Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Encoding Gauge of Fermionic Variational Quantum Algorithms: Classical Simulability is Encoding-Relative, Trainability is Invariant

S. M. Yousuf Iqbal Tomal

yousuf.iqbal.tomal@g.bracu.ac.bd

BRAC University https://orcid.org/0009-0000-3391-3824

Abdullah Al Shafin

BRAC University https://orcid.org/0009-0007-6634-6900

Research Article

Keywords: Fermion-to-qubit encoding, Variational quantum algorithms, Classical simulability, Barren plateaus, Dynamical Lie algebra, Stabilizer R\'enyi entropy

Posted Date: July 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10405703/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

The Encoding Gauge of Fermionic Variational Quantum Algorithms: Classical Simulability is Encoding-Relative, Trainability is Invariant

S.M. Yousuf Iqbal Tomal* Abdullah Al Shafin†
Department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh

Abstract

Whether a fermionic variational quantum algorithm offers genuine advantage depends on whether its outputs resist classical computation. We show that classical simulability is encoding-dependent, not intrinsic to the circuit. Standard encodings are Clifford-related, forming a gauge group, and we classify resources by how they transform under it. Pauli weight, causal-cone width, and the entanglement setting tensor-network cost are gauge-covariant: an interacting Hamiltonian's truncated Pauli-propagation cost is linear under Jordan-Wigner but logarithmic under a tree encoding, changing the verdict at fixed ansatz. Dynamical-Lie-algebra dimension and non-stabilizerness are gauge-invariant. We prove the encoding-optimized simulation cost is bounded by the invariant resources, so encoding-robust hardness requires a large Lie algebra and large magic; covariant weight is necessary but never sufficient. These routes are sufficient certificates, not a complete characterization, so the floor is an upper bound; a full classification of representation-invariant costs remains open. We confirm both invariant routes numerically, show Lie-algebra dimension separates simulable from hard instances where weight cannot, and construct a one-parameter family where all three certificates fail at once, achieving the floor. We then establish a sharp dichotomy with trainability: because the cost landscape is encoding-invariant, gradient and barren-plateau variance are exactly gauge-invariant to machine precision. This follows from faithfulness, not Clifford-equivalence, and survives symmetry tapering, so trainability shifts only by changing what the dynamics explores, such as restricting the ansatz, never by relabeling qubits. One may re-encode a circuit to ease simulation but never to ease training. These results reframe advantage claims: an advantage that is not gauge-invariant is an artifact of the encoding.

Keywords: Fermion-to-qubit encoding · Variational quantum algorithms · Classical simulability · Barren plateaus · Dynamical Lie algebra · Stabilizer Rényi entropy

1 Introduction

1.1 The advantage question and its dependence on representation

A variational quantum algorithm [1, 2] can deliver advantage only if its outputs resist efficient classical computation. Its trainability can be destroyed by barren plateaus that flatten the optimization landscape as the system grows [3, 4, 5], by hardware noise [6], and by a proliferation of poor local minima [7]; a recent synthesis surveys these mechanisms [8]. Trainability is necessary

*Corresponding author: yousuf.iqbal.tomal@g.bracu.ac.bd; ORCiD: 0009-0000-3391-3824
†abdullah.al.shafin@g.bracu.ac.bd; ORCiD: 0009-0007-6634-6900

but not sufficient, though: the decisive axis is classical simulability [9, 10]. For fermionic problems, which include electronic structure and lattice models, the quantum circuit does not act on fermions directly. It acts on qubits through a fermion-to-qubit encoding, and the choice of encoding is a modeling convention rather than a property of the physical problem. The Jordan-Wigner transformation [11], the Bravyi-Kitaev transformation [12, 13], the parity transformation, and tree-based encodings [14] all represent the same fermionic algebra on the same number of qubits, and they all reproduce identical expectation values for identical physical inputs.

We ask a question that this freedom makes unavoidable. If two encodings represent the same problem, do they agree on whether that problem is classically simulable? The answer is no, and the structure of the disagreement is the subject of this paper.

1.2 Relation to prior work

Three facts in the literature bear directly on this question, and we are careful to separate them from our contribution.

First, the standard fermion-to-qubit encodings are linear over GF(2) [15], so they are implemented by networks of controlled-NOT gates, which are Clifford, and are therefore Clifford-equivalent; this equivalence is known to extend even to mappings in two spatial dimensions [16]. Second, the encodings differ in the Pauli weight they assign to fermionic operators: the Jordan-Wigner encoding produces operator strings of weight linear in the system size, while the Bravyi-Kitaev and tree encodings reduce this to logarithmic, the ternary tree of [14] being weight-optimal. This weight gap has been quantified empirically across many molecular Hamiltonians [17], and locality-tailored mappings reduce it further for spatially structured systems [18, 19]. Third, several leading classical simulators have cost controlled by exactly these structural quantities: truncated Pauli propagation by the operator weight [20, 21], Lie-algebraic simulation by the dynamical-Lie-algebra dimension [22, 23], and stabilizer-rank methods by the non-stabilizerness [24, 25]. These simulators are not merely formal. Related Pauli-propagation and tensor-network methods have recently matched and even surpassed flagship superconducting-processor demonstrations on classical hardware [26, 27, 28], so the encoding dependence of their cost carries practical weight. The Lie-algebraic route is sharpened by classifications of which generator sets close into a small algebra [29], and the stabilizer route by improved low-rank stabilizer decompositions [31].

Each fact is established. What has not been done, and what we provide, is to combine them into a statement about *simulability* and to draw its consequences. We treat the encodings as a gauge group acting on the simulability problem, classify the simulation resources by their behavior under this group, prove that the encoding-optimized cost is controlled by the gauge-invariant resources alone, and show that trainability is gauge-invariant while simulability is not. The practical reading is a criterion for advantage claims, stated in Section 7. None of these statements is a restatement that one encoding has lower weight than another; that observation is the starting point, not the result.

1.3 Contributions

1. A classification of the standard simulability resources under the encoding gauge: the Pauli weight and the causal-cone width are covariant, the dynamical-Lie-algebra dimension and the non-stabilizerness are invariant (Section 2.3).

2. A demonstration that the covariance changes the verdict. For an interacting fermionic ansatz, the truncated Pauli-propagation cost crosses from intractable to tractable as the encoding

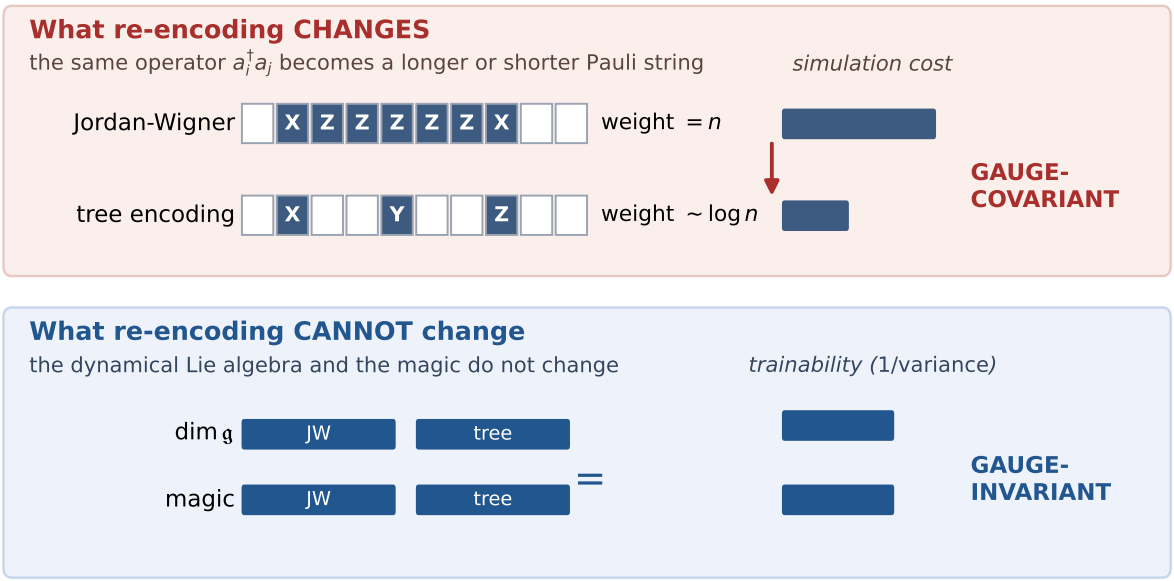


Figure 1. The encoding gauge. Top: re-encoding changes the representation. The same operator $a_i^\dagger a_j$ becomes a Pauli string of weight n under Jordan-Wigner but only $\sim \log n$ under a tree encoding, so the weight-controlled simulation cost is gauge-covariant and falls when the encoding is changed. Bottom: re-encoding leaves the dynamical-Lie-algebra dimension and the magic unchanged for every encoding, so trainability, which these invariants set, is gauge-invariant and cannot be changed by re-encoding.

changes, at a fixed physical circuit, with the change widening as the system grows (Section 4, Figure 3).

3. A floor proposition: the encoding-optimized simulation cost is bounded above by the invariant resources, so encoding-robust hardness requires both the dynamical Lie algebra and the magic to be large. We exhibit both invariant simulation routes, and we construct an explicit one-parameter interacting family for which all three certificate routes provably fail at once, so the floor is achieved (Section 5, Proposition 2).

4. A dichotomy: trainability is gauge-invariant. The gradient and the barren-plateau variance are identical across encodings to machine precision, in contrast with the covariant simulation cost (Section 6). We sharpen this to a boundary theorem: faithfulness, not Clifford-equivalence, guarantees the invariance, which is robust even to symmetry tapering; trainability moves only when the dynamically explored algebra is changed, for instance by restricting the ansatz (Proposition 3).

Figure 1 summarizes the framework: one fermionic problem, an encoding orbit related by Clifford transformations, and the split of the simulation resources into a covariant class and an invariant class, so that re-encoding can lower the simulation cost but leaves trainability unchanged.

2 The encoding gauge

2.1 Encodings and their Clifford equivalence

A fermion-to-qubit encoding on n modes is a faithful representation of the fermionic operator algebra on n qubits. We consider the standard family: Jordan-Wigner, Bravyi-Kitaev, the Bravyi-Kitaev tree, parity, and the weight-optimal ternary tree. Each is determined by an invertible linear map over GF(2) relating occupation bits to qubit bits [15], so each is realized by a controlled-NOT network and lies in the Clifford group up to single-qubit basis choices. Consequently any two encodings E and E' in this family are related by a Clifford unitary C through $E'(\cdot) = C\,E(\cdot)\,C^\dagger$.

Definition 1 (Encoding orbit)**.** *The encoding orbit of a fermionic problem is the set of qubit representations obtained from the standard encoding family. Since these are Clifford-related, the orbit is a Clifford-equivalence class, which we call the encoding gauge.*

We verified the defining property numerically: for a fixed fermionic Hamiltonian, the qubit Hamiltonians produced by the encodings in the orbit are isospectral to machine precision, and the physical expectation values of a fixed observable through a fixed ansatz are identical across the orbit to better than 10^{-15}.

2.2 Definitions

We collect the objects and the simulation resources used throughout. Let $\mathcal{P}_n$ denote the n-qubit Pauli group.

Encodings. Write the $2n$ Majorana operators of n fermionic modes as $\gamma_1, \ldots, \gamma_{2n}$. Represent each n-qubit Pauli operator, up to phase, by its binary symplectic check vector $(x \,|\, z) \in \mathrm{GF}(2)^{2n}$, where x and z record the X- and Z-support. A linear fermion-to-qubit encoding E is then a fixed invertible matrix $M \in \mathrm{GF}(2)^{2n\times 2n}$ that sends the check vector of each γ_a to M times that vector; invertibility makes E a faithful representation. The Jordan-Wigner encoding is the map

$$\gamma_{2j-1} \;\mapsto\; \Big(\textstyle\prod_{k<j} Z_k\Big) X_j, \qquad \gamma_{2j} \;\mapsto\; \Big(\textstyle\prod_{k<j} Z_k\Big) Y_j. \tag{1}$$

Such a matrix is realized by a network of controlled-NOT gates, so any two encodings E, E' in the standard family obey $E'(\cdot) = C\,E(\cdot)\,C^\dagger$ for a Clifford unitary C (Section 2).

Pauli weight W. For $P = \bigotimes_j P_j \in \mathcal{P}_n$ the weight $|P|$ is the number of tensor factors with $P_j \neq I$. The weight of an operator is the largest weight among the Pauli strings of its expansion. The cost of truncated Pauli propagation scales as $n^{O(W)}$ [20].

Causal-cone width ν. For an observable, the backward causal cone is the set of qubits that can influence it through the circuit; its width ν is the number of qubits in that set. Contracting the circuit inside the cone costs $2^{O(\nu)}$ [9].

Dynamical Lie algebra $\mathfrak{g}$. For an ansatz $U(\theta) = \prod_k e^{-i\theta_k H_k}$ with Hermitian generators $\{H_k\}$,

$$\mathfrak{g} = \langle\, iH_1, \ldots, iH_K \,\rangle_{\mathrm{Lie}} \tag{2}$$

is the real Lie algebra spanned by the generators and their nested commutators; $\dim \mathfrak{g}$ is its real dimension. Lie-algebraic simulation costs $\mathrm{poly}(\dim \mathfrak{g})$ [22, 23].

Non-stabilizerness (magic) M_2. For a pure state $|\psi\rangle$ the second stabilizer Rényi entropy is

$$M_2(|\psi\rangle) = -\log_2\Big(2^{-n} \sum_{P\in\mathcal{P}_n} \langle\psi|P|\psi\rangle^4\Big), \tag{3}$$

which is non-negative and vanishes exactly on stabilizer states [25]. Stabilizer-rank simulation costs $2^{O(t)}$ in a magic measure t such as the stabilizer rank [24]; t and M_2 are both minimized on stabilizer states, and we use M_2 as the computed witness.

Truncation weight W^*. The Heisenberg observable $O(\theta) = U^\dagger(\theta)\, O\, U(\theta) = \sum_P a_P(\theta)\, P$ is propagated keeping only terms with $|P| \le W$. The absolute-error truncation weight W^* is the least W for which the truncated estimate of $\langle O \rangle$ lies within a fixed tolerance ϵ of the exact value.

2.3 Resources transform either covariantly or invariantly

A quantity attached to the qubit representation is gauge-invariant if it is preserved by Clifford conjugation, and gauge-covariant otherwise. The simulation resources split cleanly.

The **dynamical-Lie-algebra dimension** and any **magic monotone**, including the stabilizer rank and the stabilizer Rényi entropy, are Clifford-invariant, hence gauge-invariant. The **Pauli weight**, the **causal-cone width**, and the **entanglement** across a cut are not preserved by Clifford conjugation. A Clifford map can lengthen or shorten operator strings, and it can create or destroy entanglement (a single controlled-NOT turns a product state into a Bell pair). These three are therefore gauge-covariant. The last of these matters because the cost of tensor-network contraction is set by entanglement; it is therefore a covariant resource, not a blind spot of the framework (Section 7). The following lemma makes the split precise.

Lemma 1 (Resource transformation under the gauge)**.** *Let $E'(\cdot) = C\, E(\cdot)\, C^\dagger$ for a Clifford unitary C. Then (i) the dynamical-Lie-algebra dimension is invariant; (ii) the second stabilizer Rényi entropy is invariant, $M_2(C|\psi\rangle) = M_2(|\psi\rangle)$; and (iii) the Pauli weight is in general not invariant.*

Proof. (i) Conjugation by C is a Lie-algebra automorphism, $C[A,B]C^\dagger = [CAC^\dagger, CBC^\dagger]$, so the algebra (2) is carried to $C\mathfrak{g}C^\dagger$, an isomorphic copy of equal dimension. (ii) A Clifford conjugation permutes $\mathcal{P}_n$ up to sign, $C^\dagger P C = \pm P'$ with $P \mapsto P'$ a bijection of $\mathcal{P}_n$; the sign cancels in the fourth power in (3), so $\sum_P \langle\psi|C^\dagger P C|\psi\rangle^4 = \sum_{P'} \langle\psi|P'|\psi\rangle^4$ and $M_2(C|\psi\rangle) = M_2(|\psi\rangle)$. (iii) A controlled-NOT staircase carries the weight-one operator Z_1 to the weight-n string $Z_1 Z_2 \cdots Z_n$, so weight is not preserved. □

The causal-cone width is covariant for the same reason as part (iii): the backward cone of an observable is the union of the supports of its back-propagated Pauli terms, and re-encoding changes those supports, so a non-local encoding such as Jordan-Wigner produces wider cones than a tree encoding. Figure 2 shows the structural origin of both covariant resources. The same operator is a full Pauli-Z string spanning the register on the Jordan-Wigner chain, but a single short root-to-leaf path in a tree, so its weight falls from n to $O(\log n)$ (panels (a) and (b)). For the same reason the backward causal cone of an observable is wide under the long-range Jordan-Wigner gates but narrow under the local gates of a tree encoding (panels (c) and (d)). Table 1 summarizes the split together with the classical method each resource controls.

Table 1. Classification of simulability resources under the encoding gauge.

Resource	Classical method	Cost	Under the gauge
Pauli weight W	truncated Pauli propagation	$n^{O(W)}$	covariant
Causal-cone width ν	light-cone contraction	$2^{O(\nu)}$	covariant
Entanglement S	tensor-network contraction	$2^{O(S)}$	covariant
Lie-algebra dimension $\dim \mathfrak{g}$	Lie-algebraic simulation	$\mathrm{poly}(\dim \mathfrak{g})$	invariant
Non-stabilizerness t	stabilizer-rank simulation	$2^{O(t)}$	invariant

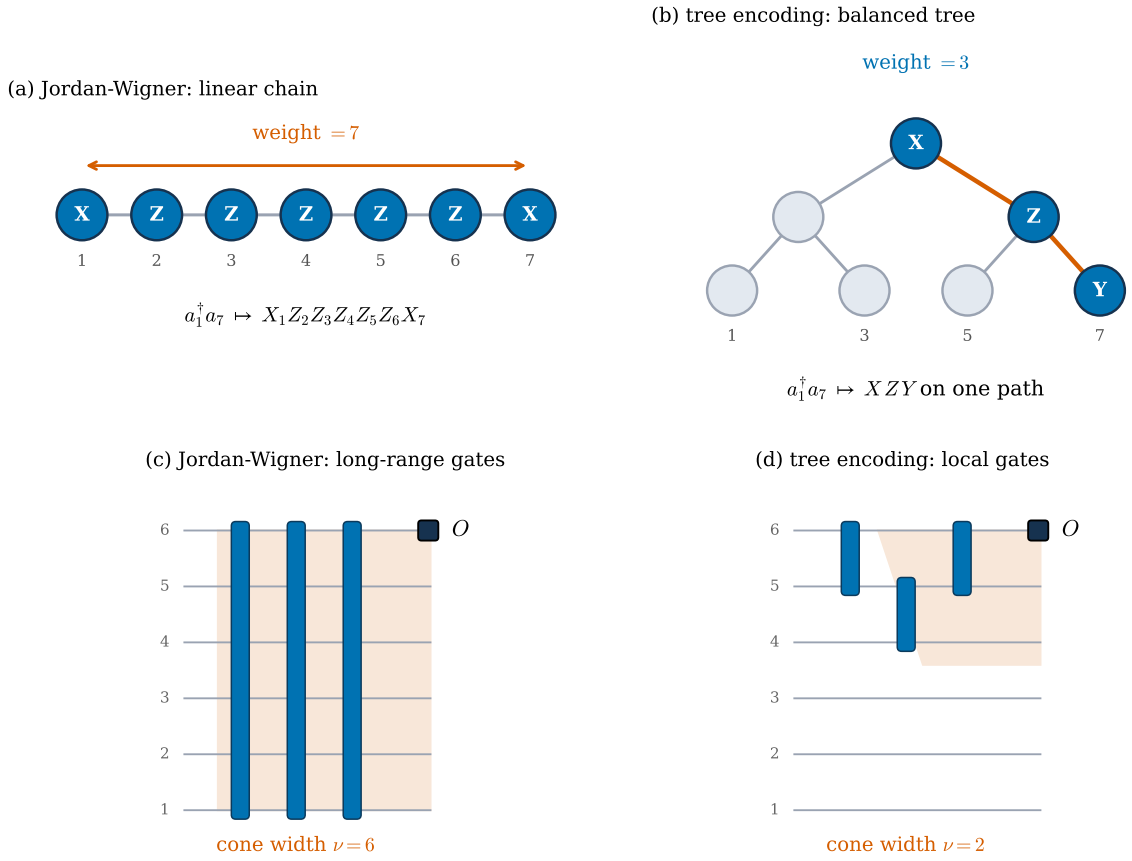


Figure 2. The structural origin of covariance. Top: the same operator $a_i^\dagger a_j$ is a full Pauli-Z string of weight n along the Jordan-Wigner chain (a), but a single short root-to-leaf path of weight $O(\log n)$ in a tree encoding (b). Bottom: because Jordan-Wigner gates act on long strings, the backward causal cone of an observable covers the whole register (c), whereas the local gates of a tree encoding keep it narrow (d). The two covariant resources shown here, the Pauli weight and the causal-cone width, trace to the same fact: Jordan-Wigner is non-local while tree encodings are local.

We confirmed the classification directly. Across the orbit, the dynamical-Lie-algebra dimension and the stabilizer Rényi entropy of fixed states are identical to machine precision, while the maximum Pauli weight of the encoded Hamiltonian differs by amounts that grow with the system size. The validation of the Lie-algebra computation against two independent closed-form algebras, $\mathfrak{u}(L)$ for nearest-neighbor free fermions and $\mathfrak{su}(n)$ for all-to-all hopping, is reported in Section 5.

3 Models and numerical methods

Encodings and reference states. We use the standard linear fermion-to-qubit encodings, Jordan-Wigner, Bravyi-Kitaev, the Bravyi-Kitaev tree, and parity, each applied to the Hamiltonian, to the ansatz generators, and to the Hartree-Fock reference state, with the reference encoded consistently in the same representation as the circuit. For each encoding the Hartree-Fock state is the computational-basis state whose encoded number operators reproduce the target occupations.

Problem families. Three families appear. The first two share a generic all-to-all form. With Gaussian random couplings t_{ij} and V_{ij},

$$H_{\text{free}} = \sum_{i<j} t_{ij}(a_i^\dagger a_j + a_j^\dagger a_i), \qquad H_{\text{int}} = H_{\text{free}} + \sum_{i<j} V_{ij}\, n_i n_j, \tag{4}$$

the free model H_{free} being quadratic and hence Gaussian, and H_{int} its interacting extension. These are the $\lambda = 0$ and $\lambda \neq 0$ ends of the family (8) studied in Section 5. The third family is molecular electronic structure, the standard second-quantized Hamiltonian

$$H_{\text{mol}} = \sum_{pq} h_{pq}\, a_p^\dagger a_q + \tfrac{1}{2} \sum_{pqrs} h_{pqrs}\, a_p^\dagger a_q^\dagger a_r a_s, \tag{5}$$

with one- and two-electron integrals h_{pq} and h_{pqrs} in a chosen basis. In every case the variational circuit is a Hamiltonian-variational ansatz built from the Hermitian generators of the corresponding problem, with each generator Trotterized into elementary Pauli rotations.

The structured ansatz of Figure 3(b). The circuit-level measurement uses a fixed non-local ansatz, specified here in full. Its generator set is

$$\mathcal{G} = \{\, a_i^\dagger a_{i+n/2} + a_{i+n/2}^\dagger a_i \,\}_{i=0}^{n/2-1} \;\cup\; \{\, n_i n_{i+1} \,\}_{i \text{ even}}, \tag{6}$$

which combines long-range hops of range $n/2$ with local density-density terms. The ansatz is applied in two Trotter layers with independent per-generator angles drawn uniformly from $[0.4, 0.85]$ under a fixed random seed, the observable is the number operator $n_{n/2}$, and the reference is the half-filling Hartree-Fock state ($n/2$ occupied modes) encoded per representation. The absolute-error tolerance is $\epsilon = 0.05$ and the hardness threshold is $W = 2\log_2 n$, the scale at which the $n^{O(W)}$ cost of truncated Pauli propagation crosses from polynomial to quasi-polynomial; the qualitative separation between the encodings does not depend on this constant. The long range of the hopping generators is what makes the Jordan-Wigner strings non-local and the truncation weight covariant.

Pauli weight and truncated propagation. The weight-controlled simulator back-propagates an observable in the Heisenberg picture through the sequence of Pauli rotations, discarding at each step any Pauli term whose weight exceeds a cutoff W. The absolute-error truncation weight W^* is the smallest cutoff at which the truncated estimate reproduces the exact expectation value to a fixed tolerance. The exact reference and the propagation engine were cross-checked against an independent statevector simulation, agreeing to a relative error below 10^{-15}, which is what makes W^* well defined.

Dynamical Lie algebra. The dimension $\dim \mathfrak{g}$ is that of the real Lie algebra generated by the encoded ansatz generators under commutation. We compute it two ways: a dense closure over operator matrices, used for validation at small sizes, and a sparse closure over Pauli strings that avoids exponential-size matrices and reaches substantially larger systems. The two agree exactly where both run, and both reproduce the closed-form dimensions $\mathfrak{u}(L) = L^2$ and $\mathfrak{su}(n) = n^2 - 1$.

Non-stabilizerness. Magic is quantified by the second stabilizer Rényi entropy of the output state. As a check it vanishes on stabilizer states and reproduces the analytic value on single-qubit magic states.

Trainability. Gradients are obtained by central finite differences of the exact expectation value, itself validated against the statevector simulation. The barren-plateau diagnostic is the variance of a cost derivative over uniformly random parameters, averaged over the most active parameters and many random samples to control the estimator noise.

Validation. All computations are carried out in IEEE double precision (float64), so the "machine-precision" agreements quoted throughout are at the 10^{-12}–10^{-15} level set by accumulated floating-point rounding. Every component is checked against independent ground truth before any result is reported: the encodings are isospectral; the encoded Hartree-Fock energy matches the stored molecular value to better than 10^{-12}; the Pauli engine matches the statevector simulation; $\dim \mathfrak{g}$ matches the closed-form algebra dimensions; and the magic measure matches its analytic value.

4 Simulation cost is gauge-covariant

4.1 Resource-level separation

The Pauli weight of the encoded Hamiltonian is the simplest covariant resource, and its behavior is unambiguous. For a generic interacting fermionic Hamiltonian with all-to-all two-body terms, the maximum Pauli weight under the Jordan-Wigner encoding equals the system size n, while under the Bravyi-Kitaev and tree encodings it grows only logarithmically. Figure 3(a) shows this separation out to $n = 30$: the Jordan-Wigner and parity curves coincide on the line of slope one, and the Bravyi-Kitaev curve sits near the logarithm. Because the truncated Pauli-propagation cost scales as $n^{O(W)}$, this is already an exponential separation in the simulation cost between two representations of the same problem. The all-to-all two-body structure is not an artifice of the example. Molecular electronic-structure Hamiltonians are all-to-all in the orbital basis, and long-range fermionic systems such as the Sachdev-Ye-Kitaev model and trapped-ion simulators share it. The Jordan-Wigner weight of such Hamiltonians is genuinely linear in n, which is the documented reason the Bravyi-Kitaev and tree encodings were introduced [13, 17, 14].

4.2 Circuit-level change of verdict

The resource-level separation propagates to the realized simulation cost. We measure the absolute-error truncation weight W^* defined in Section 3, the smallest Pauli weight at which truncated propagation reproduces an observable through the ansatz to fixed accuracy, with the Hartree-Fock reference encoded consistently in each representation.

Figure 3(b) reports W^* as a function of system size for the structured non-local ansatz specified in Section 3, over the directly computed range $6 \le n \le 16$. Under the Jordan-Wigner encoding W^* rises with the system size and crosses the tractability threshold $2\log_2 n$, while under the Bravyi-Kitaev encoding it stays below. For every size $n \ge 12$ in the range shown the verdict therefore splits,

$$W^*_{\mathrm{JW}} \;>\; 2\log_2 n \;\ge\; W^*_{\mathrm{BK}}, \tag{7}$$

so the same physical circuit is intractable for truncated Pauli propagation under Jordan-Wigner and parity yet tractable under Bravyi-Kitaev. The gap widens with system size across the computed range.

We note the scope of this measurement. A large W^* certifies that truncated Pauli propagation is expensive, not that every classical method fails. The verdict that changes across the orbit is the verdict of this specific, weight-controlled method, which is exactly the method whose cost is gauge-covariant. A second limitation is intrinsic to the quantity itself. Computing W^* requires

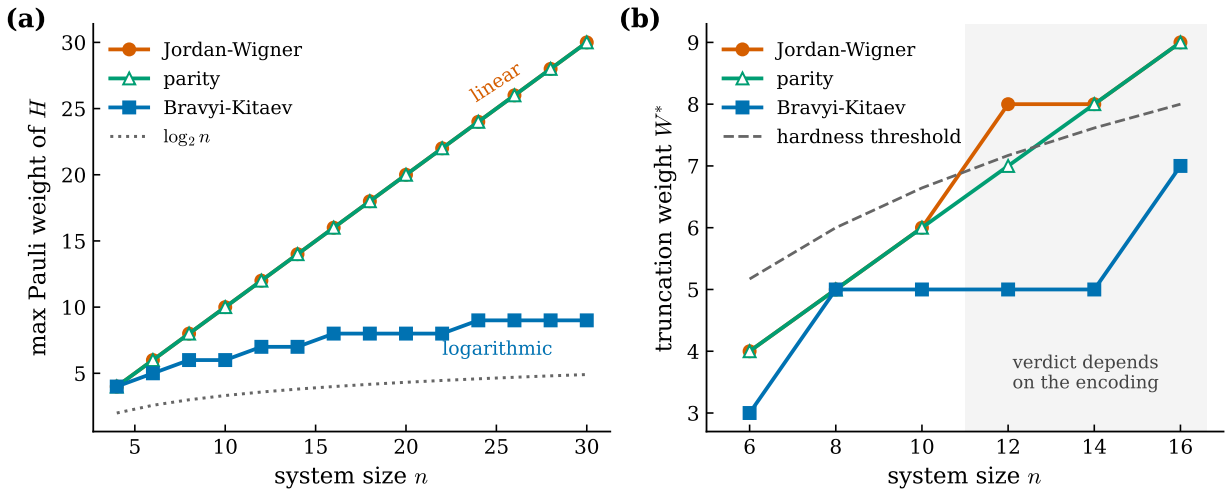


Figure 3. Simulation cost is gauge-covariant. (a) Maximum Pauli weight of a generic interacting Hamiltonian: linear under Jordan-Wigner and parity, logarithmic under Bravyi-Kitaev. (b) Absolute-error truncation weight W^* of the number operator $n_{n/2}$ through the structured nonlocal ansatz of Section 3 (long-range hops $a_i^\dagger a_{i+n/2}$ + h.c. and local terms $n_i n_{i+1}$, two Trotter layers, tolerance $\epsilon = 0.05$, threshold $2\log_2 n$), with the half-filling Hartree-Fock reference encoded per representation, over the directly computed range $6 \le n \le 16$. Jordan-Wigner and parity cross the hardness threshold while Bravyi-Kitaev stays below it, so the simulability verdict depends on the encoding in the shaded region ($n \ge 12$). W^* is deterministic, the least integer cutoff meeting the tolerance ϵ, and carries no statistical uncertainty.

the exact expectation value as a reference, so the flip can be measured only where exact classical simulation is still feasible. The regime in which the encoding gap grows largest is also the regime in which that exactness runs out. We therefore probe the onset of the flip at reachable sizes, and the structured family is built to enter it while remaining exactly simulable.

4.3 A molecular instance

The covariant and invariant resources both appear on real electronic-structure Hamiltonians. Table 2 reports the dihydrogen molecule in the STO-3G and 6-31G bases, with the half-filling Hartree-Fock reference encoded per representation and the same observable convention and tolerance as above. The invariant resources, the dynamical-Lie-algebra dimension and the stabilizer Rényi entropy, are identical across the encodings to the computed precision, as is the physical expectation value, while the covariant truncation weight is ordered Bravyi-Kitaev $\le$ Jordan-Wigner, the same direction as in Figure 3. At these small, exactly simulable sizes the weight gap is not yet wide enough to force a change of tractability verdict, which the structured family of Figure 3(b) exhibits only for $n \ge 12$; the molecular data therefore confirm the gauge split itself rather than a change of verdict.

Table 2. A molecular instance. The gauge split on the dihydrogen molecule (STO-3G and 6-31G bases, half-filling Hartree-Fock reference, $\epsilon = 0.05$, threshold $2\log_2 n$). The invariant resources $\dim \mathfrak{g}$ and the second stabilizer Rényi entropy M_2, and the physical expectation $\langle O \rangle$, are identical across encodings; the covariant truncation weight W^* is ordered Bravyi-Kitaev $\leq$ Jordan-Wigner. The invariants are tabulated where the dense computation is feasible ($n = 4$) and hold by Lemma 1 at $n = 8$.

Instance	Encoding	W^*	$\dim \mathfrak{g}$	M_2	$\langle O \rangle$
H_2 / STO-3G ($n = 4$)	Jordan-Wigner	4	16	0.182	+0.675
	Bravyi-Kitaev	4	16	0.182	+0.675
	parity	4	16	0.182	+0.675
H_2 / 6-31G ($n = 8$)	Jordan-Wigner	8	invariant		+0.262
	Bravyi-Kitaev	7	invariant		+0.262
	parity	8	invariant		+0.262

We also tested a larger molecule. Lithium hydride in the STO-3G basis uses $n = 12$ qubits, and we ran it with a UCCSD ansatz. The physical expectation value is again invariant across the encodings to better than 10^{-15}. A sweep over operating points shows that the truncated-propagation verdict is genuinely encoding-dependent: at one operating point the circuit is simulable under Jordan-Wigner and parity yet hard under Bravyi-Kitaev. The clean directional flip of Figure 3(b), in which a tree encoding always lowers the cost, does not appear at this size. The Jordan-Wigner and Bravyi-Kitaev truncation weights track each other closely for this instance, consistent with the directional flip requiring the long-range structure of the synthetic family rather than a small molecule at exactly-simulable sizes.

4.4 A condensed-matter instance

The framework is not specific to chemistry. The one-dimensional Fermi-Hubbard model is a canonical interacting fermionic system, and it foregrounds the invariant axis. Table 3 reports it at half filling for $L = 2$ and $L = 3$ sites, that is $n = 4$ and $n = 6$ qubits. The dynamical-Lie-algebra dimension and the ansatz magic are identical across Jordan-Wigner, Bravyi-Kitaev, and parity, and we cross-validated $\dim \mathfrak{g}$ against an independent sparse Pauli-string closure. The model is local, so the Jordan-Wigner weight stays small and bounded as the system grows. The Bravyi-Kitaev weight is in fact the larger of the two. That makes a useful point on its own: the cost-minimizing encoding is problem-dependent and does not always favour a tree. Hardness here is certified by the invariant dynamical-Lie-algebra dimension, already 343 at $n = 6$, and not by the weight. This complements a molecular Hamiltonian, whose large Jordan-Wigner weight is the covariant resource that a tree encoding can cut.

Table 3. A condensed-matter instance. The one-dimensional Fermi-Hubbard model ($t = 1$, $U = 3$, half filling) under the encoding orbit. The dynamical-Lie-algebra dimension $\dim \mathfrak{g}$ and the second stabilizer Rényi entropy M_2 of the ansatz output are identical across Jordan-Wigner, Bravyi-Kitaev, and parity. The maximum Pauli weight is small and bounded because the model is local, and Bravyi-Kitaev does not reduce it.

L	n	weight (JW / BK / parity)	$\dim \mathfrak{g}$ (invariant)	M_2 (invariant)
2	4	3 / 3 / 4	24	0.947
3	6	3 / 5 / 4	343	1.673

5 The gauge floor: invariants bound the encoding-optimized cost

5.1 Statement and proof

The covariant resources can be reduced by changing the encoding. The invariant resources cannot. This asymmetry has a precise consequence. For an encoding E in the orbit, classical simulation admits at least the Pauli-weight, Lie-algebra, and magic routes of Table 1, so the cost under E along these routes is the minimum of the three (the remaining covariant routes, causal cone and entanglement, only lower it further; Section 7). Write $W_{\mathrm{opt}} = \min_E W_E$ for the encoding-optimized Pauli weight, and recall that the Lie-algebra dimension $\dim \mathfrak{g}$ and the magic t are the same for every E.

Proposition 1 (Gauge floor)**.** *The encoding-optimized simulation cost satisfies*

$$\mathrm{cost}^* \;=\; \min\Big(n^{O(W_{\mathrm{opt}})},\; \mathrm{poly}(\dim \mathfrak{g}),\; 2^{O(t)} \Big),$$

and therefore

(i) $\mathrm{cost}^* \le \min\left(\mathrm{poly}(\dim \mathfrak{g}),\, 2^{O(t)}\right)$*, independently of the Pauli weight;*

(ii) if cost^* *is super-polynomial, then* $\dim \mathfrak{g}$ *is super-polynomial and* t *is super-logarithmic.*

Proof. Under a fixed E the problem is simulable if any single route is efficient, so its cost is the minimum of the three route costs, $\min(n^{O(W_E)}, \mathrm{poly}(\dim \mathfrak{g}), 2^{O(t)})$. The Lie-algebra and magic terms are gauge-invariant by Lemma 1, hence constant over the orbit, while the weight term is minimized at W_{opt}. Minimizing over E gives the stated expression. Part (i) drops the weight term from a minimum. Part (ii) is the contrapositive of (i): if either invariant term were small, cost^* would be small. □

The reading is that the gauge-covariant weight is necessary but never sufficient for hardness. An instance can be hard for truncated Pauli propagation in one encoding, yet simulable, because an encoding-invariant route closes it. Genuine, encoding-robust hardness is certified only by the invariants. Figure 4 renders this as a partition of instances by the route that certifies them. A problem is simulable if it lies in any of the three regions, and the example instances make the asymmetry concrete: free fermions sit in the low-dimension region and a Clifford-angle circuit in the low-magic region whatever their weight, while only the weight boundary moves under re-encoding, so re-encoding can carry a high-weight Jordan-Wigner instance into the Pauli-simulable region. An instance outside all three regions is an advantage candidate.

5.2 Both invariant routes, exhibited

We make Proposition 1 concrete by exhibiting each invariant route as a certificate that the weight misses.

The Lie-algebra route. A free-fermion problem with long-range hopping has Jordan-Wigner Pauli weight equal to n, which a weight-only reading would call hard. Yet its generators close into $\mathfrak{su}(n)$, of dimension $n^2 - 1$, so Lie-algebraic simulation is exact and polynomial. Here the Pauli weight does not distinguish this simulable problem from a genuinely interacting one of the same weight; the Lie-algebra dimension does. Using a sparse Lie-closure over Pauli strings, which avoids exponential-size matrices, we confirmed $\dim \mathfrak{g} = n^2 - 1$ exactly out to $n = 24$, where the dense matrix method is already infeasible, while the Jordan-Wigner weight grows only linearly (Fig. 5).

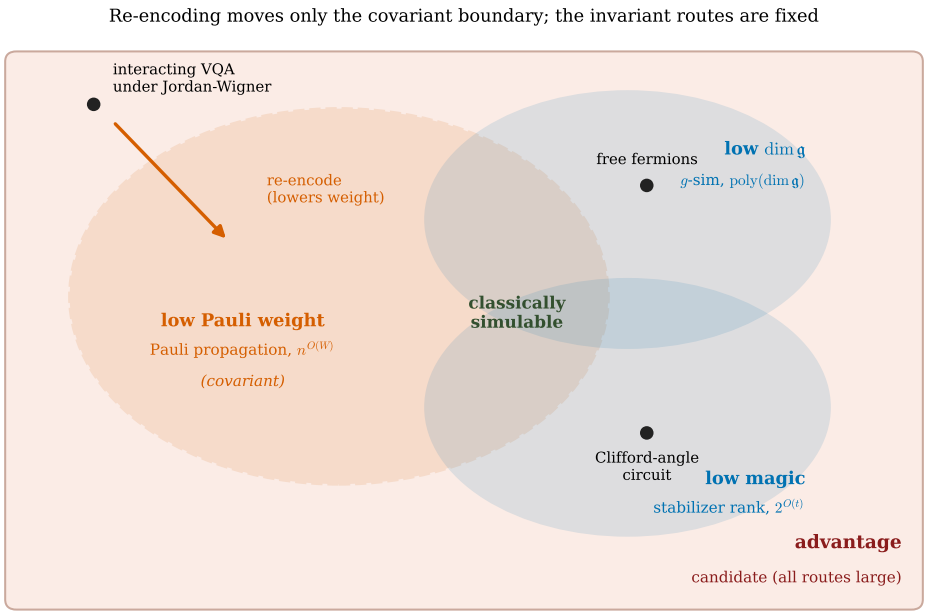


Figure 4. The simulability routes as a partition of instances. A problem is classically simulable if it falls in any region: small Pauli weight (truncated Pauli propagation), small dynamical-Lie-algebra dimension (Lie-algebraic simulation), or small non-stabilizerness (stabilizer-rank simulation). Free fermions lie in the low-dimension region and a Clifford-angle circuit in the low-magic region, irrespective of weight. Only the weight boundary (dashed) moves under re-encoding, so re-encoding can carry a high-weight instance into the Pauli-simulable region; the invariant routes are fixed, and an instance outside all three is an advantage candidate.

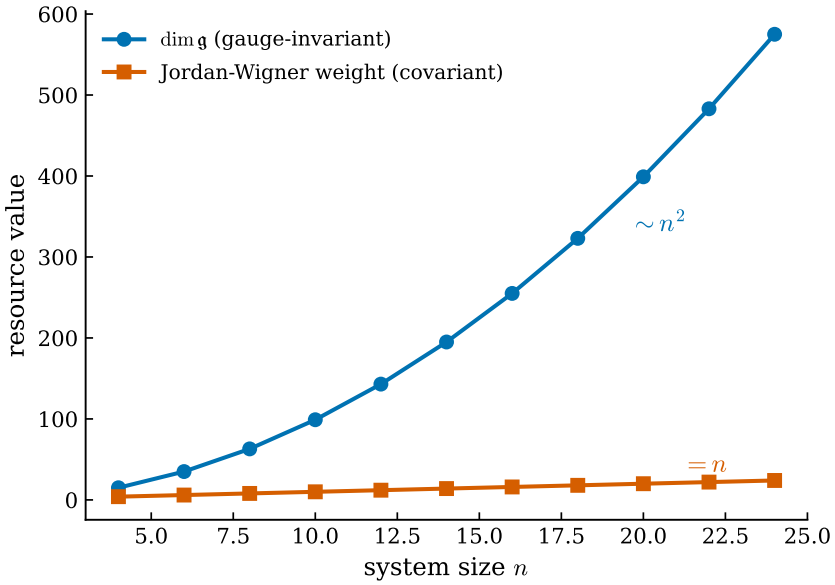


Figure 5. The Lie-algebra route at scale. For long-range free fermions the encoding-invariant dynamical-Lie-algebra dimension is polynomial, matching $n^2 - 1 = \dim \mathfrak{su}(n)$ exactly, while the Jordan-Wigner Pauli weight is linear. The problem is therefore simulable in every encoding despite its weight.

The magic route. Take an interacting ansatz whose Pauli weight and Lie-algebra dimension are both large, and evaluate it at parameter values for which every elementary rotation is a Clifford gate. The output from the stabilizer reference is then a stabilizer state, its stabilizer Rényi entropy is zero, and it is exactly simulable, although neither the weight nor the Lie-algebra route is cheap. The stabilizer Rényi entropy is gauge-invariant, so this certificate holds in every encoding. Away from the Clifford locus the entropy is nonzero and identical across the orbit.

Both demonstrations rely on the gauge-invariance of $\dim \mathfrak{g}$ and the magic, which we validated against independent ground truth: the Lie-algebra computation reproduces $\mathfrak{u}(L)$ for free-fermion chains and $\mathfrak{su}(n)$ for all-to-all hopping, and the stabilizer Rényi entropy reproduces its analytic single-qubit value and vanishes on stabilizer states.

5.3 An explicit family that achieves the floor

The two routes above are exhibited separately. A single parameterized family shows the floor itself being achieved, with the polynomial Lie algebra of the free case and the large invariants of the interacting case as the two ends of one tunable model. On n modes define

$$H(\lambda) = \sum_{i<j} t_{ij}(a_i^\dagger a_j + a_j^\dagger a_i) \;+\; \lambda \sum_{i<j} V_{ij}\, n_i n_j, \tag{8}$$

with all-to-all couplings and the Hamiltonian-variational ansatz built from its Hermitian terms. Because the dynamical Lie algebra is fixed by the generator set and not by the couplings, λ acts as a switch on whether the density-density generators are present.

Proposition 2 (A family that achieves the gauge floor)**.** *For the family* (8)*:*

(i) at $\lambda = 0$ the generators are the all-to-all hopping terms, which close into $\mathfrak{su}(n)$; hence $\dim \mathfrak{g} = n^2 - 1 = O(n^2)$ and Lie-algebraic simulation is exact and efficient in every encoding, even though the Jordan-Wigner weight is n;

(ii) for $\lambda \neq 0$ the density-density generators make the ansatz non-Gaussian, so its dynamical Lie algebra is no longer the polynomial free-fermion algebra; across the sizes we compute, $\dim \mathfrak{g}$ grows far beyond $n^2 - 1$ (Table 4), consistent with the rarity of polynomial-dimensional dynamical Lie algebras established by the classification of Wiersema et al. [29]. At a generic operating point the non-stabilizerness M_2 of the output is bounded away from zero and grows with n (the Clifford locus, where it vanishes, has measure zero).

Consequently, for $\lambda \neq 0$ and at a fixed physical circuit, the covariant weight is n under Jordan-Wigner but $O(\log n)$ under a tree encoding, while both invariants are large: all three certificate routes of Proposition 1 fail simultaneously for the instances we exhibit, and the gauge floor is achieved. The instance is an advantage candidate under the three certificates. It is not a proof of hardness, since methods outside the three routes are not bounded by the floor.

Proof. (i) The Hermitian hopping operators $a_i^\dagger a_j + a_j^\dagger a_i$ are single-particle (Gaussian) generators whose real Lie closure is $\mathfrak{su}(n)$, of dimension $n^2 - 1$; Lie-algebraic simulation then costs $\mathrm{poly}(\dim \mathfrak{g})$ by Table 1, while the long-range hop $a_1^\dagger a_n + \text{h.c.}$ carries a full Jordan-Wigner Z-string and so has weight n. (ii) Adjoining the density-density terms $n_i n_j$ makes the ansatz non-Gaussian, so its dynamical Lie algebra is no longer the polynomial free-fermion algebra; the closure grows far beyond $n^2 - 1$, as Table 4 shows (66 and 918 at $n = 4, 6$ against 15 and 35). The Trotterized circuit at generic angles is non-Clifford, hence $M_2 > 0$, with the Clifford locus a measure-zero set of angles. The floor claim is Proposition 1 applied at fixed circuit, the weight being reducible across the orbit while the two invariants are not. □

We computed both ends of the family directly (Table 4). At $\lambda = 0$ the dynamical-Lie-algebra dimension matches $n^2 - 1$ exactly; at $\lambda \neq 0$ it jumps to 66 at $n = 4$ and 918 at $n = 6$, far above the polynomial trajectory, while the generic-point magic grows from 1.65 to 2.66 and the Jordan-Wigner weight stays at n with the Bravyi-Kitaev weight below it. Every quantity is identical across the encodings except the weight.

Table 4. The family (8) achieves the floor for $\lambda \neq 0$. At $\lambda = 0$ the dynamical Lie algebra is polynomial ($n^2 - 1$) and the Lie route closes the instance; for $\lambda \neq 0$ both invariants are large while the covariant weight is reducible (Jordan-Wigner n vs. Bravyi-Kitaev $< n$). All listed quantities are gauge-invariant except the weight.

n	$\dim \mathfrak{g}$ ($\lambda = 0$)	$\dim \mathfrak{g}$ ($\lambda \neq 0$)	JW weight	BK weight	M_2 (generic)
4	15 ($= n^2 - 1$)	66	4	4	1.65
6	35 ($= n^2 - 1$)	918	6	5	2.66

6 Trainability is gauge-invariant: a dichotomy

The simulation cost is gauge-covariant. Trainability behaves oppositely.

Lemma 2 (Gauge-invariance of the cost landscape)**.** *For a fixed fermionic ansatz, observable, and reference state, the cost* $L_E(\theta) = \langle \mathrm{ref}|_E\, U_E^\dagger(\theta)\, O_E\, U_E(\theta)\, |\mathrm{ref}\rangle_E$ *does not depend on the encoding* E*. Consequently the gradient* ∇L *and the variance* $\mathrm{Var}_\theta[\partial_k L]$ *are gauge-invariant.*

Proof. Each encoding is a $*$-isomorphism of the operator algebra and maps the generators H_k, the observable O, and the reference state consistently. A $*$-isomorphism preserves expectation values, so $L_E(\theta)$ equals the single function $\langle \mathrm{ref}|U^\dagger(\theta) O U(\theta)|\mathrm{ref}\rangle$ for every E. Differentiation and the variance over θ act on this common function. □

Figure 6 expresses Lemma 2 as a commutative diagram: the fermionic data reaches the cost $L(\theta)$ along either encoding, the two qubit representations differ by a Clifford map, and the two paths agree, so the cost is the same in every encoding.

We verified this directly. The full gradient vector of an interacting ansatz is identical across encodings to better than 10^{-12}, with nonzero parameter-dependent entries, so the agreement is not an artifact of vanishing gradients. The barren-plateau diagnostic, the variance of a cost derivative over random parameters, has ratio exactly one between encodings at every system size we tested (Fig. 7). Lemma 2 explains why the agreement is exact rather than approximate. The mechanism is also consistent with the Lie-algebraic theory of barren plateaus, in which $\mathrm{Var}_\theta[\partial_k L] = \Theta(1/\dim \mathfrak{g})$ [32, 33]; combined with Lemma 1(i), that the Lie-algebra dimension is gauge-invariant, this gives a second and independent proof that the barren-plateau scaling cannot be changed by re-encoding. Appendix A derives this scaling exactly for the free-fermion family, where $\dim \mathfrak{g}$ appears as a Weingarten denominator. The scaling matches that prediction for the free-fermion family, where the product of the gradient variance and the Lie-algebra dimension stays of order unity across the sizes we tested, consistent with the $\mathcal{C}/(n^2 - 1)$ law of Appendix A up to its $O(n^{-3})$ corrections and the residual sampling noise.

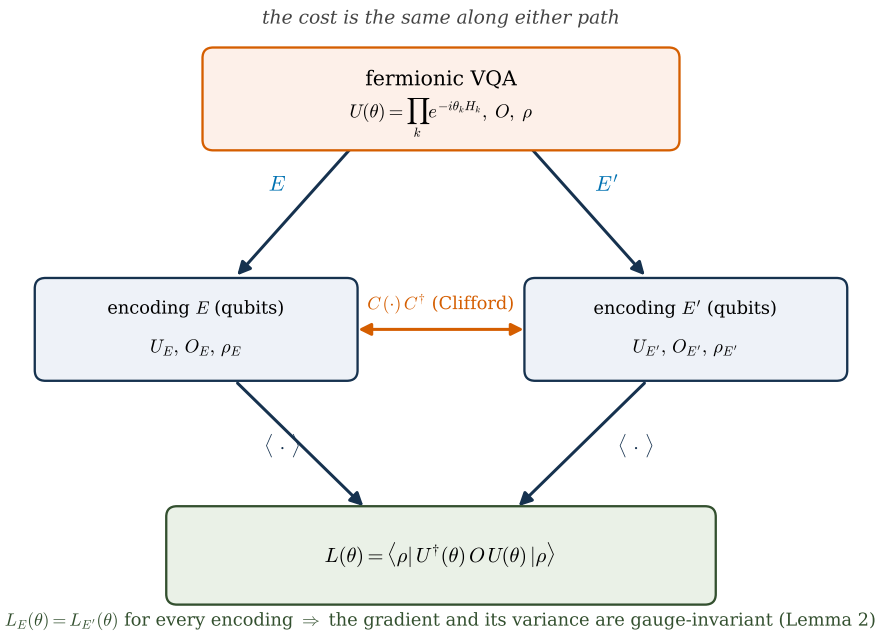


Figure 6. The cost is encoding-invariant. Each encoding E is a $*$-isomorphism that maps the ansatz, observable, and reference consistently, so the two paths from the fermionic data to the cost $L(\theta)$ agree. The cost, and hence the gradient and the barren-plateau variance, are therefore identical for every encoding.

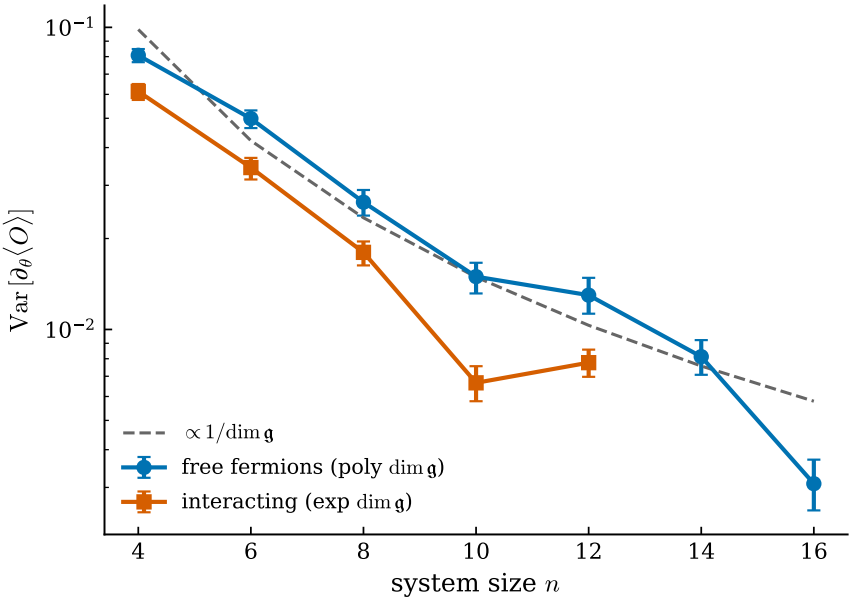


Figure 7. Gauge-invariance of trainability. Variance of a cost derivative over random parameters, as a function of system size. The variance is identical across encodings to the displayed precision, in contrast with the encoding-dependent simulation cost of Figure 3, and it tracks the inverse of the gauge-invariant dynamical-Lie-algebra dimension. Error bars are 95% bootstrap confidence intervals over $S = 400$ random parameter sets; the estimator averages the variance over the most active gradient components.

The two axes therefore transform oppositely under the encoding gauge. One can re-encode a fermionic circuit to lower its simulation cost, as Sections 4 and 5 show. One cannot re-encode it to raise its trainability. A barren plateau is a property of the physical problem, not of its qubit representation.

6.1 Faithfulness is the boundary

The proof of Lemma 2 used only that an encoding is a faithful $*$-isomorphism, not that it is linear or Clifford. This isolates the property responsible for trainability-invariance, namely faithfulness, and with it the precise condition under which any change of representation can move trainability at all. The relevant object is the dynamical Lie algebra *as it acts on the subspace the dynamics explores.* A reference that respects a symmetry of the generators lives in a single sector $\mathcal{H}_S$, which the ansatz never leaves; the barren-plateau variance is then governed by the dimension of the restricted algebra $\rho_S(\mathfrak{g})$ on $\mathcal{H}_S$, not by the dimension of $\mathfrak{g}$ on the full Fock space.

Proposition 3 (Faithfulness and the trainability boundary)**.** *Let the ansatz explore the subspace $\mathcal{H}_S$, on which its dynamical Lie algebra acts as $\rho_S(\mathfrak{g})$; the barren-plateau variance is set by* $\dim \rho_S(\mathfrak{g})$.

(i) Every faithful $*$-representation of the fermionic operator algebra (every encoding in the standard orbit, and more generally any faithful encoding, including nonlinear and ancilla-assisted ones) preserves $\rho_S(\mathfrak{g})$ and the entire cost landscape, hence the gradient and the barren-plateau variance.

(ii) Symmetry tapering onto a sector the reference already occupies is an isometry of $\mathcal{H}_S$: it carries $\rho_S(\mathfrak{g})$ to an isomorphic copy on fewer qubits and reproduces the cost $L(\theta)$ identically, so it leaves the gradient and the variance exactly unchanged. It is a trainability no-op, even though it lowers the dimension of $\mathfrak{g}$ on the full *Fock space.*

(iii) The variance changes only under a move that changes $\rho_S(\mathfrak{g})$ itself: tapering a symmetry the reference does not *respect, so that dynamics exploring several sectors is confined to one, or restricting the ansatz's reachable set (an equivariant gate family). In the first case the variance rises by the ratio of the explored dimensions.*

Proof. (i) A faithful $*$-representation Φ is injective and preserves products, adjoints, and expectation values, so it restricts to a Lie-algebra isomorphism on $\rho_S(\mathfrak{g})$ and reproduces $L(\theta)$ for every θ. This is the argument of Lemma 2, which used faithfulness alone. (ii) Tapering is an isometry $V : \mathcal{H}_S \to (\mathbb{C}^2)^{\otimes(n-k)}$; mapping the ansatz, observable, and reference by V leaves $L(\theta) = \langle \mathrm{ref}|U^\dagger(\theta) O U(\theta)|\mathrm{ref}\rangle$ unchanged term by term, so its gradient and variance are identical. (iii) If instead the reduction is non-faithful on the explored algebra, because the reference had support on sectors that tapering discards, or because the gate set is restricted, then $\rho_S(\mathfrak{g})$ and its dimension change, and with them the variance. □

We verified all three cases with exact, number-conserving dynamics (Table 5). Re-encoding leaves the gradient variance identical across the orbit (ratio 1). For the interacting ansatz the dynamical Lie algebra acts reducibly: its dimension is 918 on the full Fock space at $n = 6$ but only 400 on the half-filling sector (66 versus 36 at $n = 4$), so the sector value, not the full-Fock value, governs a Hartree-Fock-reference VQE. Tapering that respected symmetry is then a no-op: the gradient variance is unchanged to machine precision (ratio 1.0000 at $n = 4, 6$). Only when the reference breaks the symmetry does tapering bite. Confining dynamics that explored all sectors

to the half-filling sector raises the variance by the dimension ratio, measured as 2.29 at $n = 6$, matching $918/400 = 2.30$.

Table 5. Trainability is invariant under both standard representational moves. The interacting ansatz's dynamical Lie algebra acts reducibly: its dimension on the full Fock space, $\dim \mathfrak{g}$, far exceeds its dimension on the half-filling sector, $\dim \rho_S(\mathfrak{g})$, and the latter governs a symmetry-respecting VQE. Re-encoding (faithful) and tapering of the respected symmetry (an isometry of the sector) both leave the gradient variance unchanged; only a symmetry-breaking reference, subsequently tapered, changes it, by the dimension ratio. Exact number-conserving dynamics.

	$n = 4$	$n = 6$
$\dim \mathfrak{g}$ (full Fock)	66	918
$\dim \rho_S(\mathfrak{g})$ (half-filling sector)	36	400
Var ratio, re-encoding (JW/BK/parity)	1	1
Var ratio, tapering, Hartree-Fock reference	1.0000	1.0000
Var ratio, tapering, symmetry-breaking reference	2.88	2.29 ($\approx$918/400)

The operational rule is therefore stronger than a one-sided dichotomy. *Both* free relabelings of the problem are trainability-neutral: re-encoding among Jordan-Wigner, Bravyi-Kitaev, and tree encodings, and qubit tapering [30] of a symmetry the reference respects. A barren plateau survives all of them. Trainability moves only when one changes what the dynamics actually explores. That happens by restricting the ansatz, for instance through an equivariant gate set, or by tapering a symmetry the reference did not respect. This also reconciles the framework with reports that symmetry-aware ansätze can ease training: the gain comes from shrinking the explored algebra, not from relabeling the qubits.

7 Discussion

7.1 A criterion for advantage claims

The results give a simple test. A claimed quantum advantage for a fermionic ansatz that rests on a covariant resource, such as a large Pauli weight or a wide causal cone in a particular encoding, is not encoding-robust: a different representation of the same problem may admit an efficient classical route. An advantage that is to survive must be gauge-invariant, which by Proposition 1 means a large dynamical-Lie-algebra dimension together with large non-stabilizerness. We stress that these two conditions are necessary, not sufficient: Proposition 1 bounds only those three routes (Pauli weight, Lie-algebra dimension, and non-stabilizerness), so even when both invariants are large another classical method, such as tensor-network contraction, may still simulate the instance (Section 7, limitations). The criterion is thus a filter that rules advantage out, not a certificate that establishes it. We recommend that simulability and advantage analyses for fermionic problems report the encoding-optimized cost, or at minimum the invariant resources, rather than quantities tied to a single encoding.

7.2 Scope and limitations

The named simulation routes are sufficient certificates, not a complete characterization, so the encoding-optimized cost is an upper bound on the true classical cost rather than a hardness proof. Two methods deserve explicit mention. Tensor-network contraction uses matrix-product, tree, or

projected-entangled-pair states, and it is frequently the strongest classical competitor for molecular and lattice instances, having recently matched and surpassed flagship hardware demonstrations [26, 27]. Its cost is set by the entanglement of the represented state. Entanglement is itself gauge-covariant: a Clifford re-encoding can create or destroy it, so the qubit entanglement of a fixed fermionic state depends on the encoding. Tensor-network cost therefore sits on the covariant axis alongside the Pauli weight and the causal-cone width (Table 1), not in a blind spot of the framework: like them it shifts under re-encoding and, like them, it cannot certify encoding-robust hardness. Adjoining it as a further route only lowers the encoding-optimized cost, so the floor of Proposition 1 is unaffected. The bound $\text{cost}^* \leq \min(\text{poly}(\dim \mathfrak{g}), 2^{O(t)})$ and its contrapositive continue to hold. The honest residual scope is that the floor is an upper bound assembled from the routes we name; a complete classification of which efficiently computable cost functionals are gauge-invariant and which are covariant, with these resources as instances, we leave open. Separately, the dequantization program for variational models [10, 9] shows that provable trainability often coincides with classical simulability through structural conditions that overlap with our gauge-invariant resources, notably a polynomial dynamical Lie algebra; our floor makes the encoding-independence of that overlap explicit, but it does not subsume the dequantization results, which also exploit data and measurement structure that the floor does not track. The orbit we consider is the standard linear encoding family; nonlinear or problem-tailored encodings could enlarge it and lower the optimized cost further, which only strengthens the direction of our conclusion. Finally, the change of verdict in Section 4 is a statement about the weight-controlled method in the operating regimes we probe; the resource-level separation of Figure 3(a) and the floor of Section 5 are independent of operating regime. A note on scope closes the section. The framework is encoding-theoretic, and the physical instances we report illustrate its reach across chemistry and condensed matter rather than establishing new results in either. The covariant verdict flip is a long-range phenomenon. It is isolated cleanly in the synthetic structured family, whose all-to-all form mirrors the structure of real molecular and long-range fermionic Hamiltonians, and it does not appear in small molecules (Table 2) or in the local Hubbard model (Table 3) at exactly-simulable sizes. The invariant axis behaves differently. The dynamical-Lie-algebra dimension and the magic are identical across encodings in every instance we tested, physical and synthetic alike. Neither table claims a new chemistry or condensed-matter result. Both confirm that the encoding-dependence of cost and the encoding-invariance of the invariants are properties of the representation, not of the model.

7.3 Outlook

Treating the representation as a gauge degree of freedom suggests that other circuit quantities should be sorted by their behavior under it, and that the encoding-optimized cost is the natural invariant to track when assessing advantage. The dichotomy between a covariant simulability axis and an invariant trainability axis also clarifies why trainability filters, useful as they are, cannot by themselves settle the advantage question.

A Gradient variance of the free-fermion family

We derive the inverse-dimension scaling directly for the free-fermion family, a self-contained complement to the cited Lie-algebraic law. Take the all-to-all free-fermion ansatz of Section 5, whose generators close into $\mathfrak{g} = \mathfrak{su}(n)$ with $\dim \mathfrak{g} = n^2 - 1$. A number-conserving Gaussian unitary acts on the n single-particle modes as an element $V \in U(n)$ through $U^\dagger a_i U = \sum_j V_{ji} a_j$. For the Hartree-Fock reference $|\psi\rangle$ with occupied set S, $|S| = m$, and the observable $O = n_p$, the cost reduces

exactly to

$$L = \langle\psi|U^\dagger n_p U|\psi\rangle = \sum_{i\in S} |V_{pi}|^2. \tag{9}$$

The second moments of V are evaluated with the $U(n)$ Weingarten formula at order two,

$$\int_{U(n)} V_{i_1 j_1} V_{i_2 j_2} V^*_{i'_1 j'_1} V^*_{i'_2 j'_2}\, dV = \sum_{\sigma,\tau\in S_2} \mathrm{Wg}(\tau\sigma^{-1}, n)\, \delta_{i_1 i'_{\sigma(1)}} \delta_{i_2 i'_{\sigma(2)}}\, \delta_{j_1 j'_{\tau(1)}} \delta_{j_2 j'_{\tau(2)}}, \tag{10}$$

with the S_2 Weingarten weights $\mathrm{Wg}(e, n) = \frac{1}{n^2-1}$ and $\mathrm{Wg}((12), n) = \frac{-1}{n(n^2-1)}$, which are the denominators that appear below.

Proposition 4 (Free-fermion gradient variance)**.** *Assume the ansatz forms a 2-design on the dynamical group, equivalently that V is Haar-distributed on $U(n)$. Then*

$$\mathrm{Var}_\theta[\partial_\mu L] = \frac{\mathcal{C}}{n^2-1} + O(n^{-3}),$$

where $\mathcal{C} = \Theta(1)$ is set by the filling m/n and the generator, and $n^2 - 1 = \dim \mathfrak{g}$. By Lemma 1(i) this variance is encoding-invariant.

Proof. By (9) the cost is quadratic in the entries of V, and the first Haar moment $\mathbb{E}[V_{pi} V^*_{pj}] = \delta_{ij}/n$ gives $\mathbb{E}[L] = m/n$. A parameter θ_μ generates a single-particle rotation, so $\partial_\mu L$ is again quadratic in V and $V^\dagger$ with $\mathbb{E}[\partial_\mu L] = 0$. Its second moment is evaluated with (10), whose two-fold contractions carry the denominators $n^2 - 1$ and $n(n^2 - 1)$; the leading term is therefore $\mathcal{C}/(n^2 - 1)$, with $\mathcal{C}$ collecting the $\Theta(1)$ overlaps of O and the reference one-body density matrix with $\mathfrak{g}$. Since $n^2 - 1 = \dim \mathfrak{su}(n) = \dim \mathfrak{g}$, the variance scales as $1/\dim \mathfrak{g}$, matching the general estimate [32, 33] and the numerics of Section 6. Encoding-invariance follows because $\dim \mathfrak{g}$ is invariant by Lemma 1(i). □

The dynamical-Lie-algebra dimension thus enters as the Weingarten denominator $n^2 - 1$, which is why the measured product Var × $\dim \mathfrak{g}$ stays of order unity (the constant $\mathcal{C}$), up to the $O(n^{-3})$ corrections.

Data and code availability

The complete source code, the self-validating test suite, and the scripts that reproduce every figure and table in this paper are openly available at https://github.com/tom32bit/Encoding-Gauge-of-Fermionic-Variational-Quantum-Algorithms and archived with a permanent identifier at Zenodo (DOI: 10.5281/zenodo.21059513). The repository includes the figure-generating scripts, the molecular instances of Table 2, and the logged numerical outputs underlying Figures 3–7.

Acknowledgments

The authors thank colleagues at BRAC University for helpful discussions.

Conflict of interest

The authors declare no competing interests.

References

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nat. Commun. **5**, 4213 (2014).

[2] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, "The variational quantum eigensolver: a review of methods and best practices," Phys. Rep. **986**, 1 (2022).

[3] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," Nat. Commun. **9**, 4812 (2018).

[4] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost function dependent barren plateaus in shallow parametrized quantum circuits," Nat. Commun. **12**, 1791 (2021).

[5] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, "Connecting ansatz expressibility to gradient magnitudes and barren plateaus," PRX Quantum **3**, 010313 (2022).

[6] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, "Noise-induced barren plateaus in variational quantum algorithms," Nat. Commun. **12**, 6961 (2021).

[7] E. R. Anschuetz and B. T. Kiani, "Quantum variational algorithms are swamped with traps," Nat. Commun. **13**, 7760 (2022).

[8] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, "Barren plateaus in variational quantum computing," Nat. Rev. Phys. **7**, 174 (2025).

[9] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, "Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing," arXiv:2312.09121 (2023).

[10] E. Gil-Fuster, C. Gyurik, A. Pérez-Salinas, and V. Dunjko, "On the relation between trainability and dequantization of variational quantum learning models," arXiv:2406.07072 (2024).

[11] P. Jordan and E. Wigner, "Über das Paulische Äquivalenzverbot," Z. Phys. **47**, 631 (1928).

[12] S. B. Bravyi and A. Yu. Kitaev, "Fermionic quantum computation," Ann. Phys. **298**, 210 (2002).

[13] J. T. Seeley, M. J. Richard, and P. J. Love, "The Bravyi-Kitaev transformation for quantum computation of electronic structure," J. Chem. Phys. **137**, 224109 (2012).

[14] Z. Jiang, A. Kalev, W. Mruczkiewicz, and H. Neven, "Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning," Quantum **4**, 276 (2020).

[15] M. Steudtner and S. Wehner, "Fermion-to-qubit mappings with varying resource requirements for quantum simulation," New J. Phys. **20**, 063010 (2018).

[16] Y.-A. Chen and Y. Xu, "Equivalence between fermion-to-qubit mappings in two spatial dimensions," PRX Quantum **4**, 010326 (2023), arXiv:2201.05153.

[17] A. Tranter, P. J. Love, F. Mintert, and P. V. Coveney, "A comparison of the Bravyi-Kitaev and Jordan-Wigner transformations for the quantum simulation of quantum chemistry," J. Chem. Theory Comput. **14**, 5617 (2018).

[18] F. Verstraete and J. I. Cirac, "Mapping local Hamiltonians of fermions to local Hamiltonians of spins," J. Stat. Mech. **2005**, P09012 (2005).

[19] C. Derby, J. Klassen, J. Bausch, and T. Cubitt, "Compact fermion to qubit mappings," Phys. Rev. B **104**, 035118 (2021).

[20] A. Angrisani, A. Schmidhuber, M. S. Rudolph, M. Cerezo, Z. Holmes, and H.-Y. Huang, "Classically estimating observables of noiseless quantum circuits," arXiv:2409.01706 (2024).

[21] M. S. Rudolph, E. Fontana, Z. Holmes, and L. Cincio, "Classical surrogate simulation of quantum systems with LOWESA," arXiv:2308.09109 (2023).

[22] M. L. Goh, M. Larocca, L. Cincio, M. Cerezo, and F. Sauvage, "Lie-algebraic classical simulations for quantum computing," arXiv:2308.01432 (2023).

[23] R. Somma, H. Barnum, G. Ortiz, and E. Knill, "Efficient solvability of Hamiltonians and limits on the power of some quantum computational models," Phys. Rev. Lett. **97**, 190501 (2006).

[24] S. Bravyi and D. Gosset, "Improved classical simulation of quantum circuits dominated by Clifford gates," Phys. Rev. Lett. **116**, 250501 (2016).

[25] L. Leone, S. F. E. Oliviero, and A. Hamma, "Stabilizer Rényi entropy," Phys. Rev. Lett. **128**, 050402 (2022).

[26] J. Tindall, M. Fishman, E. M. Stoudenmire, and D. Sels, "Efficient tensor network simulation of IBM's Eagle kicked Ising experiment," PRX Quantum **5**, 010308 (2024).

[27] T. Begušić, J. Gray, and G. K.-L. Chan, "Fast and converged classical simulations of evidence for the utility of quantum computing before fault tolerance," Sci. Adv. **10**, eadk4321 (2024).

[28] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, "A polynomial-time classical algorithm for noisy random circuit sampling," in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing* (2023), p. 945.

[29] R. Wiersema, E. Kökcü, A. F. Kemper, and B. N. Bakalov, "Classification of dynamical Lie algebras of 2-local spin systems on linear, circular and fully connected topologies," npj Quantum Inf. **10**, 110 (2024).

[30] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, "Tapering off qubits to simulate fermionic Hamiltonians," arXiv:1701.08213 (2017).

[31] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, "Simulation of quantum circuits by low-rank stabilizer decompositions," Quantum **3**, 181 (2019).

[32] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. Ortiz Marrero, M. Larocca, and M. Cerezo, "A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits," Nat. Commun. **15**, 7172 (2024).

[33] E. Fontana, D. Herman, S. Chakrabarti, N. Kumar, R. Yalovetzky, J. Heredge, S. H. Sureshbabu, and M. Pistoia, "Characterizing barren plateaus in quantum ansatze with the adjoint representation," Nat. Commun. **15**, 7171 (2024).