

The lattice’s teeth: primitive reciprocal axes govern which Fourier components neural cellular automata can grow

Daniel Olliver

Independent researcher, Auckland, New Zealand
daniel@danielolliver.com

July 2026

Abstract

Neural cellular automata (NCA) are trained on square grids by default. I treat the lattice as an experimental variable and report two results. First, a methodological one. In matched cross-lattice comparisons, snapping periodic wave targets independently per wavevector onto each lattice’s reciprocal box, then min-max normalizing, sets a three-wave target’s variance to one of two narrow classes (near 1.5/36 or 1.5/20.25) depending on whether rounding preserves the wave-set’s closure. In a 310-training study this produced a 44% “advantage” of hexagonal over square substrates that reproduced in 12/12 restarts and survived an operator-conditioning control, while, outside twelve trainings in a single panel that had genuinely learned (the seed of section 4), every wave training on every lattice had in fact converged to the predict-the-mean solution or stopped short of it: test error 0.97 to 1.62 times its own target variance, never explaining more than 3% of it. The advantage was the ratio of the two variances. I document the failure, the diagnostic (per-condition loss/variance), and the repair. Second, with closure enforced and a per-component readout, a lattice-level constraint on learnability: a trained NCA grows those Fourier components of its target whose wavevectors lie on a primitive reciprocal-lattice axis of its substrate, with amplitude recovery 0.55–0.88 in restart means (per-restart minimum 0.50), and fails to grow all others, recovery at most 0.11 in restart means (per-restart maximum 0.21), including components exactly representable along non-primitive (diagonal) directions. Cell scores (explained variance) follow the variance shares of the aligned components (on the two Bravais substrates of the calibration study). In the tested family, a substrate therefore hosts a full pattern class only if its point-group symmetry admits the class. A pre-registered out-of-sample battery — 96 trainings across honeycomb, grid-size and wavelength, learned-kernel, misalignment, and data-budget panels, predictions frozen before the first training — then tested the constraint quantitatively and partially refuted its frozen form: the pre-registered decider, predicted-versus-measured R^2 at least 0.8, failed ($R^2 = -0.39$), because the tooth has an angular half-width near 6 degrees rather than the calibrated 0.5, its snap-strain clause is inert, aligned recovery is graded by data budget and by the target’s mode order rather than fixed, and the honeycomb gates on its embedding box’s reciprocal axes rather than its six-fold stencil symmetry. What survives out-of-sample is the gate itself: growth stays direction-selective, no geometry-forbidden component grows, and far-off-axis components remain dead (recovery 0.00); the exactly-representable-diagonal cell is a calibration result the battery did not re-probe. The constraint is real and soft, not sharp.

1 Introduction

The Growing-NCA line (Mordvintsev et al., 2020) trains local rules to grow and repair target patterns. The work runs on square grids inherited from image tensors. Where the grid is addressed at all, the aim

is to remove its influence: isotropic and steerable rules, mesh and graph domains. Whether the lattice itself — the substrate the automaton runs on — limits what a trained rule can grow does not appear to have been asked.

The classical ancestor of the question is Othmer and Scriven (1971): on a discrete lattice, reaction-diffusion instabilities select their modes from the lattice spectrum. That result concerns mode selection under fixed dynamics. It says nothing about the learnability of modes by a trained rule. On the learning side, spectral bias (Rahaman et al., 2019) orders learnability by frequency magnitude and is isotropic in k ; it cannot distinguish directions at fixed $|k|$.

That a substrate’s point group restricts which pattern classes a fixed local rule can host is classical: lattice spectra select the unstable modes of reaction-diffusion on discrete supports (Othmer and Scriven, 1971), square lattice gases cannot produce the isotropic hydrodynamics hexagonal ones can (Frisch et al., 1986), and an isotropic automaton yields hexagonal spot arrays that a classical square lattice cannot (Schepers and Markus, 1992). Direction-dependent learnability on a fixed pixel grid is documented too (Ortiz-Jiménez et al., 2020). What is reported here is the component-level learnability counterpart: an optimizer minimizing the error of a single Fourier component that its substrate represents exactly, at zero snap strain, still cannot grow it.

Section 2 documents a normalization artifact specific to cross-lattice comparisons of periodic targets. A preprocessing step manufactured a large, restart-stable, control-resistant learning effect, and the diagnostic that catches it is one number per condition. Sections 3–5 establish and then stress-test the positive result: a component-level primitive-axis constraint on what NCA can learn to grow.

2 A normalization artifact in cross-lattice comparisons

The original experiment compared matched NCA on square and hexagonal lattices growing periodic targets: a hexatic set (three unit cosines at 120 degrees) and a tetratic set (two at 90 degrees). Arbitrary wavevectors do not fit a finite torus, so each wave was snapped to the nearest point of the lattice’s reciprocal box, and the summed field was min-max normalized to $[0, 1]$. Perception operators, parameter count (3,944), seeds, and budgets were matched.

The hexagonal lattice appeared to grow the hexatic target with 40.6%, 44.2%, and 43.6% lower test error at three data budgets, in twelve of twelve restarts. A conditioning control (a square-lattice stencil tuned to lower plane-wave anisotropy than the hexagonal stencil) recovered none of the gap. A five-panel battery of shear ladders and learned-kernel controls appeared to confirm a symmetry-matching law.

A grid-size and wavelength robustness panel then reversed the direction of the effect at two of four cells, and a pre-registered seed sweep found every cell deterministic: 8/8 or 0/8, with per-seed test losses agreeing to within 0.0006. An audit followed.

Two facts account for everything but one battery panel. First, in 250 of the family’s 262 wave-target trainings (the other 48 of the 310 were non-periodic controls), test error tracked the model’s own target variance: loss/variance 0.97–1.08 at standard and abundant budgets, within 0.6% in the headline cells, and 1.03–1.62 at scarce budgets, where models were simply undertrained. These models had converged to, or stopped short of, the predict-the-mean solution; none had learned any wave target. (Non-periodic control targets trained normally in the same harness, with loss/variance ratios down to 0.009, so the training machinery was not at fault.) The twelve exceptions all sat in the battery’s orientation panel, the only genuine wave learning in the family, and they are the seed of sections 3–4. At the 30-degree hexatic target, whose three wavevectors all fall on the hexagonal lattice’s reciprocal axes and whose rounded triads close on both lattices, the hexagonal models reached explained variance 0.64–0.71 with

genuine restart scatter; the square models reached 0.26–0.28 there and 0.25–0.28 at the 0-degree hexatic, consistent with growing a single axis-aligned component (section 4). That signal survived the audit and set the design of the repaired instrument. Second, the compared lattices were not scored against equal targets. Rounding the three hexatic wavevectors independently sometimes preserves their closure (sum zero) and sometimes breaks it. A closed triad’s field has range $[-1.5, 3]$; a broken one, $[-3, 3]$; after min-max normalization the variances are $1.5/20.25 = 0.0741$ and $1.5/36 = 0.0417$. Which value each lattice draws at a given grid size, wavelength, and orientation is arithmetic. The headline effect is the identity $1 - 0.04182/0.07420 = 43.6\%$, computed from the two measured test errors, each equal to its arm’s realized target variance (0.04179 and 0.07407, within 0.3% of the class values above). The tetratic null that had seemed meaningful was forced: a two-wave set has no closure to break, both lattices draw variance 0.0625 exactly, and the comparison ties. The conditioning control was uninformative for the same reason: a stencil cannot change the variance of a target.

I retracted the symmetry-matching claim across the project record. The two ingredients of the artifact are standard knowledge (per-condition normalization is hazardous; MSE equal to variance means R^2 zero). Their compound, a deterministic two-level variance dichotomy produced by rounding and presenting as a robust cross-condition learning effect, cost this project five confirmatory experiments, and I have not found it reported. The diagnostic is one column: loss divided by own-target variance, per condition. Compared conditions that all sit at 1.0 differ only in preprocessing. Figure 2 shows both halves: the two-variance mechanism, and the one-column diagnostic over all 310 trainings.

3 Repaired instrument

Closure is enforced by construction: the third hexatic wavevector is defined as minus the sum of the other two. All hexatic cells on all lattices then share the 0.074-class variance and all tetratic cells 0.0625; the dichotomy cannot arise.

The score is explained variance, $1 - \text{MSE}/\text{Var}$, so the predict-the-mean solution scores zero on every lattice.

The readout is per-component. Each snapped wave is projected out of the grown field (equivalent to reading its FFT bin on Bravais lattices; the projection form remains exact on non-Bravais substrates such as the honeycomb used in section 5’s battery). For each component I record the realized angle and the snap strain (the angular deviation of the snapped wavevector from the ideal direction it was rounded from); the angular distance to the nearest primitive reciprocal-lattice axis of the substrate — the directions of the lattice’s primitive, i.e. shortest nonzero, reciprocal vectors, so the square’s $(1, 1)$ diagonal reciprocal direction is a lattice direction but not a primitive axis — is derived from the realized angle.

The rest of the protocol is unchanged from the original study: density-matched bases (square; hexagonal with $a = \sqrt{2/\sqrt{3}}$), four fixed perception operators (identity, zero-sum Laplacian, two gradients) sampled at each lattice’s true nearest neighbours, 3,944 parameters, cell-wise MLP update with fire rate 0.5, 72 steps per episode, $N = 96$ torus, 1,500 Adam iterations at batch 8, 16 held-out evaluation seeds, 4 restarts per cell. The state has 8 channels; the target is channel 0 read after 72 steps. Episodes start from a random soup confined to the central half-width of the grid; there is no sample pool, damage, or alive-masking as in the canonical Growing-NCA recipe, and each section-4 cell trains on 16 soup seeds. Cell values in section 4 are means over the four restarts; per-restart ranges are reported where they matter.

One provenance note: the repaired study was designed around the twelve surviving trainings of section 2, and its 30-degree hexatic target is anchored wavevector-for-wavevector to theirs by a pre-registered gate, with a frozen falsifier requiring that signal to reproduce. Section 4’s central cell is therefore a sharpened reproduction of that finding rather than an independent discovery.

4 Result: a primitive-axis constraint on growable components

Twelve cells: both lattices; hexatic targets at orientations 0, 15, and 30 degrees; tetratic at 0, 22.5, and 45. The hexagonal lattice’s primitive reciprocal axes lie at $30 + 60k$ degrees, the square’s at 0 and 90 (wave and axis directions are taken mod 180, so 270 degrees lies on the 90-degree axis).

Hexagonal lattice, hexatic target: at 30 degrees, where all three wavevectors lie on axes, the score is 0.704 (restarts 0.67–0.72) and the three components recover 0.55, 0.78, and 0.78 of their target amplitude. At 0 and 15 degrees, no component lies on an axis. Nothing grows: mean scores at or below 0.016, with the largest single restart reaching 0.057 by growing one off-axis component to 0.20.

Square lattice, same hexatic targets: at 0 and 30 degrees the square grows exactly the one component that falls on a square axis; at 15 degrees no component falls on an axis and nothing grows (score -0.002). The 0-degree wave of the 0-degree triad recovers 0.79; the 270-degree wave of the 30-degree triad recovers 0.79; the two cells score 0.265 and 0.264, tracking from below that component’s one-third share of the target variance.

Tetratic target at 0 degrees: the square lattice grows both components (0.58, 0.59; score 0.582). The hexagonal lattice grows exactly the 90-degree component, which lies on one of its axes (0.88; score 0.439, tracking that component’s one-half variance share). The one-sided interaction reported in the original study does not survive the repaired instrument. With aligned orientations and a variance-referenced score, each lattice grows its matched class, symmetrically.

The decisive cell is tetratic at 45 degrees on the square lattice. Both wavevectors snap strain-free onto the square’s diagonal (1, 1) reciprocal direction. They are exactly representable, and they do not grow: recovery 0.00. Exact representability is not the gate; primitive axes are. Figure 3 plots the gate across all twelve cells; Figure 4 shows the targets themselves.

Across the 30 component readouts in the study (120 across the four restarts): components within 0.5 degrees of a primitive axis recover 0.55–0.88 in restart means, per-restart minimum 0.50; all others recover at most 0.11 in restart means, per-restart maximum 0.21. The per-restart bands frozen before the run (on-axis recovery at least 0.6; components five or more degrees off-axis at most 0.2) missed on both edges (0.50 against 0.6; 0.21 against 0.2), and the frozen decision rule accordingly returned “mixed — report per-band, no headline”. The per-band report is the staircase above: the qualitative dichotomy is unambiguous in all twelve cells; the quantitative bands were set too tight, and section 5’s misalignment ladder measures the tooth’s half-recovery half-width at approximately 6 degrees. Since no orientation places three 120-degree-separated waves on the square lattice’s axes, nor two 90-degree-separated waves on the hexagonal lattice’s, in the tested family of Bravais substrates a substrate hosts a complete pattern class only when its point group matches the class; section 5’s honeycomb shows the operative axes are those the perception stencil samples, which for a Bravais lattice coincide with the point group. This is the defensible content of “symmetry match” in NCA: not a training subsidy, a hard constraint on which components can grow at all.

5 Pre-registered battery [COMPLETE 96/96; predictions were frozen before the first training; frozen decider FALSIFIED, gate survives softened]

The constraint predicts scores from lattice geometry alone. Before running anything further I froze a prediction rule calibrated only on section 4 (a component grows iff it lies within 0.5 degrees of a primitive axis with snap strain under 0.5 degrees; a growing component recovers 0.7; predicted cell score = 0.7 times the aligned variance share) and printed predicted scores for all 24 new cells, 96 trainings, before the first ran. Panels:

- Honeycomb, embedded brick-wall style: hexagonal (three-bond) connectivity drawn on a rectangular grid, so six-fold bond directions sit inside a rectangular box. Brick-wall is the standard embedding in hexagonal image processing and is known there to be a coarse, geometrically distorting approximation (Middleton and Sivaswamy, 2005; Rummelt and Wilson, 2011); it is used here because it is standard, and what it does to learnability is the measurement. It is the only substrate where the constraint's two possible readings diverge, and they diverge in opposite directions across the panel's cells. Reading by explicit axis sets: the stencil / point-group reading (axes {30, 90, 150} degrees, what section 4's wording predicts) gives 0.70 at the 30-degree hexatic (all three components grow) and 0.35 at the 0-degree tetratic; the box / embedding reading (the rectangular grid's reciprocal axes {0, 90} degrees) gives 0.23 at the 30-degree hexatic (only its 270-degree component) and 0.70 at the 0-degree tetratic. Both interpretations were committed to in writing before the run; the result below selected the {0, 90} box reading.
- Grid-size and wavelength robustness at four (N, λ) combinations, the axis of attack that exposed section 2. The predictions are per-cell geometric. Any cell with all components off-axis scoring above 0.25 falsifies the rule.
- Learned perception kernels on the square lattice (+36 parameters): whether learning can create hostability. Prediction: learned kernels can strengthen an already-aligned component but cannot make any off-axis component growable.
- A misalignment ladder: hexatic triads rotated 0.25 to 6.75 degrees off the hex axis, giving realized per-component axis distances from 0 (at 0.25 degrees snap strain) up to roughly 9 degrees, measuring the constraint's angular width.
- Data budgets 1, 4, and 64 training seeds: whether the gate is budget-independent.

Frozen decider: predicted-versus-measured R^2 at least 0.8 over all 24 rung-2d cells, and no component violating its class by more than 0.3 recovery; the pre-registered falsifier is R^2 below 0.5 or three or more class violations, reported per-cell either way.

The battery completed at 96 trainings (24 cells $\times$ 4 restarts; 101 raw records, of which five were benign re-appends of an already-recorded cell/restart on the same host and are dropped first-per-key, changing no cell mean by more than 0.002). **The frozen decider failed.** Predicted-versus-measured R^2 over the 24 cells is **-0.386** (-0.387 with the duplicate records retained); even the most charitable reading, squared Pearson correlation, is 0.20, below the 0.5 falsifier line. Class violations number 46 under the analysis code's implemented criterion and 70 under the strict pre-registration wording, against a falsifier threshold of three. By its own pre-registered terms the frozen quantitative rule — a 0.5-degree

binary axis gate with a 0.5-degree strain clause and a flat 0.7 recovery — is **wrong or incomplete**. The rule made a falsifiable a-priori prediction for every cell before any of these 96 trainings ran. The out-of-sample data refute the quantitative rule and sharpen the qualitative constraint.

Per-cell record (predicted / measured cell score, mean of 4 restarts). As required, reported for every cell regardless of verdict.

Panel	Cell	Predicted	Measured
D (budget)	hex hexatic-30, $n_{\text{train}} = 1$	+0.70	+0.147
D	hex hexatic-30, $n_{\text{train}} = 4$	+0.70	+0.613
D	hex hexatic-30, $n_{\text{train}} = 64$	+0.70	+0.677
D	square tetratic-0, $n_{\text{train}} = 1$	+0.70	+0.117
D	square tetratic-0, $n_{\text{train}} = 4$	+0.70	+0.530
D	square tetratic-0, $n_{\text{train}} = 64$	+0.70	+0.562
F (honeycomb)	honey hexatic-0	+0.00	+0.231
F	honey hexatic-30	+0.70	+0.137
F	honey tetratic-0	+0.35	+0.389
F	honey tetratic-45	+0.00	−0.002
M (learned kernels)	square-learn hexatic-0	+0.23	+0.288
M	square-learn hexatic-30	+0.23	+0.289
R (size / lambda)	hex hexatic-30, $N = 64$, $\lambda = 8$	+0.70	+0.855
R	hex hexatic-30, $N = 96$, $\lambda = 12$	+0.70	+0.814
R	hex hexatic-30, $N = 96$, $\lambda = 6$	+0.70	+0.467
R	hex hexatic-30, $N = 128$, $\lambda = 8$	+0.70	+0.437
R	square tetratic-0, $N = 64$, $\lambda = 8$	+0.70	+0.797
R	square tetratic-0, $N = 96$, $\lambda = 12$	+0.70	+0.699
R	square tetratic-0, $N = 96$, $\lambda = 6$	+0.70	+0.479
R	square tetratic-0, $N = 128$, $\lambda = 8$	+0.70	+0.392
T (misalignment)	hex hexatic, realized axis-dist ~ 0 deg ($\theta = 30.25$)	+0.70	+0.680
T	hex hexatic, $\sim 0/4.7/0$ deg, strain 2.25 ($\theta = 32.25$)	+0.00	+0.620
T	hex hexatic, ~ 4.3 deg ($\theta = 33.25$)	+0.00	+0.523
T	hex hexatic, ~ 8.9 deg ($\theta = 36.75$)	+0.00	−0.001

What survives out-of-sample: the gate, qualitatively. No component the geometry forbids and that is genuinely far off-axis ever grew. The 8.9-degree misalignment cell (T, $\theta = 36.75$) recovers exactly 0.000 ± 0.000 across all four restarts and all three components; the honeycomb tetratic-45 cell, off every candidate axis, scores -0.002 ; the square diagonal of section 4 stays at 0.00. Growth remains strictly direction-selective, and the sanity anchors reproduce section 4 through the new instrument (the on-axis misalignment anchor T $\theta = 30.25$ scores 0.680 against 2c’s hex hexatic-30 at 0.704; the budget-saturated square cell D $n_{\text{train}} = 64$ scores 0.562 against 2c’s square tetratic-0 at 0.582). *Post-hoc, so labelled:* on the 19 cells the rule classifies correctly the discrete predictions give $R^2 +0.209$ (against the global -0.386), and on the 15 cells that are both correctly classified and adequately budgeted (dropping the four data-starved D $n_{\text{train}} = 1$ and 4 cells) R^2 is $+0.67$. Both are subsets chosen after seeing the data, and the symmetric diagnostic is that no post-hoc refit rescues the rule over the full battery: the best single axis-width (4.31 degrees) reaches $R^2 +0.040$, refitting the two class levels reaches $+0.198$, and amending the classification over all 24 cells reaches $+0.212$ — none crosses the 0.5 falsifier line. A correct classification survives the discrete score metric, and the global negative R^2 reflects genuine

misclassification rather than a metric artifact (target variances are pinned per floor and kind, scores recompute exactly as $1 - \text{MSE}/\text{Var}$, and there is no route by which predict-the-mean can masquerade as learning; see section 2’s diagnostic).

What the frozen rule got wrong: the gate is soft and its amplitude is graded. The refutation is calibration on four axes, each localized to the panel built to probe it.

Tooth width (panel T). The gate is not 0.5 degrees wide; its half-recovery half-width is near 6 degrees, an order of magnitude wider than the frozen calibration. Cell scores fall as realized axis-distance grows — 0.680 at ~ 0 degrees, 0.620 at the mixed 0/4.7-degree cell, 0.523 at 4.3 degrees, 0.000 at 8.9 degrees — but components 4.3 to 4.7 degrees off-axis recover 0.34 to 0.66, against the frozen “at most 0.2 beyond 4 degrees.” The cliff is real but far out: partial growth through ~ 5 degrees, hard zero by ~ 9 . The pre-registered P-T band $\text{rec}(\geq 4 \text{ deg}) \leq 0.2$ is refuted; the reported half-width is ~ 6 degrees (bracketed between the 4.7-degree partial-growth and 8.9-degree hard-zero cells).

The strain clause is inert (panel T). Two cells ($\theta = 32.25$ and 33.25) were predicted all-dead, and both grew (0.620 and 0.523). In $\theta = 32.25$, two of three components snap exactly onto the axis (axis-distance 0.00) and were called “no” only by the strain clause; they recover 0.45 and 0.76 at strain 2.25 degrees. Strain up to 2.5 degrees costs at most 0.08 recovery. These two cells also trip the separately pre-registered gate-rule falsifier (any all-predicted-no-grow cell scoring at least 0.25): both exceed it, for the same reason the global decider fired.

Recovery is graded, not a flat 0.7 (panels D and R). The gate — which components can grow — is budget- and size-independent, but the amplitude actually achieved is not. Escape from predict-the-mean is graded with data budget and saturates by roughly four training seeds (D: hex 0.147, 0.613, 0.677 at $n_{\text{train}} = 1, 4, 64$; square 0.117, 0.530, 0.562), with no bimodality — every restart at $n_{\text{train}} = 1$ is already positive, so the low score reflects partial amplitude on partially-grown components rather than a failure to grow. Across grid size and wavelength (R), scores collapse not on N or λ separately but on their ratio, the target’s mode order $m = N/\lambda$ (waves per box): recovery is highest at the lowest mode order (~ 0.85 at $m \sim 8$) and lowest at the highest (~ 0.44 at $m \sim 16$) on both floors, monotone in m , with the $N = 128/\lambda = 8$ and $N = 96/\lambda = 6$ cells scoring alike despite sharing neither N nor λ . Every component in every R cell is partially recovered (no dead components anywhere in the panel) with tight restart spread — an undertraining-like graded-amplitude signature at fixed training steps, not a gate or geometry failure. (Speculation, so labelled: a coherence-horizon or credit-assignment limit of local updates at high mode order; at fixed step budget this is observationally indistinguishable from insufficient budget at high m .)

Honeycomb: the gate is the embedding box, not the six-fold stencil (panel F). The honeycomb was the only substrate where the constraint’s two readings diverge, and both were committed to in writing before the run. The pre-registered decider is honey hexatic-30: at or above 0.5 the gate is the stencil point group; at or below 0.35 with only the third component recovered it is box commensuration. Measured: **0.137, with only the 270-degree component recovered (0.407, against 0.001 and 0.001 for the other two)** — the box-commensuration branch, cleanly. The frozen rule had encoded the stencil reading (predicting hexatic-30 grows 3/3 at 0.70); the data reject it. The embedding rectangular box’s reciprocal axes (0 and 90 degrees) quantitatively explain all four F cells: hexatic-0 = 0.231 is its 0-degree box-axis component (recovery 0.695) over three; hexatic-30 = 0.137 is its 270-degree box-axis component alone; tetratic-0 = 0.389 grows its stencil-forbidden 0-degree box-axis component (0.554) while its 90-degree component is suppressed to 0.224; tetratic-45 sits off both and stays at zero. Binary grow/dead class is correct 10/10 components under the box reading. Post-hoc, box-y (90/270-degree) components recover at roughly half the amplitude of box-x components — a directional suppression on the brick-wall parity-

offset direction rather than a gate effect. On the honeycomb the honeycomb’s own six-fold point group did not gate growth; the axes its brick-wall stencil actually samples did (discussed in section 7).

Panels that passed as frozen. Learned perception kernels (M) confirm the prediction: off-axis components recover under 0.0005 across all eight records (nothing minted, two orders of magnitude under the 0.2 band), while the aligned component strengthens from the fixed-kernel baseline (0.86 against 2c’s 0.79, non-overlapping restart ranges, cell scores 0.288/0.289 against 2c’s 0.264/0.265). Conditioning polishes a hostable component; it does not mint one. Budget grading direction (P-D) holds. The one panel that resolves a pre-declared ambiguity, the honeycomb, resolves it — to box commensuration.

The frozen rule’s classification survives where growth is far on- or off-axis; its numbers are wrong in between and in amplitude. The constraint is real; it is a soft, budget- and size-graded gate with a tooth half-width near 6 degrees and, on non-Bravais substrates, an axis set fixed by the sampling stencil rather than the intrinsic point group — not the sharp 0.5-degree binary the calibration cell had suggested. Figure 1 plots predicted against measured for all 12 calibration (2c) and 24 out-of-sample (2d) cells.

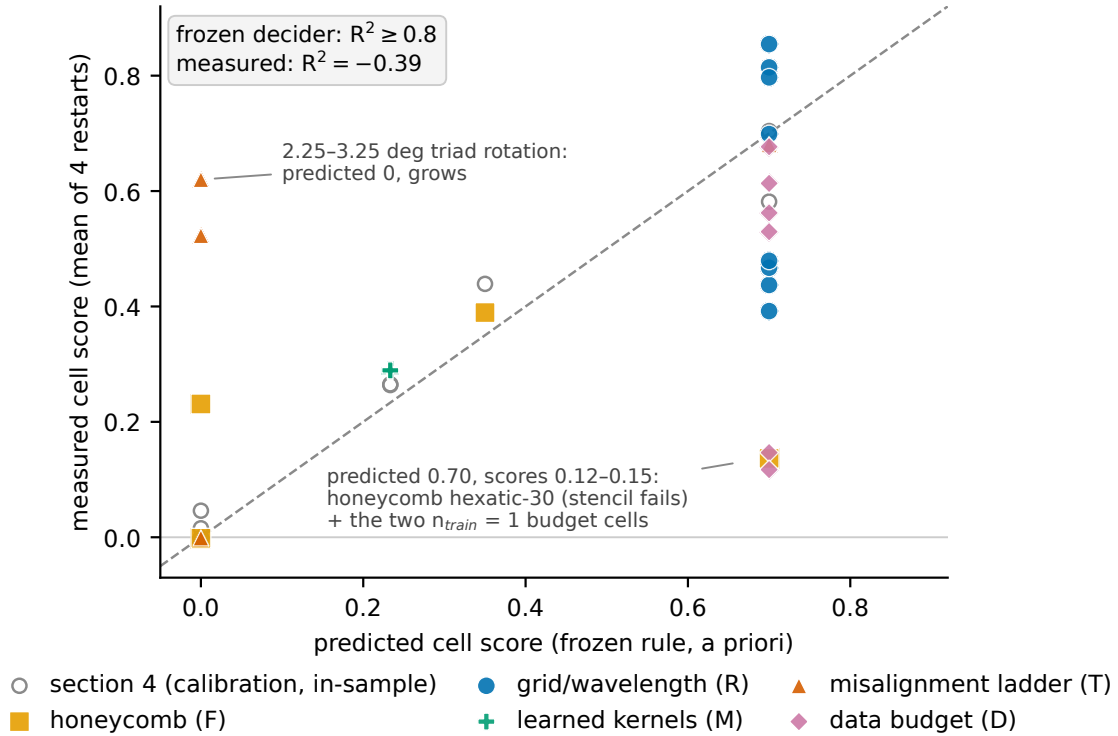


Figure 1: Predicted-versus-measured cell scores for the frozen primitive-axis rule, over the 12 calibration cells (rung-2c, section 4; open circles) and the 24 pre-registered out-of-sample cells (rung-2d battery, section 5; filled, by panel). Prediction is pure lattice geometry, frozen before any training: a component grows at recovery 0.7 within 0.5 degrees of a primitive reciprocal axis at snap strain below 0.5 degrees, else 0; cell score = 0.7 times the aligned variance share. Dashed line is $y = x$. Calibration $R^2 = 0.96$; out-of-sample $R^2 = -0.39$ (deduped 24 cells; the pre-registered decoder required at least 0.8, so it is falsified). The out-of-sample scatter is structured: components predicted dead by the 0.5-degree strain clause but snapped onto an axis grow to 0.5–0.6 (above the line at predicted 0); the honeycomb grows its embedding-box-axis component the stencil reading calls dead; and the flat 0.7 level overpredicts low-budget and high-mode-order cells (below the line at predicted 0.70). The qualitative gate survives — every geometry-forbidden far-off-axis component sits at the origin — while the quantitative rule is refuted on width, strain, and level. Built by `scripts/make_teeth_fig1.py` from the completed rung-2c + rung-2d records.

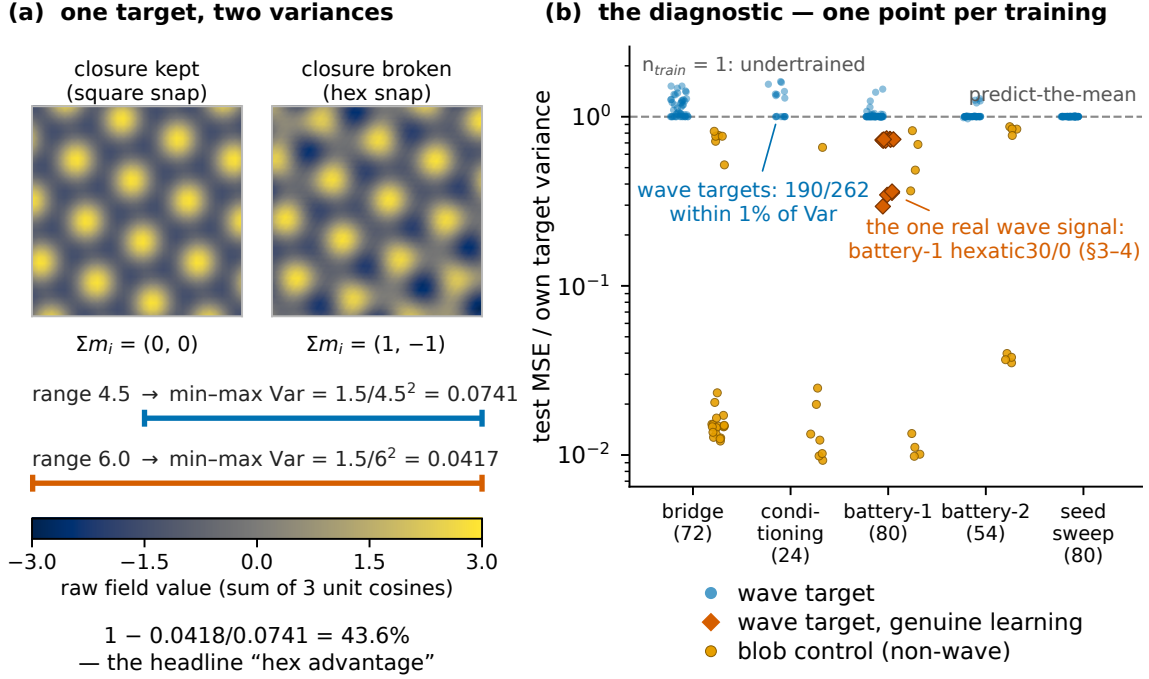


Figure 2: The normalization artifact (section 2). (a) The same three-cosine hexatic target (ideal angles 10/130/250 deg, $\lambda = 8$) snapped independently per wavevector onto each lattice’s reciprocal box ($N = 96$ torus). The square snap happens to preserve triad closure (sum of the integer mode vectors = $(0, 0)$): the raw field has range $[-1.5, 3]$, so min-max normalization leaves variance $1.5/4.5^2 = 0.0741$. The hex snap breaks closure (sum = $(1, -1)$): range $[-3, 3]$, variance $1.5/6^2 = 0.0417$. Which value a lattice draws is arithmetic in (basis, N , λ , orientation); the headline “hex advantage” is the identity $1 - 0.0418/0.0741 = 43.6\%$. (b) The one-column diagnostic over all 310 trainings of the original five-experiment study: test MSE divided by the model’s own target variance, one point per training. Every converged wave-target training (190 of 262) sits within 1% of 1.0, the predict-the-mean solution; points above the line are undertrained low-budget cells (mostly $n_{\text{train}} = 1$). Non-wave (blob) controls in the same harness reach ratios of 0.009, so the training machinery was not at fault. The only genuine wave learning in the family (diamonds) is battery-1’s hexatic-30/0 cells — the signal the repaired instrument of section 3 isolates. Conditioning and learned-kernel records store no target variance; it is taken from square-lattice records with identical snapped targets.

6 Related work

Othmer and Scriven (1971) and Turing-on-networks (Nakao and Mikhailov, 2010; Muolo et al., 2024): lattice spectra select the modes of fixed reaction-diffusion dynamics. The same restriction appears in automata: an isotropic cellular automaton produces hexagonal spot arrays a classical square lattice cannot (Schepers and Markus, 1992). No learning, no distinction among representable modes, no matched cross-lattice comparison. The present result is the learnability counterpart; an adversarial literature audit repeated on 2026-07-27 found no prior statement of it.

Spectral bias (Rahaman et al., 2019; Xu et al., 2019): learnability ordered by $|k|$, direction-blind. It does not predict the collapse of exactly-representable off-axis components at fixed $|k|$. Direction does matter to learnability elsewhere: neural anisotropy directions (Ortiz-Jiménez et al., 2020) show a CNN’s separability depends on feature orientation on a fixed pixel grid. That is an architectural bias at one substrate; the effect here varies with the substrate the same architecture runs on.

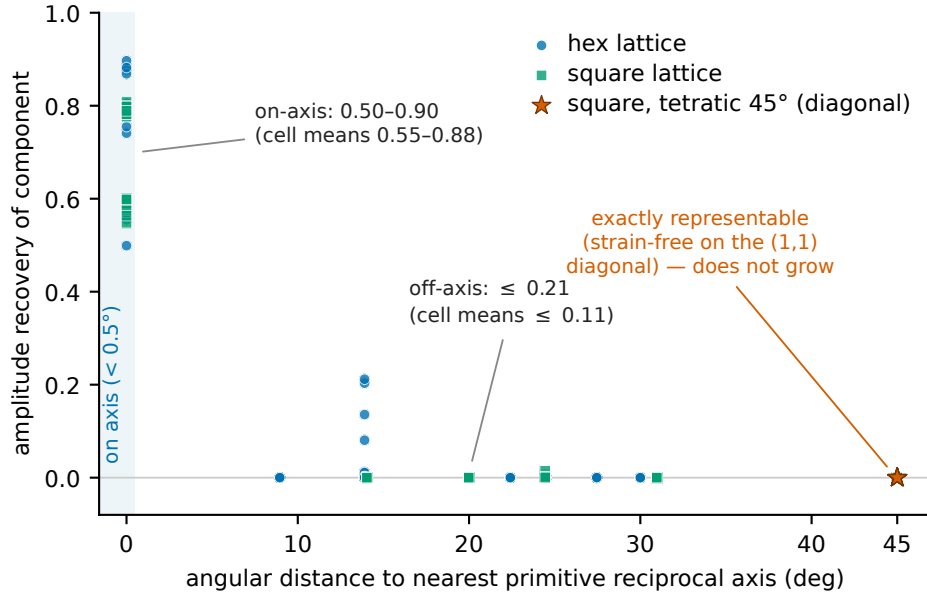


Figure 3: The primitive-axis gate (section 4). Amplitude recovery of every target Fourier component in the repaired study (12 cells $\times$ 4 restarts = 48 trainings; 120 component readouts) versus the angular distance of the component’s realized wavevector to the nearest primitive reciprocal-lattice axis of its substrate (hex: $30 + 60k$ deg; square: 0 and 90 deg; distances mod 180, inversion-equivalent). Components on a primitive axis (< 0.5 deg, shaded band) recover 0.50–0.90 of their target amplitude (cell means 0.55–0.88); every off-axis component stays at or below 0.21 (cell means ≤ 0.11). The design contains no components between 0.5 and 8.9 deg; section 5’s misalignment ladder probes that gap — and finds partial recovery through ~ 5 deg, placing the tooth’s half-recovery half-width near 6 deg. Stars: the decisive square/tetratic-45 cell, whose two components snap strain-free onto the square lattice’s (1,1) diagonal — exactly representable, 45 deg from a primitive axis, and not grown (recovery 0.00). Representability is not the gate; primitive axes are.

NCA and lattices: within the Growing-NCA line (Mordvintsev et al., 2020), IsoNCA (Mordvintsev et al., 2022) and steerable NCA (Randazzo et al., 2023) make the rule isotropic on a fixed square grid; Hexells (Mordvintsev, 2021) is a hexagonal demonstration; MeshNCA (Pajouheshgar et al., 2024) and E(n)-equivariant graph NCA (Gala et al., 2023) generalize the domain and remove lattice dependence by construction. None measures which Fourier components a lattice permits a trained rule to grow. A recent high-resolution NCA method paper (Pajouheshgar et al., 2025, arXiv 2506.22899) contains no commensuration analysis. Neither do the field’s two current reviews: Hartl et al. (2025) surveys NCA across biology and AI, and Spitznagel and Keuper (2026) systematically evaluates perception modules — the component of the architecture closest to section 7’s mechanism — without addressing lattice geometry, reciprocal axes, or spectral constraints. Kvalsund and Stovold (2026) is the closest current use of Fourier analysis on trained NCA, taking Lyapunov and Fourier spectra of attractor dynamics on a square-grid NCA; that is a temporal spectrum of the dynamics, not a measurement of which spatial components a substrate permits a rule to grow. One scope condition matters here: the constraint is established for fixed-phase targets scored by pixel MSE. NCA trained with translation-invariant texture losses (Niklasson et al., 2021) grow periodic patterns at arbitrary orientations on square grids; whether the gate reappears once the loss demands a globally registered phase, as here, versus a phase-free statistic, is exactly the boundary this result draws.

Adjacent mechanisms: checkerboard artifacts in learned upsamplers (Odena et al., 2016) imprint lattice frequencies via stride overlap, not target commensuration. Hexagonal sampling theory (Mersereau,

1979) characterizes representable frequencies; section 4’s diagonal cell shows representability is not the operative condition. In classical cellular automata, lattice point group already limits hostable physics: square HPP lattice gases cannot produce isotropic hydrodynamics where hexagonal FHP can (Frisch et al., 1986); that line concerns emergent continuum symmetry under fixed rules, not component-level learnability.

Methods side: per-condition normalization hazards and the MSE-equals-variance baseline are textbook individually. The compound of section 2 appears unreported.

7 Discussion and limitations

A candidate mechanism, stated as interpretation: growing a global periodic pattern from a localized seed requires the rule to build and maintain a phase reference. On a primitive reciprocal axis the wave is constant along one of the substrate’s nearest-neighbour lattice directions. The pattern is then one-dimensional in lattice coordinates, and a cell need only hold a fixed phase offset along a single stencil-sampled line. Off axis, including along the square diagonal, the constant-phase direction is a lattice line the stencil does not sample: the pattern varies along every sampled neighbour direction, and no single-line registration is available for the rule to check and correct. This reading carries a confound: on the tested lattices a wavevector lies on a primitive reciprocal axis exactly when its constant-phase direction is a stencil-sampled neighbour line, so lattice point group and perception geometry cannot be separated here. A Moore-neighbourhood or learned-kernel training on the square diagonal target would separate them; section 5’s learned-kernel panel does not include that cell, so the separation is a first follow-up. The misalignment and learned-kernel panels of section 5 bear on this reading but were not designed to prove it; whether the trained kernels themselves encode the single-line registration is untested.

Section 5’s honeycomb bears directly on this confound and partially resolves it. The honeycomb, embedded brick-wall style on a rectangular grid, does separate point group from stencil geometry: its intrinsic six-fold point group and the axes its rectangular stencil actually samples diverge, and growth followed the stencil / embedding-box axes (the box-commensuration branch of the pre-registered P-F decider — honey hexatic-30 = 0.137 with only the box-axis component recovered) rather than the six-fold symmetry. This is direct evidence for the perception-geometry reading over the pure point-group reading, so I report it as a partial in-paper result rather than a follow-up. It rests on one non-Bravais substrate. The clean Moore-versus-diagonal separation on the square lattice remains the cited follow-up. A post-hoc detail supports the reading: on the honeycomb the box-y (parity-offset) components recover at roughly half the box-x amplitude, the direction carrying the brick-wall’s two-site offset.

Read through solid-state conventions the box outcome is the less surprising of the two branches, which is worth saying plainly given that the frozen rule had encoded the other one. A reciprocal lattice is a property of the underlying Bravais lattice and does not depend on how many sites the unit cell carries (Ashcroft and Mermin, 1976, ch. 4–5); a brick-wall honeycomb on a rectangular grid is a rectangular Bravais lattice with a two-site basis, whose reciprocal axes are $\{0, 90\}$ degrees. The measured gate follows those axes. I did not reason this way when freezing the prediction — the pre-registered decider gave the stencil reading equal standing and the battery chose between them — but it means the result agrees with the textbook account of the substrate rather than standing against it.

The finite tooth width bears on the mechanism too. The phase-registration reading predicts a sharp gate, yet the battery measured a half-recovery half-width near 6 degrees, with partial growth through ~ 5 degrees and hard zero by ~ 9 . The reading must accommodate that width: near-axis components admit partial single-line registration that degrades gradually with misalignment rather than switching off at

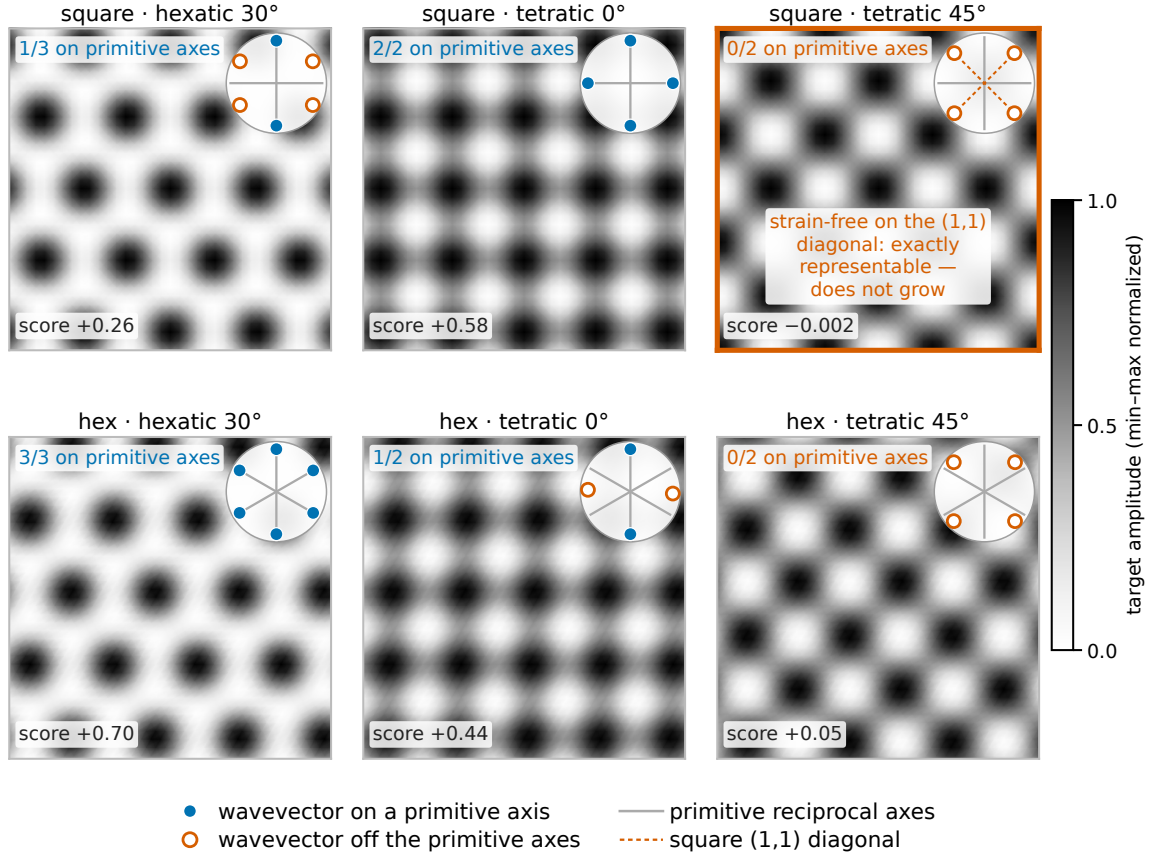


Figure 4: Lattice/target gallery (sections 3–4). Closure-enforced snapped wave targets ($\lambda = 8$, $N = 96$, min-max normalized, rendered in real space; 4.5-wavelength crop) for both substrates at hexatic 30 deg, tetratic 0 deg, and tetratic 45 deg. Compass insets show the substrate’s primitive reciprocal axes (grey lines) and the realized wavevector directions ($\pm k$) of each target component: filled if within 0.5 deg of a primitive axis, open otherwise. Panel scores are explained variance, mean of 4 restarts. Scores follow the aligned-component variance share (1/3 per hexatic component, 1/2 per tetratic): 3/3 aligned scores 0.70, 2/2 scores 0.58, 1/2 scores 0.44, 1/3 scores 0.26, 0/n scores about 0. The outlined cell is the discriminator: square/tetratic-45’s wavevectors lie strain-free on the square’s (1,1) diagonal (dashed) — exactly representable yet off the primitive axes — and the cell does not grow.

the exact axis. The snap-strain clause proved inert (strain up to 2.5 degrees costs at most 0.08 recovery), so strain is not a second gate variable.

Limitations: one dynamics family (Growing-NCA-style stochastic update, without the canonical recipe’s sample pool, damage, or alive-masking), one wavelength per cell, single-channel targets, four restarts per cell, a single stochastic fire rate (0.5; synchronous update untested), and periodic domains — the torus is what quantizes the reciprocal box, and the constraint’s form on open boundaries is untested. A further limitation the battery exposed: the gate (which components can grow) is budget- and size-independent, but the amplitude achieved is graded — by data budget (escape from predict-the-mean saturates by ~ 4 training seeds), by the target’s mode order $m = N/\lambda$ (recovery ~ 0.85 at $m \sim 8$ falling to ~ 0.44 at $m \sim 16$ on both floors), and by direction (on hexagonal floors the index-diagonal component underperforms its single-axis siblings, a deficit growing with m). This is plausibly a coherence-horizon or credit-assignment limit of local updates at high mode order at fixed training budget (labelled speculation), and it means a single “recovery 0.7” is a calibration-point value rather than a constant. The

mechanism paragraph is post-hoc. The scope of the claim is the tested family; whether the constraint binds other trained local systems (learned PDE solvers, learned reaction-diffusion) is open.

A companion study in the same research programme finds the same commensuration structure in self-organized (non-learned) pattern formation on these lattices. The parallel is suggestive, and it is kept out of this paper’s claims.

8 Reproducibility

All experiments ran under decision bands frozen in script docstrings before any data existed, and instrument gates (synthetic-truth recovery tests) preceded every analysis. Every script states as a hard rule that no result is written to findings before its run completes, and analysis code refuses to emit verdicts until the full pre-registered record count exists. The project maintains an artifact ledger, currently twelve entries, of which the retraction in section 2 is the largest. The section-5 partial falsification is not an artifact — it is a pre-registered prediction refuted by out-of-sample data, recorded as a prediction-versus-outcome entry, and the ledger count stands at twelve. The section-5 battery completed at 96 trainings (24 cells $\times$ 4 restarts) recorded across 101 JSONL lines, five of which are benign re-appends of an already-recorded cell/restart (deduped first-per-key; maximum score delta 0.014; no cell mean shifts by more than 0.002 and no verdict changes); the calibration study (section 4) is 48 trainings (12 cells $\times$ 4 restarts). The frozen prediction table (`--gate`) and the analysis verdict (`--analyze`, printing the R^2 and the per-record violation count) are reproducible read-only entry points; Figure 1 is generated from the completed rung-2c and rung-2d records. Scripts and per-training JSONL records accompany the paper. The integrity tooling used throughout has been extracted as a small package (prereg), to be released; its worked example is section 2.

References

- Ashcroft, N. W., & Mermin, N. D. (1976). Solid State Physics. Holt, Rinehart and Winston. (Chapters 4–5, reciprocal lattice and lattices with a basis.)
- Frisch, U., Hasslacher, B., & Pomeau, Y. (1986). Lattice-gas automata for the Navier–Stokes equation. *Physical Review Letters*, 56(14), 1505–1508. doi:10.1103/PhysRevLett.56.1505.
- Gala, G., Grattarola, D., & Quaghebeur, E. (2023). E(n)-equivariant Graph Neural Cellular Automata. arXiv:2301.10497.
- Hartl, B., Levin, M., & Pio-Lopez, L. (2025). Neural cellular automata: applications to biology and beyond classical AI. arXiv:2509.11131. doi:10.48550/arXiv.2509.11131.
- Kvalsund, M.-K., & Stovold, J. (2026). Stability and Geometry of Attractors in Neural Cellular Automata. arXiv:2604.12720. doi:10.48550/arXiv.2604.12720.
- Mersereau, R. M. (1979). The processing of hexagonally sampled two-dimensional signals. *Proceedings of the IEEE*, 67(6), 930–949. doi:10.1109/PROC.1979.11356.
- Middleton, L., & Sivaswamy, J. (2005). Hexagonal Image Processing: A Practical Approach. Springer. doi:10.1007/1-84628-203-9.
- Mordvintsev, A., Randazzo, E., Niklasson, E., & Levin, M. (2020). Growing Neural Cellular Automata. *Distill*. doi:10.23915/distill.00023. <https://distill.pub/2020/growing-ca>

- Mordvintsev, A. (2021). Hexells. ACM SIGGRAPH 2021 Art Gallery. doi:10.1145/3450507.3457438. Demo: <https://znah.net/hexels/>
- Mordvintsev, A., Randazzo, E., & Fouts, C. (2022). Growing Isotropic Neural Cellular Automata. arXiv:2205.01681.
- Muolo, R., Giambagli, L., Nakao, H., Fanelli, D., & Carletti, T. (2024). Turing patterns on discrete topologies: from networks to higher-order structures. *Proceedings of the Royal Society A*, 480(2302), 20240235. doi:10.1098/rspa.2024.0235.
- Nakao, H., & Mikhailov, A. S. (2010). Turing patterns in network-organized activator-inhibitor systems. *Nature Physics*, 6, 544–550. doi:10.1038/nphys1651. arXiv:1005.1986.
- Niklasson, E., Mordvintsev, A., Randazzo, E., & Levin, M. (2021). Self-Organising Textures. *Distill*. doi:10.23915/distill.00027.003. <https://distill.pub/selforg/2021/textures/>
- Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and Checkerboard Artifacts. *Distill*. doi:10.23915/distill.00003. <https://distill.pub/2016/deconv-checkerboard/>
- Ortiz-Jiménez, G., Modas, A., Moosavi-Dezfooli, S.-M., & Frossard, P. (2020). Neural Anisotropy Directions. *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. arXiv:2006.09717.
- Othmer, H. G., & Scriven, L. E. (1971). Instability and dynamic pattern in cellular networks. *Journal of Theoretical Biology*, 32(3), 507–537. doi:10.1016/0022-5193(71)90154-8.
- Pajouheshgar, E., Xu, Y., Mordvintsev, A., Niklasson, E., Zhang, T., & Süssstrunk, S. (2024). Mesh Neural Cellular Automata. *ACM Transactions on Graphics*, 43(4). doi:10.1145/3658127. arXiv:2311.02820.
- Pajouheshgar, E., Xu, Y., Abbasi, A., Mordvintsev, A., Jakob, W., & Süssstrunk, S. (2025). Neural Cellular Automata: From Cells to Pixels. arXiv:2506.22899.
- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the Spectral Bias of Neural Networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, PMLR 97, 5301–5310. arXiv:1806.08734.
- Randazzo, E., Mordvintsev, A., & Fouts, C. (2023). Growing Steerable Neural Cellular Automata. arXiv:2302.10197.
- Rummelt, N. I., & Wilson, J. N. (2011). Array set addressing: enabling technology for the efficient processing of hexagonally sampled imagery. *Journal of Electronic Imaging*, 20(2), 023012. doi:10.1117/1.3589306.
- Schepers, H. E., & Markus, M. (1992). Two types of performance of an isotropic cellular automaton: stationary (Turing) patterns and spiral waves. *Physica A*, 188(1–3), 337–343. doi:10.1016/0378-4371(92)90277-W.
- Spitznagel, M., & Keuper, J. (2026). A New Kind of Network? Review and Reference Implementation of Neural Cellular Automata. arXiv:2604.24990. doi:10.48550/arXiv.2604.24990.
- Xu, Z.-Q. J., Zhang, Y., Luo, T., Xiao, Y., & Ma, Z. (2019). Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks. arXiv:1901.06523. (*Communications in Computational Physics*, 28(5), 1746–1767, 2020.)