Gauge Structure and Architectural Inductive Bias for Axial Parity Codes: A Theorem and a Negative Result

Brian Guarraci
Galatheus
brian@galatheus.com

May 2026

Abstract

We study the design of learned decoders for sparse-binary voxel compression with an axial parity code. We prove that the codespace of axial-line, plane, and global parity on the 3-torus is exactly the image of the discrete third partial derivative $\partial_x\partial_y\partial_z$ over $\mathbb{F}_2$, with closed-form rank $3n^2 - 3n + 1$, identifying the parity codec as a discrete-$\mathbb{Z}_2$ instance of a scalar-charge fracton gauge theory and placing it in the same algebraic family as surface-code quantum error correction. The proof is elementary once the lattice is read through its tensor-product structure, and the theorem caught a real indexing bug in our previously deployed parity matrix (reported rank 2016 instead of the correct 2977 at $n = 32$).

The natural architectural reading of this theorem — parameterise the network's output as $x = \partial_x\partial_y\partial_z\phi$ and predict the gauge potential ϕ — fails decisively. Across 14 configurations spanning capacity, training time, and several optimisation tricks, this *phi-prior* decoder plateaus at validation IoU ≈ 0.29 versus ≈ 0.48 for an unconstrained baseline at matched configuration; the gap is far larger than seed or checkpoint-selection noise. A soft U(1) latent-gauge variant does better (≈ 0.51), but a controlled ablation shows its gauge machinery is not load-bearing, and a generic multi-scale autoencoder with no gauge structure of any kind wins outright (0.572 ± 0.004 over three seeds). The principal lesson is methodological: a known mathematical structure characterises *what is reachable*, not the right inductive bias for a learned decoder, and enforcing exact algebraic structure architecturally can actively harm optimisation — a concrete coding-theoretic instance of the broader finding that hard equivariance/constraint enforcement often underperforms a soft prior. We report validation metrics on ShapeNet-32^3 and flag a held-out, multi-seed cross-domain study (including surface-code QEC) as future work.

1 Introduction

The parity codec for sparse-binary voxel compression encodes a 32^3 binary occupancy field by transmitting (a) a syndrome consisting of parity sums along axis-aligned lines, planes, and the volume total, and (b) optional spectral side-information (low-frequency DCT coefficients of the voxel field). At the receiver, a learned network reconstructs the voxel field given the syndrome and side-information. This is a clean instance of *syndrome-conditioned reconstruction on a regular lattice* and shares its abstract structure with surface-code quantum error correction, where stabiliser measurements play the role of parity checks.

Within the companion paper family, this note is the algebraic and negative-control member. The compression paper asks which checks help a learned proposal; the surface-code and X-cube papers test geometry-aware neural decoders on quantum codes. Here we prove the algebraic structure behind the axial-parity codec and then test the tempting architectural conclusion one

might draw from that proof. The main lesson is that a correct mathematical description of the constraint set is not, by itself, a recipe for a good learned decoder.

We pursue three connected goals:

1. **Characterise the codespace algebraically.** What is the structure of the set of voxel fields satisfying every parity check? We give a theorem (§3) stating that the codespace is exactly the image of the discrete third partial derivative over $\mathbb{F}_2$. This identifies the parity codec as a discrete-$\mathbb{Z}_2$ instance of Pretko's scalar-charge fracton gauge theory [1] and provides a constructive parameterisation: every codeword is the discrete-XOR of eight corner values of an underlying scalar potential ϕ.
2. **Test whether the theorem prescribes a useful architecture.** The most direct interpretation is to predict ϕ with a neural network and emit $x = \partial_x\partial_y\partial_z\phi$ as a fixed output operator. We call this the *phi-prior* decoder. We find it fails decisively (§5) — across capacity, training time, and several optimisation tricks, it plateaus at ≈ 0.29 IoU versus an unconstrained baseline's ≈ 0.48 at matched compute.
3. **Find what does work.** A multi-scale latent autoencoder (§6) — encoder $\to 8^3$ latent $\to$ SharpNet upsample, with no gauge structure of any kind — beats both the phi-prior and the unconstrained baseline. A controlled ablation isolates the cause: the gauge machinery (both the hard $\mathbb{F}_2$ operator and the soft U(1) regularisers) is not the active ingredient; multi-scale processing is.

The intended contribution is therefore a theorem, a negative result, and a controlled ablation, organised around the question of when a known mathematical structure should and should not be enforced architecturally in a learned decoder. A natural follow-up — whether the same architectural lesson transfers across the algebraic family the theorem unifies (in particular to surface-code QEC) — is discussed as a deferred, multi-seed cross-domain study in §7.

2 Background

2.1 Axial parity codec

Let $\Lambda = (\mathbb{Z}/n)^3$ be a periodic 3-cubic lattice with voxel field $x : \Lambda \to \mathbb{F}_2$. The codec defines three families of parity checks on x:

- **Line.** $L^a_{\alpha\beta}(x) = \sum_t x_{t\hat{e}_a+\alpha\hat{e}_b+\beta\hat{e}_c} \pmod 2$ for each axis $a \in \{x, y, z\}$ and perpendicular coordinates (α, β).
- **Plane.** $P^a_\alpha(x)$ summing over the plane perpendicular to a at coord α.
- **Global.** $G(x) = \sum_v x_v$.

These checks are stacked into a parity matrix $H : \mathbb{F}_2^{n^3} \to \mathbb{F}_2^M$. For $n = 32$ the bug-fixed implementation has rank 2977. The codespace is $\mathcal{C}(n) = \ker H$.

The encoder transmits the syndrome Hx together with a low-frequency DCT field of x (variable bit budget; we use levels $k \in \{4, 5, 6, 7\}$ giving $\{2744, 5832, 10648, 17576\}$ bits respectively). The decoder constructs a canonical base solution x_{base} from the syndrome and a learned residual $r = x \oplus x_{\text{base}}$; the residual lies in the codespace by construction since x and x_{base} share the same syndrome. The learning task is to predict r from x_{base}, the syndrome, and (optionally) the DCT side-info.

2.2 Surface-code QEC

The surface code is a $\mathbb{Z}_2$ stabiliser code on a 2D toric lattice [5]; its decoder must infer the logical-error class given the syndrome of stabiliser measurements (a 3D detector grid when stacked across rounds). It is a $\mathbb{Z}_2$ 1-form gauge theory in the Wegner–Kitaev sense [4, 5], one rung below the fracton gauge theory the parity codec instantiates. Our deployed surface-code decoder is a flat-resolution 3D ResNet [7].

2.3 Notation

$(\partial_a \phi)_v := \phi_v + \phi_{v+\hat{e}_a} \pmod 2$ is the discrete forward difference over $\mathbb{F}_2$ in direction a.

3 Theorem: codespace structure of the axial parity code

Theorem 1. *For all $n \geq 1$ on the 3-torus $\Lambda = (\mathbb{Z}/n)^3$,*

$$\mathcal{C}(n) = \mathrm{image}\big(\partial_x \partial_y \partial_z\big) \subseteq \mathbb{F}_2^{n^3}, \qquad \mathrm{rank}_{\mathbb{F}_2}(H) = 3n^2 - 3n + 1.$$

Proof. Write $V = \mathbb{F}_2^n$ and identify $\mathbb{F}_2^{n^3} = V \otimes V \otimes V$, the three factors indexed by the x, y, z coordinates. The one-direction forward difference acts on a single tensor factor: $\partial_x = \partial \otimes I \otimes I$, $\partial_y = I \otimes \partial \otimes I$, $\partial_z = I \otimes I \otimes \partial$, where $\partial = I + S$ and S is the cyclic shift on V. Hence $T := \partial_x \partial_y \partial_z = \partial \otimes \partial \otimes \partial$.

Step 1 (codespace as an orthogonal complement). Each line, plane, and global check equals the $\mathbb{F}_2$ inner product of x with the indicator vector of an axis-parallel line, an axis-normal plane, or the whole lattice, respectively. Thus $\mathcal{C}(n) = \ker H = R^{\perp}$, where $R \subseteq \mathbb{F}_2^{n^3}$ is the span of all check indicators. A plane is the $\mathbb{F}_2$-sum of the parallel lines it contains, and the global vector is the sum of parallel planes, so R is already spanned by the axis-parallel line indicators alone.

Step 2 (the span of lines). The indicator of the x-line at transverse coordinates (β, γ) is $\mathbf{1} \otimes \delta_\beta \otimes \delta_\gamma$, where $\mathbf{1} \in V$ is the all-ones vector and δ_j are standard basis vectors. As (β, γ) range over $[n]^2$ the vectors $\delta_\beta \otimes \delta_\gamma$ span $V \otimes V$, so the x-lines span $\langle \mathbf{1} \rangle \otimes V \otimes V$. This is exactly $K_x := \{f : \partial_x f = 0\}$, the fields constant along x. Likewise the y- and z-lines span K_y and K_z, so $R = K_x + K_y + K_z$.

Step 3 (dimension). Choose any basis B of V with $\mathbf{1} \in B$. Then K_x, K_y, K_z are each spanned by a subset of the common tensor basis $B \otimes B \otimes B$, hence are coordinate subspaces, so inclusion–exclusion for dimensions is exact. With $\dim K_a = n^2$, $\dim(K_a \cap K_b) = n$ (fields constant in two axes), and $\dim(K_x \cap K_y \cap K_z) = 1$ (global constants),

$$\dim R = 3n^2 - 3n + 1.$$

Therefore $\mathrm{rank}_{\mathbb{F}_2}(H) = \dim R = 3n^2 - 3n + 1$ and $\dim \mathcal{C}(n) = n^3 - (3n^2 - 3n + 1)$.

Step 4 (kernel of T). For a tensor product of operators acting on separate factors, $\ker(\partial \otimes \partial \otimes \partial) = \ker \partial \otimes V \otimes V + V \otimes \ker \partial \otimes V + V \otimes V \otimes \ker \partial$. On the n-cycle, $\ker \partial = \langle \mathbf{1} \rangle$ (a vector killed by $I + S$ is shift-invariant, hence constant). Thus $\ker T = K_x + K_y + K_z = R$, and $\dim \mathrm{image}(T) = n^3 - \dim \ker T = n^3 - (3n^2 - 3n + 1) = \dim \mathcal{C}(n)$.

Step 5 (equality). For any ϕ, every axis-aligned line/plane/global sum of $T\phi = \partial_x \partial_y \partial_z \phi$ telescopes to 0 (mod 2) under the periodic boundary, since each ϕ_v enters such a sum an even number of times; hence $\mathrm{image}(T) \subseteq \mathcal{C}(n)$. Two subspaces, one contained in the other and of equal finite dimension, coincide. Therefore $\mathcal{C}(n) = \mathrm{image}(T)$. □

Empirical verification. We verified the theorem by brute-force $\mathbb{F}_2$ Gaussian elimination on the full check matrix at $n \in \{4, 6, 8\}$ (`verify_fracton_theorem.py`), obtaining $\mathrm{rank}(H) \in \{37, 91, 169\}$ in exact agreement with $3n^2 - 3n + 1$, $\mathrm{rank}(T) \in \{27, 125, 343\}$ in agreement with $n^3 - (3n^2 - 3n + 1)$, and $\dim \mathcal{C}(n) = \mathrm{rank}(T)$ with random ϕ producing codewords that satisfy every check (50/50 per n), confirming $\mathcal{C}(n) = \mathrm{image}(T)$ as a subspace. The $n = 32$ arithmetic gives $3 \cdot 1024 - 96 + 1 = 2977$.

The bug the theorem caught. The previously deployed parity matrix reported rank 2016 at $n = 32$ rather than 2977. Running its construction through the same verifier yields rank $\in \{28, 66, 120\}$ at $n \in \{4, 6, 8\}$ — strictly smaller than $\{37, 91, 169\}$, hence a *different, lower-rank* code (a duplicate-axis indexing error, not a re-presentation of the same checks). The closed-form rank is thus also a practical correctness check for the codec's bit-budget accounting.

Interpretation. The codespace is the discrete-$\mathbb{F}_2$ analogue of Pretko's continuum scalar-charge fracton gauge theory [1]: the codeword x is the third-derivative "field strength" of an underlying potential ϕ. The conserved charges of the fracton matter sector — total charge (G), planar charges (P^a_α), and line charges ($L^a_{\alpha\beta}$) — are exactly the parity checks of the codec. Surface-code QEC corresponds to one rung up the fracton hierarchy: a $\mathbb{Z}_2$ 1-form gauge theory whose Gauss law is the standard four-weight stabiliser. The fracton↔code dictionary is itself established [1, 2, 3]; our contribution is the specific identification of the axial-parity codec as the $\partial_x\partial_y\partial_z$ field-strength theory and the closed-form rank that follows.

4 Experimental setup

All experiments use ShapeNet-32^3 [6] binary voxels with the bug-fixed parity check matrix giving syndrome of dimension $M = 2977$. Training target is the residual $r = x \oplus x_{\text{base}}$, which lies in the codespace. Loss is binary cross-entropy on r, optionally with positive-class weighting. The reported metric is intersection-over-union (IoU) on the reconstructed voxel field $\hat{x} = x_{\text{base}} \oplus \hat{r}$, evaluated on a held-back validation set (not a disjoint test split; see §9).

Unless stated otherwise: train subset 2,048 samples, validation 512 samples, batch size 8, AdamW with lr 5×10^{-4}, cosine LR schedule. We report best validation IoU achieved during training, and multi-seed mean ± stdev where indicated.

Architectures compared:

- **x-prior** (baseline): ResNet3D at varying width/depth, predicts r directly via flat-resolution 3D convolution at constant spatial resolution, syndrome conditioning via MLP.
- **phi-prior:** same backbone, output reinterpreted as ϕ; final operator is the bilinear soft-XOR of 8 corner ϕ's with output bias init at -5 to break XOR symmetry.
- **gauge transporter:** U(1) latent-lattice gauge with parallel-transport reconstruction from a seed voxel, plus auxiliary curvature/transport regularisers.
- **latent autoencoder:** encoder → 8^3 latent → SharpNet decoder, no gauge structure.

5 Hard architectural enforcement plateaus far below baseline

Across 14 experiments varying capacity, training time, and optimisation tricks (output bias init, positive-class weighting, residual escape hatch, straight-through estimator, multi-scale capacity), the phi-prior decoder plateaus at val_iou ≈ 0.29 vs the unconstrained x-prior at 0.48 at matched parameter count and training budget.

The plateau is robust to capacity (5M = 12M), training duration (20 = 30 epochs), positive weighting (sweet spot at pos_weight= 19), and data scale (2K = 8K). The phi-prior occasionally exhibits a phase-transition jump from ~ 0.13 to ~ 0.21 around epoch 12 under the right hyperparameters but always saturates at ~ 0.29 soon after. **Architectural enforcement of the codespace is therefore not a viable inductive bias on this task.**

The reason is mechanical: the bilinear soft-XOR over 8 corners has gradient bounded by 1 but multiplicatively attenuated through the chain, and the network must simultaneously modulate 8 voxel-position ϕ values to flip a single residual bit — a non-local credit-assignment problem that the BCE loss cannot drive efficiently when the target distribution is sparse (most residuals are 0).

6 What works instead: a generic multi-scale prior

If hard gauge enforcement is the wrong inductive bias, what is the right one? A generic multi-scale latent autoencoder — encoder downsampling $32^3 \to 8^3$ with C_{latent} channels, then a SharpNet decoder upsampling $8^3 \to 32^3$ with residual blocks at each scale, and *no gauge structure*

variant	params	epochs	pos_weight	val_iou
x-prior	5.0M	20	19	**0.481**
x-prior	12.0M	20	19	0.489
x-prior	17.9M	20	19	0.474
phi-prior, tanh-product XOR	1.7M	3	1	0.002
phi-prior, bilinear + bias init -5	1.7M	5	1	0.021
phi-prior, +pos_weight= 19	5.0M	20	19	0.283
phi-prior, 12M	12.0M	20	19	0.276
phi-prior, 30 epochs	5.0M	30	19	**0.291**
phi-prior, 8K train data	5.0M	25	19	0.292
phi-prior, residual escape hatch	5.0M	5	1	0.001
phi-prior, STE hard-XOR	1.7M	10	1	0.001

Table 1: Phi-prior decoder vs x-prior baseline (best validation IoU, single-seed per cell). The phi-prior plateau at ≈ 0.29 IoU is robust across capacity, training duration, and optimisation; the unconstrained x-prior at matched compute reaches ≈ 0.48. The robustness of the gap across 14 cells, not any single comparison, is the evidence.

of any kind — beats both the phi-prior and the unconstrained x-prior at matched compute (Table 2).

architecture	gauge structure	params	val_iou
phi-prior (hard $\mathbb{F}_2$, every voxel)	hard	5.0M	0.283
x-prior (unconstrained ResNet3D)	none	5.0M	0.481
soft U(1) transporter	soft, coarse latent	4.4M	0.509
— auxiliary gauge losses removed	soft, coarse latent	4.4M	0.518
multi-scale autoencoder	**none**	16M	0.572 ± 0.004

Table 2: Architecture comparison (best validation IoU on 512 held-back ShapeNet samples). The autoencoder row is the mean $\pm$ stdev over seeds $\{42, 43, 44\}$; the other rows are single-seed (the phi-prior plateau is corroborated by the 14-run sweep in Table 1). Hard gauge enforcement is worst; soft gauge structure helps modestly but its machinery is not load-bearing; the gauge-free multi-scale autoencoder wins.

Two controlled observations isolate the cause. First, the soft U(1) transporter with all auxiliary curvature/transport/latent-reconstruction losses *removed* (val_iou 0.518) slightly outperforms the full version (0.509): the gauge regularisers are not load-bearing. Second, dropping the parallel-transport operator entirely — leaving a plain multi-scale autoencoder — improves things further (4.1M: 0.554; 16M: 0.572). The win is multi-scale processing, not gauge structure.

A single-seed bottleneck-size sweep indicates an 8^3 latent is near-optimal: a tighter 4^3 bottleneck loses information (0.562 at 32M) and a looser 16^3 underconstrains (0.542 at 3M). This is consistent with the interpretation that the bottleneck implements an implicit low-frequency/smoothness prior, matched to the sparse, surface-concentrated structure of the occupancy field — precisely the property the gauge constraint is orthogonal to. We treat the bottleneck-sweep numbers as single-seed exploratory evidence, not multi-seed claims.

7 A deferred cross-domain question

The gauge theorem places axial-parity compression, surface-code QEC, and X-cube fracton QEC in a common constraint-system family: the surface code is a $\mathbb{Z}_2$ 1-form gauge theory, while the parity codec is a discrete-$\mathbb{Z}_2$ scalar-charge fracton gauge theory. It is natural to ask whether the multi-scale prior that wins here also helps for QEC syndrome decoding, where the syndrome is

a dense 3D detector volume rather than a sparse occupancy field.

Preliminary single-seed experiments suggest the opposite: aggressive spatial bottlenecks appear to destroy the per-detector-cell information that dense surface-code syndromes carry, where a deployed flat-resolution convolution preserves it. We deliberately do *not* report these as a result. A defensible cross-domain claim requires a disjoint test split, multiple seeds, and matched baselines decoded on the same shots — the standard we apply to the compression numbers above is not yet met on the QEC side. We state only the hypothesis the preliminary runs suggest: that the multi-scale prior's advantage is specific to sparse signals with a smoothness structure and does not extend to dense per-cell syndromes. Testing it rigorously is future work.

8 Discussion

A theorem characterises feasibility, not inductive bias. The gauge theorem tells us *what is reachable* (the codespace), not *how to reach it* (the architecture). Architectural enforcement of the theorem's constructive parameterisation hurt; the right inductive bias for this task is unrelated to the gauge structure and is well-served by a generic multi-scale architecture. The compression codec's existing DCT side-information implements a related smoothness prior explicitly, and the multi-scale autoencoder appears to learn one implicitly via its 8^3 bottleneck.

Relation to soft-constraint priors. The phi-prior failure is a concrete coding-theoretic instance of a more general phenomenon: hard architectural enforcement of an exact algebraic or equivariance structure can underperform a soft prior or an unconstrained pathway. Finzi et al.'s residual pathway priors [8] make this point for equivariance constraints and propose converting hard constraints into soft penalties; our soft U(1) transporter is the analogous move in this setting, and the ablation shows even its soft gauge losses are not what carries the win. The lesson is not that the gauge structure is wrong — it is exactly correct as a characterisation of the codespace — but that exactness of the constraint and usefulness as an inductive bias are different properties.

Open questions. (i) Can the gauge structure be exploited as a *post-hoc* projection or a *soft penalty* (rather than as architectural enforcement) without the optimisation cost we observed? (ii) The theorem generalises to higher dimensions (the codespace of axial parity on the d-torus is the image of $\partial_1 \cdots \partial_d$, with rank obtainable by the same tensor argument); does the architectural lesson generalise too? (iii) Does the sparse-with-smoothness vs dense-per-cell distinction sketched in §7 hold up under a rigorous multi-seed cross-domain test, and where do quantum LDPC codes with sparser stabilisers fall?

9 Limitations

We are explicit about scope. (1) All reported IoU is *validation* IoU on a held-back ShapeNet-32^3 split, not a disjoint held-out test split; the multi-seed autoencoder result uses seeds $\{42, 43, 44\}$, while the x-prior, phi-prior, and transporter rows and the bottleneck-size sweep are single-seed exploratory measurements. (2) The matched-compute comparison in Table 2 is matched in training budget but not exactly in parameter count across all rows; the headline conclusion (hard enforcement worst, gauge machinery not load-bearing, multi-scale wins) rests on gaps far larger than seed noise and on the directionality of the two ablations rather than on a single point estimate. (3) Results are at 32^3 resolution on a single object dataset. (4) The cross-domain QEC discussion (§7) is a hypothesis, not a measured result.

10 Conclusion

We proved a clean algebraic characterisation of the parity codec's codespace — it is exactly the image of the discrete third derivative $\partial_x\partial_y\partial_z$ over $\mathbb{F}_2$, with closed-form rank $3n^2 - 3n + 1$ — connecting it to fracton gauge theory and to surface-code QEC, and we used it to catch a real indexing bug in a deployed parity matrix. We then tested the natural architectural translation of the theorem (predicting the gauge potential and applying the third derivative as a fixed output operator) and found it catastrophically worse than an unconstrained baseline. The actual win on the compression task comes from a generic multi-scale autoencoder with no gauge structure, and a controlled ablation shows the gauge machinery is not the active ingredient.

The principal lesson is methodological: a known mathematical structure characterises feasibility but does not prescribe the right inductive bias; architectural choice should be governed by the optimisation landscape and signal structure of the specific task. We provide a verified theorem, a clean negative example of constraint enforcement, and the ablation that localises why it fails.

Reproducibility. All experiments are reproducible from the released code. In the family repository, the theorem verifier is grouped with the X-cube fracton code, and the gauge negative-result experiments are preserved as their own reproducible submodule. Run-by-run configuration and results, including discarded approaches, are catalogued in `experiments.tsv`. The theorem verifier (`verify_fracton_theorem.py`) reproduces the rank claims at $n \in \{4, 6, 8\}$ in ~ 20 s on CPU. See the companion `REPRODUCIBILITY.md` for exact commands.

References

[1] M. Pretko. Subdimensional particle structure of higher rank U(1) spin liquids. *Phys. Rev. B*, 95:115139, 2017.

[2] S. Vijay, J. Haah, and L. Fu. Fracton topological order, generalized lattice gauge theory, and duality. *Phys. Rev. B*, 94:235157, 2016.

[3] J. Haah. Local stabilizer codes in three dimensions without string logical operators. *Phys. Rev. A*, 83:042330, 2011.

[4] F. Wegner. Duality in generalized Ising models and phase transitions without local order parameters. *J. Math. Phys.*, 12:2259, 1971.

[5] A. Y. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003. arXiv:quant-ph/9707021 (1997).

[6] A. X. Chang et al. ShapeNet: An Information-Rich 3D Model Repository. arXiv:1512.03012, 2015.

[7] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016.

[8] M. Finzi, G. Benton, and A. G. Wilson. Residual pathway priors for soft equivariance constraints. In *NeurIPS*, 2021. arXiv:2112.01388.