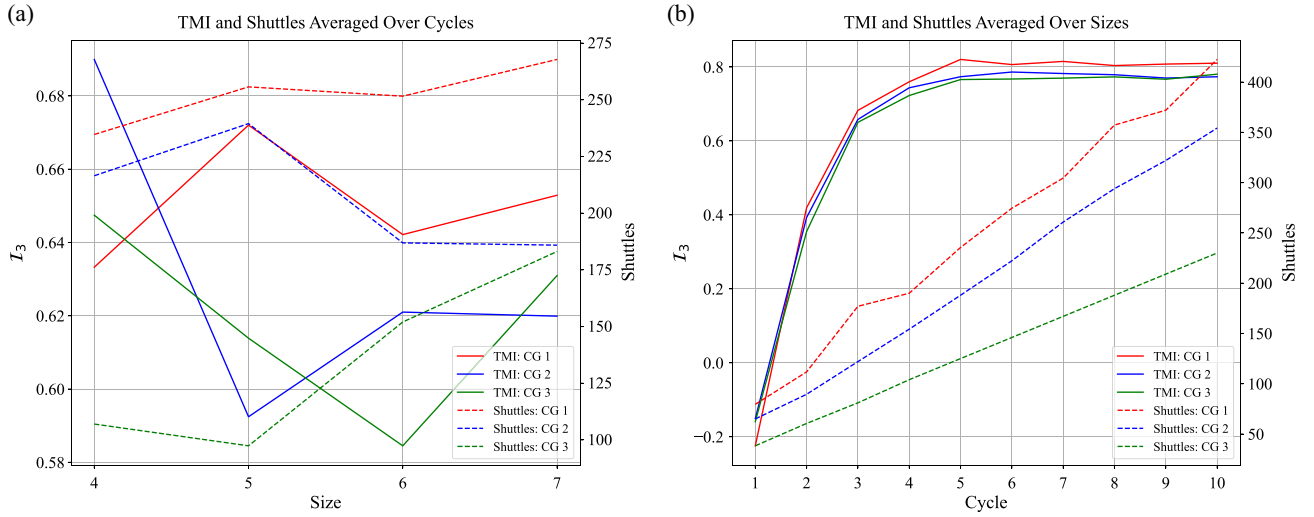
TABLE II.   Averaged ESP, logical success rate, and shuttle operations, as reported in Figs. 6 and 7, over all cycles of each size and CG.

	Ξ (%)	Logical success rate (p_s)	Shuttle count
CG_1	1.8812	0.2754	252.45
CG_2	2.0923	0.2785	207.20
CG_3	2.4441	0.2812	134.88
Size 4	2.1463	0.2782	186.07
Size 5	2.1891	0.2793	197.53
Size 6	2.1513	0.2783	196.83
Size 7	2.0700	0.2777	212.27

FIG. 8.   Results from the tensor-network simulation of the $[\![4,1,2]\!]$ surface error-detecting code, under (a) cycle averaging and (b) averaging over all sizes for a particular CG. 2,000 Monte Carlo trials were taken per data point. Tripartite mutual information is generally shown to become more positive in both (a) and (b), as additional shuttles are added to the compiled circuit; this is consistent with the $\mathcal{I}_3$, which is observed to leave the volume law (QEC) phase at approximately the second cycle of stabilizer measurements, entering the area-law (disentangling) phase.

from single-qubit operations [101]. While these deleterious effects can be mitigated experimentally to some extent in single-qubit operations, crosstalk from two-qubit gates remains a significant challenge and is not well understood in general [101,154]. Given these issues, we aim to provide a preliminary investigation of crosstalk using a naive model; we simulated the impact of crosstalk originating from two-qubit interactions using the same tensor-network approach as utilized in Fig. 8.

The error model employed for the crosstalk simulations takes the following form, in addition to the noise parameters from $\mathcal{E}_{\mathrm{TN}}$:

(a) After a two-qubit gate CU, we apply a CPHASE(ζ) gate with probability $p_{\mathrm{cross}} = 10(p_1^r/3\xi)$. Either the control or target qubit of the CPHASE(ζ) is designated from among the operands of the CU gate, while the other qubit is randomly selected from the remaining, nonoperational qubits.

One may inquire as to why we included the term ξ in our modeling of the crosstalk. This term is used in order to take into account the connectivity differences on average across all CGs tested, as it is known that crosstalk is more prevalent in highly connected devices [153,154]. Additionally, our crosstalk simulation adheres to one of the criteria on the definition of a crosstalk error [101]; namely, our model violates the *locality condition*, since one of the operands is scrambled to another random qubit with some probability. Using these model parameters, we analyzed the impact of crosstalk

across a spectrum of gate errors, specifically using the probabilities $p_1^r = \{0.01, 0.03, 0.05, 0.08\}$, as discussed in Sec. II F.

In Fig. 9, the $\mathcal{I}_3$ results from tensor-network simulations based on the same compiled circuits are shown, conducted over the range of previously specified error rates for two scenarios: (a) without crosstalk and (b) with crosstalk. In each subfigure, the four groups of three colored lines (red, green, and blue) correspond to each error rate in the range, with the lowest rate at the bottom (opaque) and the highest error rate at the top (translucent and faded). For reference, the lower set of lines in Fig. 9(a) essentially mirrors the results displayed in Fig. 8(a). Notably, as we go from the lower error rates (at the bottom line group) in the range to the highest (top line group), the $\mathcal{I}_3$ growth rate progressively declines, indicating a nonlinear relationship between error rates and $\mathcal{I}_3$.

Since it could be difficult to clearly distinguish subtle, but nonetheless important, differences in Fig. 9, we have additionally placed the average $\mathcal{I}_3$ received from all error rates (per CG and lattice size) in Table III. Taking stock of the data from Table III and Fig. 9, we can surmise that, as the error rates are increased with crosstalk, the resulting values calculated show a severely weakened advantage of CG3 over the other two connectivities; this was not observed so prominently in our simulations without crosstalk. Indeed, at higher error rates, the $\mathcal{I}_3$ calculated for CG3 even becomes worse than that of the other two CGs, indicating that the advantage of more local connectivity in CG3 is counterbalanced by the introduction of more crosstalk.
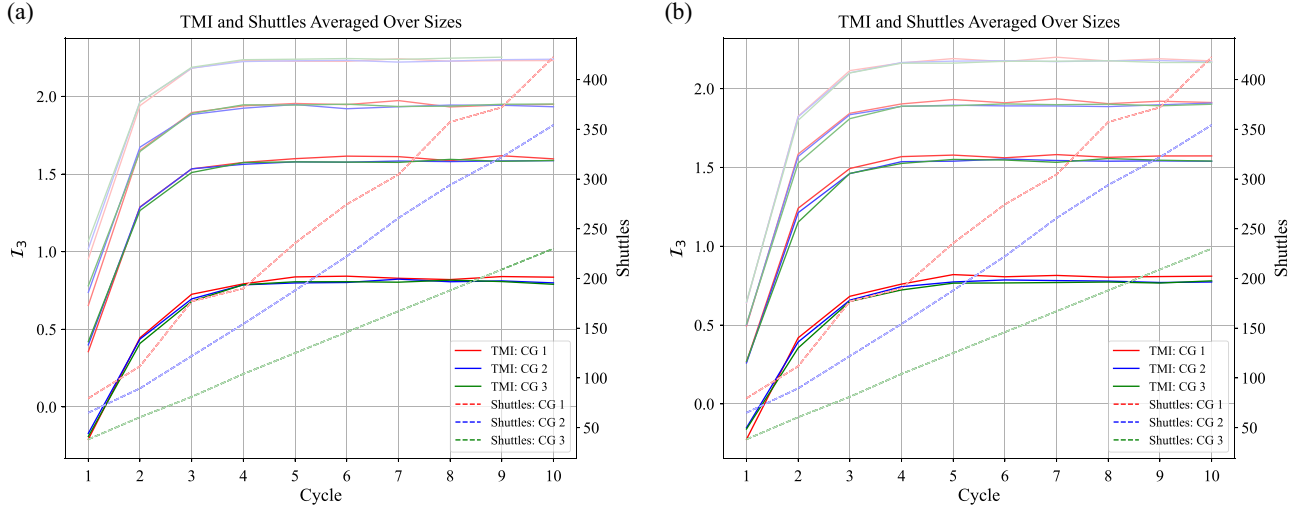
FIG. 9.   Results from the tensor-network simulation of the $[\![4, 1, 2]\!]$ surface error-detecting code for a range of error rates, (a) with crosstalk (parameterized by p_1^r after two-qubit gates) and (b) without crosstalk.

IV. DISCUSSION

There are many points of discussion here. Firstly, let us consider the simulation for generating the AME(6,2) state and subsequent measurement of the Bell operator, in addition to the ESP. In particular, one may ask why we have chosen to graph $\langle B \rangle$, ESP, and the shuttle count together on one vertical axis, as we have done in Fig. 5; after all, the actual absolute values of the data we obtained lie staunchly on different orders of magnitude, and, a simple glance at the equations governing ESP and $\langle B \rangle$ seems to imply that no *a priori* relationship between these metrics exists; in this sense, it may appear that we have manufactured a correlation where in fact none exists. However, the parameters of the simulation were constructed in a way such that the only significant allowed change to every circuit per trial is the shuttle count. In addition, the major theme of our simulations involves the notion of constructing a *hierarchy of*

comparison, instead of establishing absolute values for our measures. In this way, our purpose centers on establishing relations to assess the quality of a spin-qubit architecture. As such, we graphed these three measures together in order to highlight a trend; namely, that the change in shuttle count throughout the simulation wrought changes that are observable in the Bell operator and ESP, and we observe exactly this behavior throughout our investigation.

Next, one may critique the minor differences present in the ESP and Bell operator values reported in Fig. 5; in particular, one may state that the degree of change in $\langle B \rangle$ is not significant, and that furthermore, the ESP values themselves are very close to zero. In this work, we chose to analyze architectures on the basis of their entanglement properties; quantum entanglement structure is fragile in and of its own accord, and previous studies have shown that even small deviations in the Bell operator can signify large changes in qubit-qubit correlations, in addition

TABLE III.   The upper-half table shows the average $\mathcal{I}_3$ by coupling graph for various error rates, both with and without crosstalk; subsequently, the lower-half table displays the average $\mathcal{I}_3$ by size for various error rates, also with and without crosstalk.

| CG	Error Rate 0.01		Error Rate 0.03		Error Rate 0.05		Error Rate 0.08	
	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk
1	0.6759	0.6501	1.4380	1.4000	1.7841	1.7337	2.0688	1.9844
2	0.6589	0.6308	1.4270	1.3728	1.7839	1.7161	2.0778	1.9777
3	0.6517	0.6192	1.4263	1.3684	1.7928	1.7118	2.0909	1.9725
Size	**Error Rate 0.01**		**Error Rate 0.03**		**Error Rate 0.05**		**Error Rate 0.08**	
	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk	Crosstalk	No Crosstalk
4	0.6856	0.6616	1.4628	1.4123	1.8067	1.7439	2.0916	1.9984
5	0.7054	0.6720	1.4686	1.4327	1.8111	1.7591	2.0836	1.9999
6	0.6546	0.6316	1.4187	1.3743	1.7720	1.7180	2.0681	1.9735
7	0.6624	0.6364	1.4263	1.3828	1.7808	1.7225	2.0696	1.9759
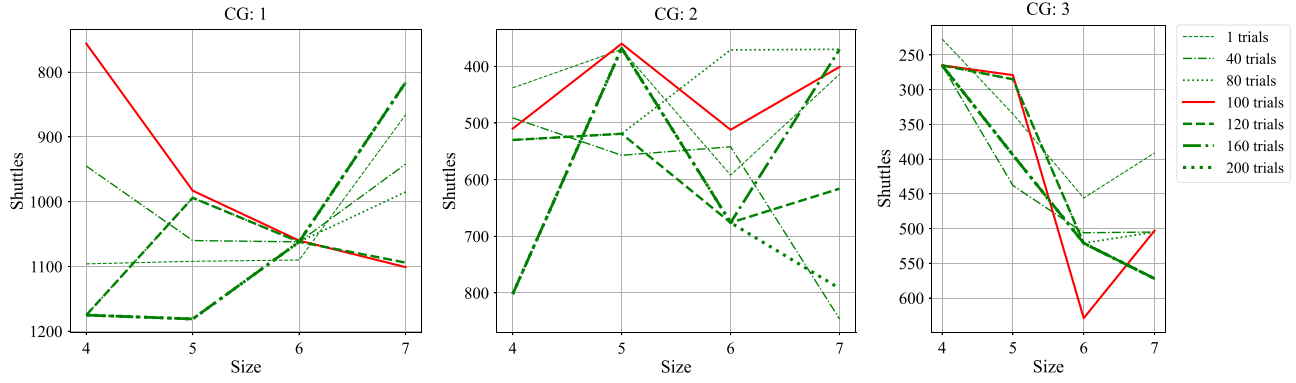
FIG. 10. Shuttle counts resulting from compiling the AME(6,2) circuit for each CG, using a range of initial placement optimization trials from SABRE, are shown. The red line represents the shuttle counts from simulations in Fig. 1. Due to SABRE's limitations in optimizing shuttle operations, the resulting shuttles differ significantly. Typically, larger CG sizes allow initial placement algorithms to achieve more effective initial placements. However, in this case, we observe a potential trend in the opposite direction even with more trials.

to the general structure of multipartite entanglement [29]. This effect is also present in the other metrics chosen, as very small differences in the $\mathcal{I}_3$ (or in the bipartite mutual information) can signal the advent of measurement-induced phase transitions [51,52,93,94]; we will comment more on this shortly.

In Fig. 5(b), we observed incongruent behavior of the measures for CG_2; although all of the metrics are clearly correlated with one another, we discussed in Sec. III B one possible reason as to why CG_2's results fluctuate in comparison to those of the other architectures tested. In particular, this discrepancy can be explained by considering the initial placement algorithm SABRE. To better demonstrate the effect of initial placement with SABRE on the same circuit simulations of Fig. 5, a range of optimization trials (in between [1,200] trials) were tested, and the resulting shuttle counts were plotted in Fig. 10. Here, we define a *trial* in SABRE as a simultaneous three-step search process; more details on this can be found in Ref. [107]. As is evident, a higher or lower number of trials does not always equate with a more or less favorable shuttle count (even if the qubit density decreases with larger sizes), since SABRE is not optimized for shuttle operations present in spin-qubit devices. As explained in Sec. II E, the results can, therefore, vary significantly, regardless of how many trials are taken. Consequently, one can still conclude that compilation methods can have a significant influence on the values of all measures utilized in our studies; this agrees with the findings in the literature [109,110].

Regarding the simulation results shown in Figs. 6 and 7, the close correspondence between both sets of data suggests that small quantum error-detection experiments with up to five cycles of stabilizer measurements could be possible using any of the devices considered in this work. CG_3 outright obtains the highest value for both the ESP

and logical success rate in this cycle range, with CG_2 and CG_1 falling in line with lower logical success rate and ESP values. By examining Figs. 7(a) and 7(b) more closely, we observe the same partial ordering as per the values of both of the measures captured, with CG_3 showing an advantage over the other two as expected. However, for certain sizes, both the metrics of CG_2 and CG_3 converge closely. The main message from this observation is that, when averaged over many cycles, the measured ESP and logical success rates for CG_2 and CG_3 do not differ heavily; this is significant, as it implies (as in Fig. 5) that for near-term experiments with highly entangled quantum states, it is not necessarily more beneficial to fabricate a device with higher local connectivity. Instead, under the assumption of utilizing quantum compilation methods, we note that realizing a properly-compiled quantum circuit on CG_2 in fact approaches the values of CG_3 in several key metrics that we have studied. We also observe this for Figs. 8 and 9, although as the crosstalk effect is increased, we notice that the $\mathcal{I}_3$ calculated for CG_2 and CG_3 become even worse than CG_1.

Looking more deeply at the ESP and logical success rate, it becomes clear that the ESP approaches 0 relatively quickly while p_s converges to approximately 0.25. As explained in Sec. II D, ESP stands out from the other metrics due to its worst-case approach to estimating performance and with better computational efficiency. In contrast, the logical success rate p_s is calculated over numerous state-vector simulation trials; one can expect that, under the conditions affecting our noise model, projective measurements should repeatedly project the logical state back into the codespace, as was discussed in Sec. III C. Although both metrics scale differently with larger circuits, their relative performance remains correlated, as previously discussed, making them equally reliable for their respective use cases.

We also remark that the noise models utilized for the tensor-network simulations of Figs. 8 and 9 indicate a distinct noise model from those of Secs. III B and III C. Although the inclusion of idling errors do not affect the main conclusions from our results, the crosstalk present in our other error model does affect the conclusion of our study. This effect can be seen in Fig. 9(a); as the error rate is incremented, the $\mathcal{I}_3$ grows nonlinearly, per cycle. At the higher error rates (shown in faded colors), we discern that the three size-averaged CGs converge after smaller and smaller cycle numbers; this indicates the effects of crosstalk at higher error probabilities, since all other noise effects in our simulations are accounted for. Taken at face value, these results further imply fundamental limitations to the size and depth of quantum circuits that the selected architectures will be able to perform.

Regarding the results shown in Fig. 8, we mentioned the possible appearance of a *measurement-induced phase transition* (MIPT) from volume- to area-law phases. While it is known that the $\mathcal{I}_3$ generally is a good indicator of such MIPTs, we assert that more investigation is needed in order to ascertain this supposition; for a start, the size of the circuits tested (in this case, the surface-code circuit) could be scaled with the size of the device, in order to maintain roughly equal qubit densities. Additionally, both the entanglement entropy and the $\mathcal{I}_3$ should exhibit the prototypical peaks consistent with MIPTs, and our results do not confirm this trend at the moment, even though other works have observed the existence of this phenomenon in the surface code [95]. We leave the exploration and confirmation of MIPTs in such prospective architectural studies for future work.

The results from Fig. 9 and Table III support our findings that, in the presence of crosstalk, there are underlying tradeoffs with regard to adding local connectivity to a spin-qubit device. In particular, it is evident that, even at the error rate 0.03, the size-averaged $\mathcal{I}_3$ are the same for both CG_2 and CG_3. As we increase the error rate for the crosstalk simulations, CG_3 attains the worst $\mathcal{I}_3$ of the three connectivity graphs. These same results are visible, even without crosstalk, but become accentuated as we add crosstalk to error model $\mathcal{E}_{TN}$. As such, we can deduce that, in the presence of the naive crosstalk model used here, our results indicate a disadvantage when utilizing the most highly connected CGs. It is interesting to remark that we have also probed the same simulations, but with single-qubit crosstalk parameters added; again, the same tendency is again present, albeit to larger degree.

Surprisingly, the CG-averaged size results from Table III imply that lattice size 2×6 on average achieves consistently the lowest $\mathcal{I}_3$. One may ordinarily suspect that the largest device, the 2×7 array, would achieve the most favorable $\mathcal{I}_3$, as all qubits in the circuit are relatively isolated from one another; however, we do not observe this in our results. One intuitive reason may have to do with the qubit density of the array. Consider, for example, an empty spin-qubit array; as we start to add qubits, the degree to which they can be correlated with one another is quite limited, as it is known that entanglement correlations exponentially decay over small distances when exposed to environmental effects [155]. Notwithstanding, we will eventually approach a *critical point* of filling the lattice, after which it becomes progressively more and more difficult to perform shuttling operations, and the benefit of doing so in a spin-qubit device fades. This phenomenon is known in percolation theory as the *percolation threshold*, and separates classical "phases" of particle behavior into those which interact independently versus those which are strongly dependent upon the interactions of neighboring particles or clusters on the lattice [156]. For a Euclidean square lattice, the site-percolation threshold is known to be $p_{\text{perc}} \approx 0.5927$, for which there is an approximately 1% difference with the qubit filling in our 2×6 array (i.e., $7/12 \approx 0.583$) [156]. In light of this, it seems reasonable to conclude that, for the seven-qubit circuits that we have tested, a 2×6 bilinear array may yield better $\mathcal{I}_3$ results. One may be able to formalize this concept more, by looking concretely into *critical exponents* and how they affect the quality of entanglement on the device as it realizes the quantum circuit; more investigation is needed into this, which we leave for future work. In the most realistic case, a pragmatic experimentalist would consider a 2×6 array as a subset of a larger spin-qubit array, permitting the system to be always operated at an optimal filling. This condition allows for the exploration of optimally sized lattices for circuit execution in future spin-qubit architectures.

We would like to briefly note that our simulations, at first glance, do not seem scalable [with the notable exception of ESP, which scales as $\mathcal{O}(g)$, where g is the number of gates in a given circuit]. One reason for this issue may be due to necessity of large-scale Monte Carlo simulations, tensor-network contraction, or calculations such as matrix diagonalization, which are known to scale at worst $\mathcal{O}(n^3)$ for dense matrices (where n is the dimension of the matrix), and at best $\mathcal{O}(n)$ for sparse matrices. However, our methods could be substantially improved by utilizing optimized tensor-network contraction techniques, such as those from Refs. [157,158]. We leave such optimizations for future work.

Finally, one may be tempted to establish a global precedent of lattice sizes, given the overall implications of our work; after all, only one set of results from our work [Fig. 5(b)] does not immediately agree with the conclusions drawn from the rest. Accepted as is, we cannot claim that a particular lattice size of spin-qubit device would be more or less advantageous than another. As stated before, the major reason for this added nuance to our results lies in the fact that the initial placement algorithm, SABRE, is not optimized for all of the architectural features that typically characterize spin-qubit technologies. If an initial

placement algorithm for spin-qubit architectures can be developed with provable guarantees for solution quality (such as a specialized version of [159] for spin-qubit technology), we believe that the results of our work should be revisited. However, in spite of the complications discussed above, we can safely conclude that circuit compilation does greatly influence the four metrics that we have proposed, in order to study the connectivity trade-offs between future architectural designs. In light of our demonstration, it is highly probable that, by utilizing appropriate quantum circuit-compiling techniques, it is possible to achieve better metric values for more sparsely connected devices. In particular, our results strongly suggest that, while CG_2 and CG_3 indeed may not have such different entanglement properties, the extra connectivity of CG_2 and CG_3 do not appear to be worth the fabrication effort, especially in light of the convincing results evinced by our crosstalk simulations. These outright worse results were prominent in almost all of the data obtained, especially in the limit of both higher cycle numbers of stabilizer measurements, as well as higher lattice sizes (i.e., lower qubit densities) when we employed our crosstalk model.

V. CONCLUSION

In this work, we have presented a framework based on quantum information-theoretic and compilation-based measures for methodically evaluating the entanglement properties of prospective quantum architectures. More specifically, we have utilized the following entanglement measures: the average *Bell operator* $\langle B \rangle$; the *logical success rate* p_s for the smallest error-detecting surface code; and the *tripartite mutual information* $\mathcal{I}_3$ (again for the $[\![4, 1, 2]\!]$ surface code). We also proposed a modification to the *estimated success probability*, a known metric in quantum compilation, in order to take into account the effects of decoherence in certain quantum devices stemming from *thermal relaxation*. We demonstrated the usefulness of our techniques by realizing an architectural study which profiles the structure of entanglement of generated quantum many-body states. Indeed, it is supposed that more local connectivity is immediately beneficial towards executing highly entangled quantum circuits. Surprisingly, we find that, under the assumptions of the noise models chosen, it is possible to approach comparable qubit-qubit correlations, ESP, logical success rates p_s, and $\mathcal{I}_3$ values to the most highly connected quantum devices in our set, by utilizing appropriate techniques from spin-qubit quantum compilation with SpinQ [11] and beSnake [53], as well as the advantage conferred through the usage of shuttling. Our results suggest that for small-scale spin-qubit experiments, more device connectivity does not necessarily guarantee an improvement in the quality of quantum entanglement arising from circuit-prepared quantum states; this narrative was strongly apparent in our $\mathcal{E}_{TN}$ error model

simulations, as well as for all other simulations (albeit to a somewhat lessened degree, which we attribute mainly to the lack of a specialized initial placement algorithm). Central to our approach was the incorporation of expected device characteristics, such as gate and shuttling durations and expected dominant noise sources.

Our results come in the wake of several recent works attempting to evaluate the device-level fitness of spin-qubit architectures for near-term quantum error-correction experiments [37,39,99]. Modification of our methodology for the direct simulation of such error-correction schemes is straightforward, and in this way, our framework can assist with device prototyping and design, without the financial, manufacturing, and time overhead costs normally ascribed to the development of new experimental quantum devices. Indeed, the framework proposed can be utilized to test further spin-qubit architectures, outside of the connectivities tested in this work. Moreover, by modifying parameters of the simulation such as gate duration or the parameters of the compiler, one may be able to investigate other qubit technologies as well, such as trapped ion, neutral atom, and superconducting devices. One may be able to simulate more nuanced environmental hardware noise sources using a *density-matrix* simulation [160]; we leave this for future exploration.

As we remarked in Sec. IV, several future directions are possible. We address a few of these ideas below:

(a) As we have investigated a simple quantum error-detection code's logical success rate, it would also be useful to simulate directly the logical success rate of various transversal logical operations in the $[\![4, 1, 2]\!]$ surface code, as was experimentally realized in Ref. [48] for superconducting qubit technologies. This concept would help to inform what logical-level success rates can be expected for similar experiments in near-term spin-qubit devices.

(b) The development of a specialized spin-qubit initial placement algorithm would help to substantiate and bolster many of the conclusions that we drew from our study, as we have mentioned in Sec. IV.

(c) Reference [161] attempted to address the question of connectivity among different types of qubit technologies, focusing on the implementation of various quantum algorithms in trapped-ion and superconducting devices available at the time. As the semiconductor spin-qubit community scales up their devices, it may be useful to revisit this work and benchmark again, with near-fault-tolerant logical-level algorithms and specialized compilation techniques (such as those used in our work).

(d) Larger circuits could be leveraged in order to investigate the presence of an MIPT, which we preliminarily have observed in our results with tripartite mutual information. By scaling the size of a

commensurate error-correction code, one may be able to investigate the asymptotic threshold properties that emerge, as in Ref. [99] (although such efforts may be problematic, as the definition of a threshold in QEC is contingent on the existence of arrays in the asymptotic limit). Scaling the size of our simulations would be expensive computationally; however, there are many methods by which one can optimize tensor-network simulations for large-scale quantum circuit investigations [157,158].

(e) It has been suggested in Ref. [162] that multiple device cores could be connected via shuttling- and microwave-based module interlinks [163,164]; it would be fascinating to see whether our observations hold in this modular regime.

(f) We have neglected gate parallelization in this work, in order to focus on the unique characteristics of spin-qubit devices: namely, finite-connectivity architectures and qubit density, combined with allowed shuttling operations. In the future, one could benefit from investigating parallelization protocols in this vein.

(g) Lastly, one could easily expand this study towards even more practical techniques related to near-fault-tolerant experiments, such as other types of quantum error-correction codes [37,165–167] or *magic-state distillation* [83,168,169], ingredients that are needed for large-scale, fault-tolerant quantum computing.

The architectural conclusions that we have arrived at in this study are specifically bound to the parameters that have been chosen for each error model. Although we strongly suspect that our observations hold in more general circumstances, future work will be needed in order to ascertain the specificity of our results for spin-qubit platforms. As an example, the ratio of the shuttling time to the coherence time may have a significant impact on the inferences arrived at.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available because of legal restrictions preventing unrestricted public distribution. The data are available from the authors upon reasonable request.

APPENDIX: SUCCESS RATES FOR SELECTED SIMULATIONS WITH PAULI ERROR RATES

In the results shown in Figs. 6 and 7, the success rate is calculated as the frequency of 000 or 111 syndrome measurement results, relative to the total number of Monte Carlo simulation trials. In Fig. 11, we present individual success rates alongside the corresponding Pauli error rates for each of the three CGs of size 6. The differences observed in the figures are subtle; however, with increased connectivity, we see a reduction in shuttle operations and a modest increase in the overall success rate.

Examining the Pauli error rates more closely, we noticed that Z and Y errors increase slightly up to approximately cycle 4, after which they stabilize. Conversely, Y errors exhibit a more pronounced increase until reaching a plateau at the same level as X errors. Thus, if Z and X error rates are relatively stable, Y errors appear to be the primary factor impacting the logical success rate. This is evident from the distinct trend between Y errors and the success rate. These findings suggest that the combination of this circuit with the specific architectures and error model is particularly susceptible to Y errors, and with our framework, we are able to explicitly identify and extract this vulnerability, providing valuable insights into the system's behavior.

From a more detailed perspective, Z errors are relatively low because they occur exclusively with a 100 stabilizer measurement. Additionally, the error model is negatively biased towards Z gates due to the effects of decoherence:
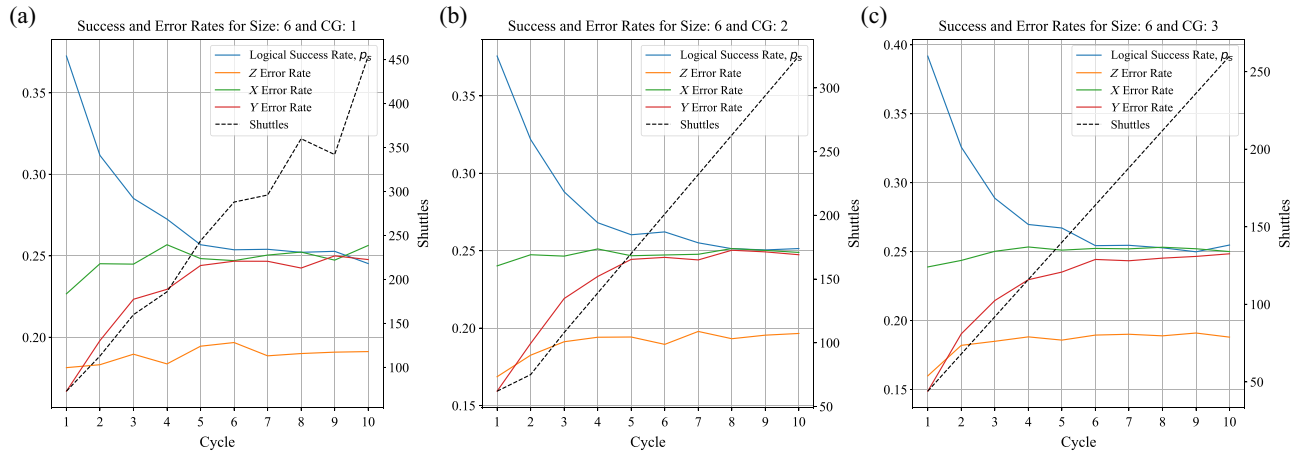
FIG. 11. Success rates for select simulations with Pauli error rates. (a) Architecture with CG_1 of size 2×6. (b) Architecture with CG_2 of size 2×6. (c) Architecture with CG_3 of size 2×6.

Z errors introduced from the decoherence-induced errors can be canceled by subsequent Z errors occurring from the depolarizing model. In contrast, X and Y errors arise under different conditions: X errors occur when the stabilizer measurement is 010 or 001, while Y errors manifest in measurements of 110 or 101 which explains their higher error rate compared to Z errors. For the X errors, we see that, at the beginning of the simulation, the amount of relative X errors is similar to the proportion dictated by depolarizing noise; however, as the simulation progresses, the injection of other Pauli errors can cause nontrivial error mixing, thus resulting in slightly larger error rates of X errors than may be normally expected in the depolarizing noise framework.

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010), ISBN 978-1-107-00217-3.

[2] M. Enríquez, I. Wintrowicz, and K. Życzkowski, in *Journal of Physics: Conference Series*, Vol. 698 (IOP Publishing, 2016), p. 012003.

[3] O. Gühne and G. Tóth, Entanglement detection, Phys. Rep. **474**, 1 (2009).

[4] D. Bluvstein, H. Levine, G. Semeghini, T. T. Wang, S. Ebadi, M. Kalinowski, A. Keesling, N. Maskara, H. Pichler, and M. Greiner *et al.*, A quantum processor based on coherent transport of entangled atom arrays, Nature **604**, 451 (2022).

[5] S. A. Moses *et al.*, A race-track trapped-ion quantum processor, Phys. Rev. X **13**, 041052 (2023).

[6] M. De Smet, Y. Matsumoto, A.-M. J. Zwerver, L. Tryputen, S. L. de Snoo, S. V. Amitonov, A. Sammak, N. Samkharadze, Ö. Gül, and R. N. Wasserman *et al.*, High-fidelity single-spin shuttling in silicon, arXiv:2406.07267.

[7] L. R. Schreiber and H. Bluhm, Toward a silicon-based quantum computer, Science **359**, 393 (2018).

[8] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Phys. Rev. Lett. **106**, 180504 (2011).

[9] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, Phys. Rev. A **85**, 042311 (2012).

[10] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Self-consistent quantum process tomography, Phys. Rev. A **87**, 062119 (2013).

[11] N. Paraskevopoulos, F. Sebastiano, C. G. Almudever, and S. Feld, Spinq: Compilation strategies for scalable spin-qubit architectures, ACM Trans. Quantum Comput. **5**, 1 (2023).

[12] N. Paraskevopoulos, D. Hamel, A. Sarkar, C. G. Almudever, and S. Feld, Arta: Automating design space exploration of spin qubit architectures, arXiv:2407.18151.

[13] L. Schmid, D. F. Locher, M. Rispler, S. Blatt, J. Zeiher, M. Müller, and R. Wille, Computational capabilities and compiler development for neutral atom quantum processors–connecting tool developers and hardware experts, Quantum Sci. Technol. **9**, 033001 (2024).

[14] S. Nishio, Y. Pan, T. Satoh, H. Amano, and R. V. Meter, Extracting success from IBM's 20-qubit machines using error-aware compilation, ACM Journal on Emerging Technologies in Computing Systems (JETC) **16**, 1 (2020).

[15] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, Validating quantum computers using randomized model circuits, Phys. Rev. A **100**, 032328 (2019).

[16] A. Wack, H. Paik, A. Javadi-Abhari, P. Jurcevic, I. Faro, J. M. Gambetta, and B. R. Johnson, Quality, speed, and scale: three key attributes to measure the performance of near-term quantum computers, arXiv:2110.14108.

[17] R. Blume-Kohout and K. C. Young, A volumetric framework for quantum computer benchmarks, Quantum **4**, 362 (2020).

[18] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, and D. A.

Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[19] W. van der Schoot, R. Wezeman, N. M. Neumann, F. Phillipson, and R. Kooij, Q-score max-clique: The first quantum metric evaluation on multiple computational paradigms, arXiv:2302.00639.

[20] M. Bandic, C. G. Almudever, and S. Feld, Interaction graph-based characterization of quantum benchmarks for improving quantum circuit mapping techniques, Quantum Mach. Intell. **5**, 40 (2023).

[21] M. A. Steinberg, S. Feld, C. G. Almudever, M. Marthaler, and J.-M. Reiner, Topological-graph dependencies and scaling properties of a heuristic qubit-assignment algorithm, IEEE Trans. Quantum Eng. **3**, 1 (2022).

[22] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, Gate set tomography, Quantum **5**, 557 (2021).

[23] B. Apak, M. Bandic, A. Sarkar, and S. Feld, in *International Conference on Computational Science* (Springer, 2024), p. 235.

[24] T. Tomesh, P. Gokhale, V. Omole, G. S. Ravi, K. N. Smith, J. Viszlai, X.-C. Wu, N. Hardavellas, M. R. Martonosi, and F. T. Chong, in *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)* (IEEE, 2022), p. 587.

[25] N. Quetschlich, L. Burgholzer, and R. Wille, Mqt bench: Benchmarking software and design automation tools for quantum computing, Quantum **7**, 1062 (2023).

[26] T. Proctor, K. Young, A. D. Baczewski, and R. Blume-Kohout, Benchmarking quantum computers, arXiv:2407.08828.

[27] W. Liu, F. Wang, H. Lin, and J. Shang, A user-centric quantum benchmarking test suite and evaluation framework, Quantum Inf. Process. **22**, 397 (2023).

[28] T. Lubinski, S. Johri, P. Varosy, J. Coleman, L. Zhao, J. Necaise, C. H. Baldwin, K. Mayer, and T. Proctor, Application-oriented performance benchmarks for quantum computing, IEEE Trans. Quantum Eng. **4**, 1 (2023).

[29] A. Cervera-Lierta, J. I. Latorre, and D. Goyeneche, Quantum circuits for maximally entangled states, Phys. Rev. A **100**, 022342 (2019).

[30] Z. Raissi, A. Teixidó, C. Gogolin, and A. Acín, Constructions of k-uniform and absolutely maximally entangled states beyond maximum distance codes, Phys. Rev. Res. **2**, 033411 (2020).

[31] Z. Raissi, C. Gogolin, A. Riera, and A. Acín, Optimal quantum error correcting codes from absolutely maximally entangled states, J. Phys. A: Math. Theor. **51**, 075301 (2018).

[32] F. Huber and M. Grassl, Quantum codes of maximal distance and highly entangled subspaces, Quantum **4**, 284 (2020).

[33] D. Miller, K. Levi, L. Postler, A. Steiner, L. Bittel, G. A. L. White, Y. Tang, E. J. Kuehnke, A. A. Mele, S. Khatri, L. Leone, J. Carrasco, C. D. Marciniak, I. Pogorelov, M. Guevara-Bertsch, R. Freund, R. Blatt, P. Schindler, T. Monz, M. Ringbauer, and J. Eisert, Experimental measurement and a physical interpretation of quantum shadow enumerators, arXiv:2408.16914 [quant-ph].

[34] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, 2013).

[35] T.-K. Hsiao, P. Cova Fariña, S. D. Oosterhout, D. Jirovec, X. Zhang, C. J. van Diepen, W. Lawrie, C.-A. Wang, A. Sammak, and G. Scappucci *et al.*, Exciton transport in a germanium quantum dot ladder, Phys. Rev. X **14**, 011048 (2024).

[36] O. Crawford, J. Cruise, N. Mertig, and M. Gonzalez-Zalba, Compilation and scaling strategies for a silicon quantum processor with sparse two-dimensional connectivity, Npj Quantum Inf. **9**, 13 (2023).

[37] A. Siegel, A. Strikis, and M. Fogarty, Towards early fault tolerance on a $2 \times n$ array of qubits equipped with shuttling, PRX Quantum **5**, 040328 (2024).

[38] X. Xue, *Performance benchmarking of silicon quantum processors*, Ph.D. thesis, TU Delft, 2022.

[39] G. Üstün, A. Morello, and S. Devitt, Single-step parity check gate set for quantum error correction, Quantum Sci. Technol. **9**, 035037 (2024).

[40] C. Jones, M. A. Fogarty, A. Morello, M. F. Gyure, A. S. Dzurak, and T. D. Ladd, Logical qubit in a linear array of semiconductor quantum dots, Phys. Rev. X **8**, 021058 (2018).

[41] A. Saraiva and S. D. Bartlett, The dawn of error correction with spin qubits, Nat. Mater. **22**, 157 (2023).

[42] F. Van Riggelen, W. Lawrie, M. Russ, N. Hendrickx, A. Sammak, M. Rispler, B. Terhal, G. Scappucci, and M. Veldhorst, Phase flip code with semiconductor spin qubits, Npj Quantum Inf. **8**, 124 (2022).

[43] B. Hetényi and J. R. Wootton, Tailoring quantum error correction to spin qubits, Phys. Rev. A **109**, 032433 (2024).

[44] K. Takeda, A. Noiri, T. Nakajima, T. Kobayashi, and S. Tarucha, Quantum error correction with silicon spin qubits, Nature **608**, 682 (2022).

[45] Z. Cai, M. A. Fogarty, S. Schaal, S. Patomäki, S. C. Benjamin, and J. J. L. Morton, A silicon surface code architecture resilient against leakage errors, Quantum **3**, 212 (2019).

[46] C. D. Hill, M. Usman, and L. C. L. Hollenberg, An exchange-based surface-code quantum computer architecture in silicon, arXiv:2107.11981 [quant-ph].

[47] A. Borras, A. Plastino, J. Batle, C. Zander, M. Casas, and A. Plastino, Multiqubit systems: Highly entangled states and entanglement distribution, J. Phys. A: Math. Theor. **40**, 13407 (2007).

[48] J. F. Marques, B. Varbanov, M. Moreira, H. Ali, N. Muthusubramanian, C. Zachariadis, F. Battistel, M. Beekman, N. Haider, and W. Vlothuizen *et al.*, Logical-qubit operations in an error-detecting surface code, Nat. Phys. **18**, 80 (2022).

[49] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, N. Lacroix, G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Repeated quantum error detection in a surface code, Nat. Phys. **16**, 875 (2020).

[50] A. Higuchi and A. Sudbery, How entangled can two couples get? Phys. Lett. A **273**, 213 (2000).

[51] A. Zabalo, M. J. Gullans, J. H. Wilson, S. Gopalakrishnan, D. A. Huse, and J. Pixley, Critical properties of the measurement-induced transition in random quantum circuits, Phys. Rev. B **101**, 060301 (2020).

[52] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse, and V. Khemani, Entanglement phase transitions

in measurement-only dynamics, Phys. Rev. X **11**, 011030 (2021).

[53] N. Paraskevopoulos, C. G. Almudever, and S. Feld, besnake: A routing algorithm for scalable spin-qubit architectures, IEEE Trans. Quantum Eng. **5** (2024).

[54] J. Yoneda, W. Huang, M. Feng, C. H. Yang, K. W. Chan, T. Tanttu, W. Gilbert, R. Leon, F. Hudson, and K. Itoh *et al.*, Coherent spin qubit transport in silicon, Nat. Commun. **12**, 4114 (2021).

[55] F. van Riggelen-Doelman, C.-A. Wang, S. L. de Snoo, W. I. Lawrie, N. W. Hendrickx, M. Rimbach-Russ, A. Sammak, G. Scappucci, C. Déprez, and M. Veldhorst, Coherent spin qubit shuttling through germanium quantum dots, Nat. Commun. **15**, 5716 (2024).

[56] T. Struck, M. Volmer, L. Visser, T. Offermann, R. Xue, J.-S. Tu, S. Trellenkamp, Ł. Cywiński, H. Bluhm, and L. R. Schreiber, Spin-EPR-pair separation by conveyor-mode single electron shuttling in Si/SiGe, Nat. Commun. **15**, 1325 (2024).

[57] D. Gottesman, Theory of quantum secret sharing, Phys. Rev. A **61**, 042311 (2000).

[58] M. Hillery, V. Bužek, and A. Berthiaume, Quantum secret sharing, Phys. Rev. A **59**, 1829 (1999).

[59] D. Gottesman, *Stabilizer Codes and Quantum Error Correction* (California Institute of Technology, 1997).

[60] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence, J. High Energy Phys. **2015**, 1 (2015).

[61] R. J. Harris, N. A. McMahon, G. K. Brennen, and T. M. Stace, Calderbank-Shor-Steane holographic quantum error-correcting codes, Phys. Rev. A **98**, 052301 (2018).

[62] M. Steinberg, S. Feld, and A. Jahn, Holographic codes from hyperinvariant tensor networks, Nat. Commun. **14**, 7314 (2023).

[63] T. Farrelly, R. J. Harris, N. A. McMahon, and T. M. Stace, Tensor-network codes, Phys. Rev. Lett. **127**, 040507 (2021).

[64] M. Steinberg, J. Fan, R. J. Harris, D. Elkouss, S. Feld, and A. Jahn, Far from perfect: Quantum error correction with (hyperinvariant) evenbly codes, arXiv:2407.11926.

[65] M. Hein, J. Eisert, and H. J. Briegel, Multiparty entanglement in graph states, Phys. Rev. A–At., Mol., Opt. Phys. **69**, 062311 (2004).

[66] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, Ann. Phys. (N. Y) **349**, 117 (2014).

[67] W. Helwig, Absolutely maximally entangled qudit graph states, arXiv:1306.2879.

[68] D. Goyeneche, D. Alsina, J. I. Latorre, A. Riera, and K. Życzkowski, Absolutely maximally entangled states, combinatorial designs, and multiunitary matrices, Phys. Rev. A **92**, 032316 (2015).

[69] M. A. Nielsen, Cluster-state quantum computation, Rep. Math. Phys. **57**, 147 (2006).

[70] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel, Bell inequalities for graph states, Phys. Rev. Lett. **95**, 120405 (2005).

[71] F. Huber, O. Gühne, and J. Siewert, Absolutely maximally entangled states of seven qubits do not exist, Phys. Rev. Lett. **118**, 200502 (2017).

[72] F. Huber, C. Eltschka, J. Siewert, and O. Gühne, Bounds on absolutely maximally entangled states from shadow inequalities, and the quantum Macwilliams identity, J. Phys. A: Math. Theor. **51**, 175301 (2018).

[73] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's theorem, quantum theory and conceptions of the universe* (Springer, 1989), p. 69.

[74] Z. Raissi, Modifying method of constructing quantum codes from highly entangled states, IEEE Access **8**, 222439 (2020).

[75] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[76] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[77] A. Y. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. (N. Y) **303**, 2 (2003).

[78] S. B. Bravyi and A. Y. Kitaev, Quantum codes on a lattice with boundary, arXiv:quant-ph/9811052.

[79] R. Acharya, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, N. Astrakhantsev, and J. Atalaya *et al.*, Quantum error correction below the surface code threshold, arXiv:2408.13687.

[80] A. Asfaw, A. Megrant, C. Jones, C. Gidney, D. Bacon, D. Debroy, D. Kafri, E. Lucero, H. Neven, and J. Hilton *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676 (2023).

[81] Z. Chen, K. J. Satzinger, J. Atalaya, A. N. Korotkov, A. Dunsworth, D. Sank, C. Quintana, M. McEwen, R. Barends, and P. V. Klimov *et al.*, Exponential suppression of bit or phase flip errors with repetitive error correction, arXiv:2102.06132.

[82] D. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, Surface code quantum computing by lattice surgery, New J. Phys. **14**, 123011 (2012).

[83] S. Bravyi and A. Kitaev, Universal quantum computation with ideal Clifford gates and noisy ancillas, Phys. Rev. A **71**, 022316 (2005).

[84] Y. Tomita and K. M. Svore, Low-distance surface codes under realistic quantum noise, Phys. Rev. A **90**, 062320 (2014).

[85] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, The XZZX surface code, Nat. Commun. **12**, 2172 (2021).

[86] M. Vasmer and D. E. Browne, Three-dimensional surface codes: Transversal gates and fault-tolerant architectures, Phys. Rev. A **100**, 012312 (2019).

[87] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Tailoring surface codes for highly biased noise, Phys. Rev. X **9**, 041031 (2019).

[88] Y.-L. Wang, Planar k-uniform states: A generalization of planar maximally entangled states, Quantum Inf. Process. **20**, 271 (2021).

[89] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, Quantum error correction in scrambling dynamics and measurement-induced phase transition, Phys. Rev. Lett. **125**, 030505 (2020).

[90] P. Sierant, M. Schirò, M. Lewenstein, and X. Turkeshi, Measurement-induced phase transitions in (d+1)-dimensional stabilizer circuits, Phys. Rev. B **106**, 214316 (2022).

[91] R. Fan, S. Vijay, A. Vishwanath, and Y.-Z. You, Self-organized error correction in random unitary circuits with measurement, Phys. Rev. B **103**, 174309 (2021).

[92] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, Rev. Mod. Phys. **82**, 277 (2010).

[93] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, Phys. Rev. B **100**, 134306 (2019).

[94] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, Phys. Rev. X **9**, 031009 (2019).

[95] J. Behrends, F. Venn, and B. Béri, Surface codes, quantum circuits, and entanglement phases, Phys. Rev. Res. **6**, 013137 (2024).

[96] H. Ollivier and W. H. Zurek, Quantum discord: A measure of the quantumness of correlations, Phys. Rev. Lett. **88**, 017901 (2001).

[97] G. Vidal and R. Tarrach, Robustness of entanglement, Phys. Rev. A **59**, 141 (1999).

[98] N. Quetschlich, L. Burgholzer, and R. Wille, Predicting good quantum circuit compilation options, arXiv:2210.08027.

[99] J. Helsen, M. Steudtner, M. Veldhorst, and S. Wehner, Quantum error correction in crossbar architectures, Quantum Sci. Technol. **3**, 035005 (2018).

[100] P. Murali, J. M. Baker, A. Javadi-Abhari, F. T. Chong, and M. Martonosi, in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems* (2019), p. 1015.

[101] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, Detecting crosstalk errors in quantum information processors, Quantum **4**, 321 (2020).

[102] *Numeric underflows* can happen when there is a loss of accuracy in numerical calculations if ESP becomes smaller than the smallest positive representable value in a programming language's floating-point arithmetic.

[103] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).

[104] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta, Quantum computing with Qiskit, arXiv:2405.08810 [quant-ph].

[105] C. Developers, Cirq (2024).

[106] R. S. Smith, M. J. Curtis, and W. J. Zeng, A practical quantum instruction set architecture, arXiv:1608.03355 [quant-ph].

[107] G. Li, Y. Ding, and Y. Xie, in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems* (2019), p. 1001.

[108] M. Steinberg, M. Bandić, S. Szkudlarek, C. G. Almudever, A. Sarkar, and S. Feld, Lightcone bounds for quantum circuit mapping via uncomplexity, Npj Quantum Inf. **10**, 113 (2024).

[109] J. Kusyk, S. M. Saeed, and M. U. Uyar, Survey on quantum circuit compilation for noisy intermediate-scale quantum computers: Artificial intelligence to heuristics, IEEE Trans. Quantum Eng. **2**, 1 (2021).

[110] S. Khandavilli, I. Palanisamy, M. V. Nguyen, T. V. Le, T. N. Nguyen, and T. N. Dinh, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01 (2023), p. 89.

[111] A. Katabarwa, K. Gratsea, A. Caesura, and P. D. Johnson, Early fault-tolerant quantum computing, PRX Quantum **5**, 020101 (2024).

[112] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[113] J. M. Taylor, H.-A. Engel, W. Dür, A. Yacoby, C. M. Marcus, P. Zoller, and M. D. Lukin, Fault-tolerant architecture for quantum computation using electrically controlled semiconductor spins, Nat. Phys. **1**, 177 (2005).

[114] L. Vandersypen, H. Bluhm, J. Clarke, A. Dzurak, R. Ishihara, A. Morello, D. Reilly, L. Schreiber, and M. Veldhorst, Interfacing spin qubits in quantum dots and donors–hot, dense, and coherent, Npj Quantum Inf. **3**, 1 (2017).

[115] A. Zwerver, T. Krähenmann, T. Watson, L. Lampert, H. C. George, R. Pillarisetty, S. Bojarski, P. Amin, S. Amitonov, and J. Boter *et al.*, Qubits made by advanced semiconductor manufacturing, Nat. Electron. **5**, 184 (2022).

[116] P. Steinacker, N. D. Stuyck, W. H. Lim, T. Tanttu, M. Feng, A. Nick, S. Serrano, M. Candido, J. D. Cifuentes, and F. E. Hudson *et al.*, A 300 mm foundry silicon spin qubit unit cell exceeding 99% fidelity in all operations, arXiv:2410.15590.

[117] H. C. George, M. T. Madzik, E. M. Henry, A. J. Wagner, M. M. Islam, F. Borjans, E. J. Connors, J. Corrigan, M. Curry, and M. K. Harper *et al.*, 12-spin-qubit arrays fabricated on a 300 mm semiconductor manufacturing line, arXiv:2410.16583.

[118] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, Semiconductor spin qubits, Rev. Mod. Phys. **95**, 025003 (2023).

[119] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. Vandersypen, Spins in few-electron quantum dots, Rev. Mod. Phys. **79**, 1217 (2007).

[120] T. Watson, S. Philips, E. Kawakami, D. Ward, P. Scarlino, M. Veldhorst, D. Savage, M. Lagally, M. Friesen, and S. Coppersmith *et al.*, A programmable two-qubit quantum processor in silicon, Nature **555**, 633 (2018).

[121] M. Veldhorst, H. Eenink, C.-H. Yang, and A. S. Dzurak, Silicon CMOS architecture for a spin-based quantum computer, Nat. Commun. **8**, 1 (2017).

[122] R. Li, L. Petit, D. P. Franke, J. P. Dehollain, J. Helsen, M. Steudtner, N. K. Thomas, Z. R. Yoscovits, K. J. Singh, and S. Wehner *et al.*, A crossbar network for silicon quantum dot qubits, Sci. Adv. **4**, eaar3960 (2018).

[123] C. D. Hill, E. Peretz, S. J. Hile, M. G. House, M. Fuechsle, S. Rogge, M. Y. Simmons, and L. C. Hollenberg, A surface code quantum computer in silicon, Sci. Adv. **1**, e1500707 (2015).

[124] J. M. Boter, J. P. Dehollain, J. P. Van Dijk, Y. Xu, T. Hensgens, R. Versluis, H. W. Naus, J. S. Clarke, M. Veldhorst, and F. Sebastiano *et al.*, Spiderweb Array: A sparse spin–qubit array, Phys. Rev. Appl. **18**, 024053 (2022).

[125] M. Künne, A. Willmes, M. Oberländer, C. Gorjaew, J. D. Teske, H. Bhardwaj, M. Beer, E. Kammerloher, R. Otten,

and I. Seidler *et al.*, The spinbus architecture for scaling spin qubits with electron shuttling, Nat. Commun. **15**, 4977 (2024).

[126] B. Buonacorsi, Z. Cai, E. B. Ramirez, K. S. Willick, S. M. Walker, J. Li, B. D. Shaw, X. Xu, S. C. Benjamin, and J. Baugh, Network architecture for a topological quantum computer in silicon, Quantum Sci. Technol. **4**, 025003 (2019).

[127] H. Bluhm and L. R. Schreiber, in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2019), p. 1.

[128] M. De Michielis, E. Ferraro, E. Prati, L. Hutin, B. Bertrand, E. Charbon, D. J. Ibberson, and M. F. Gonzalez-Zalba, Silicon spin qubits from laboratory to industry, J. Phys. D: Appl. Phys. **56**, 363001 (2023).

[129] N. P. De Leon, K. M. Itoh, D. Kim, K. K. Mehta, T. E. Northup, H. Paik, B. Palmer, N. Samarth, S. Sangtawesin, and D. W. Steuerman, Materials challenges and opportunities for quantum computing hardware, Science **372**, eabb2823 (2021).

[130] D. P. Franke, J. S. Clarke, L. M. Vandersypen, and M. Veldhorst, Rent's rule and extensibility in quantum computing, Microprocess. Microsyst. **67**, 1 (2019).

[131] B. Paquelet Wuetz, P. Bavdaz, L. Yeoh, R. Schouten, H. Van Der Does, M. Tiggelman, D. Sabbagh, A. Sammak, C. G. Almudever, and F. Sebastiano *et al.*, Multiplexed quantum transport using commercial off-the-shelf CMOS at sub-kelvin temperatures, Npj Quantum Inf. **6**, 1 (2020).

[132] S. Pauka, K. Das, R. Kalra, A. Moini, Y. Yang, M. Trainer, A. Bousquet, C. Cantaloube, N. Dick, and G. Gardner *et al.*, A cryogenic interface for controlling many qubits, arXiv:1912.01299.

[133] L. C. Camenzind, S. Geyer, A. Fuhrer, R. J. Warburton, D. M. Zumbuhl, and A. V. Kuhlmann, A hole spin qubit in a fin field-effect transistor above 4 kelvin, Nat. Electron. **5**, 178 (2022).

[134] N. W. Hendrickx, W. I. Lawrie, M. Russ, F. van Riggelen, S. L. de Snoo, R. N. Schouten, A. Sammak, G. Scappucci, and M. Veldhorst, A four-qubit germanium quantum processor, Nature **591**, 580 (2021).

[135] A. Chatterjee, P. Stevenson, S. De Franceschi, A. Morello, N. P. de Leon, and F. Kuemmeth, Semiconductor qubits in practice, Nat. Rev. Phys. **3**, 157 (2021).

[136] 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, Silicon quantum electronics, Rev. Mod. Phys. **85**, 961 (2013).

[137] D. Loss and D. P. DiVincenzo, Quantum computation with quantum dots, Phys. Rev. A **57**, 120 (1998).

[138] M. Veldhorst, C. Yang, J. Hwang, W. Huang, J. Dehollain, J. Muhonen, S. Simmons, A. Laucht, F. Hudson, and K. M. Itoh *et al.*, A two-qubit logic gate in silicon, Nature **526**, 410 (2015).

[139] D. Zajac, T. Hazard, X. Mi, K. Wang, and J. R. Petta, A reconfigurable gate architecture for Si/SiGe quantum dots, Appl. Phys. Lett. **106**, 223507 (2015).

[140] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, and Y. Hoshi *et al.*, A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%, Nat. Nanotechnol. **13**, 102 (2018).

[141] F. van Riggelen-Doelman, C.-A. Wang, S. L. de Snoo, W. I. L. Lawrie, N. W. Hendrickx, M. Rimbach-Russ, A. Sammak, G. Scappucci, C. Déprez, and M. Veldhorst, Coherent spin qubit shuttling through germanium quantum dots, Nat. Commun. **15** (2024).

[142] X. Xue, M. Russ, N. Samkharadze, B. Undseth, A. Sammak, G. Scappucci, and L. M. Vandersypen, Quantum logic with spin qubits crossing the surface code threshold, Nature **601**, 343 (2022).

[143] A. Noiri, K. Takeda, T. Nakajima, T. Kobayashi, A. Sammak, G. Scappucci, and S. Tarucha, Fast universal quantum gate above the fault-tolerance threshold in silicon, Nature **601**, 338 (2022).

[144] K. Takeda, A. Noiri, T. Nakajima, L. C. Camenzind, T. Kobayashi, A. Sammak, G. Scappucci, and S. Tarucha, Rapid single-shot parity spin readout in a silicon double quantum dot with fidelity exceeding 99%, Npj Quantum Inf. **10** (2024).

[145] S. G. Philips, M. T. Madzik, S. V. Amitonov, S. L. de Snoo, M. Russ, N. Kalhor, C. Volk, W. I. Lawrie, D. Brousse, and L. Tryputen *et al.*, Universal control of a six-qubit quantum processor in silicon, Nature **609**, 919 (2022).

[146] I. Heinz and G. Burkard, Crosstalk analysis for single-qubit and two-qubit gates in spin qubit arrays, Phys. Rev. B **104**, 045420 (2021).

[147] S. Patomäki, M. Gonzalez-Zalba, M. Fogarty, Z. Cai, S. Benjamin, and J. Morton, Pipeline quantum processor architecture for silicon spin qubits, Npj Quantum Inf. **10**, 31 (2024).

[148] J. Gray, quimb: A python package for quantum information and many-body calculations, J. Open Source Softw. **3**, 819 (2018).

[149] D. Bhatnagar, M. Steinberg, D. Elkouss, C. G. Almudever, and S. Feld, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 1 (IEEE, 2023), p. 63.

[150] R. Chao and B. W. Reichardt, Quantum error correction with only two extra qubits, Phys. Rev. Lett. **121**, 050502 (2018).

[151] C. Chamberland and M. E. Beverland, Flag fault-tolerant error correction with arbitrary distance codes, Quantum **2**, 53 (2018).

[152] S. Vittal, A. Javadi-Abhari, A. W. Cross, L. S. Bishop, and M. Qureshi, Flag proxy networks: Tackling the architectural, scheduling, and decoding obstacles of quantum ldpc codes, arXiv:2409.14283.

[153] P. Murali, D. C. McKay, M. Martonosi, and A. Javadi-Abhari, in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems* (2020), p. 1001.

[154] P. Parrado-Rodríguez, C. Ryan-Anderson, A. Bermudez, and M. Müller, Crosstalk suppression for fault-tolerant quantum error correction with trapped ions, Quantum **5**, 487 (2021).

[155] T. Yu and J. Eberly, Sudden death of entanglement, Science **323**, 598 (2009).

[156] X. Feng, Y. Deng, and H. W. J. Blöte, Percolation transitions in two dimensions, Phys. Rev. E **78**, 031136 (2008).

[157] J. Gray and S. Kourtis, Hyper-optimized tensor network contraction, Quantum **5**, 410 (2021).

[158] J. Gray and G. K.-L. Chan, Hyperoptimized approximate contraction of tensor networks with arbitrary geometry, Phys. Rev. X **14**, 011009 (2024).

[159] X. Xu, S. C. Benjamin, and X. Yuan, Variational circuit compiler for quantum error correction, Phys. Rev. Appl. **15**, 034068 (2021).

[160] T. E. O'Brien, B. Tarasinski, and L. DiCarlo, Density-matrix simulation of small surface codes under current and projected experimental noise, Npj Quantum Inf. **3**, 39 (2017).

[161] N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, Experimental comparison of two quantum computing architectures, Proc. Natl. Acad. Sci. **114**, 3305 (2017).

[162] H. Jnane, B. Undseth, Z. Cai, S. C. Benjamin, and B. Koczor, Multicore quantum computing, Phys. Rev. Appl. **18**, 044064 (2022).

[163] T. A. Baart, M. Shafiei, T. Fujita, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen, Single-spin CCD, Nat. Nanotechnol. **11**, 330 (2016).

[164] J. Dijkema, X. Xue, P. Harvey-Collard, M. Rimbach-Russ, S. L. de Snoo, G. Zheng, A. Sammak, G. Scappucci, and L. M. Vandersypen, Two-qubit logic between distant spins in silicon, arXiv:2310.16805.

[165] J. Fan, M. Steinberg, A. Jahn, C. Cao, and S. Feld, Overcoming the zero-rate hashing bound with holographic quantum error correction, arXiv:2408.06232.

[166] J. Old, M. Rispler, and M. Müller, Lift-connected surface codes, Quantum Sci. Technol. **9**, 045012 (2024).

[167] J. Roffe, D. R. White, S. Burton, and E. Campbell, Decoding across the quantum low-density parity-check code landscape, Phys. Rev. Res. **2**, 043423 (2020).

[168] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, Experimental magic state distillation for fault-tolerant quantum computing, Nat. Commun. **2**, 169 (2011).

[169] P. S. Rodriguez, J. M. Robinson, P. N. Jepsen, Z. He, C. Duckering, C. Zhao, K.-H. Wu, J. Campo, K. Bagnall, and M. Kwon *et al.*, Experimental demonstration of logical magic state distillation, arXiv:2412.15165.