Technical supplement

A sealed signed-holonomy benchmark on K_4

1 Generators, basepoints and the invariant

Write a unit quaternion as $(w, \mathbf{v})$ and use Hamilton multiplication

$$(w, \mathbf{v})(z, \mathbf{u}) = (wz - \mathbf{v} \cdot \mathbf{u},\ w\mathbf{u} + z\mathbf{v} + \mathbf{v} \times \mathbf{u}).$$

The standard faithful unitary representation gives

$$(w, x, y, z) \longmapsto \begin{pmatrix} w + ix & y + iz \\ -y + iz & w - ix \end{pmatrix}, \qquad \tfrac{1}{2} \operatorname{tr} U = w.$$

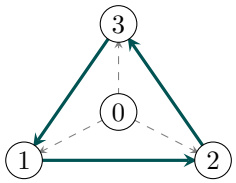
The six stored slots are $01, 02, 03, 12, 13, 23$, oriented from the lower vertex to the higher. At root zero transport the three chords back along the spanning tree. Define

$$A = U_{02}^{-1} U_{12} U_{01}, \quad B = U_{03}^{-1} U_{23} U_{02}, \quad C = U_{03}^{-1} U_{13} U_{01}.$$

Each transforms by the same conjugation at vertex zero. Cancelling adjacent inverse pairs gives

$$C^{-1} B A = U_{01}^{-1} (U_{13}^{-1} U_{23} U_{12}) U_{01}.$$

The scalar parts of both sides coincide with the target h. Each rooted chord here has scalar part zero. Writing their imaginary parts as a, b, c gives $h = c \cdot (b \times a)$. Indeed BA has scalar part $-b \cdot a$ and vector part $b \times a$, while C^{-1} has vector part $-c$. Swapping A and B flips the signed volume. Reversing the loop inverts the complete product, which preserves its scalar part.



The solid arrows give the target face $1 \to 2 \to 3 \to 1$. The dashed edges form a spanning tree rooted at 0. All six physical links enter the fixture.

Unit-quaternion conjugation on the imaginary subspace acts by a proper rotation. A reflection of c across the plane spanned by a, b reverses the oriented volume while preserving the Gram matrix. Taking the three rooted traces discards even more information. These distinctions motivate the feature and dephasing controls on the following pages.

Methods

2 Constructing paired finite and continuous tests

Finite catalog. Fix the root tree to identity and choose one signed imaginary basis unit for each of the three chords, using every unsigned axis exactly once. Root-preserving conjugation by Q_8 flips two of the three axis signs. Every fixed permutation therefore has two conjugacy classes distinguished by $h = \pm 1$. Exhaustive assignment of four vertex frames gives $8^4 = 4{,}096$ assignments per representative. Simultaneous central inversion acts trivially. The checker verifies exactly two assignments per resulting presentation and hence 2,048 distinct presentations per orbit.

Split	Structures	Orbits	Rows	Index of axis permutation
Training	3	6	12,288	0, 1, 4
Validation	1	2	4,096	2
Test	2	4	8,192	3, 5

Indices follow lexicographic order of permutations of (i, j, k). Each split contains equal label counts. The verifier reconstructs each root-conjugacy class and checks disjoint presentations, structures and orbit identifiers. It compares all six edge histograms and all fifteen two-edge histograms across classes. These checks constrain what marginal features can reveal. A histogram of all six edges can retain higher-order correlations, so the verification also measures its empirical within-split lookup ceiling. On the finite test this ceiling is 62.21%, computed using that split's labels solely as a diagnostic.

Continuous panel. Seed 20261127 generates triples of independent random unit directions. Accept a triple with $0.12 < |h| < 0.85$ and with the maximum absolute off-diagonal dot product strictly between 0.15 and 0.95. With $\hat{n}$ parallel to $a \times b$, pair (a, b, c) with $(a, b, c - 2(c \cdot \hat{n})\hat{n})$. This preserves the three dot products while negating the signed volume. After 256 accepted pairs, apply independent Haar vertex frames separately to each of the 512 rows and draw a second complete view. A normalized four-component standard Gaussian supplies each Haar unit quaternion.

Unit of comparison. The reflection partner changes the label while preserving Gram geometry. An alternate frame keeps the label and geometry fixed. The pair rank is the fraction for which the positive member receives the larger score, with half credit on ties. Frame agreement measures equality of thresholded decisions on a geometry's two views. These denominators are 256 pairs and 512 geometries respectively. Repeated frames supply robustness checks, not additional independent geometries.

Models

3 Graph predictors, invariant features and a ququart

Generic predictors. The raw MLP flattens six quaternion links into 24 coordinates. The edge bag shares a four-coordinate edge encoder across six links, averages its outputs and applies a scalar head. The MPNN and TokenGT both tokenize four vertices plus twelve directed edges, with reverse payloads given by quaternion conjugation. Vertex input is a zero scalar. The two-block, width-32 graph models use directed, edge-aware message aggregation or four-head softmax token attention. A mean over all tokens feeds layer normalization and a linear head.

Rooted predictors. The three scalar chord parts and the three distinct pairwise vector dot products occupy six shared feature slots. The implementation rounds them to 10^{-10} to suppress float residue of analytically identical Grams. The seventh slot contains zero for the Gram-only arm, and $c \cdot (b \times a)$ for the oriented arm. Their equal-width MLPs have two hidden layers and GELU activations. The oriented feature already encodes the target sign before fitting.

Cyclic gate. Represent $1, i, j, k$ by the four real amplitudes of one ququart. Left multiplication by a unit quaternion is an orthogonal 4×4 matrix $L(q)$ and hence a unitary over complex amplitudes. With $X|j\rangle = |j + 1 \bmod 4\rangle$ and $Z|j\rangle = i^j|j\rangle$, the simulator expands each gate as

$$L(q) = \sum_{r,s=0}^{3} c_{rs} X^r Z^s, \qquad c_{rs} = \tfrac{1}{4} \sum_{j=0}^{3} L(q)_{j+r,j}\, i^{-sj},$$

with row subscripts modulo four. The word applies signed slots $(4, 6, -5)$ in traversal order to $|0\rangle$. Its real return amplitude equals h. A hardware measurement of population gives h^2; a phase-referenced controlled-unitary interference measurement could recover the signed amplitude. Here all gates and readouts run in a classical simulation. The sixteen coefficients expand one operator, independently of the gate count of a possible physical synthesis.

Fixed basis dephasing. Set $p = e_0$ and update $p \leftarrow |L(q)|^{\circ 2} p$ after each link. The feature is $2p_0 - 1$, and the learned version calibrates it with an affine head, as does the coherent arm. On Q_8, left gates are signed permutations and the completed word is ± 1, hence $p_0 = 1$ for either class. In continuous frames the chosen dephasing basis can affect the answer. A threshold of zero predicts the negative class on an exact tie.

Evaluation

4 Independent scoring and interpretation

	Q_8 test		Continuous SU(2) stress		
Method	Accuracy	Worst orbit	Accuracy	Pair rank	Agreement
Raw MLP	100.00	100.00	65.89	73.70	56.64
Edge bag	49.96	20.41	50.78	50.00	61.33
Directed MPNN	49.93	11.25	48.31	49.35	84.51
TokenGT	50.14	1.19	49.54	49.09	97.59
Rooted Gram MLP	50.00	0.00	50.00	50.00	100.00
Oriented MLP	100.00	100.00	97.98	100.00	100.00
Cyclic ququart + affine	100.00	100.00	100.00	100.00	100.00
Dephased + affine	50.00	0.00	50.00	48.83	100.00
Reference traces	50.00	0.00	50.00	50.00	100.00
Ordered Wilson	100.00	100.00	100.00	100.00	100.00
Dephased return	50.00	0.00	49.41	48.83	97.27

Values are percentages. Learned rows average the three registered seeds. Worst orbit is computed within each seed before averaging. Pair rank awards half credit for tied mirror scores. Agreement compares decisions on two frames of the same geometry. A constant classifier can agree perfectly.

The raw MLP and both sign-informed fitted arms attained 100% on the finite test. Under continuous stress, the raw MLP fell to 65.89% with 56.64% frame agreement. The oriented MLP ranked every reflection pair correctly but classified 97.98% of individual presentations at its fixed threshold. The coherent ququart and the classical Wilson product both recovered every sign. The learned dephased head classified 50.00% on both tests. Across seeds the raw MLP's primary continuous accuracy ranged from 65.625% to 66.211%; the oriented MLP ranged from 97.27% to 98.83%. The control replay's largest absolute score discrepancy across all four roles was 2.23×10^{-16} or less for ordered Wilson and 5.56×10^{-16} or less for dephasing.

Fitting and seal. AdamW uses learning rate 10^{-3}, weight decay 10^{-4}, batch 256 and 20 epochs. Seeds 20261129, 20261130 and 20261201 fix matched epoch shuffles for each arm. Each run takes 960 updates. The earliest minimum validation binary cross-entropy selects the checkpoint. The freezing stage reads only prediction inputs and outputs four score archives before separate verification. The verifier checks checkpoint identities and validation choice, reconstructs controls with complex 2×2 matrices and an independent population transition, and scores the sealed vectors. The learned forwards receive no independent re-execution in that verifier.

Reproduction. From the repository root run the scripts described in the source bundle's `README.md`. Use a fresh run directory, the fixed fixture manifest and the sealed prediction hash. The table generator accepts only the checked receipt. The three seeds quantify optimization variability under a fixed construction. The competing arms have different inductive inputs and parameter budgets. No interval for arbitrary connections, financial markets or quantum hardware follows from these denominators. **Sources.** M. Favoni, A. Ipp, D. I. Müller and D. Schuh, *Lattice gauge equivariant convolutional neural networks*, Phys. Rev. Lett. 128, 032003 (2022). The finite four-vertex witness records the original motivating example and exact checks. The present task concerns a finite graph.