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.

Machine Learning the Phase Structure and Algebraic Invariants of Clifford- and Grassmann-Valued Cellular Automata

fuat kaan mirza

kaan.mirza@khas.edu.tr

Kadir Has University

Research Article

Keywords: Clifford algebra, Grassmann (exterior) algebra, cellular automata, machine learning, experimental mathematics, conjecture formulation, grade entropy, phase transitions.

Posted Date: July 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10394430/v1

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

Additional Declarations: No competing interests reported.

Machine Learning the Phase Structure and Algebraic Invariants of Clifford- and Grassmann-Valued Cellular Automata

Fuat Kaan Mirza

Department of Electrical Electronics Engineering, Faculty of Engineering and Natural Sciences,
Kadir Has University, Cibali 34083, Istanbul, Turkiye

Abstract

The application of data-science methods to the discovery of mathematical structure, from the string landscape to topological invariants, representation theory and automated conjecture generation, has grown into a distinct branch of experimental mathematics. In the spirit of that programme, and of the interests of this collection, we ask what machine learning can reveal about Clifford and Grassmann algebras when those algebras are used to drive a discrete dynamical system. We introduce a family of normalized multivector-valued cellular automata whose sites carry states in $\mathrm{Cl}(p,q)$ or in the exterior algebra $\Lambda(\mathbb{R}^3)$, and treat the automaton as a generator of algebra-labeled dynamical data. Using grade-resolved algebraic statistics, grade energy and entropy, a grade-mixing index, a stabilized Clifford energy and perturbation-growth descriptors, we (i) compute two-parameter phase diagrams for $\mathrm{Cl}(2,0), \mathrm{Cl}(3,0), \mathrm{Cl}(1,2), \mathrm{Cl}(1,3)$ and the nilpotent Grassmann algebra $\Lambda(\mathbb{R}^3)$, exposing ordered, periodic, chaotic grade-mixing, algebraically constrained and intermittent regimes; (ii) show that grade entropy and an effective Lyapunov exponent act as coincident order parameters; and (iii) demonstrate that unsupervised learning recovers the phase organization directly from trajectory descriptors. Crucially for the experimental-mathematics setting, supervised learning recovers the generating algebra and its metric-signature class from dynamical observables alone. Under a genuinely blind protocol, using eight descriptors computed in Euclidean coefficient space that contain neither the algebra dimension nor the Clifford metric, and folds grouped so that every algebra at a parameter coordinate shares a fold, the signature class is recovered with 87% accuracy (chance 33%) and a definite-versus-indefinite classifier trained on one algebra pair transfers to the other at 81–94% balanced accuracy (chance 50%); the dynamics carry a learnable, transferable algebraic fingerprint, although a normalization control shows that the metric normalization contributes part of it. Finally, symbolic regression proposes a compact candidate scaling relation for the onset of chaos, $\lambda_{\mathrm{eff}} \approx 0.41\, H_{\mathrm{spec}}^{3.33}$, and a depth-three decision tree reproduces all five phase labels with 98.3% fidelity as an interpretable surrogate. A matched-protocol comparison across scalar, complex, quaternionic, Clifford and Grassmann automata shows that phase diversity is not monotone in algebra dimension. A normalization control, however, shows that the collapse of the indefinite algebras is produced by the interaction of the indefinite bilinear form with the normalization step rather than by the signature of the product itself: under a common Euclidean coefficient normalization the indefinite algebras recover a phase diversity comparable to the definite ones.

Keywords: Clifford algebra · Grassmann (exterior) algebra · cellular automata · machine learning · experimental mathematics · conjecture formulation · grade entropy · phase transitions.

1 Introduction

In recent years, data-science methods, including deep learning, data mining, and symbolic and statistical inference, have been applied to the fundamental structures of mathematics itself. The programme began in mathematical physics with deep-learning explorations of the string landscape and Calabi–Yau data [1–4], and has since reached algebraic geometry (learning topological invariants and efficient Gröbner bases), algebra and representation theory (neural classifiers distinguishing finite simple groups), combinatorics and graph theory (detecting Hamiltonian cycles, estimating Laplacian eigenvalues) [5, 6, 11], and symbolic manipulation, where systems such as the Ramanujan Machine and SciNet generate new identities and physical representations [7, 8, 10]. Together with the long-term goal of automated theorem proving in Voevodsky's univalent foundations [6, 9], these developments have established machine learning as an instrument for experimental mathematics and conjecture formulation. In light of the interests of this collection, the application of machine learning to patterns in Clifford and Grassmann algebras warrants direct investigation; recent work in the same collection has learned Clifford invariants of Coxeter elements [34] and discovered invariants of braids [35], and Clifford and quaternion neural architectures continue to be studied directly [36].

Clifford (geometric) algebras $\mathrm{Cl}(p, q)$ and the Grassmann (exterior) algebra $\Lambda(\mathbb{R}^n)$ unify scalars, vectors, bivectors and higher blades into a single graded associative structure, and have become a natural language for physics and, increasingly, for machine learning, where Clifford, quaternionic and Clifford-group-equivariant neural layers exploit the grade structure as an inductive bias [13–21]. Yet these algebras have largely been used inside learning machines rather than being themselves the object of machine-learning study. Here we take the complementary view: we let a Clifford- or Grassmann-valued dynamical system generate algebra-labeled data, and ask what a learner can discover about the underlying algebra from that data alone.

Our vehicle is a multivector-valued cellular automaton (CA). Classical CA are among the simplest discrete dynamical systems capable of universal computation and of the full range of ordered, periodic, chaotic and complex behavior [22–24], but almost always with scalar or vector states. We make the state a multivector $X_i(t) \in \mathrm{Cl}(p, q)$ (or $\Lambda(\mathbb{R}^3)$) and let the local rule act through the algebra's geometric (or exterior) product. The central hypotheses are that (a) such automata possess phase structure that scalar observables cannot resolve but grade-resolved algebraic statistics can; (b) the generating algebra and its metric-signature invariants are recoverable from the dynamics by machine learning; and (c) the resulting data support the automated formulation of compact, interpretable conjectures. Our contributions are:

- a family of normalized multivector CA over $\mathrm{Cl}(p, q)$ and $\Lambda(\mathbb{R}^3)$, with a single interpretable five-parameter local rule (Section 3);
- grade entropy, a grade-mixing index, a stabilized Clifford energy and perturbation-growth descriptors as algebra-sensitive observables (Section 4);
- two-parameter phase diagrams for five algebras of increasing dimension, signature and degeneracy, with grade entropy and an effective Lyapunov exponent as coincident order parameters (Sections 6.2–6.3);
- unsupervised recovery of the phase organization and, at the core of the experimental-mathematics contribution, the supervised recovery of the generating algebra and its metric-signature class from dynamical descriptors, $95.8\%/97.3\%$ with the grade-component-excluded set $\mathcal{D}_{14}$, and $72.2\%/87.2\%$ under a genuinely blind, dimension- and metric-free protocol (Sections 7.1–7.3);
- symbolic-regression and decision-tree conjecture formulation of the chaos-onset law and the rule-based phase-label surrogate (Section 7.5); and
- a matched-protocol comparison across scalar, complex, quaternionic, Grassmann and Clifford dynamics, together with a common-normalization control that identifies the normalization convention, not the signature of the product alone, as the mechanism behind the collapse of the indefinite algebras (Section 6.5).

The remainder of the paper follows the pipeline of Fig. 1. Sections 2 and 3 define the algebras and the automaton; Section 4 introduces the grade-resolved observables and the phase criteria; Section 5 fixes the simulation and learning protocol; Section 6 reports the dynamical phase structure; Section 7 reports the machine-learning discovery of phases, algebraic invariants and conjectures; and Sections 8 and 9 present the discussion and conclusions, respectively.

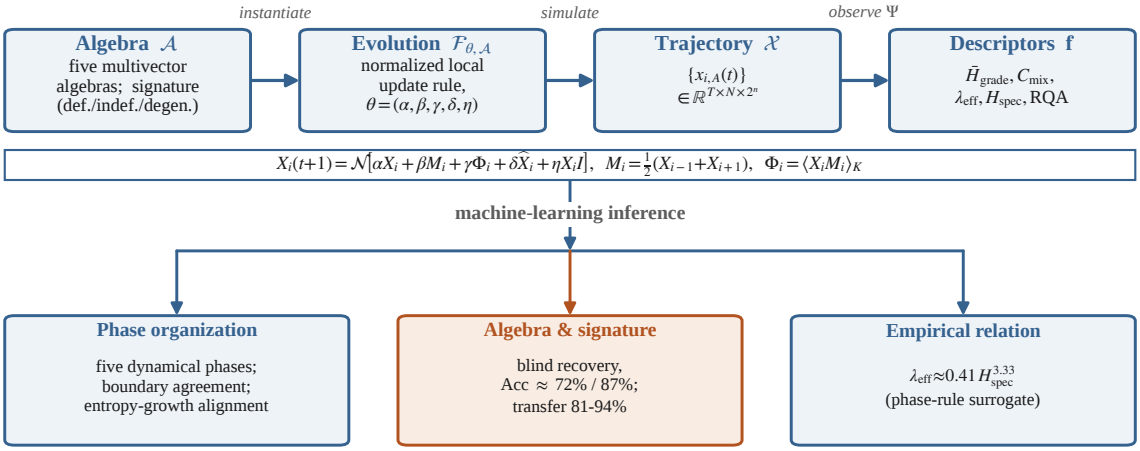


Figure 1: Graphical overview of the study. Five multivector algebras of varying signature drive a normalized cellular automaton; grade-resolved observables summarize each trajectory; a rule-based scheme and two order parameters organize the dynamics into phases; and unsupervised, supervised and symbolic learning then recover the phase diagrams, the generating algebra and its signature invariant, and compact conjectures.

2 Clifford and Grassmann algebraic background

Let $V = \mathbb{R}^{p,q,r}$ be an n-dimensional real vector space, $n = p + q + r$, equipped with a symmetric bilinear form g whose associated quadratic form is

$$Q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 + 0 \cdot \sum_{i=p+q+1}^{n} x_i^2, \tag{1}$$

so that p directions are positive, q negative and r null. The real geometric algebra $\mathcal{G} = \mathcal{G}(p, q, r)$ is the associative unital algebra generated by an orthonormal basis $\{e_1, \ldots, e_n\}$ subject to the Clifford relation

$$e_i e_j + e_j e_i = 2\, g_{ij}, \qquad g_{ij} = \begin{cases} +1, & i = j \le p, \\ -1, & p < i = j \le p + q, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

For $r = 0$ Eq. (2) defines the real Clifford algebra $\mathrm{Cl}(p, q)$, whose classification, center and periodicity structure are standard [37]; for $p = q = 0$ every generator is nilpotent, $e_i^2 = 0$, and Eq. (2) reduces to the anticommutation $e_i e_j = -e_j e_i$ of the Grassmann (exterior) algebra $\Lambda(\mathbb{R}^n)$, whose product we write by juxtaposition or by the wedge $\wedge$.

A basis of blades is indexed by subsets $A = \{a_1 < \cdots < a_k\} \subseteq \{1, \ldots, n\}$ through $e_A = e_{a_1} e_{a_2} \cdots e_{a_k}$ with $e_\emptyset = 1$, and a general multivector reads

$$X = \sum_{A \subseteq \{1,\ldots,n\}} x_A\, e_A, \qquad x_A \in \mathbb{R}, \qquad \dim \mathcal{G} = 2^n. \tag{3}$$

The geometric product of two blades is $e_A e_B = \varepsilon(A, B)\, e_{A \triangle B}$, where $A \triangle B$ is the symmetric difference and $\varepsilon(A, B) \in \{-1, 0, +1\}$ collects the reordering sign together with the metric contractions of repeated generators; the

product of two multivectors is therefore bilinear with structure constants $C_{AB}^{\ D}$,

$$(XY)_D = \sum_{A,B} C_{AB}^{\ D}\, x_A\, y_B, \qquad C_{AB}^{\ D} = \varepsilon(A,B)\,\delta_{D,\,A\triangle B}. \tag{4}$$

The algebra is $\mathbb{Z}$-graded, $\mathcal{G} = \bigoplus_{k=0}^{n} \mathcal{G}^{(k)}$, and the grade-$k$ projector retains the blades of cardinality k,

$$\langle X\rangle_k = \sum_{|A|=k} x_A\, e_A, \qquad X = \sum_{k=0}^{n} \langle X\rangle_k. \tag{5}$$

The scalar (k=0), vector (k=1), bivector (k=2) and top (k=n) grades are of particular interest, the last spanned by the pseudoscalar

$$I = e_1 e_2 \cdots e_n, \qquad I^2 = (-1)^{n(n-1)/2}(-1)^q \quad (r=0), \tag{6}$$

which for the algebras of Table 1 gives $I^2 = -1$ for the four Clifford algebras and $I^2 = 0$ for $\Lambda(\mathbb{R}^3)$.

Three linear (anti)involutions structure the dynamics. Acting grade-wise, the grade involution, the reversion and the Clifford conjugation are

$$\widehat{X} = \sum_k (-1)^k \langle X\rangle_k, \qquad \widetilde{X} = \sum_k (-1)^{k(k-1)/2} \langle X\rangle_k, \qquad \bar{X} = \widehat{\widetilde{X}} = \sum_k (-1)^{k(k+1)/2} \langle X\rangle_k, \tag{7}$$

with $\widehat{XY} = \widehat{X}\,\widehat{Y}$ and $\widetilde{XY} = \widetilde{Y}\,\widetilde{X}$; they are the algebraic analogues of parity, time reversal and conjugation. The reversion induces a scalar product and a quadratic form on the whole algebra,

$$\langle X, Y\rangle = \langle X\widetilde{Y}\rangle_0 = \sum_A s_A\, x_A\, y_A, \qquad \|X\|_{\mathrm{Cl}}^2 = \langle X\widetilde{X}\rangle_0 = \sum_A s_A\, x_A^2, \tag{8}$$

where $s_A = \langle e_A \widetilde{e_A}\rangle_0 \in \{-1, 0, +1\}$ is the metric sign of blade A. For Euclidean signatures $s_A = +1$ and Eq. (8) is positive-definite; for indefinite signatures it is sign-indefinite; and for the Grassmann algebra $s_A = 0$ on every blade containing a null generator, so that $\|X\|_{\mathrm{Cl}}^2 = x_\emptyset^2$ degenerates to the scalar part. To obtain a well-defined, blow-up-free magnitude in all cases we use the stabilized norm

$$\|X\|_\epsilon = \begin{cases} \sqrt{\left|\langle X\widetilde{X}\rangle_0\right| + \epsilon}\,, & r = 0 \ \text{(Clifford)}, \\ \sqrt{\sum_A x_A^2 + \epsilon}\,, & r > 0 \ \text{(Grassmann)}, \end{cases} \qquad \epsilon = 10^{-8}, \tag{9}$$

and the normalization map $\mathcal{N}[X] = X/\|X\|_\epsilon$. Equation (9) is a metric-dependent normalization with null-cone regularization. For a definite signature it maps the state onto a sphere; for an indefinite signature the pseudo-norm level sets are non-compact, and the stabilizer ϵ regularizes the map near the null cone. It preserves the grading of Eq. (5).

Five multivector algebras are used throughout the paper, and they are chosen so that the three structural properties that could plausibly matter for the dynamics, namely the dimension, the signature and the presence of null generators, vary as independently as possible within the selected eight- and sixteen-dimensional families. Table 1 lists them with their generators, their dimension, their grade content and their signature class. Two comparison structures make this analysis possible. $\mathrm{Cl}(3,0)$, $\mathrm{Cl}(1,2)$ and $\Lambda(\mathbb{R}^3)$ all have dimension eight and differ only in signature, definite, indefinite and degenerate respectively, which is what allows the same-dimension control of Section 7.3. $\mathrm{Cl}(2,0)$ and $\mathrm{Cl}(1,3)$ bracket that group from below and above in dimension. The scalar, complex and quaternionic systems used as baselines in Section 6.5 are lower-dimensional baseline systems, the real scalars ($D = 1$), the complex numbers $\mathrm{Cl}(0,1)$ and the quaternions $\mathrm{Cl}(0,2)$, all nondegenerate low-dimensional special cases, and are not listed separately.

To keep the exposition self-contained, Table 2 collects the principal notation used in the remainder of the paper before the dynamics are defined.

The schematic makes explicit that every ingredient of the update acts through the algebra: the neighborhood coupling and the geometric product mix grades, while the involution and pseudoscalar terms probe parity and chirality. All observables introduced below are computed on the grade decomposition previewed in panel (b).

Table 1: The five multivector algebras used in this study, with the algebraic invariants a learner is later asked to recover. I^2 is the square of the pseudoscalar (computed from the engine); $\mathfrak{z}$ is semisimplicity. The three algebras of dimension eight differ only in signature class, which makes the same-dimension control of Section 7.3 possible. All four Clifford algebras share $I^2 = -1$; a future $I^2 = +1$ algebra would be needed to test contrasting pseudoscalar-square classes. The scalar, complex and quaternionic systems of Section 6.5 are baselines, not members of this set.

Algebra	(p, q, r)	n	$\dim \mathcal{A}$	Signature	I^2	Role in the comparison
$\mathrm{Cl}(2,0)$	$(2,0,0)$	2	4	definite	-1	minimal definite Clifford system
$\mathrm{Cl}(3,0)$	$(3,0,0)$	3	8	definite	-1	eight-dim. definite control
$\mathrm{Cl}(1,2)$	$(1,2,0)$	3	8	indefinite	-1	eight-dim. indefinite control
$\Lambda(\mathbb{R}^3)$	$(0,0,3)$	3	8	degenerate	0	eight-dim. degenerate control
$\mathrm{Cl}(1,3)$	$(1,3,0)$	4	16	indefinite	-1	sixteen-dim. indefinite system

Table 2: Principal notation.

Symbol	Meaning
$\mathrm{Cl}(p, q)$	Clifford algebra of signature (p, q)
$\Lambda(\mathbb{R}^n)$	Grassmann (exterior) algebra
$X_i(t)$	multivector state at site i, time t
$\langle X \rangle_k$	grade-k part of X
$\widehat{X},\ \widetilde{X},\ \bar{X}$	grade involution, reversion, conjugation
$I = e_1 \cdots e_n$	pseudoscalar
$(\alpha, \beta, \gamma, \delta, \eta)$	memory, diffusion, product, parity, chirality
$\mathcal{N}[\cdot]$	stabilized normalization

3 Multivector-valued cellular automata as a data generator

Each site $i \in \{1, \ldots, N\}$ of a one-dimensional lattice with periodic boundary $X_{i+N}(t) = X_i(t)$ carries a state $X_i(t) \in \mathcal{G}$. The nearest-neighbor set is $\mathcal{N}_i = \{i-1, i+1\}$ and the local algebra-valued mean is

$$\overline{X}_i(t) = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} X_j(t) = \tfrac{1}{2}\big(X_{i-1}(t) + X_{i+1}(t)\big). \tag{10}$$

The local update is the normalized multivector map

$$X_i(t{+}1) = \mathcal{N}\big[\, \alpha X_i + \beta\, \overline{X}_i + \gamma\, \Phi_i + \delta\, \widehat{X}_i + \eta\, X_i I \,\big], \tag{11}$$

with $\mathcal{N}$ given by Eq. (9), involution $\widehat{X}_i$ from Eq. (7), pseudoscalar I from Eq. (6), and the geometric-product interaction

$$\Phi_i(t) = \big\langle X_i(t)\, \overline{X}_i(t) \big\rangle_K, \qquad \langle\, \cdot\, \rangle_K = \sum_{k \in K} \langle\, \cdot\, \rangle_k. \tag{12}$$

The full-grade coupling $K = \{0, 1, \ldots, n\}$, for which $\Phi_i = X_i \overline{X}_i$, is used throughout; grade-restricted choices such as $K = \{0\}$, $\{2\}$ or $\{0, 2\}$ isolate scalar, bivector or even interactions. The five real parameters

$$\theta = (\alpha, \beta, \gamma, \delta, \eta) \in \mathbb{R}^5 \tag{13}$$

control memory, diffusion, geometric-product nonlinearity, parity forcing through Eq. (7), and pseudoscalar chirality through Eq. (6), respectively. Random initial states are drawn coefficient-wise and normalized,

$$x_{i,A}(0) \sim \mathcal{N}(0, 1) \text{ i.i.d.}, \qquad X_i(0) \leftarrow \mathcal{N}[X_i(0)], \tag{14}$$

while structured single-grade initializations

$$X_i(0) = \mathcal{N}\big[\langle X_i(0) \rangle_{k_0}\big], \qquad k_0 \in \{0, 1, 2\}, \tag{15}$$

test whether the regimes obtained from Eq. (14) are genuine attractors rather than artefacts of initialization. Treating $(\mathcal{G}, \theta, \text{seed})$ as input and the spacetime array $\{X_i(t)\}_{i,t}$ as output, Eq. (11) is a controllable generator of algebra-labeled dynamical data, the substrate for the learning experiments below.

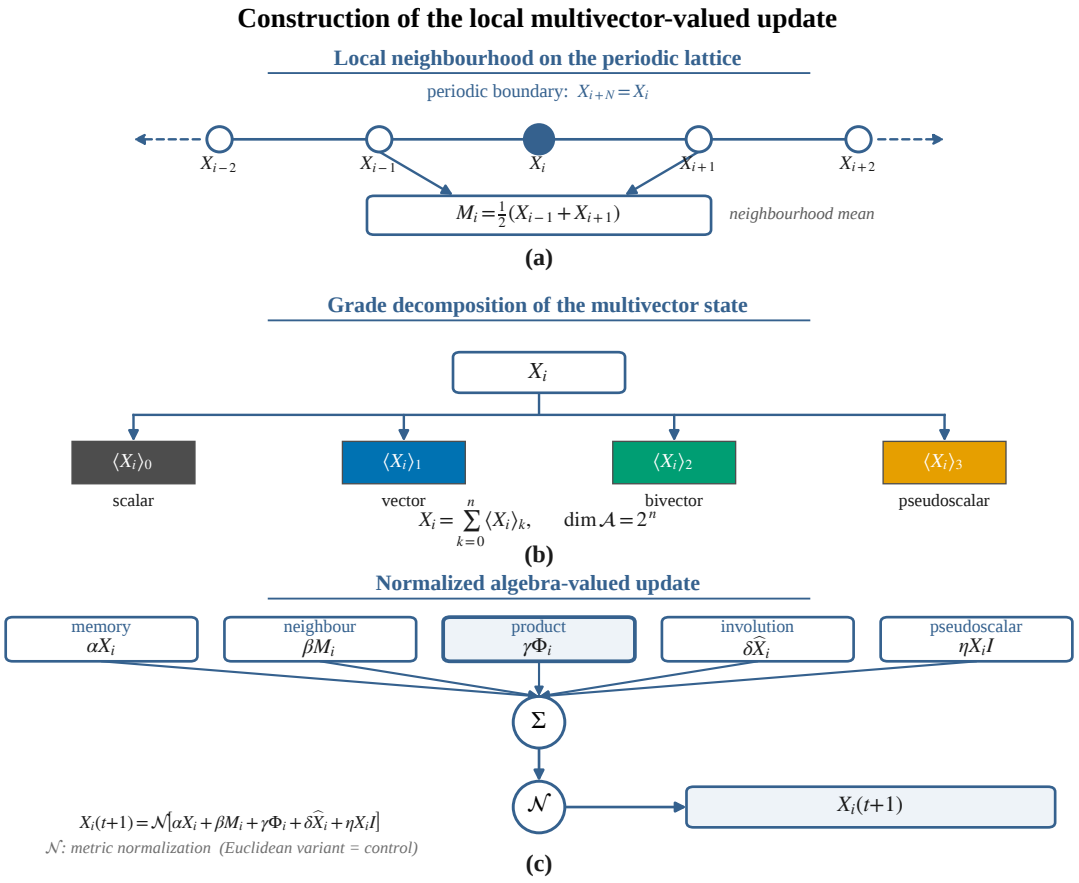


Figure 2: Schematic of a multivector-valued cellular automaton. Each site carries a state $X_i(t) \in \mathrm{Cl}(p,q)$ or $\Lambda(\mathbb{R}^3)$; the local update combines memory, diffusive neighbor coupling, a geometric/exterior-product nonlinearity, grade involution and a pseudoscalar rotation, followed by the stabilized normalization $\mathcal{N}[\cdot]$.

4 Grade-resolved observables and phase criteria

A scalar magnitude cannot resolve grade content, so we construct observables that are sensitive to the grading of Eq. (5). Unless noted, time averages $\langle\cdot\rangle_t$ run over the stationary window $t \in \{T_{\mathrm{tr}}+1,\dots,T\}$ with $T' = T - T_{\mathrm{tr}}$, and $\mathbf{x}_i(t) \in \mathbb{R}^{2^n}$ denotes the coefficient vector of $X_i(t)$.

The grade-k energy density and its normalized distribution over grades are

$$E_k(t) = \frac{1}{N}\sum_{i=1}^{N} \big\|\langle X_i(t)\rangle_k\big\|^2 = \frac{1}{N}\sum_{i}\sum_{|A|=k} x_{i,A}(t)^2, \tag{16}$$

$$p_k(t) = \frac{E_k(t)}{\sum_{\ell=0}^{n} E_\ell(t) + \epsilon}, \qquad \sum_{k=0}^{n} p_k(t) = 1. \tag{17}$$

From the grade distribution of Eq. (17) we form the normalized grade (Shannon) entropy and the grade-mixing, or Gini–Simpson, index

$$\bar{H}_{\mathrm{grade}}(t) = -\frac{1}{\log(n+1)}\sum_{k=0}^{n} p_k(t)\log p_k(t) \in [0,1], \tag{18}$$

$$C_{\mathrm{mix}}(t) = 1 - \sum_{k=0}^{n} p_k(t)^2 \in \Big[0, \tfrac{n}{n+1}\Big], \tag{19}$$

which vanish under single-grade dominance and are maximal under uniform grade mixing. The stabilized Clifford energy of Eq. (8) and its temporal fluctuation are

$$E_{\mathrm{Cl}}(t) = \frac{1}{N}\sum_i \big|\langle X_i(t)\widetilde{X_i(t)}\rangle_0\big|, \qquad \sigma_E = \mathrm{std}_{t>T_{\mathrm{tr}}} E_{\mathrm{Cl}}(t). \tag{20}$$

Spatial order is quantified by the coefficient-vector correlation function and the correlation length ξ obtained from its exponential decay,

$$C(r) = \Big\langle \frac{1}{N}\sum_i \mathbf{x}_i(t)\cdot\mathbf{x}_{i+r}(t)\Big\rangle_t, \qquad C(r)\sim e^{-r/\xi}, \tag{21}$$

$$\xi = -\Big(\tfrac{d}{dr}\log \widehat{C}(r)\Big)^{-1}, \qquad \widehat{C}(r) = C(r)/C(0), \tag{22}$$

while temporal order is measured by the normalized autocorrelation

$$A(\tau) = \frac{\big\langle \sum_i \mathbf{x}_i(t)\cdot\mathbf{x}_i(t+\tau)\big\rangle_t}{\big\langle \sum_i \|\mathbf{x}_i(t)\|^2\big\rangle_t}. \tag{23}$$

Sensitivity to initial conditions is estimated from two replicas $X^{(A)}$ and $X^{(B)}(0) = X^{(A)}(0)+\Delta$ with $\|\Delta\| \ll 1$, through the mean Clifford distance and the finite-time (effective Lyapunov) exponent

$$D(t) = \frac{1}{N}\sum_i \big\|X_i^{(A)}(t) - X_i^{(B)}(t)\big\|_\epsilon, \tag{24}$$

$$\lambda_{\mathrm{eff}} = \frac{1}{t_2-t_1}\log\frac{D(t_2)}{D(t_1)}, \qquad (t_1,t_2) = (5,50). \tag{25}$$

Spacetime complexity is captured by the two-dimensional discrete Fourier transform of each grade image $S_k(i,t) = \|\langle X_i(t)\rangle_k\|$, its normalized power spectrum, and the grade-averaged spectral entropy

$$\widehat{S}_k(\omega,q) = \sum_{t,i} S_k(i,t)\, e^{-\mathrm{i}(\omega t+qi)}, \qquad P_k(\omega,q) = \frac{|\widehat{S}_k(\omega,q)|^2}{\sum_{\omega,q}|\widehat{S}_k(\omega,q)|^2}, \tag{26}$$

$$H_{\mathrm{spec}} = \frac{1}{(n{+}1)\log(NT')}\sum_{k=0}^{n}\Big(-\sum_{\omega,q} P_k(\omega,q)\log P_k(\omega,q)\Big). \tag{27}$$

Regularity is quantified by recurrence quantification analysis [25, 40]. With the global state vector $Z(t) = \big(\mathbf{x}_1(t),\ldots,\mathbf{x}_N(t)\big)$, threshold ε and Heaviside step Θ, the recurrence matrix and its descriptors are

$$R_{ab} = \Theta\big(\varepsilon - \|Z(a)-Z(b)\|\big), \qquad \mathrm{RR} = \frac{1}{T'^2}\sum_{a,b} R_{ab}, \tag{28}$$

$$\mathrm{DET} = \frac{\sum_{\ell\ge\ell_{\min}} \ell\, P(\ell)}{\sum_{\ell\ge1}\ell\, P(\ell)}, \quad \mathrm{LAM} = \frac{\sum_{v\ge v_{\min}} v\, P(v)}{\sum_{v\ge1} v\, P(v)}, \quad \mathrm{ENTR} = -\sum_\ell p(\ell)\log p(\ell), \tag{29}$$

where $P(\ell)$ and $P(v)$ are the histograms of diagonal and vertical line lengths.

Each parameter point is summarized by the feature vector

$$f(\theta) = \big(\langle \bar{H}_{\mathrm{grade}}\rangle, \sigma_{\bar{H}_{\mathrm{grade}}}, \langle C_{\mathrm{mix}}\rangle, \langle E_{\mathrm{Cl}}\rangle, \sigma_E, \lambda_{\mathrm{eff}}, H_{\mathrm{spec}}, \xi, \mathrm{RR}, \mathrm{DET}, \mathrm{LAM}, \mathrm{ENTR}, p_0,\ldots,p_n\big), \tag{30}$$

whose grade-component-excluded subset, all entries except the grade probabilities p_k, is used for the algebra-recovery experiments. Averaging over the M initial conditions, the label uncertainty at a parameter point follows from the empirical fractions $\pi_c(\theta)$ assigned to phase c,

$$U(\theta) = -\sum_c \pi_c(\theta)\log\pi_c(\theta), \tag{31}$$

Table 3: The exact phase-assignment rule. Rules are evaluated sequentially and the first satisfied condition returns the label; there is no tie-breaking ambiguity. One universal threshold set is used for every algebra. All quantities are the stationary averages of Table 4 and $p_{\max} = \max_k p_k$. Appendix A gives the implementation.

Order	Condition	Returned phase
1	$\lambda_{\text{eff}} < -0.02$ and $\sigma_E < 0.02$ and $A_\infty > 0.60$	Ordered
2	$\lambda_{\text{eff}} > 0.03$ and $\langle \bar{H}_{\text{grade}} \rangle > 0.70$	Chaotic
3	$p_{\max} > 0.65$ and $\lambda_{\text{eff}} \leq 0.03$	Constrained
4	$\lambda_{\text{eff}} \leq 0.02$ and ($A_{\text{osc}} > 0.25$ or DET > 0.85)	Periodic
5	$\lvert\lambda_{\text{eff}}\rvert \leq 0.05$ and ($\sigma_H > 0.03$ or LAM > 0.50)	Intermittent
6	fallback: $\lambda_{\text{eff}} > 0.02$	Chaotic
7	fallback: otherwise	Constrained

Table 4: Authoritative descriptor dictionary. The full vector $f(\theta)$ has $14 + (n + 1)$ entries and is padded to 19 for pooling across algebras ($n \leq 4$). "n-dep." marks a descriptor whose definition contains the number of generators explicitly. The last three columns state which task uses which subset: the phase surrogate uses all 19; the algebra and signature tasks use the grade-component-excluded set $\mathcal{D}_{14}$, and the stricter n-free set $\mathcal{D}_{11}$ is used for the controls of Table 7.

Descriptor	Eq.	n-dep.	in phase rule	phase task	$\mathcal{D}_{14}$	$\mathcal{D}_{11}$
$\langle \bar{H}_{\text{grade}} \rangle$	(18)	yes	yes	yes	yes	no
σ_H	(18)	yes	yes	yes	yes	no
C_{mix}	(19)	yes	no	yes	yes	no
E_{Cl}	(20)	no	no	yes	yes	yes
σ_E	(20)	no	yes	yes	yes	yes
λ_{eff}	(25)	no	yes	yes	yes	yes
H_{spec}	(27)	no	no	yes	yes	yes
ξ	(22)	no	no	yes	yes	yes
RR	(29)	no	no	yes	yes	yes
DET	(29)	no	yes	yes	yes	yes
LAM	(29)	no	yes	yes	yes	yes
ENTR	(29)	no	no	yes	yes	yes
A_∞	(23)	no	yes	yes	yes	yes
A_{osc}	(23)	no	yes	yes	yes	yes
$p_0, \ldots, p_n$	(17)	yes	yes	yes	no	no

and is largest near phase boundaries. The rule-based classification into ordered, periodic, chaotic, algebraically constrained and intermittent phases combines Eqs. (18), (19), (25), (27) and (29) through the thresholds of Table 3.

Different learning tasks in this paper see different subsets of these observables, and the distinction matters enough that we state it once, in full, rather than reconstructing it in each subsection. Table 4 is the single authoritative feature list of the paper. It records, for every descriptor, the equation that defines it, whether its definition contains the number of generators n, whether it participates in the phase-assignment rule, and which of the three feature sets it belongs to: the full vector used by the phase tasks, the grade-component-excluded set $\mathcal{D}_{14}$ used for the algebra and signature tasks, and the stricter n-free set $\mathcal{D}_{11}$ used for the dimension controls of Section 7.3. Two entries of the table carry the caveats that motivated it. The grade entropy and its dispersion are normalized by $\log(n + 1)$ and the grade-mixing index is bounded by $n/(n + 1)$, so all three encode the dimension of the algebra implicitly and cannot appear in an experiment that claims to be dimension-neutral; and DET, LAM, σ_E, A_∞, A_{osc}, λ_{eff}, $\langle \bar{H}_{\text{grade}} \rangle$, σ_H and the p_k all enter the phase-assignment rule, so they cannot be used to corroborate that rule independently.

The exact phase-assignment algorithm, with its thresholds, its evaluation order and its tie-breaking, is transcribed from the implementation in Appendix A; the qualitative descriptions in Table 3 are a summary of it, not its definition.

5 Simulation and learning protocol

Phase diagrams span $(\beta, \gamma) \in [-2, 2]^2$ at $\alpha = 1, \delta = \eta = 0$ on a 31×31 grid for all five algebras, with $M = 3$ random initial conditions per point, $N = 96$ sites, $T = 250$ steps, transient $T_{\text{tr}} = 130$ and a 70-step replica run for λ_{eff}. Secondary pseudoscalar (γ, η) and involution (γ, δ) sweeps for $\text{Cl}(3, 0)$ use 25×25 grids with $M =$

2. The scalar, complex and quaternionic baselines of Section 6.5 are run on exactly the same 31×31 grid with the same $M = 3$ seeds and the same horizons as the Clifford and Grassmann sweeps, so that the cross-algebra comparison is free of any resolution or sampling confound; the complex and quaternionic systems are realized as $\mathrm{Cl}(0,1)$ and $\mathrm{Cl}(0,2)$ under the identical rule. Observables are time-averaged over $t > T_{\mathrm{tr}}$ and the reported phase is the majority label over initial conditions. For learning we pool the per-trajectory descriptors; unsupervised analysis uses PCA, UMAP [26], spectral clustering and Gaussian-mixture models, supervised phase and invariant classification uses random forests [28], gradient-boosted trees, an RBF support-vector machine and a multilayer perceptron, and conjecture formulation uses genetic-programming symbolic regression [29] and shallow decision trees. All results are five-fold cross-validated.

Before learning, the descriptor matrix is standardized column-wise to zero mean and unit variance,

$$\tilde{f}_m = \frac{f_m - \mu_m}{\sigma_m}, \qquad \mu_m = \tfrac{1}{P}\sum_j f_m^{(j)}, \quad \sigma_m^2 = \tfrac{1}{P}\sum_j \big(f_m^{(j)} - \mu_m\big)^2, \tag{32}$$

over the P pooled trajectories. Linear structure is exposed by principal component analysis, which diagonalizes the empirical covariance $\Sigma = \frac{1}{P}\tilde{F}^{\top}\tilde{F}$ and projects onto its leading eigenvectors,

$$\Sigma\, u_a = \varrho_a\, u_a, \quad \varrho_1 \geq \varrho_2 \geq \cdots, \qquad z_a = \tilde{f}^{\top} u_a, \qquad \mathrm{EVR}_a = \varrho_a \Big/ \sum_b \varrho_b. \tag{33}$$

Cluster quality against the rule-based labels is measured by the silhouette coefficient and by the adjusted Rand and normalized mutual information indices,

$$\mathrm{sil}(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{34}$$

$$\mathrm{ARI} = \frac{\sum_{ij}\binom{n_{ij}}{2} - \big[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\big]/\binom{P}{2}}{\frac{1}{2}\big[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\big] - \big[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\big]/\binom{P}{2}}, \qquad \mathrm{NMI} = \frac{2\, I(U;V)}{H(U) + H(V)}, \tag{35}$$

where $a(i)$ and $b(i)$ are the mean intra- and nearest-cluster distances, n_{ij} is the label-cluster contingency table, and I, H denote mutual information and entropy. Random-forest splits minimize the Gini impurity, and the impurity-decrease importance of feature m aggregates its weighted reductions over all nodes ν that split on it,

$$G(\mathcal{S}) = 1 - \sum_c \rho_c^2, \qquad \mathrm{Imp}(m) = \sum_{\nu:\,\mathrm{split}(m)} \frac{|\mathcal{S}_\nu|}{|\mathcal{S}|}\, \Delta G_\nu, \tag{36}$$

with ρ_c the class fraction in node set $\mathcal{S}$. Finally, symbolic regression searches a space $\mathcal{P}$ of expression trees over $\{+, -, \times, \div, \sqrt{\ }, \log\}$ for the program that minimizes a parsimony-penalized error,

$$g^{\star} = \arg\min_{g \in \mathcal{P}} \frac{1}{|\mathcal{D}|} \sum_{(x,y)\in\mathcal{D}} \big(g(x) - y\big)^2 + \lambda_c\, |g|, \tag{37}$$

where $|g|$ is the tree complexity and λ_c the parsimony coefficient.

The numerical settings that fix the lattice, the horizon, the sweep resolution and the perturbation-growth estimate are collected in Table 5, which applies to every run reported below unless stated otherwise.

Table 5: Complete simulation and validation settings. All sweeps share the update rule, the lattice, the horizons and the labeling rule; they differ only in the plane swept, the grid and the seed count.

Setting	Value
Lattice, boundary	$N = 96$ sites, periodic
Horizon, transient	$T = 250$, $T_{\mathrm{tr}} = 130$
Primary sweep (β, γ)	31×31 grid on $[-2, 2]^2$, $M = 3$ seeds, $\alpha = 1$, $\delta = \eta = 0$
Secondary sweeps (γ, η), (γ, δ)	25×25, $M = 2$, $\alpha = 1$, $\beta = 0$
Common-normalization control	21×21, $M = 2$, all five algebras
Baselines (scalar, $\mathbb{C}$, $\mathbb{H}$)	identical to the primary sweep
Replica perturbation	$\delta_0 = 10^{-6}$, 70-step replica run
Finite-time perturbation-growth window	$(t_1, t_2) = (5, 50)$
Recurrence threshold	$\varepsilon = 0.15 \max_{a,b} \|Z(t_a) - Z(t_b)\|$ (fixed fraction of the maximum distance; RR varies)
Recurrence line handling	$\ell_{\min} = v_{\min} = 2$; line of identity retained for RR and LAM, removed (Theiler window 1) for DET and ENTR
Normalization stabilizer	$\epsilon = 10^{-8}$
Cross-validation	five-fold GroupKFold; grouped by (β, γ) for the blind cross-algebra recovery (all algebras at a coordinate share a fold), by (algebra, β, γ) for phase classification and the $\mathcal{D}_{14}/\mathcal{D}_{11}$ comparison
Robustness design	blocked: five contiguous strips of the γ axis
Random forest	250 trees, balanced subsample weighting
Other models	see Appendix E

6 Dynamical phase structure of multivector automata

This section characterizes the automata as dynamical systems, before any learning is applied. We first establish, from representative trajectories, that grade-resolved fields support five qualitatively distinct regimes; we then map those regimes over the interaction plane for all five algebras and quantify the uncertainty of the labels; we show that grade entropy and the effective exponent act as coincident order parameters and corroborate the classification with independent spectral and recurrence diagnostics; we probe the pseudoscalar and grade-involution terms through dedicated sweeps; and we close with a matched-protocol comparison across algebraic systems that isolates the role of the metric signature.

6.1 Representative multivector dynamics

The five representative parameter points used in this subsection are not tuned by hand. Each is a cell of the sweep grid of Section 6.2 whose majority label is the phase in question and whose 3×3 neighborhood carries the same label, so that the trajectories and the phase diagrams agree by construction. Four points belong to $\mathrm{Cl}(3, 0)$; the ordered phase does not occur anywhere in the $\mathrm{Cl}(3, 0)$ plane, so its representative is taken from $\Lambda(\mathbb{R}^3)$.

Figure 3 shows the fields. The ordered regime is frozen and confines its energy to the scalar grade; the periodic regime is vertically striped and phase-locked with all four grades active; the constrained regime supports narrow, grade-selective filaments on a scalar background; the intermittent regime is laminar with localized bursts; and the chaotic regime is broadband across every grade. Because one absolute scale is used, the near-extinction of grades 1 to 3 in the ordered row is a statement about the dynamics and not an artefact of per-panel normalization.

Figure 4 follows the same runs at the level of the observables. Grade probabilities collapse onto a single grade in the ordered and constrained phases, equilibrate near the uniform value in the periodic and chaotic phases, and drift on long timescales in the intermittent phase. Both $\bar{H}_{\mathrm{grade}}$ and C_{mix} reach a stationary value well inside the transient window. Five trajectories cannot by themselves justify the choice of $T_{\mathrm{tr}} = 130$ out of $T = 250$; they are consistent with it, and the sweep-wide check is deferred to Appendix B. The stabilized Clifford energy stays within 10^{-8} of unity throughout (Appendix B), so the normalization is a numerical device and not a dynamical degree of freedom; it is reported as a stability diagnostic, not as an order parameter.

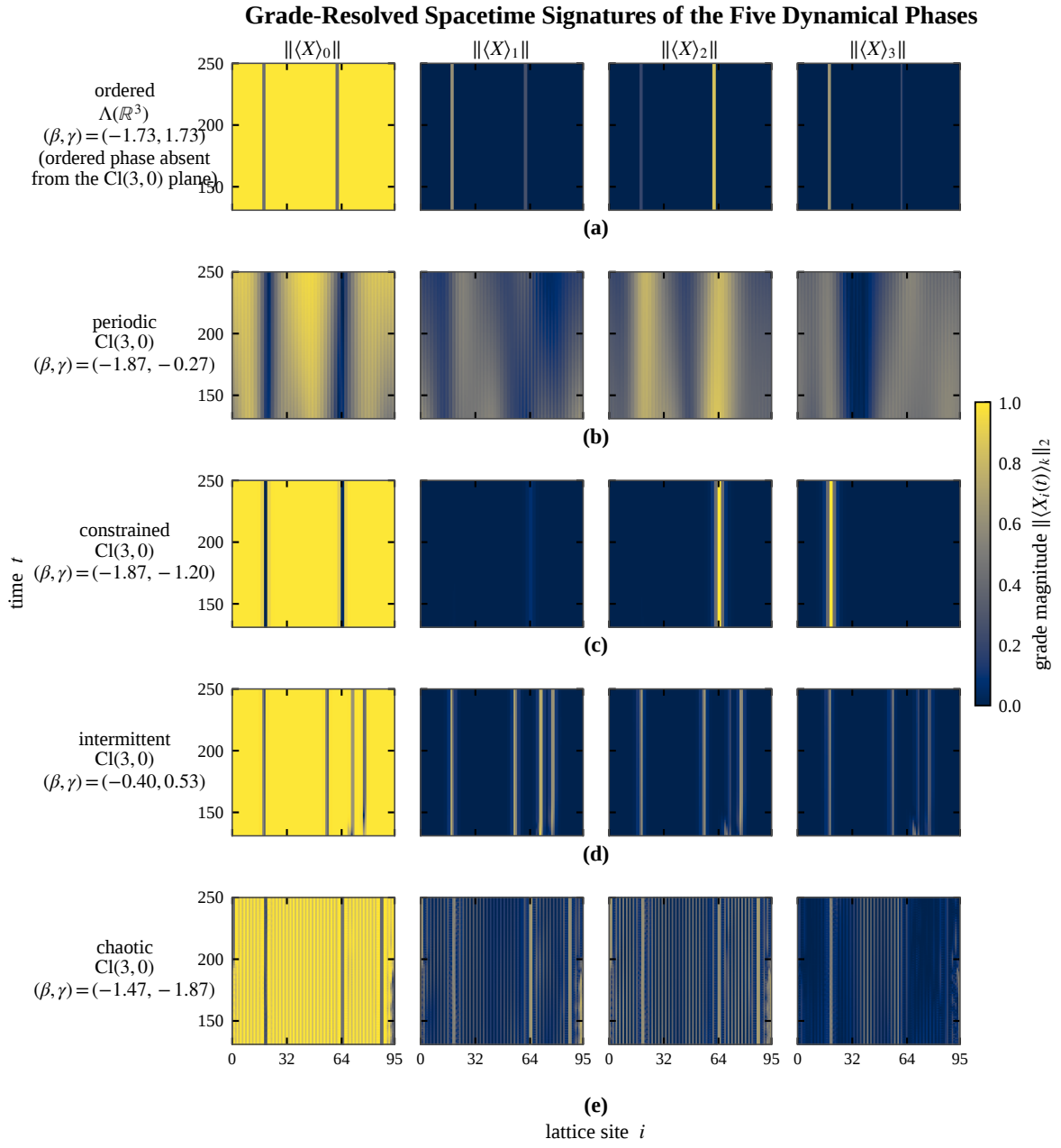


Figure 3: Grade-resolved spacetime fields for the five phases, on the stationary window $t = 131, \ldots, 250$ of the $N = 96$ lattice specified in Table 5. Rows are phases, with the generating algebra and the parameter pair (β, γ) at the left; columns are the grades $k = 0, \ldots, 3$. All twenty panels share one sequential color map and one absolute scale. The scale runs over $[0, 1]$, the 99th percentile of the pooled magnitudes coinciding with the maximum, so no sample is clipped (0.00% saturated) and intensities are directly comparable across rows; the fields are in absolute units and are not independently normalized. The ordered row is $\Lambda(\mathbb{R}^3)$; the other four rows are $\mathrm{Cl}(3, 0)$.

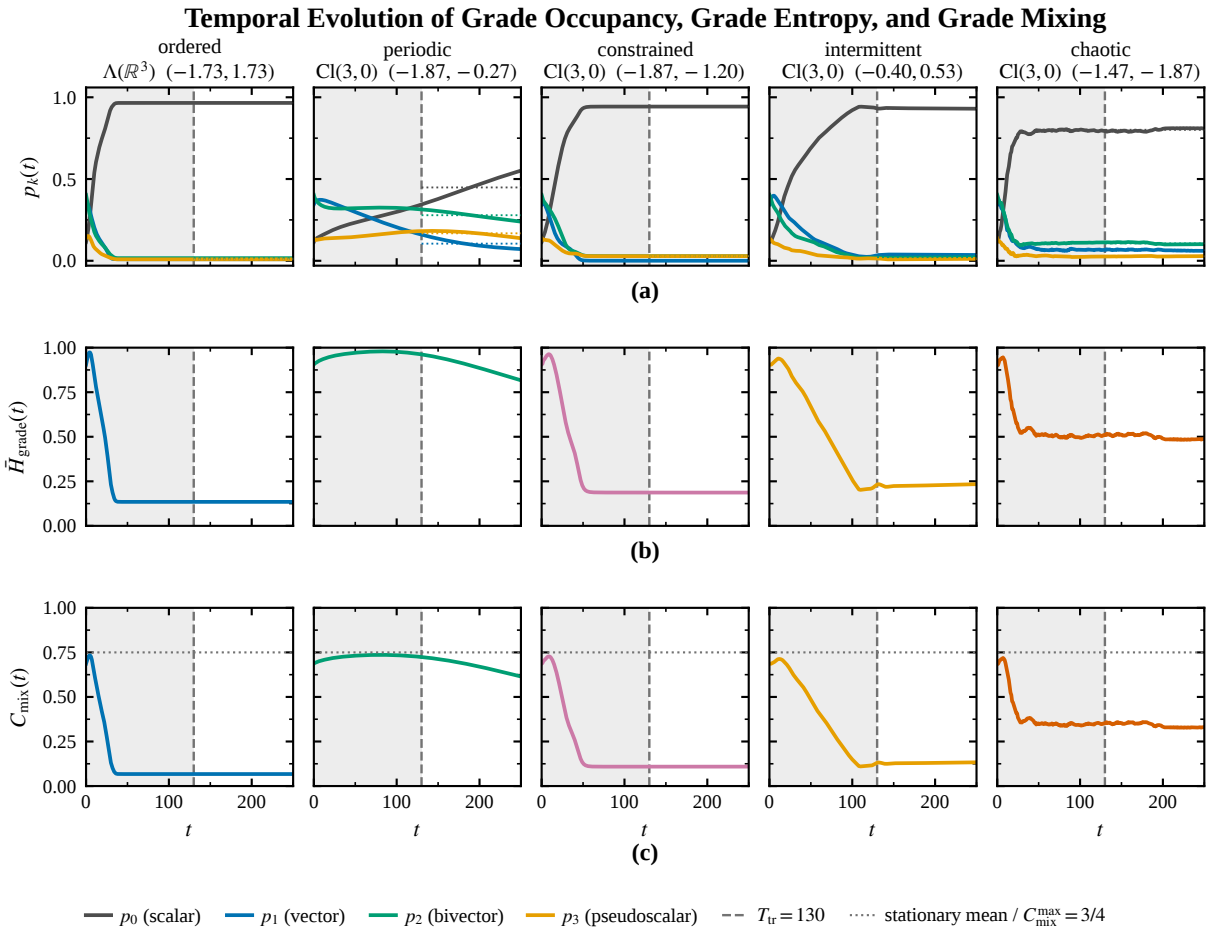


Figure 4: Trajectory-level diagnosis of the five phases (columns), for the runs of Fig. 3. (a) grade probabilities $p_k(t)$, Eq. (17), with dotted lines at the stationary means over $t > T_{\mathrm{tr}}$; (b) grade entropy $\bar{H}_{\mathrm{grade}}(t)$, Eq. (18); (c) grade-mixing index $C_{\mathrm{mix}}(t)$, Eq. (19), with a dotted reference line at the theoretical maximum $3/4$ for a four-grade algebra. Shading marks the discarded transient $t \leq T_{\mathrm{tr}} = 130$ and the dashed line is T_{tr}. Grade colors are fixed throughout the paper. Both order parameters settle well before the end of the transient, which justifies the choice of horizon.

Table 6: Stationary phase fractions over the (β, γ) plane. Entries are fractions of the 961 grid cells. A zero means that the phase is never the majority label, not that it was not evaluated.

Algebra	Ordered	Periodic	Constrained	Intermittent	Chaotic
$\mathrm{Cl}(2,0)$	0.000	0.458	0.127	0.037	0.378
$\mathrm{Cl}(3,0)$	0.000	0.083	0.509	0.019	0.389
$\mathrm{Cl}(1,2)$	0.000	0.003	0.061	0.002	0.933
$\mathrm{Cl}(1,3)$	0.000	0.001	0.084	0.000	0.915
$\Lambda(\mathbb{R}^3)$	0.012	0.166	0.658	0.009	0.154

6.2 Phase diagrams across five algebras

We now leave the single trajectory behind and map the whole interaction plane. The sweep covers $(\beta, \gamma) \in [-2, 2]^2$ at $\alpha = 1$, $\delta = \eta = 0$ on a 31×31 grid, that is 961 parameter points, with $M = 3$ independent initial conditions at each point, for each of the five algebras; the label of a cell is the majority label over its three seeds. Two questions are asked of this map, and they are answered by the two objects that follow. The first is quantitative: how much of each plane does each phase occupy? Table 6 answers it with the phase fractions, which underpin the subsequent cross-algebra comparisons, and which we therefore give before the maps themselves rather than after them. The second is geometric: how are the phases arranged? Figure 5 answers that, and the arrangement turns out to be as informative as the fractions, because the wedges, tongues and boundaries that appear are what the order parameters of Section 6.3 will later be shown to track.

With the fractions in hand, we turn to the geometry. Figure 5 shows the five planes side by side on identical axes and with the identical color code, together with a composition panel that renders the rows of Table 6 as stacked bars, so that the numerical and the geometric readings of the same sweep can be compared at a glance.

The planes are structurally different. $\mathrm{Cl}(2,0)$ is dominated by broad periodic wedges along the diagonals with chaotic corners; $\mathrm{Cl}(3,0)$ replaces most of that periodic territory with a constrained, grade-selective region; the indefinite algebras $\mathrm{Cl}(1,2)$ and $\mathrm{Cl}(1,3)$ are almost entirely chaotic, with chaotic fractions of 0.933 and 0.915; and the degenerate $\Lambda(\mathbb{R}^3)$ is the only algebra in which the ordered phase occurs at all. The organizing feature in every plane is the sign and magnitude of γ, which sets the strength of the geometric or exterior product, in interaction with the neighbor coupling β.

Figure 6 shows where the rule-based labeling is fragile. Interior cells are almost always unanimous, whereas boundary cells carry mean normalized disagreement between 0.13 and 0.54, an order of magnitude larger. With $M = 3$ seeds this quantity measures the consistency of the label across initial conditions and is not a statistical uncertainty; we therefore call it seed-wise label disagreement throughout.

The residual ambiguity of the classification is therefore confined to the transition regions, which is exactly where the moderate unsupervised agreement reported in Section 7.1 originates.

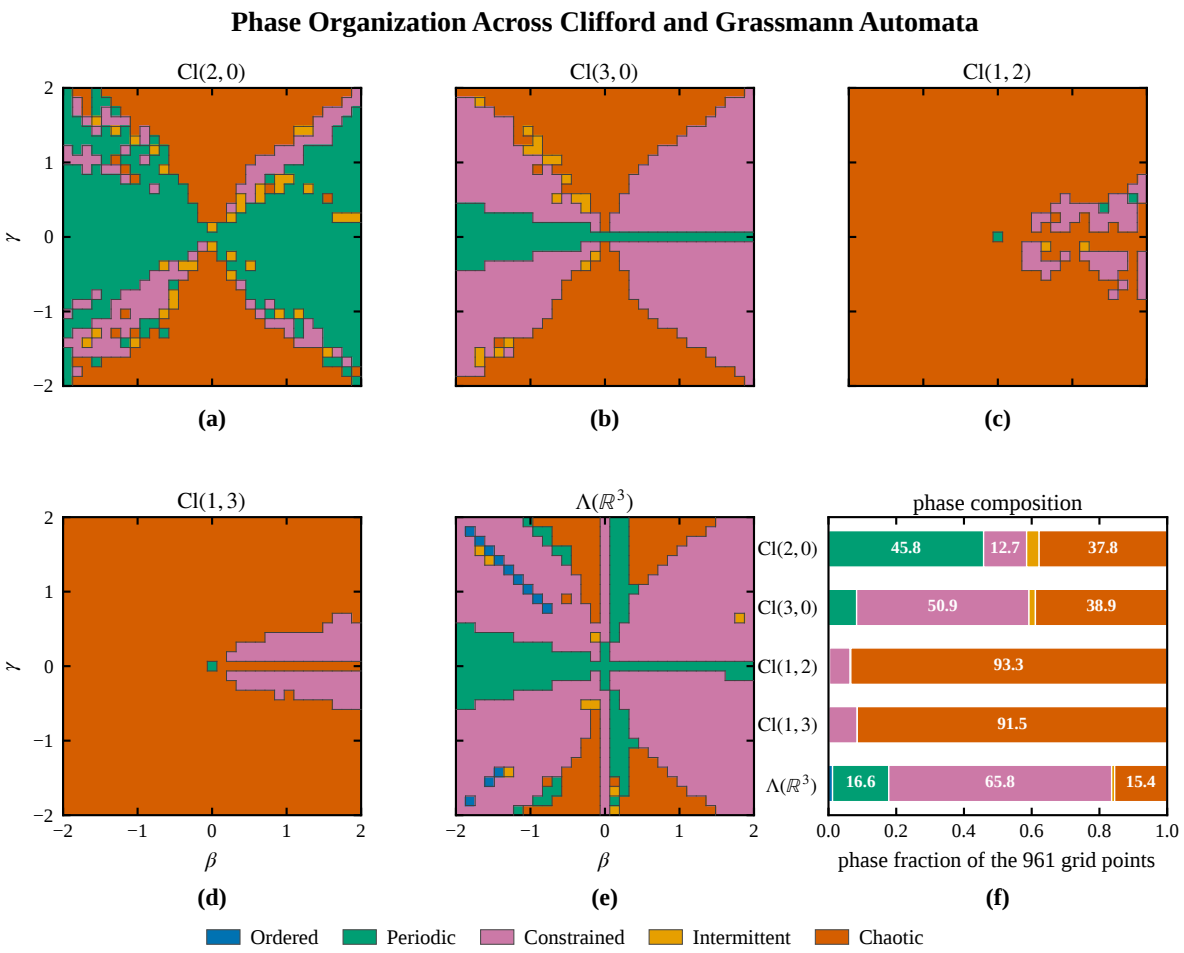


Figure 5: Phase diagrams in the (β, γ) plane for (a) $\mathrm{Cl}(2,0)$, (b) $\mathrm{Cl}(3,0)$, (c) $\mathrm{Cl}(1,2)$, (d) $\mathrm{Cl}(1,3)$ and (e) $\Lambda(\mathbb{R}^3)$, at $\alpha = 1$, $\delta = \eta = 0$, on a 31×31 grid with $M = 3$ seeds per cell. Color is the majority label over the three seeds, drawn without interpolation, with thin neutral lines separating cells of different label so that the map stays readable in greyscale. (f) phase composition, computed from the majority label of each of the 961 cells and not from pooling individual trajectories, in the order of the legend. The five-phase palette is fixed throughout the paper.

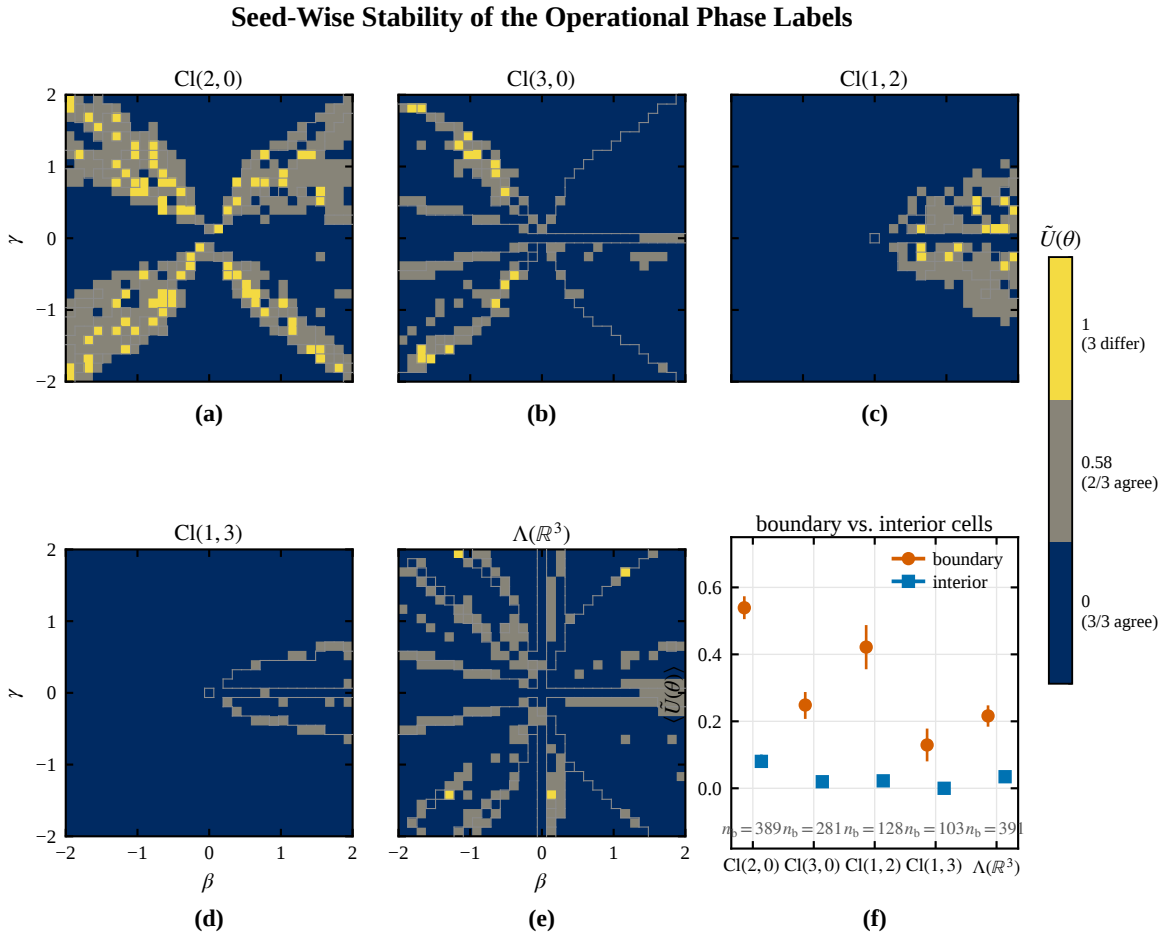


Figure 6: Normalized seed-wise label disagreement $\tilde{U}(\theta) = U(\theta)/\log M$, Eq. (31), with $M = 3$, so that $\tilde{U} \in [0, 1]$ and only three values are attainable: 0 when all three seeds agree, 0.58 when two of three agree and 1 when all three differ. The scale is discrete for that reason, not a continuously estimated field. (a)–(e) the five algebras, on the identical geometry and axes as Fig. 5, with the phase boundaries of that figure overlaid in thin grey. (f) mean $\tilde{U}$ over boundary cells, defined as cells with at least one differently labeled neighbor and counted as n_{b}, and over interior cells, with 95% bootstrap confidence intervals over cells. With only three seeds this is a measure of label consistency, not of statistical confidence.

6.3 Grade entropy and the effective exponent as coincident order parameters

The central dynamical hypothesis of this work is that a purely algebraic quantity, the stationary grade entropy, and a purely dynamical one, the finite-time perturbation-growth exponent, identify the same transition. Figure 7 tests this by placing the two fields directly above one another for all five algebras.

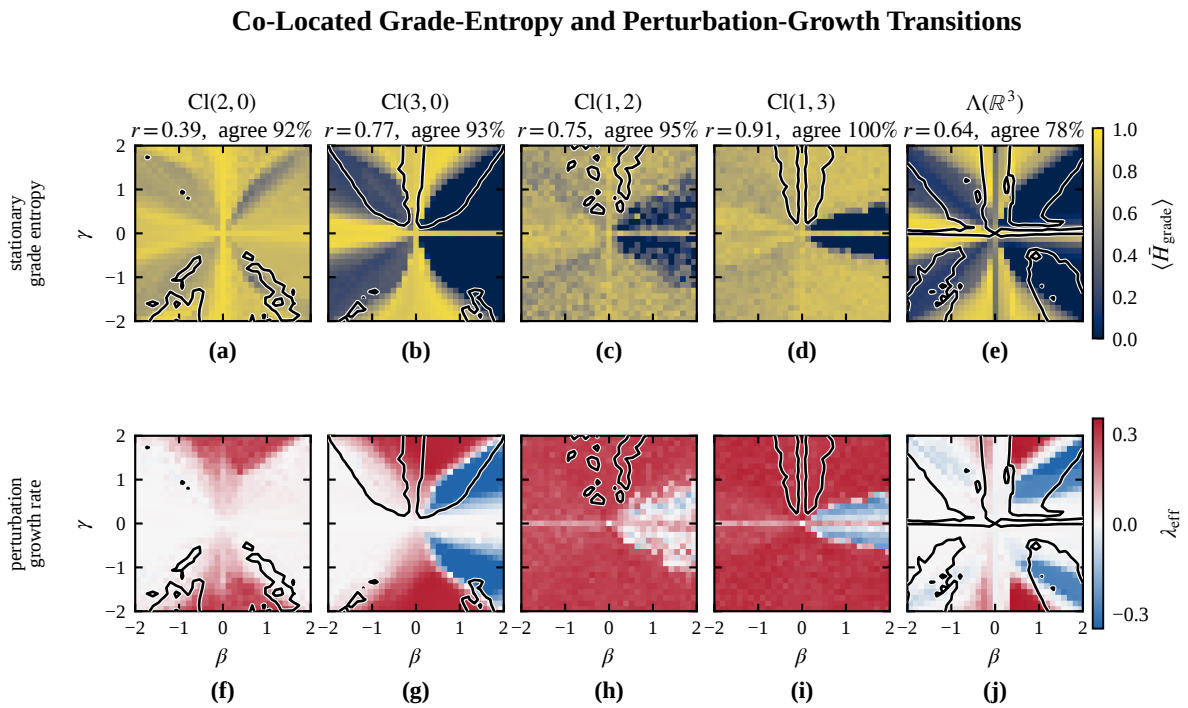


Figure 7: Co-location of the grade-entropy and perturbation-growth transitions. Top row (a)–(e): stationary grade entropy $\langle \bar{H}_{\mathrm{grade}} \rangle$, Eq. (18), on the fixed sequential map over $[0, 1]$, with the plane mean quoted above each panel. Bottom row (f)–(j): the finite-time exponent λ_{eff}, Eq. (25), fitted over $(t_1, t_2) = (5, 50)$, on the fixed diverging map, centered exactly at zero and clipped symmetrically at ± 0.35. The identical $\lambda_{\mathrm{eff}} = 0$ contour is drawn on both rows. Columns are the five algebras in the order used throughout. λ_{eff} is a finite-time effective exponent, not an asymptotic Lyapunov exponent.

The zero contour of λ_{eff} tracks the sharp gradient of the entropy field in every algebra: low-entropy grade-constrained regions lie inside the contour and high-entropy grade-mixing regions lie outside it. The visual impression is supported quantitatively. Over the 4805 cells the two fields correlate at Pearson $r = 0.67$, with per-algebra values of 0.39 for $\mathrm{Cl}(2, 0)$, 0.77 for $\mathrm{Cl}(3, 0)$, 0.75 for $\mathrm{Cl}(1, 2)$, 0.91 for $\mathrm{Cl}(1, 3)$ and 0.64 for $\Lambda(\mathbb{R}^3)$ (bootstrap 95% intervals over cells are narrower than ± 0.04 in every case), and a single threshold on the entropy field reproduces the sign of λ_{eff} for 91% of the cells pooled, ranging from 78% for $\Lambda(\mathbb{R}^3)$ to 99.9% for $\mathrm{Cl}(1, 3)$. We therefore describe the two quantities as co-located transition indicators. The correlations are strong but not perfect, and no finite-size scaling or critical-exponent analysis has been performed, so we do not claim coincident critical behavior in the statistical-mechanical sense.

A clarification is required here, because an earlier formulation of this argument overstated the independence of these diagnostics. The phase-assignment rule of Appendix A uses DET and LAM explicitly, so those two quantities cannot corroborate the labeling independently; they are shown because they display the geometry of the rule. The spectral entropy H_{spec}, the recurrence rate RR, the recurrence entropy ENTR and the correlation length ξ enter no part of the rule, and it is those quantities that provide independent corroboration: H_{spec} in Fig. 8(a) reproduces the wedge geometry of Fig. 5(b) without having contributed to it. A second caveat concerns DET itself. At the threshold used, DET saturates at or near 1.00 for four of the five representative runs, so in the present implementation DET does not discriminate between the phases; the discrimination is carried by LAM, which is 1.00 for the periodic and constrained runs and exactly 0.00 for the chaotic one, and by RR. The five recurrence matrices remain qualitatively distinct: uniformly recurrent for the frozen ordered state, a broad regular band for the periodic and constrained states, a narrower band for the intermittent state, and a sparse, fragmented structure with no vertical lines for the chaotic

state.

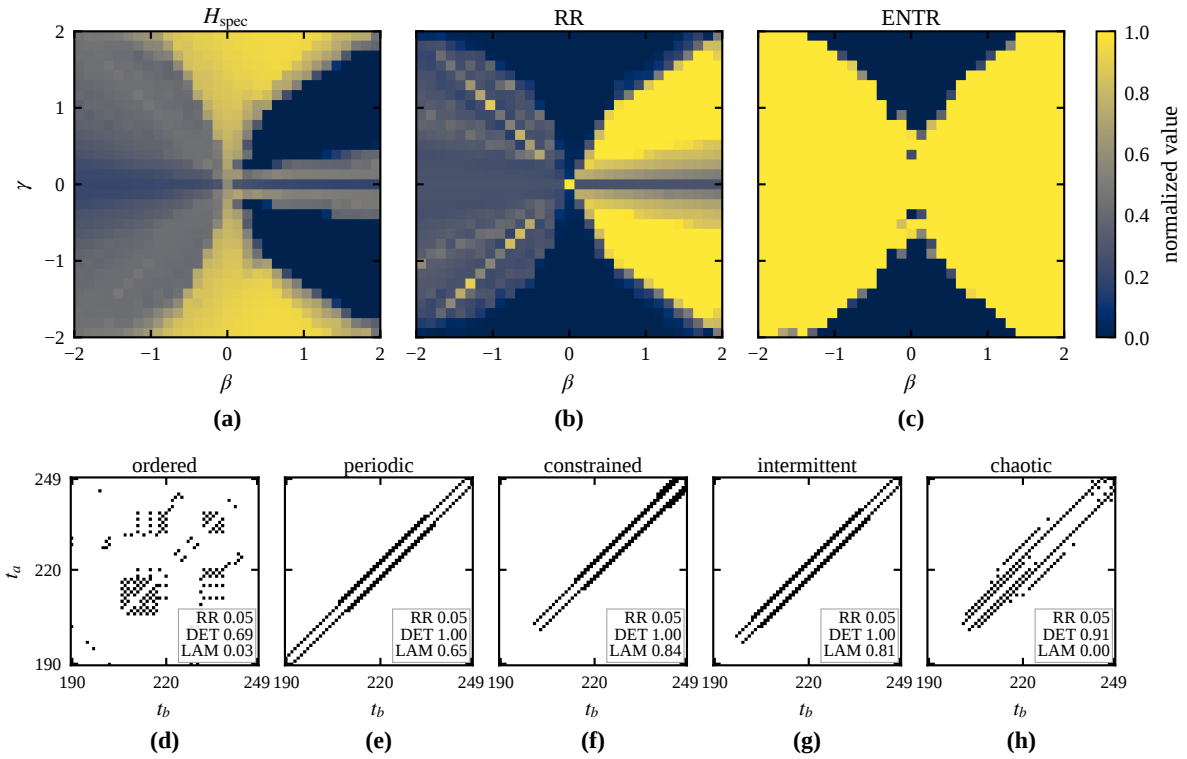


Figure 8: Spectral and recurrence diagnostics of the phase organization. (a)–(c) spacetime spectral entropy H_{spec}, recurrence rate RR and recurrence entropy ENTR over the (β, γ) plane of $\mathrm{Cl}(3,0)$; none of these three enters the phase-assignment rule of Table 3, so they are genuinely independent diagnostics. All three are normalized to $[0,1]$ and share one sequential map and one colorbar. (d)–(h) recurrence matrices $R_{ab} = \Theta(\varepsilon - \|Z(t_a) - Z(t_b)\|)$ for the five representative runs of Fig. 3 on the stationary window $190 \le t \le 249$, black for a recurrent pair and white otherwise. The threshold rule is the same in every panel and the same as the one used for the maps, $\varepsilon = 0.15 \max_{ab} \|Z(t_a) - Z(t_b)\|$, so the recurrence rate RR is free to vary and is reported with DET and LAM. The ordered state is frozen, so every pair of stationary states lies within ε and the matrix is uniformly recurrent; the chaotic run has $\mathrm{LAM} = 0.00$, that is, no laminar segments at all.

6.4 Pseudoscalar coupling and grade involution

The two remaining parameters act on grade structure rather than on amplitude: η couples the state to the pseudoscalar I, and δ adds the grade involution $\widehat{X}$, which flips the sign of the odd grades. We sweep (γ, η) at $\delta = 0$ and (γ, δ) at $\eta = 0$, both for $\mathrm{Cl}(3,0)$ with $\alpha = 1$ and $\beta = 0$, on 25×25 grids with $M = 2$ seeds.

Figure 9 shows that the pseudoscalar coupling is chirality-like: the (γ, η) phase map is nearly symmetric under $\eta \to -\eta$ and the grade entropy is almost flat, so η redistributes energy between grades without changing the order-chaos balance. The involution parameter behaves differently. Its sign selects which grade parity carries the energy, B_{eo} changing sign exactly at $\delta = 0$, and a constrained tongue opens for $\delta < 0$ that has no counterpart at $\delta > 0$. Grade parity is therefore a genuine control on the phase, whereas pseudoscalar chirality is not.

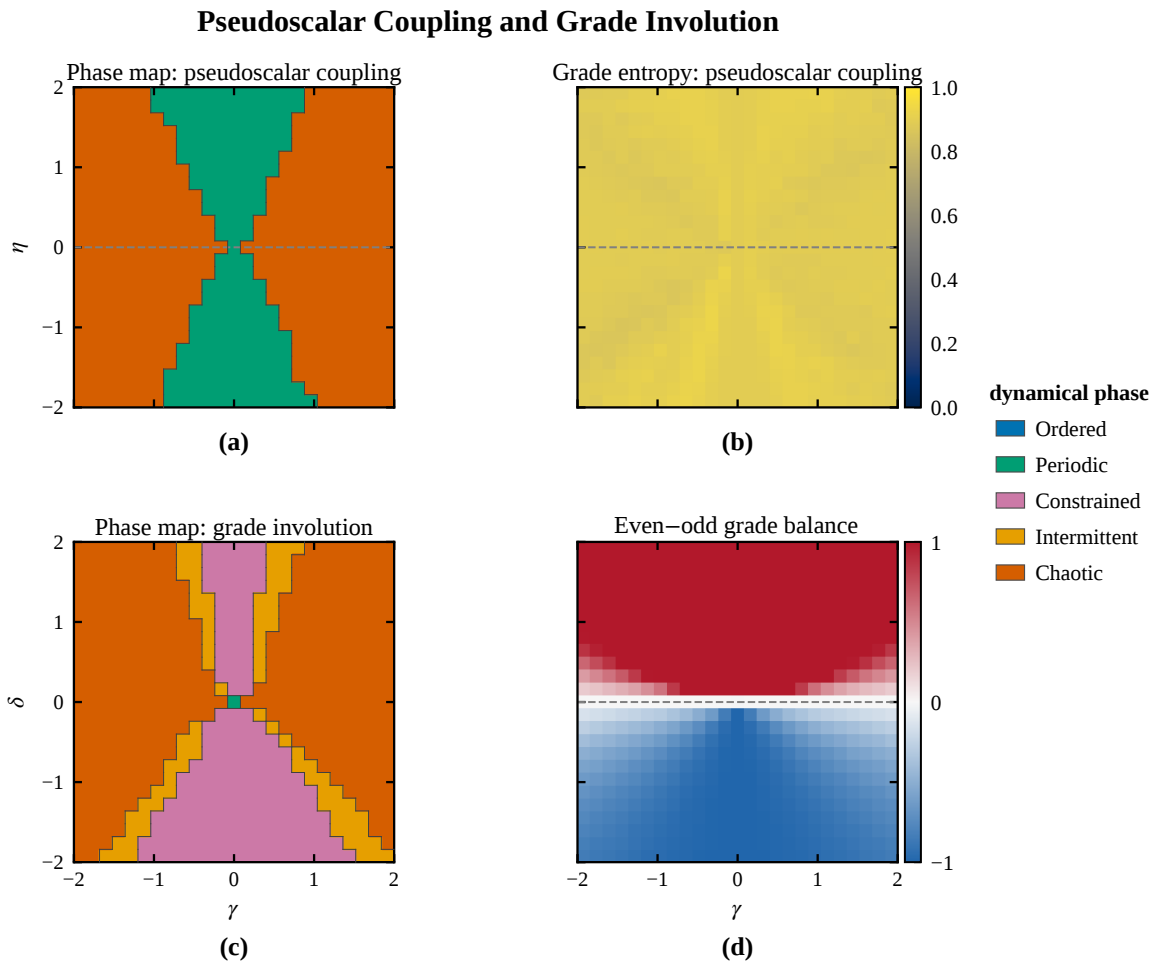


Figure 9: Pseudoscalar coupling and grade involution for $\mathrm{Cl}(3,0)$ on 25×25 grids with $M = 2$ seeds, at $\alpha = 1$ and $\beta = 0$. (a) phase map and (b) stationary grade entropy in the (γ, η) plane at $\delta = 0$; (c) phase map and (d) normalized even–odd grade-energy balance $B_{\mathrm{eo}} = (E_{\mathrm{even}} - E_{\mathrm{odd}})/(E_{\mathrm{even}} + E_{\mathrm{odd}} + \epsilon)$ in the (γ, δ) plane at $\eta = 0$, on the fixed diverging map centered at zero, so positive values denote even-grade dominance and negative values odd-grade dominance. Phases that do not occur are retained in the legend for consistency with Fig. 5.

6.5 Cross-algebra comparison and normalization controls

The comparison that follows is the one place in the paper where systems built on different algebras are placed side by side, so the protocol must be tightly controlled to avoid cross-algebra confounding. Every system, from the one-dimensional scalar automaton to the sixteen-dimensional $\mathrm{Cl}(1,3)$, is run under the identical update rule of Eq. (11), on the identical 31×31 grid over (β, γ), with the identical $M = 3$ seeds per cell, the identical horizons $T = 250$ and $T_{\mathrm{tr}} = 130$, the identical lattice $N = 96$ and the identical labeling rule of Appendix A. Nothing but the algebra changes. The scalar, complex and quaternionic baselines are realized as $D = 1$, $\mathrm{Cl}(0,1)$ and $\mathrm{Cl}(0,2)$ under the same reversion-based normalization used for every algebra. We emphasize that this is a $\mathrm{Cl}(0,1)$- and $\mathrm{Cl}(0,2)$-realized baseline under the common reversion protocol, and not the conventional complex or quaternion modulus: for $\mathrm{Cl}(0,1)$ the reversion form $\langle X\widetilde{X}\rangle_0$ evaluates to $a^2 - b^2$ for $X = a + be_1$, whereas the standard complex modulus is $a^2 + b^2$. The baselines therefore isolate the effect of the multiplication table under a fixed protocol; they are not claims about $\mathbb{C}$- or $\mathbb{H}$-valued automata under their own conventional norms. Any difference that survives this design is a difference between algebras and not between experiments.

Four indicators are compared: the number of phases that occupy more than 2% of the plane, the phase-diversity entropy, which is the Shannon entropy of the phase-fraction vector and therefore a single number summarizing a row of Table 6, the mean grade entropy, and the chaotic area fraction. The first two measure richness, the last two

measure where a system sits on the order-chaos axis. We plot the first two against the dimension of the algebra, because the hypothesis that dimension drives richness is the one a reader is most likely to arrive with, and it is the one the data reject first.

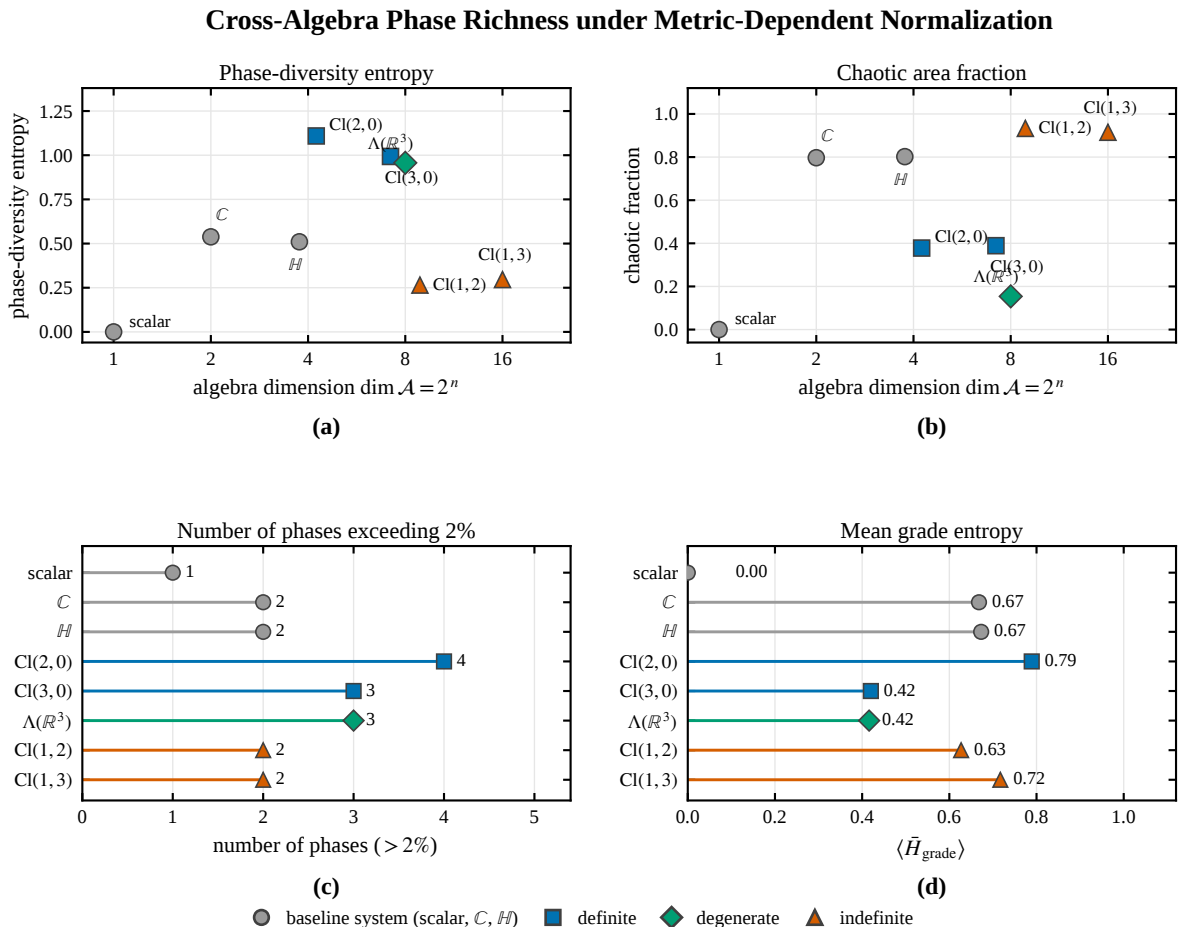


Figure 10: Cross-algebra phase richness under the metric-dependent normalization, on the matched 31×31, $M = 3$ protocol. (a) phase-diversity entropy, defined as the Shannon entropy of the phase-fraction vector of Table 6, and (b) chaotic fraction, both against $\dim \mathcal{A} = 2^n$; (c) number of phases occupying more than 2% of the plane and (d) mean grade entropy, by system. Marker shape and color both encode the signature class. At the fixed dimension 8 the diversity ranges from 0.99 and 0.96 for the definite and degenerate algebras down to 0.27 for the indefinite one, so phase diversity is not monotone in algebra dimension. The figure does not establish that the signature controls diversity; Fig. 11 tests that inference against the normalization confound.

Figure 10 shows that phase diversity is not a monotone function of algebra dimension. It rises from the scalar automaton, which collapses onto a single phase, through the complex and quaternionic automata, to the Euclidean Clifford and Grassmann automata, and then falls back for the indefinite algebras $\mathrm{Cl}(1,2)$ and $\mathrm{Cl}(1,3)$, whose planes are more than 91% chaotic and which support only two phases despite equal or larger dimension. The clearest reading is at fixed dimension. Among the three algebras of dimension eight, the definite $\mathrm{Cl}(3,0)$ reaches a diversity of 0.99 and the degenerate $\Lambda(\mathbb{R}^3)$ reaches 0.96, while the indefinite $\mathrm{Cl}(1,2)$ reaches only 0.27; and the sixteen-dimensional $\mathrm{Cl}(1,3)$, the largest algebra in the study, is poorer in phases than the four-dimensional $\mathrm{Cl}(2,0)$. Dimension is therefore not the controlling variable, and no amount of extra grades will make an indefinite automaton behave like a definite one under this rule. The natural next inference, that the metric signature is the controlling variable, is the one we tested rather than assumed, and it does not survive the control we report next.

Common-normalization control

The five algebras do not share a normalization. For a definite signature the reversion form $\langle X\widetilde{X}\rangle_0$ is positive definite and Eq. (9) divides by a genuine norm. For an indefinite signature the same form is indefinite, so it vanishes on the null cone and the stabilized quantity $\sqrt{|\langle X\widetilde{X}\rangle_0| + \varepsilon}$ can be small while the coefficient vector is large; the normalized state then acquires a large Euclidean magnitude. For the degenerate Grassmann algebra the form is identically the scalar part, and we already had to replace it by a Euclidean coefficient norm. Any comparison of phase diversity across signatures is therefore confounded with the normalization. To separate the two we re-ran the entire five-algebra sweep with one common Euclidean coefficient normalization, $\mathcal{N}[Y] = Y/\sqrt{\|Y\|_2^2 + \varepsilon}$, for every algebra, on a 21×21 grid with $M = 2$ seeds and the same horizons.

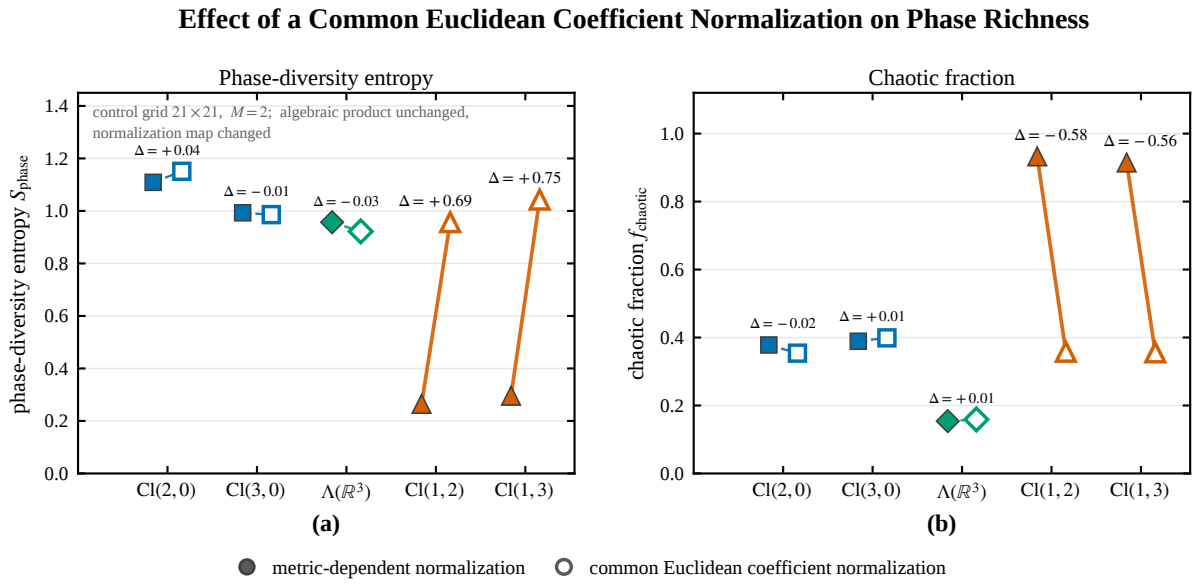


Figure 11: Common-normalization control. (a) phase-diversity entropy and (b) chaotic fraction for each algebra under the metric normalization used in the main text (filled markers) and under a common Euclidean coefficient normalization (open markers), on the matched control grid. The definite and degenerate algebras are almost unaffected. The two indefinite algebras move from a diversity of 0.27 and 0.30 to 0.95 and 1.04, and from a chaotic fraction above 0.91 to 0.36. The The control leaves the algebraic product untouched and changes only the normalization map, so the collapse of the indefinite algebras in the main text is a property of the normalization and not of the geometric product itself. The result is consistent with destabilization near the pseudo-norm null cone, a mechanism made explicit in Appendix B, Fig. 18.

The result, Fig. 11 and Table 10, is unambiguous and it corrects an inference we would otherwise have drawn. Under the common normalization the definite algebras are essentially unchanged ($\mathrm{Cl}(2,0)$: diversity $1.11 \to 1.15$; $\mathrm{Cl}(3,0)$: $0.99 \to 0.99$) and so is the degenerate one ($\Lambda(\mathbb{R}^3)$: $0.96 \to 0.92$, chaotic fraction $0.15 \to 0.16$), whereas the two indefinite algebras cease to be chaotic almost everywhere: $\mathrm{Cl}(1,2)$ moves from diversity 0.27 and chaotic fraction 0.93 to 0.95 and 0.36, and $\mathrm{Cl}(1,3)$ from 0.30 and 0.92 to 1.04 and 0.36. In other words, once every algebra is normalized in the same way, all four multivector algebras support a comparable richness of phases, and the only system that remains dynamically distinctive is the nilpotent Grassmann algebra, which stays the most ordered of the five.

The correct statement is therefore narrower than the one the uncontrolled comparison suggests. Within the tested update rule, phase diversity is more strongly associated with the normalization convention, and specifically with the behavior of the stabilized pseudo-norm near the null cone, than with the metric signature of the product as such. The null-cone mechanism is made quantitative in Appendix B: the indefinite algebras' states lie essentially on the null cone (median $\nu \approx 0$, Eq. (40)), where the metric normalization amplifies them by up to $\varepsilon^{-1/2} = 10^4$ per step. We present this as strongly consistent with the control rather than as an established cause. What survives without qualification is the negative result: phase diversity is not monotone in algebra dimension under either normalization, and the scalar, complex and quaternionic baselines are poorer in phases than every multivector algebra under both.

7 Machine learning of phases, invariants and conjectures

Having established the phase structure by direct dynamical analysis, we ask what a learner can recover from the same data. We proceed from the least to the most supervised: unsupervised embeddings that reorganize the phases without labels; a supervised classifier whose feature attribution exposes which observables carry the phase information; the recovery of the generating algebra and its metric-signature invariant from grade-component-excluded dynamical descriptors, which is the experimental-mathematics core of the work; the predictability of the asymptotic phase from short transients; and, as the culmination, the symbolic and tree-structured conjectures distilled from the data.

7.1 Unsupervised recovery of the phase organization

Feeding only trajectory descriptors, never the phase rules, into unsupervised learning already reorganizes the trajectories along the phase partition. We stress at the outset what this can and cannot show. The clustering acts on the same descriptor space from which the rule-based labels were constructed, so agreement between clusters and labels is a partial recovery of a descriptor-defined organization, not an independent discovery of the phases.

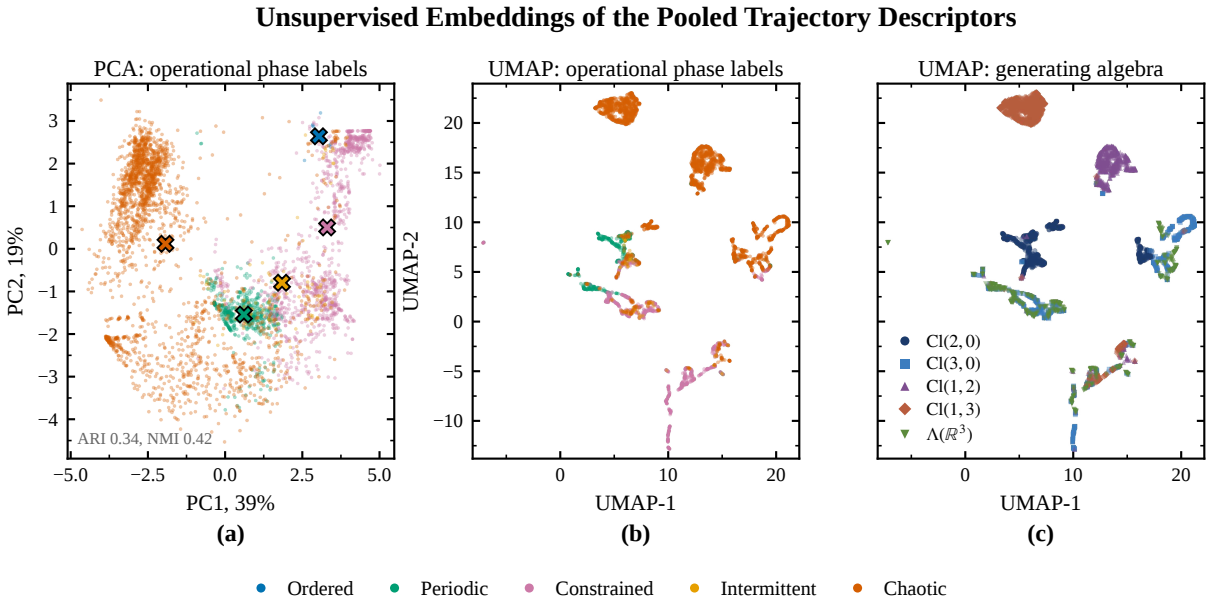


Figure 12: Unsupervised embedding of the pooled trajectory descriptors, Eq. (30), for $N_{\mathrm{traj}} = 14415$ trajectories. (a) the first two principal components, colored by rule-based phase, with crosses at the phase centroids; (b) UMAP embedding colored by phase; (c) the same UMAP embedding, identical coordinates and limits, colored by generating algebra, which is a different categorical variable and therefore carries a different palette and marker set. Points are rasterized; axes and text are vector. The organization emerges without using phase or algebra labels during the embedding; it is not, however, independent of the phase rule, because the descriptors are the same ones from which the operational labels were built. UMAP coordinates encode neighborhood organization, not metric separation. Gaussian-mixture agreement with the rule labels is ARI 0.343, NMI 0.423.

The first two principal components carry 58% of the variance and already separate chaotic from constrained and periodic trajectories; UMAP resolves islands that align both with the rule-based phases and, independently, with the generating algebra. Gaussian-mixture clustering attains an adjusted Rand index of 0.34 and a normalized mutual information of 0.42 against the rule labels. The moderate score is the expected consequence of the fuzzy boundaries quantified in Fig. 6, not a failure of discovery.

7.2 Machine-learned surrogate of the operational phase map

The supervised classifiers reproduce the rule-based labels almost perfectly. This is evidence that the operational labeling rule of Appendix A is well approximated by a nonlinear function of the descriptors, and not an independent physical validation of the five phases; the labels and the features live in the same space. What the classifier does tell us is which observables carry the discriminative information, and whether the near-perfect score survives a leakage-free validation design.

The classifier leans most heavily on the effective exponent, the scalar-grade probability, the asymptotic autocorrelation, the laminarity and the grade-mixing index, which are the quantities from which the phase criteria were built,

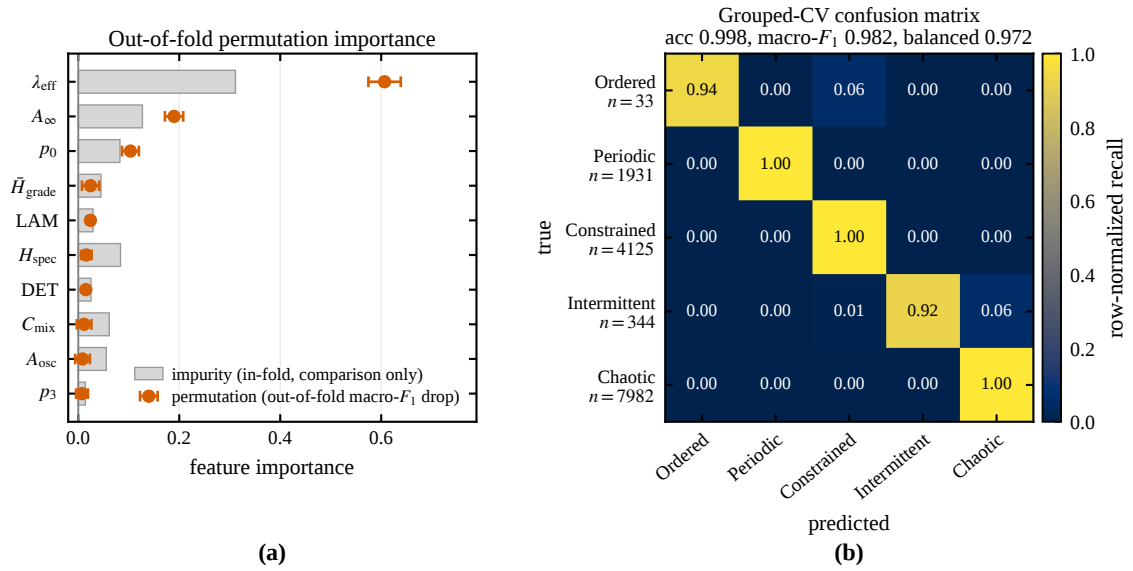


Figure 13: Interpretability and out-of-fold performance of the phase-rule surrogate. (a) the ten leading descriptors, ordered by permutation importance, computed inside each grouped validation fold as the drop in macro-averaged F_1 under feature shuffling and shown as points with the fold-to-fold standard deviation; the grey bars behind are the training impurity importance, Eq. (36), which favors high-cardinality features and is shown only for comparison. (b) row-normalized recall of the random forest over all five phases, with the class support n beside each row. The pooled set is strongly imbalanced, so macro-averaged F_1 is reported alongside accuracy.

Table 7: Controls for the recovery of algebraic structure. All rows except the first use five-fold cross-validation with folds grouped by $(\text{algebra}, \beta, \gamma)$, so that no parameter cell contributes to both training and validation; the first row repeats the five-way task under random trajectory-level splitting and is given only as a leakage comparison. $\mathcal{D}_{14}$ excludes the grade components; $\mathcal{D}_{11}$ additionally excludes every descriptor whose normalization contains n. Values are mean $\pm$ standard deviation over the five folds. The two bold rows carry the evidence against a dimension artefact.

Task	Classes	Descriptors	Validation	Chance	Accuracy	Macro-F_1
Algebra (leakage comparison)	5	$\mathcal{D}_{14}$	random	0.20	0.959 ± 0.001	0.959 ± 0.001
Algebra	5	$\mathcal{D}_{14}$	grouped	0.20	0.958 ± 0.002	0.958 ± 0.002
Algebra	5	$\mathcal{D}_{11}$	grouped	0.20	0.856 ± 0.006	0.855 ± 0.007
Signature class	3	$\mathcal{D}_{11}$	grouped	0.33	0.973 ± 0.002	0.972 ± 0.002
Same dimension, $\dim \mathcal{A} = 8$	3	$\mathcal{D}_{11}$	grouped	0.33	0.963 ± 0.004	0.963 ± 0.003

so the learned rule is interpretable in algebraic terms. The residual confusion is confined to the two rare classes that sit on phase boundaries: the intermittent class is recovered with recall 0.92, its errors leaking into the chaotic class, and the ordered class with recall 0.70, its errors leaking into the constrained class, which is the expected behavior given that an ordered state is a constrained state whose orbit has frozen. The class supports are 7982 chaotic, 4125 constrained, 1931 periodic, 344 intermittent and 33 ordered trajectories.

Because three seeds share each parameter cell and neighboring cells are strongly correlated, a random trajectory-level split can place near-duplicates on both sides of the partition. We therefore repeated the evaluation with folds grouped by $(\text{algebra}, \beta, \gamma)$, so that no cell contributes to both training and validation, and with folds blocked into contiguous strips of the γ axis, a design recommended for data with spatial dependence [41]. The scores are stable: accuracy 0.997 under random splitting, 0.998 under grouped folds and 0.994 under blocked folds, with macro-F_1 of 0.972, 0.982 and 0.902 respectively. All the numbers reported in this section, including the confusion matrix of Fig. 13, are pooled out-of-fold predictions of the grouped design: accuracy 0.998, macro-F_1 0.982 and balanced accuracy 0.972. The phase surrogate is therefore not an artefact of parameter-space leakage, although the larger fold

spread of the blocked design, macro-F_1 0.902 ± 0.097, shows that extrapolating to unseen regions of the plane is harder than interpolating within it. Every learning result reported below uses grouped folds.

Two checks separate genuine dynamical structure from circular rule reproduction. First, we predict the phase using only the descriptors that do *not* enter the assignment rule of Table 3, namely $\{H_{\mathrm{spec}}, \xi, \mathrm{RR}, \mathrm{ENTR}\}$. Under (β, γ) grouping this independent set still predicts the rule label with accuracy 0.882, above the majority-class baseline of 0.554, though its macro-F_1 of 0.631 shows that the rare classes are not well served by rule-absent descriptors alone. The phase organization therefore has support in descriptors that were not used to define it. Second, we perturb every numerical threshold of the rule by $\pm10\%$ and $\pm20\%$ and re-label the entire sweep. The labels are stable: the adjusted Rand index against the original labeling stays between 0.876 and 0.935, at most 7.4% of cells change label, and the chaotic fraction moves by less than 0.03. The five-phase partition is thus not an artefact of the particular thresholds chosen.

7.3 Recovering the generating algebra and its signature class from dynamics

This is the principal machine-learning contribution of the paper, and the one result here that is not an emulation of a rule we wrote ourselves. We ask whether a classifier can name the algebra that generated a trajectory from dynamical descriptors alone.

The descriptor set must be chosen with care, and our earlier description of it as *signature-blind* was too strong. Removing the grade probabilities p_k leaves the fourteen global descriptors of Table 4, but three of them still know the number of generators: the grade entropy and its dispersion are normalized by $\log(n+1)$, and the grade-mixing index is bounded by $n/(n+1)$. We therefore call that set *grade-component-excluded* and denote it $\mathcal{D}_{14}$, and we define a stricter set $\mathcal{D}_{11}$ that additionally removes $\bar{H}_{\mathrm{grade}}$, σ_H and C_{mix}, so that no retained descriptor contains n in its definition. The decisive control is then a task in which dimension carries no information at all: classify among $\mathrm{Cl}(3,0)$, $\mathrm{Cl}(1,2)$ and $\Lambda(\mathbb{R}^3)$, which all have $\dim \mathcal{A} = 8$.

Every matrix in Fig. 14 and every number in Table 7 comes from the same pooled out-of-fold predictions of one grouped five-fold cross-validation, so the supports, the confusion matrices and the summary metrics refer to exactly the same predictions. The matrices are near-diagonal, and the controls show that this is not an artefact of either leakage or dimension. Grouping the folds by parameter cell costs almost nothing (algebra $0.959 \to 0.958$), so the classifier is not memorizing cells. Removing the three descriptors whose normalization contains n does cost something on the five-way task ($0.958 \to 0.856$), which is exactly what one should expect: part of the five-way discrimination did rest on dimension-dependent scaling. But the signature task is untouched by that removal (0.973 with $\mathcal{D}_{14}$ against 0.973 with $\mathcal{D}_{11}$), and, decisively, on the same-dimension task, where $\mathrm{Cl}(3,0)$, $\mathrm{Cl}(1,2)$ and $\Lambda(\mathbb{R}^3)$ are indistinguishable by dimension and the learner sees only the eleven n-free descriptors, the accuracy is 0.963 and the macro-F_1 is 0.963 against a chance level of 0.333. Definite, indefinite and degenerate algebras of the same dimension leave distinguishable dynamical fingerprints, and, as the blind controls below show, they remain distinguishable when the metric is removed from the descriptors. In an inverse problem in the spirit of Kac, one can hear the signature of a Clifford algebra from the drum of its automaton, though, as we will see, the metric normalization does part of the drumming.

Is the recovery genuinely blind? Metric leakage and a transfer test

The controls above remove descriptors that contain the algebra dimension, but they do not yet remove descriptors that contain the Clifford metric. Three of the fourteen, the stabilized energy E_{Cl}, its dispersion σ_E, and the effective exponent λ_{eff}, are built on the reversion form $\langle X\widetilde{X}\rangle_0$: E_{Cl} is its site average, and λ_{eff} is fitted to a replica distance measured in that same pseudo-norm. A classifier given these could read the signature directly off the metric rather than infer it from the dynamics. We therefore define a genuinely blind set $\mathcal{D}_{\mathrm{Euc}} = \{H_{\mathrm{spec}}, \xi, \mathrm{RR}, \mathrm{DET}, \mathrm{LAM}, \mathrm{ENTR}, A_\infty, A_{\mathrm{osc}}\}$, eight descriptors that are computed entirely in Euclidean coefficient space and contain neither the dimension nor the metric, and we tighten the fold grouping to $g = (\beta, \gamma)$, so that all five algebras at one parameter coordinate share a fold and no coordinate can appear in both training and validation for any algebra.

The recovery survives, at a lower level. On $\mathcal{D}_{\mathrm{Euc}}$ with (β, γ) grouping the generating algebra is recovered with accuracy 0.722 (chance 0.20), the signature class with 0.872 (chance 0.33) and the same-dimension triple with 0.835 (chance 0.33). The drop from the metric-descriptor values (0.958, 0.973, 0.963) measures how much of the earlier performance rested on the metric; the residual, still far above chance and built on metric-free descriptors, is what the dynamics themselves carry. A stronger test isolates the signature mechanism from algebra identity: train a definite-versus-indefinite classifier on one algebra pair and test it on the other, so that the test algebras are never seen in

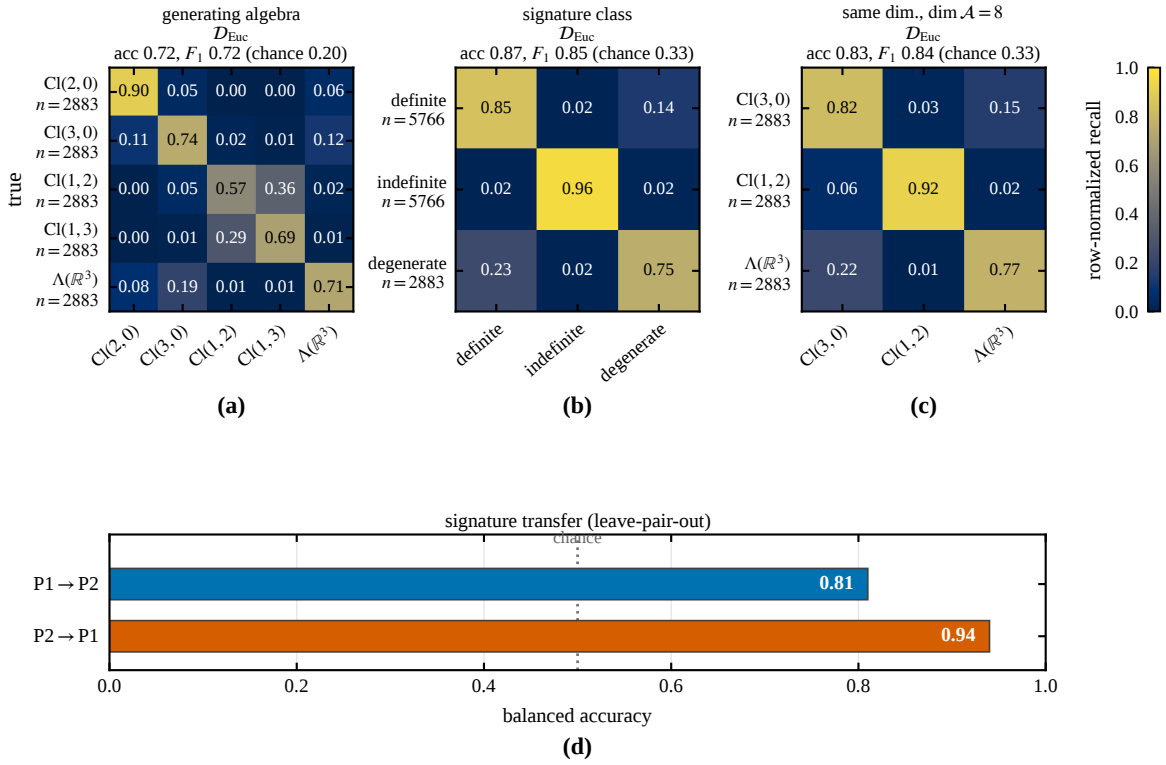


Figure 14: Blind recovery of algebraic structure from the genuinely blind descriptor set $\mathcal{D}_{\mathrm{Euc}}$, under folds grouped by (β, γ). Row-normalized recall for (a) the generating algebra among the five studied, (b) the metric-signature class and (c) the same-dimension task among the three $\dim \mathcal{A}{=}8$ algebras, with class supports n beside each row; rows are ground truth and columns predictions, on one shared sequential scale. (d) balanced accuracy of the definite-versus-indefinite transfer test, training on one algebra pair and testing on the other, where P1$= \{\mathrm{Cl}(2,0), \mathrm{Cl}(1,2)\}$ and P2$= \{\mathrm{Cl}(3,0), \mathrm{Cl}(1,3)\}$; the dotted line is the chance level 0.5. The metric-descriptor and same-dimension controls are given in Table 7.

training. Trained on $\{\mathrm{Cl}(2,0), \mathrm{Cl}(1,2)\}$ and tested on $\{\mathrm{Cl}(3,0), \mathrm{Cl}(1,3)\}$ it reaches balanced accuracy 0.81, and in the reverse direction 0.94, against a chance level of 0.5. The classifier has therefore learned a signature-related feature of the dynamics that transfers to algebras it was not trained on, not merely the identities of five labeled systems. Table 8 collects these controls.

Table 12 gives the per-class precision, recall, F_1 and support behind the two matrices of Fig. 14.

Table 8: Genuinely-blind recovery controls. $\mathcal{D}_{\mathrm{Euc}}$ is the eight-descriptor set computed in Euclidean coefficient space, containing neither the algebra dimension nor the Clifford metric; folds are grouped by $g = (\beta, \gamma)$ so that every algebra at a given coordinate shares a fold. The transfer rows train on one algebra pair and test on the other. Metric-descriptor values from Table 7 are repeated for comparison.

Task	Descriptors	Grouping	Chance	Accuracy (balanced)
Algebra (5 classes)	$\mathcal{D}_{14}$ metric	$(\mathrm{alg}, \beta, \gamma)$	0.20	0.958
Algebra (5 classes)	$\mathcal{D}_{\mathrm{Euc}}$ blind	(β, γ)	0.20	0.722 (0.722)
Signature (3 classes)	$\mathcal{D}_{11}$	$(\mathrm{alg}, \beta, \gamma)$	0.33	0.973
Signature (3 classes)	$\mathcal{D}_{\mathrm{Euc}}$ blind	(β, γ)	0.33	**0.872** (0.852)
Same dim. dim $\mathcal{A}$=8	$\mathcal{D}_{\mathrm{Euc}}$ blind	(β, γ)	0.33	0.835 (0.835)
Transfer $\{\mathrm{Cl}(2,0), \mathrm{Cl}(1,2)\} \to \{\mathrm{Cl}(3,0), \mathrm{Cl}(1,3)\}$	$\mathcal{D}_{\mathrm{Euc}}$	leave-pair-out	0.50	**0.81** (0.81)
Transfer $\{\mathrm{Cl}(3,0), \mathrm{Cl}(1,3)\} \to \{\mathrm{Cl}(2,0), \mathrm{Cl}(1,2)\}$	$\mathcal{D}_{\mathrm{Euc}}$	leave-pair-out	0.50	**0.94** (0.94)

7.4 Early-time prediction of the asymptotic phase

If the phase is decided early, a learner should be able to read it off a truncated transient. We train the three models on windows $\tau \in \{25, 50, 100, 200\}$ taken from the transient only, and predict the phase assigned from the full stationary window.

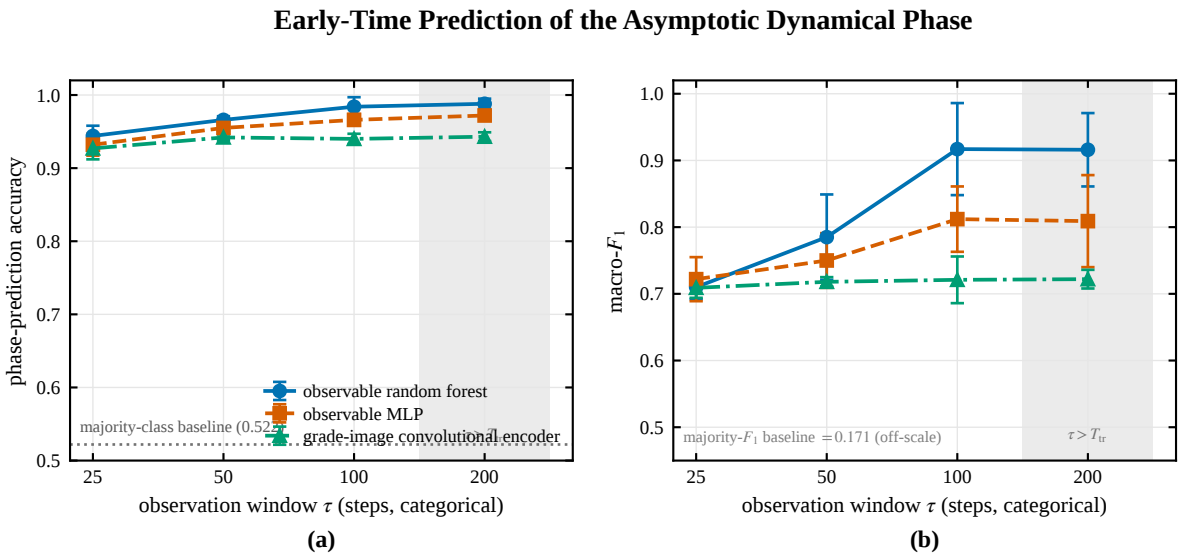


Figure 15: Early-time prediction of the asymptotic phase. Five-fold cross-validated accuracy against the observation window τ; error bars are the standard deviation across the five folds, and the four windows are plotted as categories rather than on a continuous axis. The upper panel is truncated to $[0.90, 1.00]$ to resolve the differences between the methods, and the lower panel carries the majority-class baseline 0.554 on the same horizontal axis. Accuracy rises steeply to $\tau = 100$ and then saturates.

The asymptotic phase is encoded in short transients, but the imbalance of the phase distribution means accuracy alone would flatter the result, so we report macro-F_1 beside it. Under grouped folds the observable random forest rises from accuracy 0.944 at $\tau = 25$ to 0.966 at $\tau = 50$, 0.984 at $\tau = 100$ and 0.988 at $\tau = 200$, against a majority-class baseline of 0.522; the macro-F_1 rises from 0.710 to 0.785 and then saturates at 0.917, against a baseline of 0.171. The observable multilayer perceptron and the grade-image convolutional encoder follow the same trend at a lower level. Beyond $\tau = 100$ the curves are flat rather than rising: the macro-F_1 at $\tau = 200$ lies within one fold standard deviation of its value at $\tau = 100$, so the correct statement is saturation, not continued improvement. Note that $\tau = 200$ extends past the transient cut $T_{\mathrm{tr}} = 130$ and therefore includes stationary steps, which is why we do not read the last point as evidence about the transient itself; the informative comparison is $\tau = 25 \to 100$, all of which lies strictly inside the transient.

7.5 Candidate symbolic relations and interpretable surrogates

Beyond classification, the data support the automated formulation of compact candidate relations. Genetic-programming symbolic regression, given only grade and dynamical predictors, proposes

$$\lambda_{\text{eff}} \;\approx\; a\, H_{\text{spec}}^{\,b}, \qquad a = 0.41\ [0.41, 0.42], \qquad b = 3.33\ [3.30, 3.37], \tag{38}$$

with the intervals the 95% bootstrap ranges over parameter cells; the program returned by the search is the fixed-exponent form $0.39\, H_{\text{spec}}^{3}$, which lies inside those intervals. We report this as a candidate empirical scaling relation, and three controls prevent us from claiming more. First, on grouped folds the one-parameter relation attains $R^2 = 0.66$, whereas an unconstrained linear fit in H_{spec} attains 0.82 and free cubic and quartic polynomials attain 0.82 and 0.83: the cubic law is parsimonious and interpretable, but it does not outperform an ordinary regression, so its value is conceptual rather than predictive, consistent with the caution urged by recent symbolic-regression benchmarking [42]. Second, the exponent is not universal across algebras: fitting the growth branch separately gives $b = 3.64$ for $\mathrm{Cl}(2,0)$, 3.43 for $\mathrm{Cl}(3,0)$, 2.82 for $\Lambda(\mathbb{R}^3)$, 2.24 for $\mathrm{Cl}(1,3)$ and 1.64 for $\mathrm{Cl}(1,2)$. Third, Eq. (38) is nonnegative, whereas 3474 of the 14415 trajectories have $\lambda_{\text{eff}} < 0$; the relation models the positive-growth branch, $H_{\text{spec}} \gtrsim 0.4$, and says nothing about the contracting regime. What survives these caveats is a genuine and non-obvious statement: on the growth branch, an information-theoretic quantity measured on the grade fields predicts a dynamical-systems quantity through a simple power law with an exponent near three.

Independently, a decision tree of depth three, fitted to all five phases with no class merging, distils a human-readable surrogate for the labeling rule of Appendix A,

$$\text{phase} = \begin{cases} \text{Ordered}, & \lambda_{\text{eff}} \le 0.03,\ C_{\text{mix}} \le 0.48,\ A_\infty > 0.60, \\ \text{Constrained}, & \lambda_{\text{eff}} \le 0.03,\ C_{\text{mix}} \le 0.48,\ A_\infty \le 0.60, \\ \text{Periodic}, & \lambda_{\text{eff}} \le 0.02,\ C_{\text{mix}} > 0.48, \\ \text{Intermittent}, & 0.02 < \lambda_{\text{eff}} \le 0.03,\ C_{\text{mix}} > 0.48, \\ \text{Intermittent}, & 0.03 < \lambda_{\text{eff}} \le 0.05,\ \langle \bar{H}_{\text{grade}} \rangle \le 0.70, \\ \text{Chaotic}, & 0.03 < \lambda_{\text{eff}} \le 0.05,\ \langle \bar{H}_{\text{grade}} \rangle > 0.70, \\ \text{Chaotic}, & \lambda_{\text{eff}} > 0.05, \end{cases} \tag{39}$$

built from Eqs. (19) and (25) together with the asymptotic autocorrelation A_∞, with a five-fold cross-validated accuracy of 0.983.

The two panels differ in status, and it is worth being explicit about the difference. The tree-structured surrogate approximates a labeling rule that we ourselves defined; it is an interpretable compression of our own criteria, not a discovered law of the dynamics. The leaf statistics keep the second statement honest. The ordered leaf holds only 65 of the 14415 trajectories at 51% purity, so while the surrogate does resolve all five classes, its ordered rule is its weakest part, exactly as the confusion matrix of Fig. 13 predicts. Equation (38), by contrast, relates two independently measured quantities, and it is the only statement in this paper that we would put forward as a conjecture, in the qualified sense set out above.

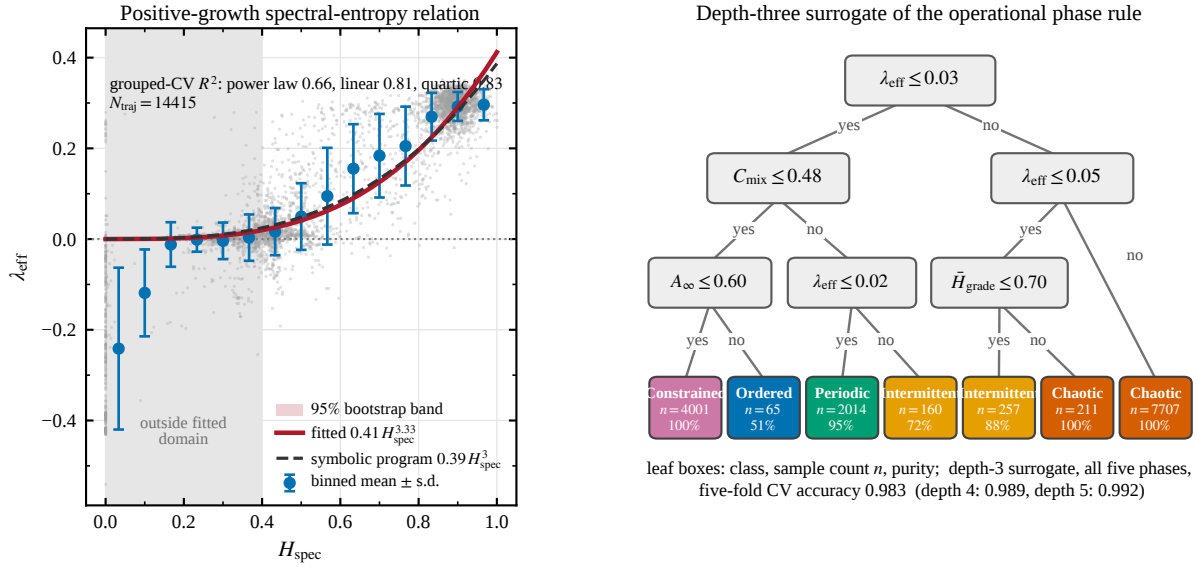


Figure 16: Automated conjecture formulation. (a) the symbolic relation of Eq. (38) against held-out trajectories, drawn as rasterized grey points, with binned means and their standard deviations; the relation models the positive-growth branch, so the negative λ_{eff} at low H_{spec} lie outside its intended range. (b) the depth-three surrogate of Eq. (39), with the sample count and the class purity at every leaf. All five phases are resolved and no class is merged away, but the ordered leaf is only 51% pure, so the surrogate separates ordered from constrained states only marginally. Deeper trees add little: 0.989 at depth four and 0.992 at depth five.

Table 9: Summary of the learning results, with the validation design stated for every row, because the designs differ. Grouped means five-fold cross-validation with folds grouped by $(\text{algebra}, \beta, \gamma)$; blocked means folds formed from contiguous strips of the γ axis. Classification metrics are accuracy and macro-averaged F_1 from pooled out-of-fold predictions. The three bold rows are the experimental-mathematics core.

Task / model	Descriptors	Validation	Metric	Result
Phase surrogate (random forest)	19	grouped	acc / macro-F_1	0.998 / 0.982
Phase surrogate (random forest)	19	blocked	acc / macro-F_1	0.994 / 0.902
Unsupervised phases (Gaussian mixture)	19	none	ARI / NMI	0.343 / 0.423
Algebra recovery (random forest)	$\mathcal{D}_{14}$	grouped	acc / macro-F_1	**0.958** / 0.958
Algebra recovery (random forest)	$\mathcal{D}_{11}$	grouped	acc / macro-F_1	0.856 / 0.855
Signature recovery (random forest)	$\mathcal{D}_{11}$	grouped	acc / macro-F_1	**0.973** / 0.972
Same-dimension recovery ($\dim \mathcal{A} = 8$)	$\mathcal{D}_{11}$	grouped	acc / macro-F_1	**0.963** / 0.963
Early prediction, τ=25 → 100 (RF)	14	grouped	accuracy	0.944 → 0.984
Early prediction, τ=25 → 100 (RF)	14	grouped	macro-F_1	0.710 → 0.917
Candidate relation, Eq. (38)	—	grouped	R^2 / RMSE	0.656 / 0.109
Phase surrogate (depth-3 tree, five classes)	$\mathcal{D}_{14}$	grouped	accuracy	0.983

Table 9 collects the learning results of this section in one place, with the validation design stated for each.

8 Discussion

The results above admit a coherent interpretation that we develop in this section. We first argue that the natural order parameter for these automata is grade content rather than any scalar amplitude, then explain how signature and degeneracy control the balance between order and chaos, next discuss what it means to recover an algebra and its

invariants from dynamics and how this places the work within the machine-learning-for-mathematics programme, and finally consider the automated conjectures, the connections to nonlinear time-series analysis, the limitations of the study and the outlook it opens.

8.1 Grade structure, order parameters and phase transitions

The central methodological message of this study is that the relevant order parameter for a multivector-valued automaton lives in the grading of the algebra rather than in any scalar amplitude. Because the update is norm-stabilized, the total Clifford energy is pinned near a constant, so a scalar magnitude is almost blind to the transitions that reorganize the dynamics. What changes across the phase boundaries is the distribution of energy over grades, and it is the grade entropy $\bar{H}_{\mathrm{grade}}$ and the grade-mixing index C_{mix} that expose this reorganization. In this sense the grade index behaves like an internal degree of freedom analogous to an isospin or a sublattice label, and a phase transition can occur purely in grade space while every scalar observable remains stationary. The strong feature-importance weight carried by $\bar{H}_{\mathrm{grade}}$, C_{mix} and the low-grade energies p_0, p_1, p_2 in the supervised classifier confirms quantitatively that these grade-resolved statistics, and not the scalar Clifford energy, are what separate the regimes. This observation is not specific to cellular automata: any normalized multivector flow, including Clifford-valued recurrent layers and geometric-algebra neural networks, is expected to exhibit dynamics whose order parameters are grade-resolved, which suggests that grade entropy is a natural diagnostic to monitor when training or analyzing such models.

8.2 Metric signature, degeneracy and phase diversity

The five algebras trace out a clear mechanistic picture of how the bilinear form governs the dynamics. For Euclidean signatures the scalar magnitude $\langle X\widetilde{X}\rangle_0$ is positive-definite, the normalization acts on a compact sphere, and the competition between the diffusive term and the geometric-product nonlinearity produces the richest multi-phase structure, with $\mathrm{Cl}(2,0)$ showing balanced periodic and chaotic domains and $\mathrm{Cl}(3,0)$ developing extended bivector and pseudoscalar constrained lobes. For indefinite signatures the magnitude becomes sign-indefinite, the level sets of the pseudo-norm are non-compact hyperboloids, and the stabilized normalization no longer confines trajectories to a bounded orbit; the result is the pervasive grade-mixing chaos seen for $\mathrm{Cl}(1,2)$ and $\mathrm{Cl}(1,3)$, which we therefore read as minimal discrete models of turbulent multivector fields in a relativistic-signature setting. The Grassmann algebra occupies the opposite extreme: because every generator is nilpotent, repeated exterior products contract the higher blades toward zero, the exterior product cannot sustain the feedback that drives chaos, and the plane is dominated by constrained and periodic order despite sharing its dimension with $\mathrm{Cl}(3,0)$. This last observation is robust: it is the one cross-algebra difference that survives the common-normalization control of Section 6.5 unchanged.

The behavior of the indefinite algebras is a different matter, and it is where we have had to correct ourselves. Under the metric normalization their planes are more than 91% chaotic, and it is tempting to read this as a statement about the signature of the geometric product. The control shows that it is not. Under a common Euclidean coefficient normalization the same algebras behave much like the definite ones, with a phase diversity of 0.95 and 1.04 against 1.15 and 0.99 for $\mathrm{Cl}(2,0)$ and $\mathrm{Cl}(3,0)$. What destabilizes them is the normalization step: the stabilized quantity $\sqrt{|\langle X\widetilde{X}\rangle_0| + \varepsilon}$ becomes small near the null cone, so states that approach it are amplified, and the amplification is what drives the map into chaos. We offer this null-cone mechanism as a hypothesis consistent with the control rather than as an established cause; distinguishing it from alternatives would require a dedicated study of the distribution of $\langle X\widetilde{X}\rangle_0$ along trajectories, which we have not performed.

Two conclusions survive without qualification. Phase diversity is not monotone in algebra dimension under either normalization, and all four multivector algebras are richer in phases than the scalar, complex and quaternionic baselines. The practical reading for the design of Clifford- and Grassmann-based neural architectures [18–20] is accordingly more specific than we first thought: the choice of (p, q, r) matters, but so does the choice of normalization, and for an indefinite signature the two cannot be varied independently.

8.3 What the learning results do and do not establish

It is worth separating five claims that are easily conflated. (i) The supervised phase classifier reaches out-of-fold accuracy 0.998; this establishes that our own labeling rule is well approximated by a smooth function of the descriptors, and nothing more, because rule and features share a space. (ii) The unsupervised embedding reorganizes the trajectories along the phase partition; because it acts on the same descriptors, this is a partial recovery of a descriptor-defined organization, not an independent discovery of the phases. (iii) The recovery of the generating algebra and of the signature class is a different kind of statement: the target is an algebraic fact about the data generator, it is nowhere used in constructing the descriptors, and it survives grouped folds, the removal of every n-dependent descriptor, and a same-dimension task. This is the result we would defend. (iv) Early-time prediction

shows that the phase is decided within the first fifth of the transient, which is a statement about the dynamics rather than about the learner. (v) The symbolic relation is a candidate empirical regularity on the growth branch, with an algebra-dependent exponent and no advantage over an ordinary regression in predictive terms; its value is that it is inspectable, not that it is accurate.

8.4 Learning algebraic structure and generating conjectures

The result most aligned with the aims of this collection is that a coarse summary of the dynamics is sufficient to reconstruct the generating algebra and its metric-signature class. With the grade-component-excluded set $\mathcal{D}_{14}$, which drops the explicit grade-component probabilities p_k but retains dimension-dependent aggregate descriptors such as the normalized grade entropy, its dispersion and the grade-mixing index, the generating algebra is recovered with 95.8% accuracy and the signature class with 97.3% accuracy under folds grouped by (algebra, β, γ). This is an inverse problem in the spirit of asking whether one can hear the shape of a drum: here one asks whether one can hear the signature of a Clifford algebra from the way it makes a lattice evolve. Crucially, the learnability is not a trivial consequence of dimension. When the three dimension-dependent descriptors are removed as well, leaving the strictly n-free set $\mathcal{D}_{11}$, the signature is still recovered at 97.3% and the algebra at 85.6%; and on the same-dimension control, where $\mathrm{Cl}(3,0)$, $\mathrm{Cl}(1,2)$ and $\Lambda(\mathbb{R}^3)$ share $\dim \mathcal{A} = 8$, the three signature classes remain separable at 96.3%. Under the strictest, genuinely blind protocol, the eight-descriptor set $\mathcal{D}_{\mathrm{Euc}}$ computed in Euclidean coefficient space, containing neither the dimension nor the metric, with folds grouped by (β, γ), the algebra is still recovered at 72.2% and the signature at 87.2%, and a definite-versus-indefinite classifier transfers across unseen algebra pairs at 81–94% balanced accuracy. The classifier must therefore exploit subtle differences in chaos onset, grade mixing, spectral content and recurrence structure that the algebra imprints on its dynamics. This places the work within the broader programme in which machine learning recovers mathematical invariants, whether topological, group-theoretic or combinatorial, from indirect observations [5, 6, 11, 12], and it suggests a concrete new target for that programme: the recovery of algebraic structure constants and signature invariants from the dynamics they generate.

The candidate relation $\lambda_{\mathrm{eff}} \approx 0.41\, H_{\mathrm{spec}}^{3.33}$ of Eq. (38) and the depth-three phase surrogate illustrate that these automata can serve as a substrate for machine-assisted conjecture generation. The cubic relation ties a dynamical-systems quantity, the finite-time perturbation growth rate, to an information-theoretic one, the spectral entropy of the grade fields, and while its held-out R^2 of 0.67 marks it as an empirical regularity rather than a proven identity, it is exactly the kind of compact, testable statement that a human researcher would then attempt to derive or refute. The tree-structured surrogate is even more transparent: the effective exponent decides at the root whether the trajectory is expansive, the grade-mixing index then separates the frozen regimes, ordered and constrained, from the recurrent ones, periodic and intermittent, and the grade entropy separates intermittency from full chaos on the expansive side. All five phases are resolved by three thresholds and no class is merged away. We are careful to call it a surrogate rather than a law: it approximates a labeling rule that we defined ourselves, and its ordered leaf is only 51% pure, whereas the cubic relation links two independently measured quantities and is therefore the one genuinely conjectural statement of the two. This mirrors the workflow that has proven productive in experimental mathematics, where symbolic and statistical learners such as the Ramanujan Machine, SciNet and modern symbolic-regression engines propose candidate relations that are subsequently examined analytically [6–8, 10, 30, 31].

8.5 Limitations and outlook

The observables used here, grade entropy, spectral entropy, correlation length and the recurrence descriptors, are close relatives of the tools developed in nonlinear time-series analysis and network science, including symbolic dynamics, visibility graphs and recurrence networks [25, 32]. These methods have extracted structure from empirical signals in domains as diverse as financial markets and vibration monitoring [33]; their effectiveness on real data suggests that richer graph- and manifold-based descriptors could sharpen the phase boundaries and the invariant-recovery accuracy reported here, and that the present framework could be extended beyond synthetic automata. Conversely, the multivector automata provide a controlled, ground-truth generator of complex signals on which such descriptors can be benchmarked, since the true phase, algebra and signature are known by construction. This two-way exchange, between algebra-driven synthetic dynamics and signal-processing descriptors, is a promising direction for both communities.

The limitations of the present study should be stated plainly. The lattice is one-dimensional with nearest-neighbor coupling, finite ($N = 96$) and run to a finite horizon ($T = 250$), so no finite-size scaling or critical-exponent analysis is possible, the reported exponents and phase fractions are effective finite-time, finite-size quantities, and we do not claim critical behavior. Only five multivector algebras are studied, and only two of the five parameters are swept at full resolution, so the extent to which the observed hierarchy persists in two dimensions, at larger lattices, and

for higher-dimensional or split signatures such as $\mathrm{Cl}(2,2)$ remains to be established. The phase labels are rule-based and therefore partly conventional, most notably near the boundaries, where the label uncertainty $U(\theta)$ is largest and the unsupervised agreement is correspondingly moderate; the thresholds of Appendix A are chosen by inspection rather than derived, and with $M = 3$ seeds the seed-wise disagreement is a consistency check rather than an uncertainty estimate, so a fully data-driven, unsupervised definition of the order parameter would be preferable. The finite-time effective Lyapunov exponent λ_{eff} is a finite-window perturbation-growth index obtained from two fitting endpoints; in cellular automata this damage-growth rate is known to differ from a tangent-space Lyapunov exponent [38, 39], and we use it only as a perturbation-propagation diagnostic. The cross-algebra comparison is normalization-dependent, as the control of Section 6.5 shows, and the null-cone mechanism we propose for the indefinite algebras remains a hypothesis. The candidate symbolic relation is an empirical regularity of modest predictive power, offered as a conjecture and not a theorem. Finally, although grouped and blocked folds show that the reported scores are not products of parameter-space leakage, the generalization of the algebra-recovery result beyond the five algebras and the single update rule studied here is untested.

Many of these limitations point directly to future work. Finite-size scaling of the grade entropy and the effective exponent would test whether the boundaries seen here sharpen into genuine transitions in the thermodynamic limit. Extending the construction to two-dimensional lattices and to larger algebras, including $\mathrm{Cl}(4,0)$, split signatures such as $\mathrm{Cl}(2,2)$, and higher exterior algebras $\Lambda(\mathbb{R}^n)$, would map the dependence of phase diversity on dimension and signature more finely and would test the generality of the algebra-identification result. On the learning side, replacing the fixed descriptors and shallow models used here with Clifford-group-equivariant sequence models would allow end-to-end discovery of order parameters that respect the algebra symmetries by construction, and would likely improve the unsupervised separation and the stability and interpretability of the symbolic relations. More ambitiously, coupling the conjecture-generation pipeline to formal proof assistants, in the direction of Voevodsky's univalent-foundations programme [9], could turn empirically discovered relations such as the chaos-onset law into candidates for formal verification. Finally, because normalized multivector flows are toy models for multivector field theories and for geometric-algebra neural networks, the diagnostics introduced here may transfer directly to the analysis of those systems.

9 Conclusion

We have recast Clifford- and Grassmann-valued cellular automata as controllable generators of algebra-labeled dynamical data and used machine learning both to chart their phase structure and to read algebraic content back out of that structure. On the dynamical side, we defined a family of normalized multivector automata with a single interpretable five-parameter local rule, introduced grade energy and entropy, a grade-mixing index, a stabilized Clifford energy and perturbation-growth descriptors as algebra-sensitive observables, and computed two-parameter phase diagrams for five algebras of increasing dimension, signature and degeneracy. These diagrams reveal ordered, periodic, chaotic grade-mixing, algebraically constrained and intermittent regimes, and they show that grade entropy and the effective Lyapunov exponent behave as coincident order parameters whose transition contours agree with independent spectral and recurrence diagnostics.

On the learning side, the experimental-mathematics core of the work is that the generating algebra and its metric-signature class are recoverable from grade-component-excluded dynamics with 95.8% and 97.3% accuracy, and that this survives the controls that matter: folds grouped by parameter cell, the removal of every descriptor whose normalization contains the number of generators, and a task restricted to three algebras of identical dimension, on which the accuracy is still 0.963 and the macro-F_1 0.963 against a chance level of 0.333. Supervised classifiers reproduce the operational phase labels with out-of-fold accuracy 0.998 and macro-F_1 0.982, which we read as evidence that the labeling rule is learnable rather than as an independent validation of the phases; the asymptotic phase is predictable from the first quarter of the transient; and symbolic regression proposes a candidate scaling relation $\lambda_{\mathrm{eff}} \approx 0.41\, H_{\mathrm{spec}}^{3.33}$ on the growth branch, whose exponent varies across algebras and which we therefore advance as a conjecture to be examined rather than as a law.

A matched-protocol comparison shows that phase diversity is not monotone in algebra dimension, and that all four multivector algebras are richer in phases than scalar, complex and quaternionic automata. A common-normalization control shows that the near-total chaos of the indefinite algebras is produced by the interaction of the indefinite bilinear form with the normalization step, not by the signature of the product alone: renormalized identically, the indefinite algebras recover a diversity comparable to the definite ones, and only the nilpotent Grassmann algebra remains dynamically distinctive. Taken together, these results position multivector-valued automata as a tractable and reproducible sandbox in which machine learning can recover algebraic structure from generated dy-

namics, a role we hope will prove useful at the interface of geometric algebra, nonlinear dynamics and experimental mathematics.

Data and code availability

The vectorized geometric/exterior-product engine, the observable suite, the sweep driver and the full learning and discovery pipeline (NumPy, SciPy, scikit-learn, UMAP, gplearn) accompany this article, with the parameter grids and random seeds needed to regenerate every figure and table.

Appendices

The appendices collect the material needed to reproduce the study without cluttering the main narrative. Appendix A gives the exact, ordered phase-assignment rule that defines the five phases; Appendix B reports the numerical-conservation diagnostic and the null-cone mechanism behind the normalization control of Section 6.5; Appendix C tabulates the full cross-algebra and common-normalization numbers; Appendix D lists the per-class machine-learning performance; and Appendix E records the exact software versions, hyperparameters and validation design.

A Exact phase-assignment algorithm

The phase label of a trajectory is assigned by the following deterministic procedure, transcribed from the implementation. The rules are evaluated strictly in the order given and the first match returns; there is therefore no tie-breaking ambiguity and no possibility of a cell satisfying two criteria. One universal threshold set is used for every algebra. Here $p_{\max} = \max_k p_k$ and all quantities are the stationary averages of Table 4.

1. if $\lambda_{\text{eff}} < -0.02$ **and** $\sigma_E < 0.02$ **and** $A_\infty > 0.60$: return ORDERED;
2. else if $\lambda_{\text{eff}} > 0.03$ **and** $\langle \bar{H}_{\text{grade}} \rangle > 0.70$: return CHAOTIC;
3. else if $p_{\max} > 0.65$ **and** $\lambda_{\text{eff}} \leq 0.03$: return CONSTRAINED;
4. else if $\lambda_{\text{eff}} \leq 0.02$ **and** ($A_{\text{osc}} > 0.25$ **or** $\text{DET} > 0.85$): return PERIODIC;
5. else if $|\lambda_{\text{eff}}| \leq 0.05$ **and** ($\sigma_H > 0.03$ **or** $\text{LAM} > 0.50$): return INTERMITTENT;
6. else: return CHAOTIC if $\lambda_{\text{eff}} > 0.02$, otherwise CONSTRAINED.

The label of a parameter cell is the majority label over the M seeds, and the seed-wise disagreement of Fig. 6 is the entropy of the seed label distribution. Two consequences of this rule are used in the main text. The descriptors σ_E, λ_{eff}, $\langle \bar{H}_{\text{grade}} \rangle$, σ_H, p_k, A_∞, A_{osc}, DET and LAM enter the rule and cannot corroborate it independently; the descriptors H_{spec}, ξ, RR and ENTR do not enter it and can. The thresholds are chosen by inspection of the representative trajectories and are not optimized; they are the principal element of arbitrariness in this study, which is why the learned surrogate of Eq. (39) is presented as a compression of the rule rather than as a discovery about the dynamics.

B Normalization diagnostic

The stabilized normalization of Eq. (9) is a numerical device, not a dynamical degree of freedom. Figure 17 reports the residual $|E_{\text{Cl}}(t) - 1|$ for the five representative runs of Fig. 3. It stays below 10^{-8} at all times and in every phase, so the stabilized Clifford energy is conserved to machine precision and cannot contribute to the phase discrimination.

We can test the null-cone mechanism proposed in Section 6.5 directly. Define the null-cone proximity

$$\nu(X) = \frac{|\langle X \widetilde{X} \rangle_0|}{\|x\|_2^2 + \varepsilon} \in [0, 1], \tag{40}$$

the ratio of the reversion form to the Euclidean coefficient norm. A definite signature has $\nu \equiv 1$; a state on the null cone of an indefinite form has $\nu \to 0$ while its coefficients remain finite. Under the metric normalization the amplification of a state relative to a definite state of the same Euclidean magnitude is $A_{\mathcal{N}}(X) = \|\mathcal{N}[X]\|_2 / \|x\|_2 = (\nu \|x\|_2^2 + \varepsilon)^{-1/2}$, which diverges as $\varepsilon^{-1/2} = 10^4$ when $\nu \to 0$. Figure 18 shows the distribution of ν over the stationary states of a representative chaotic run in each algebra. The definite algebras sit at $\nu = 1$; the indefinite algebras $\text{Cl}(1, 2)$ and $\text{Cl}(1, 3)$ have median ν of 0.003 and 0.000, that is, their states lie essentially on the null cone, so the normalization amplifies them by orders of magnitude at every step; the degenerate Grassmann algebra is intermediate. This is the mechanism behind the collapse of the indefinite algebras under the metric normalization and its disappearance under the common Euclidean normalization (Fig. 11); we present it as strongly consistent with the control rather than as a proof.

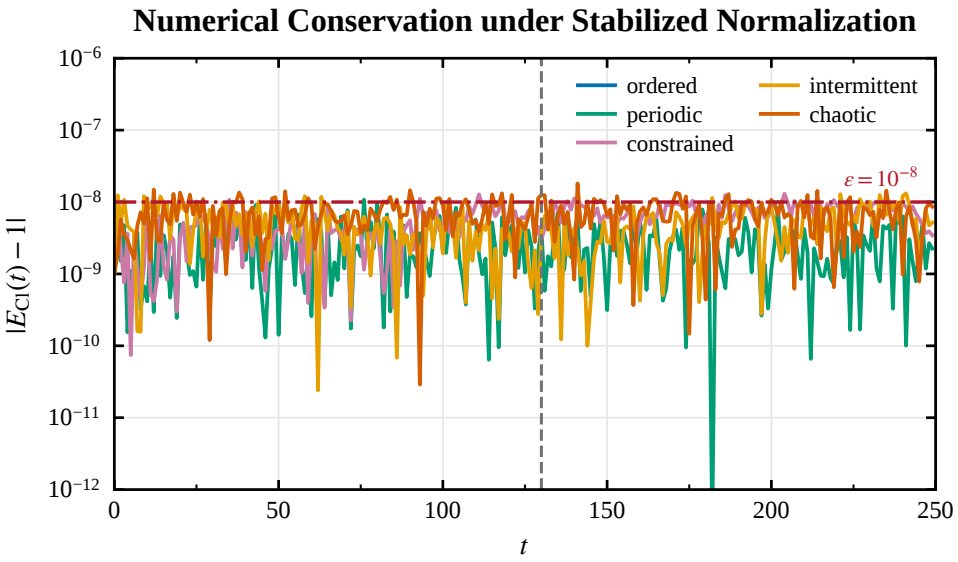


Figure 17: Normalization residual $|E_{\mathrm{Cl}}(t) - 1|$ on a logarithmic scale for the five representative runs, with the dashed line at $T_{\mathrm{tr}} = 130$. The residual never exceeds 10^{-8}.

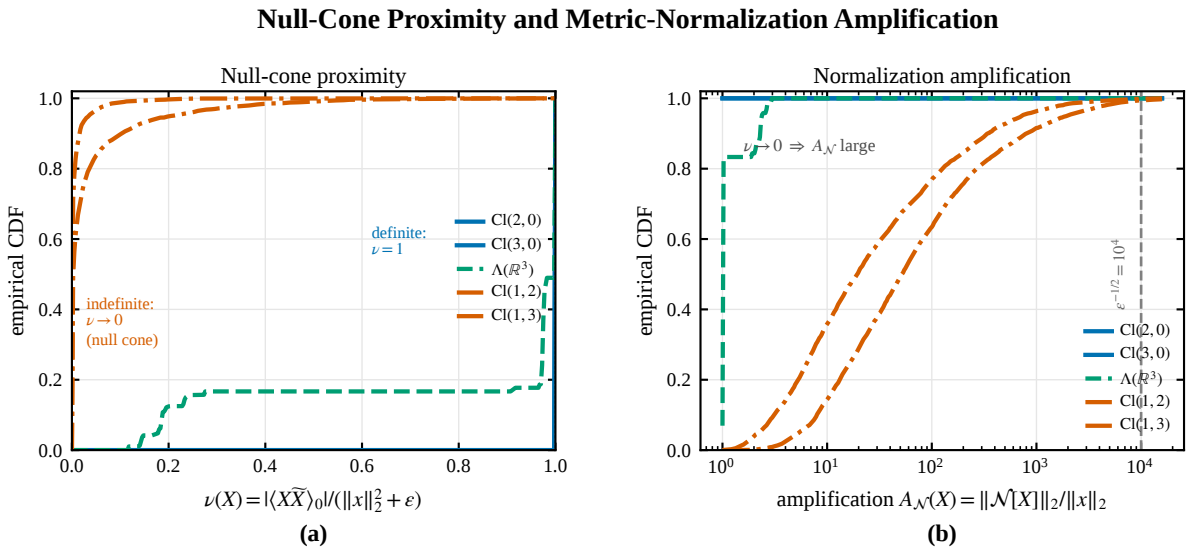


Figure 18: Empirical distribution of the null-cone proximity $\nu(X)$, Eq. (40), over the stationary states of a representative chaotic run in each algebra. Definite algebras sit at $\nu = 1$; the indefinite algebras concentrate at $\nu \approx 0$, on the null cone, where the metric normalization amplifies the state by up to $\varepsilon^{-1/2} = 10^4$.

C Cross-algebra and normalization controls

This appendix collects the tabulated numbers behind the cross-algebra comparison of Section 6.5. Table 10 reports the common-normalization control: for every algebra it places the phase-diversity entropy S_{phase} and the chaotic fraction f_{chaotic} under the metric-dependent normalization of the main text beside the same quantities under a common Euclidean coefficient normalization, on the matched 21×21, $M = 2$ control grid, so that the change ΔS, Δf

Table 10: Common-normalization control. Phase-diversity entropy S_{phase} and chaotic fraction f_{chaotic} under the metric-dependent normalization of the main text and under a common Euclidean coefficient normalization, on the matched 21×21, $M = 2$ control grid. The algebraic product is unchanged; only the normalization map differs.

		metric-dependent norm.		common Euclidean coeff. norm.		change	
Algebra	signature	S_{phase}	f_{chaotic}	S_{phase}	f_{chaotic}	ΔS	Δf
$\mathrm{Cl}(2,0)$	definite	1.11	0.38	1.15	0.35	$+0.04$	-0.02
$\mathrm{Cl}(3,0)$	definite	0.99	0.39	0.99	0.40	-0.01	$+0.01$
$\Lambda(\mathbb{R}^3)$	degenerate	0.96	0.15	0.92	0.16	-0.03	$+0.01$
$\mathrm{Cl}(1,2)$	indefinite	0.27	0.93	0.96	0.36	$+0.69$	-0.58
$\mathrm{Cl}(1,3)$	indefinite	0.30	0.92	1.04	0.36	$+0.75$	-0.56

Table 11: Phase richness across algebraic systems under the matched protocol. Phase diversity is the Shannon entropy of the phase-fraction vector.

System	signature	#phases ($> 2\%$)	phase diversity	$\langle \bar{H}_{\mathrm{grade}} \rangle$	chaotic fraction
scalar	–	1	-0.00	0.00	0.00
complex $\mathrm{Cl}(0,1)$	low-dim.	2	0.54	0.67	0.80
quaternion $\mathrm{Cl}(0,2)$	low-dim.	2	0.51	0.67	0.80
$\mathrm{Cl}(2,0)$	definite	4	1.11	0.79	0.38
$\mathrm{Cl}(3,0)$	definite	3	0.99	0.42	0.39
$\Lambda(\mathbb{R}^3)$	degenerate	3	0.96	0.42	0.15
$\mathrm{Cl}(1,2)$	indefinite	2	0.27	0.63	0.93
$\mathrm{Cl}(1,3)$	indefinite	2	0.30	0.72	0.92

isolates the contribution of the normalization map alone.

Table 11 lists the numbers plotted in Fig. 10. All systems use the identical rule, the identical 31×31 grid over (β, γ), the identical $M = 3$ seeds per cell and the identical horizons $T = 250$, $T_{\mathrm{tr}} = 130$.

D Detailed machine-learning performance

This appendix expands the aggregate recovery accuracies of Section 7.3 into per-class detail, so that the source of the residual error can be read off directly rather than inferred from the confusion matrices of Fig. 14. Three points are worth drawing out. First, the generating-algebra task is limited almost entirely by the two Euclidean Clifford algebras: $\mathrm{Cl}(3,0)$ and $\mathrm{Cl}(1,2)$ carry the lowest per-class F_1 (0.940 and 0.933), and their errors are mutual, whereas $\mathrm{Cl}(2,0)$ and $\Lambda(\mathbb{R}^3)$ are recovered at $F_1 \geq 0.969$. This is the expected behavior: the two eight-dimensional definite and indefinite algebras produce the most similar grade-resolved dynamics, so a grade-component-excluded learner confuses them more often than it confuses either with the degenerate or the low-dimensional cases. Second, the signature task is balanced across the three classes (F_1 between 0.969 and 0.976), which confirms that the signature is not being inferred through a single easy class such as the degenerate one. Third, on the same-dimension task, where dimension carries no information, precision and recall stay above 0.94 for all three algebras, so the discrimination genuinely rests on the dynamics rather than on a dimension artefact. The per-class figures are collected in Table 12.

Table 12: Per-class performance of the recovery of algebraic invariants, from pooled out-of-fold predictions of the grouped five-fold cross-validation of Table 7. Balanced accuracy is 0.958, 0.972 and 0.963 for the three tasks respectively. Supports are the full out-of-fold prediction populations and therefore match Fig. 14 exactly.

Task	class	precision	recall	F_1	support
generating algebra, $\mathcal{D}_{14}$	$\mathrm{Cl}(2,0)$	0.992	0.991	0.991	2883
	$\mathrm{Cl}(3,0)$	0.929	0.952	0.940	2883
	$\mathrm{Cl}(1,2)$	0.936	0.929	0.933	2883
	$\mathrm{Cl}(1,3)$	0.962	0.948	0.955	2883
	$\Lambda(\mathbb{R}^3)$	0.969	0.968	0.969	2883
signature class, $\mathcal{D}_{11}$	definite	0.968	0.975	0.971	5766
	indefinite	0.980	0.973	0.976	5766
	degenerate	0.970	0.968	0.969	2883
same dimension, $\mathcal{D}_{11}$	$\mathrm{Cl}(3,0)$	0.943	0.964	0.953	2883
	$\mathrm{Cl}(1,2)$	0.978	0.953	0.965	2883
	$\Lambda(\mathbb{R}^3)$	0.969	0.971	0.970	2883

E Reproducibility and methodological principles

The entire pipeline is deterministic conditional on the random seeds: for a fixed algebra, parameter vector θ and seed, the trajectory tensor $\mathcal{X}_{\mathcal{A},\theta,s}$, every descriptor in $\mathbf{f}$, and every reported metric are pure functions of the inputs, so each number in the paper can in principle be recomputed exactly. Two numerical choices protect this determinism. The automaton is integrated in IEEE double precision, because the geometric product accumulates structure-constant sums whose sign pattern is sensitive to rounding near the null cone; trajectories are then stored in single precision, which is below the threshold at which the coarse-grained descriptors change. The normalization $\mathcal{N}_\chi$ carries the regularizer $\varepsilon = 10^{-8}$ precisely so that the map is Lipschitz and globally well defined on states with $\langle X\widetilde{X}\rangle_0 \to 0$; without it the indefinite algebras would produce undefined updates on the null cone rather than the large but finite amplifications quantified in Appendix B.

The validation design follows from the dependence structure of the data rather than from convenience. The M seeds of a parameter cell and the neighbouring cells of a coordinate are not independent samples, so a naive random split would place statistically coupled points on both sides of the partition and inflate the apparent accuracy. We therefore use grouped K-fold cross-validation in which the grouping key is the unit of genuine independence: the parameter coordinate (β,γ) for the blind cross-algebra recovery, so that no coordinate is ever seen for one algebra in training and another in validation, and $(\text{algebra},\beta,\gamma)$ for the within-algebra phase and $\mathcal{D}_{14}/\mathcal{D}_{11}$ tasks. The blocked design, which partitions the γ axis into five contiguous strips, is a strictly stronger test: it forces the classifier to extrapolate across a region of parameter space rather than to interpolate within it, following the recommendation that spatially or hierarchically structured data be evaluated under blocked rather than random resampling [41]. All confusion matrices and per-class metrics are pooled out-of-fold predictions of a single such design, so supports and scores refer to exactly the same predictions. Uncertainties are bootstrap intervals over parameter cells, which quantify sampling variability across the swept plane; with only $M = 3$ seeds the seed-wise label disagreement of Appendix A is reported as a consistency check and not as a statistical confidence interval. Symbolic regression is treated as a stochastic search over the grammar of expression trees under a parsimony (description-length) prior, so its output is a candidate relation to be tested analytically, not a fitted model with a calibrated error bar.

Concretely, simulations use NumPy 1.26 and SciPy 1.11. Learning uses scikit-learn 1.6 [27] (random forests: 250 trees, `class_weight = balanced_subsample` for the imbalanced phase task; gradient boosting: histogram-based, 200 iterations; multilayer perceptron: two hidden layers of 128 and 64 units, 400 epochs; support-vector machine: RBF kernel, $C = 4$), UMAP 0.5 [26] (15 neighbors, minimum distance 0.1, cosine metric) and gplearn 0.4 for symbolic regression (population 2000, 20 generations, operators $\{+,-,\times,\div,\sqrt{},\log\}$, parsimony coefficient 10^{-3}). Standardization is fitted inside each training fold to avoid target leakage through the scaler. Seeds are $100m+7$ for the mth initial condition, and all cross-validation splits use `random_state` $= 0$. The complete code, the sweep archives and the exact package versions accompany the article.

References

[1] He, Y.-H.: Machine-learning the string landscape. Phys. Lett. B **774**, 564–568 (2017). https://doi.org/10.1016/j.physletb.2017.10.024

[2] Carifio, J., Halverson, J., Krioukov, D., Nelson, B.D.: Machine learning in the string landscape. J. High Energy Phys. **2017**(9), 157 (2017). https://doi.org/10.1007/JHEP09(2017)157

[3] He, Y.-H., Lukas, A.: Machine learning Calabi–Yau four-folds. Phys. Lett. B **815**, 136139 (2021). https://doi.org/10.1016/j.physletb.2021.136139

[4] He, Y.-H.: The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning. Lecture Notes in Mathematics, vol. 2293. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77562-9

[5] He, Y.-H., Kim, M.: Learning algebraic structures: preliminary investigations. Int. J. Data Sci. Math. Sci. **1**(1), 3–22 (2023). https://doi.org/10.1142/S2810939222500046

[6] Davies, A., Veličković, P., Buesing, L., Blackwell, S., Zheng, D., Tomašev, N., Tanburn, R., Battaglia, P., Blundell, C., Juhász, A., Lackenby, M., Williamson, G., Hassabis, D., Kohli, P.: Advancing mathematics by guiding human intuition with AI. Nature **600**, 70–74 (2021). https://doi.org/10.1038/s41586-021-04086-x

[7] Raayoni, G., Gottlieb, S., Manor, Y., Pisha, G., Harris, Y., Mendlovic, U., Haviv, D., Hadad, Y., Kaminer, I.: Generating conjectures on fundamental constants with the Ramanujan Machine. Nature **590**, 67–73 (2021). https://doi.org/10.1038/s41586-021-03229-4

[8] Iten, R., Metger, T., Wilming, H., del Rio, L., Renner, R.: Discovering physical concepts with neural networks. Phys. Rev. Lett. **124**(1), 010508 (2020). https://doi.org/10.1103/PhysRevLett.124.010508

[9] Voevodsky, V.: The origins and motivations of univalent foundations. The Institute Letter (Institute for Advanced Study), Summer 2014, pp. 8–9 (2014)

[10] Lample, G., Charton, F.: Deep learning for symbolic mathematics. In: Proceedings of the 8th International Conference on Learning Representations (ICLR) (2020)

[11] He, Y.-H.: Machine-learning mathematical structures. Int. J. Data Sci. Math. Sci. **1**(1), 23–47 (2023). https://doi.org/10.1142/S2810939222500010

[12] Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., Zdeborová, L.: Machine learning and the physical sciences. Rev. Mod. Phys. **91**(4), 045002 (2019). https://doi.org/10.1103/RevModPhys.91.045002

[13] Hestenes, D., Sobczyk, G.: Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics. Reidel, Dordrecht (1984). https://doi.org/10.1007/978-94-009-6292-7

[14] Doran, C., Lasenby, A.: Geometric Algebra for Physicists. Cambridge University Press, Cambridge (2003). https://doi.org/10.1017/CBO9780511807497

[15] Dorst, L., Fontijne, D., Mann, S.: Geometric Algebra for Computer Science: An Object-Oriented Approach to Geometry. Morgan Kaufmann, San Francisco (2007)

[16] Buchholz, S., Sommer, G.: On Clifford neurons and Clifford multi-layer perceptrons. Neural Netw. **21**(7), 925–935 (2008). https://doi.org/10.1016/j.neunet.2008.03.004

[17] Parcollet, T., Morchid, M., Linarès, G.: A survey of quaternion neural networks. Artif. Intell. Rev. **53**(4), 2957–2982 (2020). https://doi.org/10.1007/s10462-019-09752-1

[18] Brandstetter, J., van den Berg, R., Welling, M., Gupta, J.K.: Clifford neural layers for PDE modeling. In: Proceedings of the 11th International Conference on Learning Representations (ICLR) (2023)

[19] Ruhe, D., Gupta, J.K., De Keninck, S., Welling, M., Brandstetter, J.: Geometric Clifford algebra networks. In: Proceedings of the 40th International Conference on Machine Learning, PMLR **202**, 29306–29337 (2023)

[20] Ruhe, D., Brandstetter, J., Forré, P.: Clifford group equivariant neural networks. In: Advances in Neural Information Processing Systems, vol. 36, pp. 62922–62990 (2023). https://doi.org/10.52202/075280-2748

[21] Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: going beyond Euclidean data. IEEE Signal Process. Mag. **34**(4), 18–42 (2017). https://doi.org/10.1109/MSP.2017.2693418

[22] Wolfram, S.: Statistical mechanics of cellular automata. Rev. Mod. Phys. **55**(3), 601–644 (1983). https://doi.org/10.1103/RevModPhys.55.601

[23] Wolfram, S.: A New Kind of Science. Wolfram Media, Champaign (2002)

[24] Langton, C.G.: Computation at the edge of chaos: phase transitions and emergent computation. Physica D **42**(1–3), 12–37 (1990). https://doi.org/10.1016/0167-2789(90)90064-V

[25] Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007). https://doi.org/10.1016/j.physrep.2006.11.001

[26] McInnes, L., Healy, J., Saul, N., Großberger, L.: UMAP: Uniform Manifold Approximation and Projection. J. Open Source Softw. **3**(29), 861 (2018). https://doi.org/10.21105/joss.00861

[27] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)

[28] Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[29] Koza, J.R.: Genetic programming as a means for programming computers by natural selection. Stat. Comput. **4**(2), 87–112 (1994). https://doi.org/10.1007/BF00175355

[30] Udrescu, S.-M., Tegmark, M.: AI Feynman: a physics-inspired method for symbolic regression. Sci. Adv. **6**(16), eaay2631 (2020). https://doi.org/10.1126/sciadv.aay2631

[31] Schmidt, M., Lipson, H.: Distilling free-form natural laws from experimental data. Science **324**(5923), 81–85 (2009). https://doi.org/10.1126/science.1165893

[32] Mirza, F.K., Pekcan, Ö., Hekimoğlu, M., Baykaş, T.: Stock price forecasting through symbolic dynamics and state transition graphs with a convolutional recurrent neural network architecture. Neural Comput. Appl. **37**, 15855–15890 (2025). https://doi.org/10.1007/s00521-025-11325-z

[33] Mirza, F.K., Öz, U., Hekimoğlu, M., Aydemir, M.T., Pural, Y.E., Baykaş, T., Pekcan, Ö.: A novel multiscale graph signal processing and network dynamics approach to vibration analysis for stone size discrimination via nonlinear manifold embeddings and a convolutional self-attention model. Rock Mech. Rock Eng. **58**, 12769–12799 (2025). https://doi.org/10.1007/s00603-025-04779-z

[34] Chen, S., Dechant, P.-P., He, Y.-H., Heyes, E., Hirst, E., Riabchenko, D.: Machine learning Clifford invariants of ADE Coxeter elements. Adv. Appl. Clifford Algebras **34**, 20 (2024). https://doi.org/10.1007/s00006-024-01325-y

[35] Lisitsa, A., Salles, M., Vernitski, A.: Machine learning discovers invariants of braids and flat braids. Adv. Appl. Clifford Algebras **34**, 49 (2024). https://doi.org/10.1007/s00006-024-01349-4

[36] Bill, J., Cox, B.: Exploring quaternion neural network loss surfaces. Adv. Appl. Clifford Algebras **34**, 19 (2024). https://doi.org/10.1007/s00006-024-01313-2

[37] Lounesto, P.: Clifford Algebras and Spinors, 2nd edn. London Mathematical Society Lecture Note Series, vol. 286. Cambridge University Press, Cambridge (2001)

[38] Bagnoli, F., Rechtman, R., Ruffo, S.: Damage spreading and Lyapunov exponents in cellular automata. Phys. Lett. A **172**(1–2), 34–38 (1992). https://doi.org/10.1016/0375-9601(92)90185-O

[39] Tisseur, P.: Cellular automata and Lyapunov exponents. Nonlinearity **13**(5), 1547–1560 (2000). https://doi.org/10.1088/0951-7715/13/5/308

[40] Webber, C.L., Jr., Marwan, N. (eds.): Recurrence Quantification Analysis: Theory and Best Practices. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-07155-8

[41] Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography **40**(8), 913–929 (2017). https://doi.org/10.1111/ecog.02881

[42] La Cava, W., Orzechowski, P., Burlacu, B., de França, F.O., Virgolin, M., Jin, Y., Kommenda, M., Moore, J.H.: Contemporary symbolic regression methods and their relative performance. In: Vanschoren, J., Yeung, S. (eds.) Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, vol. 1 (2021)