DISENTANGLING TOPOLOGY FROM MEASURE: CERTIFIED KOOPMAN SPECTRAL LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Data-driven Koopman models learn finite-dimensional representations of nonlinear dynamics from trajectories sampled under a particular state-visitation distribution, yet deployment may change which regions of state space are visited while leaving the underlying dynamics fixed. Standard smooth dictionaries entangle this distribution-dependent fitting geometry with the transition structure of the dynamics inside the same Galerkin matrices. We ask which representation structures can separate these two sources of variation and remain reusable and certifiable under sampling shift. We show that measure-universal Gram diagonality is equivalent to pairwise disjoint support; on a compact connected state space, a dense continuous dictionary cannot retain this property, while indicators of refining generating partitions provide a canonical realization (Theorem 1). This representation-level separation yields the *Topological Koopman Certificate* (TKC), with non-asymptotic residual and pseudospectral guarantees (Theorems 2, 3), a TV-linear operator-perturbation bound with sparse support-preserving reweighting under measure shift (Theorem 4), and an $\mathcal{O}(KN)$ Krylov pseudospectral solver in place of a dense $\mathcal{O}(N^3)$ generalized eigenvalue computation. We turn the structural principle into a data-driven construction through *Generative Koopman Decoupling* (GKD, Algorithm 1), which uses either an analytic or learned preimage oracle to form forward–backward itinerary partitions. Proposition 1 links oracle error to partition resolution in the uniformly expanding regime. Controlled low-dimensional systems evaluate the certificates under matched assumptions, while Lorenz-63 and Hénon test the learned representation and oracle mechanism empirically beyond that regime. On Lorenz-63, the learned-itinerary representation attains the lowest in-distribution prediction MSE in our comparison using only about 25 occupied itineraries; on Hénon, learned- and analytic-oracle constructions give statistically indistinguishable predictions.

1 INTRODUCTION

A data-driven dynamical model is learned from finite trajectories collected under a particular state-visitation distribution and may later be evaluated when the same dynamics is visited with different frequencies. For deterministic dynamics $Y = T(X)$, such a change leaves the transition law $P(Y \mid X)$ fixed while changing the state distribution $P(X)$. This creates a representation-learning question: which parts of a learned finite-dimensional model reflect reusable dynamical structure, and which parts arise from the distribution used to fit it? Koopman learning provides a precise setting in which to study this question. Koopman operator theory (Koopman, 1931; Mezić, 2005) represents nonlinear dynamics through a linear action on observables, and data-driven methods fit a finite-dimensional operator $\hat{\mathcal{K}} = (G^{(N)})^{\dagger} A^{(N)}$ with $G_{ij} = \langle \phi_i, \phi_j \rangle_\mu$ and $A_{ij} = \langle \mathcal{K}\phi_i, \phi_j \rangle_\mu$. Choosing or learning the dictionary therefore amounts to choosing the representation in which dynamics are fitted from trajectory data, connecting EDMD, kernel methods, and deep Koopman models to representation learning and statistical operator learning (Williams et al., 2015; Klus et al., 2020; Takeishi et al., 2017; Morton et al., 2018; Lusch et al., 2018; Mardt et al., 2018; Azencot et al., 2020; Kostic et al., 2022; 2023; Meanti et al., 2023). Recent Koopman work also studies non-stationary prediction and adaptation under distribution shift (Liu et al., 2023; Selim et al., 2025). The unresolved structural question is complementary: *which representation property separates distribution-dependent fitting geometry from reusable transition structure when the underlying dynamics is fixed?* For standard

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

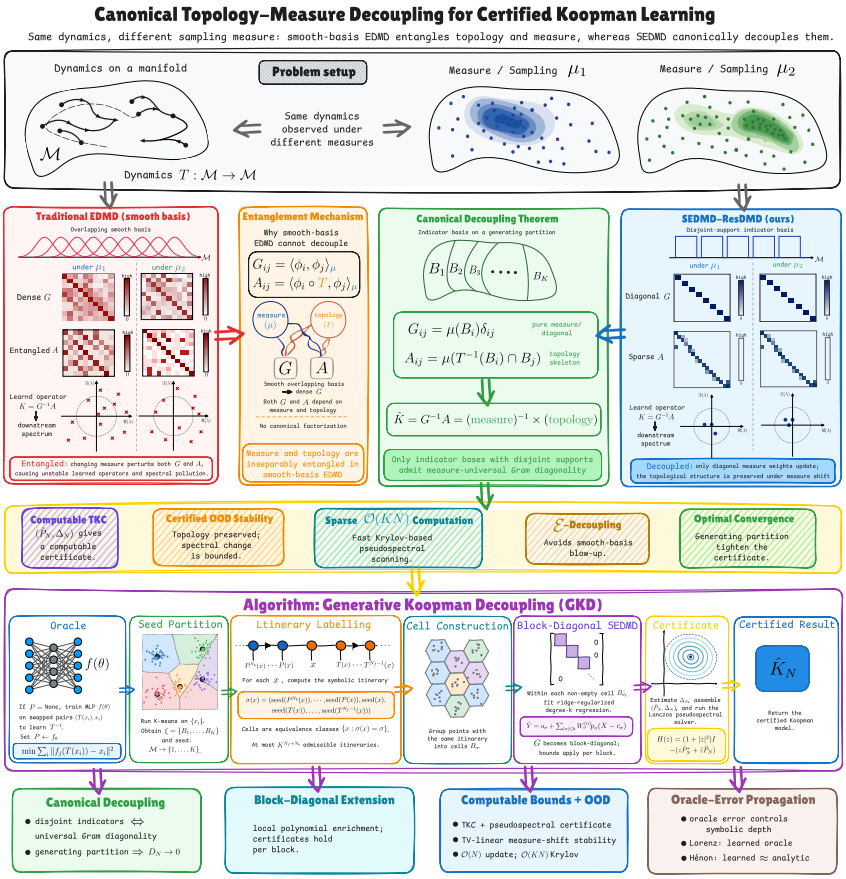


Figure 1: Representation-level separation under sampling shift. Partition indicators isolate sampling-dependent cell occupancies in the Gram geometry while preserving sparse transition support determined by the dynamics. This separation underlies GKD, TKC, and sparse measure-shift updates.

smooth dictionaries, both G and A depend on the sampling measure, so this separation is not explicit; ResDMD (Colbrook & Townsend, 2024) certifies residuals under the evaluation measure but does not by itself provide such a representation-level decomposition.

Our answer begins at the representation level. Theorem 1 characterizes exactly when the Gram geometry remains diagonal across admissible sampling measures: this occurs precisely for pairwise disjoint-support coordinates. For partition indicators, changing the sampling law changes cell occupancies while the admissible transition support remains determined by the dynamics. GKD turns this structural principle into a learned itinerary representation, while TKC quantifies the residual, pseudospectral, and finite-dimensional operator uncertainty that remains. Controlled systems test these guarantees under matched assumptions; Lorenz-63 and Hénon test whether the same representation and oracle mechanism remain useful when the partition is learned from data.

In the canonical (indicator) basis of a generating partition $\xi^{(N)} = \bigvee_{k=0}^{N-1} T^{-k}\xi$:

$$G_{ij} = \mu(B_i)\,\delta_{ij}, \qquad A_{ij} = \mu(T^{-1}(B_i) \cap B_j), \tag{1}$$

Thus G records sampling-dependent cell occupancies, whereas the sparsity pattern of A records which cell-to-cell transitions are admissible under the dynamics. This is the representation-level separation studied in the paper. Exploiting it yields the Topological Koopman Certificate (TKC) $(\tilde{P}_N, \Delta_N)$, with non-asymptotic pseudospectral control (Theorem 3), a TV-linear finite-dimensional operator-perturbation bound with sparse support-preserving reweighting (Theorem 4), and $\mathcal{O}(KN)$ Krylov pseudospectral computation. Per-cell polynomial enrichment preserves the separation block-by-block (Corollary 1). GKD then constructs the required itinerary representation using either an

analytic preimage oracle or a learned model $f_\theta \approx T^{-1}$; Proposition 1 quantifies its resolution in the uniformly expanding regime.

Our contributions are as follows. The indicator-basis SEDMD construction itself is due to (Kennedy et al., 2024; 2026); what is new here is the canonicality characterization, the certificate, the OOD bound, the unified algorithm, and the empirical match with deep methods.

C1: Generative Koopman Decoupling (GKD) (Algorithm 1). An end-to-end certified estimator with a pluggable preimage oracle, covering both analytic T^{-1} and learned f_θ regimes.

C2: Canonical Decoupling (Thm. 1). On a connected state space, an indicator dictionary is measure-universally Gram-diagonal iff its supports are a.s. disjoint; the indicator basis of a generating partition is the unique class satisfying measure-universal Gram diagonality under the stated measure class with $D_N \to 0$.

C3: Block-Diagonal Extension (Cor. 1). Polynomial enrichment yields a block-diagonal Gram; the canonical decoupling and Theorems 2, 3, 4 hold per block.

C4: Computable Pseudospectral and OOD Bounds (Thms. 2, 3, 4). Non-asymptotic inclusions estimable from one trajectory; operator perturbation under measure shift is bounded by $C_N\|\mu' - \mu\|_{\mathrm{TV}}$ with sparse certified reweighting.

C5: Oracle-Error Propagation (Prop. 1). The symbolic depth is controlled by the oracle error through $D_{N_f,N_b} \le c\max(\rho^{-(N_f+N_b)}D_0, \epsilon/(1-\rho^{-1}))$, with empirical learned-oracle behavior further studied on Lorenz-63 in Sec. 3.

Table 1: Structural comparison of Koopman estimators. ✓ = rigorous; ∼ = asymptotic only; × = absent.

Method	Computable bounds	Measure-shift stability	$\mathcal{O}(KN)$ complexity	Optimal convergence
EDMD-Fourier / polynomial (Williams et al., 2015)	∼	×	×	∼
Kernel EDMD (Klus et al., 2020)	∼	×	×	∼
ResDMD (Colbrook & Townsend, 2024)	✓	∼	×	∼
Deep Koopman (Lusch et al., 2018)	×	×	×	∼
SEDMD (Kennedy et al., 2024)	∼	×	∼	✓
GKD (this paper)	✓	✓	✓	✓

§2 develops the four theorems, Cor. 1, Algorithm 1, and Prop. 1, and §3 reports the Lorenz-63 benchmark plus supporting 1D and 2D experiments. Full proofs, scope and limitations, and additional experiments appear in the appendix.

2 Theoretical Framework

For a discrete-time measure-preserving system $(\mathcal{M}, \mathcal{B}, \mu, T)$ on a compact metric space $\mathcal{M}$, the Koopman operator $\mathcal{K} : L^2(\mathcal{M},\mu) \to L^2(\mathcal{M},\mu)$ acts by composition $\mathcal{K}f = f \circ T$. EDMD takes a finite dictionary $\mathcal{D}_N = \{\phi_i\}_{i=1}^N$ and forms $\hat{K} \triangleq (G^{(N)})^\dagger A^{(N)}$ with $G_{ij} = \langle \phi_i, \phi_j\rangle_\mu$ and $A_{ij} = \langle \mathcal{K}\phi_i, \phi_j\rangle_\mu$, while ResDMD (Colbrook & Townsend, 2024) evaluates the residual $\mathrm{res}(\lambda, g) = \|\mathcal{K}g - \lambda g\|_\mu / \|g\|_\mu$, giving $\|(\mathcal{K} - \lambda I)^{-1}\|^{-1} \le \mathrm{res}(\lambda, g)$. A generating partition ξ refined to depth N as $\xi^{(N)} = \bigvee_{k=0}^{N-1} T^{-k}\xi$ has cells $\{B_1, \dots, B_m\}$ of vanishing maximum diameter $D_N \to 0$ (Lind & Marcus, 1995), and SEDMD (Kennedy et al., 2024; 2026) uses the indicator dictionary $\{\mathbf{1}_{B_i}\}$. Fix a reference full-support Borel measure λ on M, and let $\mathcal{P}_\lambda$ denote the class of Borel probability measures absolutely continuous with respect to λ. A dictionary is λ-universally Gram-diagonal when $\langle \phi_i, \phi_j\rangle_\mu = 0$ for all $i \ne j$ and every $\mu \in \mathcal{P}_\lambda$ where Theorem 1 characterises exactly this property.

2.1 Canonical Decoupling and the Topological Koopman Certificate

Theorem 1 (Canonical Decoupling Theorem)**.** *Let $\mathcal{M}$ be a compact metric space with Borel σ-algebra.*

(a) ***Characterization.*** *A dictionary $\mathcal{D}_N = \{\phi_1, \ldots, \phi_N\} \subset L^2(M, \lambda)$ is λ-universally Gram-diagonal over $\mathcal{P}_\lambda$ if and only if its elements have pairwise λ-a.s. disjoint nonzero sets:*

$$\lambda\big(\{\phi_i \neq 0\} \cap \{\phi_j \neq 0\}\big) = 0 \qquad \forall i \neq j.$$

(b) ***Connected-Space Obstruction.*** *If $\mathcal{M}$ is compact* connected*, any countable continuous dictionary $\mathcal{D} = \{\phi_i\}_{i \geq 1} \subset C(\mathcal{M})$ whose linear span is dense in $L^2(\mathcal{M}, \lambda)$ has a finite truncation $\mathcal{D}_N$ that fails to be measure-universally Gram-diagonal: there exists an absolutely continuous Borel probability measure $\mu' \ll \lambda$ on M and indices $i \neq j$ such that $G_{ij}(\mu') \neq 0$. In particular, this rules out Fourier, polynomial, RBF, and kernel-EDMD dictionaries on intervals, balls, or smooth manifolds. The connectedness assumption is essential: on totally disconnected attractors (Cantor sets, hyperbolic basic sets) continuous dictionaries with disjoint clopen supports do exist, and on those domains smooth bases retain part of the decoupling property. We discuss the practical implications in Appendix I.*

(c) ***Canonical Form of the Indicator Basis.*** *The indicator functions $\{\mathbf{1}_{B_1}, \ldots, \mathbf{1}_{B_m}\}$ of any pointwise disjoint partition $\{B_i\}$ are Gram-diagonal for every Borel measure μ, and hence are λ-universally Gram-diagonal over $\mathcal{P}_\lambda$, with*

$$G_{ij}(\mu) = \mu(B_i)\delta_{ij}.$$

The Koopman cross-covariance satisfies

$$A_{ij}(\mu) = \langle K\mathbf{1}_{B_i}, \mathbf{1}_{B_j}\rangle_\mu = \mu(T^{-1}(B_i) \cap B_j),$$

yielding the canonical factorization $\hat{K} = G^{-1}A$ where G encodes the measure alone (topology-free diagonal $\{\mu(B_i)\}$), while A's sparsity pattern encodes solely the transition topology of T, with nonzero values carrying μ-weighted masses of cylinder-set intersections.

Sketch of proof: (a) Sufficiency follows because $\phi_i\phi_j = 0$ λ-a.e.; hence $\phi_i\phi_j = 0$ μ-a.e. for every $\mu \ll \lambda$. Necessity follows by taking μ as the normalised restriction of λ to a positive-λ-measure subset where both ϕ_i and ϕ_j are simultaneously nonzero with fixed signs. (b) If all finite truncations were universally Gram-diagonal, Theorem 1(a) would make all dictionary elements pairwise a.e. disjoint; on a connected full-support space this leaves a nonzero L^2 function orthogonal to the whole countable span, contradicting density. (c) Direct computation from $\mathbf{1}_{B_i}\mathbf{1}_{B_j} = 0$ pointwise. Full proof in Appendix A

Corollary 1 (Block-Diagonal Extension)**.** *For a partition $\{B_i\}$ with disjoint supports and polynomials $\{p_\alpha\}_{|\alpha| \leq k}$, the enriched dictionary $\{p_\alpha\, \mathbf{1}_{B_i}\}$ has* block-diagonal *Gram (one $\binom{d+k}{k} \times \binom{d+k}{k}$ block per cell, depending only on $\mu|_{B_i}$); the cell-pair sparsity pattern of A matches the bare indicator basis. Theorems 2, 3, and 4 hold per block.*

This is the theoretical basis for SEDMD-Local (per-cell affine, §3.1) and for the learned-itinerary estimator (§3.2). Since Theorem 1 establishes that G and A can be independently estimated, we now formalise the certification mechanism that exploits this independence.

Definition 1 (Topological Koopman Certificate)**.** *A* Topological Koopman Certificate *(TKC) of depth N is the computable pair* $\mathrm{TKC}_N = (\tilde{P}_N, \Delta_N)$*, where $\tilde{P}_N = G_N^{-1/2} A_N G_N^{-1/2}$ is the* Symmetrized Topological Transition Matrix *encoding the dynamics' topological skeleton, with $G_N = \mathrm{diag}(\tilde{\mu}_N(B_i))$ and $\tilde{\mu}_N$ reconstructed via the left eigenvector of A_N (Ding & Li, 1991), and Δ_N is the* computable measure distortion bound *satisfying* $\sup_{C \in \xi^{(N)} \cup \xi^{(N+1)}} |\tilde{\mu}_N(C)/\mu_{true}(C) - 1| \leq \Delta_N$.

All spectral guarantees in subsequent theorems are expressed in terms of TKC_N. As $N \to \infty$, $D_N \to 0$ and $\Delta_N \to 0$, so the certificate tightens automatically with partition refinement. Δ_N is upper-boundable from data via a Bernstein-type concentration term plus a self-consistency surrogate; see Appendix E, Remark 1. The $\mathcal{O}(M^{-1/2})$ scaling matches the experimental Figure 8(a).

2.2 Spectral Bounds and Stability under Measure Shift

The following assumptions are maintained throughout.

(GP) *Generating Partition.* ξ is a generating partition for $(\mathcal{M}, T)$ (Cornfeld et al., 1982); $D_N \to 0$.

(MD) *Bounded Measure Distortion.* The reconstructed measure $\tilde{\mu}_N$ satisfies $\Delta_N \to 0$ as defined in the TKC.

Theorem 2 (Uniform Residual Approximation; proof in Appendix B)**.** *Under (GP), (MD), for any candidate eigenpair (λ, g_N) with $g_N \in \mathcal{S}_N$: $|\hat{\text{res}}_N(\lambda, g_N) - \text{res}_{true}(\lambda, g_N)| \le \Delta_N \cdot \text{res}_{true}(\lambda, g_N) + \mathcal{O}(\Delta_N^2)$.*

Theorem 3 (Computable Pseudospectral Bounds; proof in Appendix C)**.** *Under (GP), (MD), with $\mathcal{E}_N \triangleq \tilde{\epsilon}_N(1 + \Delta_N) + \mathcal{O}(\tilde{\epsilon}_N \Delta_N^2)$ where $\tilde{\epsilon}_N = \hat{\text{res}}_N(\lambda, g_N)$, λ belongs to the $\mathcal{E}_N$-pseudospectrum of $\mathcal{K}$: $\|(\mathcal{K} - \lambda I)^{-1}\| \ge \mathcal{E}_N^{-1}$. If $\mathcal{K}$ is unitary, $\text{dist}(\lambda, \sigma(\mathcal{K})) \le \mathcal{E}_N$.*

Pseudospectrum reduces to the minimum eigenvalue of $H(z) = (\tilde{P}_N - zI)^*(\tilde{P}_N - zI)$ at each $z \in \mathbb{C}$, computable by Lanczos in $\mathcal{O}(KN)$ per grid point, total $\mathcal{O}(R^2 KN)$ for an $R \times R$ scan vs. $\mathcal{O}(R^2 N^3)$ for the dense GEVP. The canonical decoupling further yields the first rigorous OOD guarantee for Koopman operator learning, in the form of the next theorem.

Theorem 4 (OOD Stability Under Measure Shift)**.** *Let μ and μ' be two Borel probability measures on $\mathcal{M}$ with $\mu(B_i) > 0$ for all i. Let $\hat{K}_N(\mu) = G_N(\mu)^{-1} A_N(\mu)$ and $\hat{K}_N(\mu') = G_N(\mu')^{-1} A_N(\mu')$ be the depth-N SEDMD Koopman approximations under each measure. Then:*

(a) ***Topological adjacency invariance.*** *Define the topological adjacency pattern*

$$\mathcal{A}_{ij} = \mathbf{1}\{T^{-1}(B_i) \cap B_j \neq \emptyset\}.$$

Then $\mathcal{A}$ is independent of μ. Moreover, $A_{ij}(\mu) = 0$ whenever $\mathcal{A}_{ij} = 0$, and if the admissible measure μ assigns positive mass to every nonempty transition cylinder $T^{-1}(B_i) \cap B_j$, then $A_{ij}(\mu) \neq 0$ exactly when $\mathcal{A}_{ij} = 1$.

(b) ***Operator perturbation under measure shift.*** *The operator approximations satisfy*

$$\|\hat{K}_N(\mu') - \hat{K}_N(\mu)\|_F \le C_N \cdot \|\mu' - \mu\|_{TV},$$

where

$$\mu_{\min} = \min_i \{\mu(B_i), \mu'(B_i)\} > 0$$

over the finite partition cells retained by the estimator, $\|\cdot\|_{TV}$ denotes total variation distance, and

$$C_N = \frac{\sqrt{NK}}{\mu_{\min}} + \frac{\sqrt{N}\,\|A_N(\mu)\|_F}{\mu_{\min}^2},$$

with K the maximum number of nonzero entries per row of A_N (the symbolic transition connectivity). The quantity $\|A_N(\mu)\|_F$ is directly computable from the training trajectory as the Frobenius norm of the empirical transition count matrix.

(c) ***Sparse support-preserving reweighting.*** *Under a shifted measure μ', the adjacency pattern from (a) is reused, while $G_N(\mu')$ and the nonzero transition weights of $A_N(\mu')$ are reweighted from shifted samples on the fixed sparse support. The update costs $O(M' + \text{nnz}(A_N))$, with $\text{nnz}(A_N) \le NK$, and avoids dense Galerkin recomputation.*

Proof in Appendix D. The constant C_N depends on $\mu_{\min}$ and grows with partition refinement; its scaling in the Markov/Hölder regimes is discussed in Appendix F. An assumption-matched Bernoulli control in Appendix J shows that the combined certificate quantity $C_N \Delta_N$ is informative at coarse-to-moderate depths, while also exhibiting the resolution–certifiability tradeoff under further refinement. The update reuses the topology encoded by $\mathcal{A}$ and only recomputes cell/transition weights from the shifted measure, avoiding dense Galerkin recomputation.

2.3 Generative Koopman Decoupling: A Unified Algorithm

Theorems 1, 2, 3, 4 apply when the indicator representation is supported on a generating partition satisfying the stated distortion assumptions. When such a partition is known analytically, its refinement uses the preimage operator T^{-1}. When it is not available, GKD replaces this step by a learned

preimage oracle and constructs forward–backward itinerary cells from trajectory data. Proposition 1 provides a quantitative resolution guarantee for this construction in the C^1 uniformly expanding regime. For more general smooth dynamics, including Lorenz-63, we use the learned itinerary partition as an empirical predictive representation; theorem-level certification is invoked only in the regimes satisfying the stated assumptions.

Algorithm 1: Generative Koopman Decoupling (GKD)

Input: trajectory pairs $\{(x_i, T(x_i))\}_{i=1}^{M}$; seed size K; depths (N_f, N_b); oracle P (analytic T^{-1} or `None`); polynomial degree k.

Output: cell labelling $L : \mathcal{M} \to \Sigma$, block-diagonal SEDMD operator $\hat{K}_N$, certificate TKC $(\tilde{P}_N, \Delta_N)$.

1. **Oracle.** If P=`None`, train MLP f_θ on swapped pairs $(T(x_i), x_i)$ to minimize $\sum_i \|f_\theta(T(x_i)) - x_i\|^2$; set $P \leftarrow f_\theta$.
2. **Seed partition.** Run K-means on $\{x_i\}$ to obtain $\xi = \{B_1, \ldots, B_K\}$ and $\texttt{seed}: \mathcal{M} \to \{1, \ldots, K\}$.
3. **Itinerary labelling.** For each x, compute $\sigma(x) = (\texttt{seed}(P^{N_b}(x)), \ldots, \texttt{seed}(P(x)), \texttt{seed}(x), \texttt{seed}(T(x)), \ldots, \texttt{seed}(T^{N_f-1}(x)))$. Cells are equivalence classes $\{x : \sigma(x) = \sigma\}$, at most $K^{N_f+N_b}$ admissible itineraries.
4. **Block-diagonal SEDMD.** Within each non-empty cell B_σ, fit ridge-regularized degree-k regression $\hat{Y} = a_\sigma + \sum_{|\alpha|\le k} W_\sigma^{(\alpha)} p_\alpha(X - c_\sigma)$. Corollary 1 gives block-diagonal G and Theorems 2, 3, 4 per block.
5. **Certificate.** Estimate Δ_N (Appendix E); assemble $(\tilde{P}_N, \Delta_N)$ and the Lanczos pseudospectral solver on $H(z) = (\tilde{P}_N - zI)^*(\tilde{P}_N - zI)$.

The 1D/2D experiments are the analytic-oracle instance (P=T^{-1}_{exact}, k=0, small K); the Lorenz-63 experiment is the MLP-oracle instance (P=f_θ, k=1, K=4).

Proposition 1 (Oracle-Error Propagation and Optimal Depth)**.** *For C^1 uniformly expanding T with rate $\rho > 1$, oracle error $\epsilon = \|P - T^{-1}\|_\infty$, and seed diameter D_0, the GKD partition diameter satisfies*

$$D^P_{N_f,N_b} \;\le\; c \cdot \max\Big(\rho^{-(N_f+N_b)}\, D_0,\; \tfrac{\epsilon}{1-\rho^{-1}}\Big), \tag{3}$$

optimised at $N_f + N_b \sim \log(D_0/\epsilon)/\log\rho$. Past this depth, refinement is dominated by oracle error.

Sketch. Exact T^{-1} contracts by ρ^{-k}, so the depth-k refinement reduces diameter as $\rho^{-k} D_0$. Approximate P accumulates per-step error $\|P^j(x) - T^{-j}(x)\| \le \epsilon \sum_{l=0}^{j-1} \rho^{-l} \le \epsilon/(1-\rho^{-1})$, so backward refinement saturates at this scale. Combining with the exact-forward contribution $\rho^{-N_f} D_0$ and balancing yields (3) and the optimal depth. □

Proposition 1 converts the symbolic depth from a hyperparameter into a derived quantity. Given the empirical inverse-NN error $\hat{\epsilon}$, the optimal depth is $N_f + N_b \approx \log(D_0/\hat{\epsilon})/\log\rho$. The framework's quantitative claims hold inside the regime where a generating partition is given analytically or learned via the GKD oracle (L1, Prop. 1), $\mathcal{M}$ is connected (L2, Thm. 1(b)), $C_N\Delta_N$ remains controlled with $\mu_{\min}$ bounded (L3, partition-based curse of dimensionality at $d \gg 1$), and the dynamics is β-mixing for the rigorous Δ_N form (L4). Full discussion appears in Appendix I.

3 Experiments

All experiments run on a single CPU, and stochastic experiments report mean $\pm 1\sigma$ over 5 to 10 seeds. The experiments serve two complementary purposes. Controlled low-dimensional systems evaluate the quantitative certificates under matched assumptions, while Lorenz-63 and Hénon test whether the same representation and oracle mechanisms continue to deliver competitive predictive performance

in more challenging dynamical regimes; supporting 1D and 2D experiments are summarised here and detailed in Appendix J.

Five supporting 1D/2D experiments establish the framework's empirical content; full details appear in Appendix J. On the asymmetric Bernoulli map (a=0.3), the off-diagonal energy of G is identically zero for the indicator basis at every N but bounded away from zero for Fourier and polynomial bases, consistent with Theorem 1. On the smooth conjugated shift map, the sparse Krylov pseudospectral solver scales as $\mathcal{O}(KN)$ versus $\mathcal{O}(N^3)$ for the dense GEVP, verified up to N=16,384. Under extreme measure shifts (a=0.99, $s\to1$), EDMD-Fourier produces spurious eigenvalues outside r_{ess} while SEDMD-ResDMD remains stable. The residual envelope of Theorem 2 holds; under a structured measure shift, the finite-dimensional operator perturbation increases monotonically with TV distance while remaining below the local small-shift reference. The certificate distortion Δ_N decays as $\mathcal{O}(M^{-1/2})$ at N=4, 6, 8, matching the Birkhoff prediction. An assumption-matched Bernoulli check further gives $C_N\Delta_N < 0.5$ at depths 1–3, while depth 4 exposes the loss of certifiability under refinement.

3.1 Mid-dimensional Benchmark on Lorenz-63 vs. Deep Koopman

We test the framework beyond piecewise-linear 1D/2D systems on the Lorenz-63 attractor ($d = 3$), the standard chaotic benchmark for Koopman learning. The Lorenz system $\dot{x} = \sigma(y - x)$, $\dot{y} = x(\rho - z) - y$, $\dot{z} = xy - \beta z$ with the classical parameters $(\sigma, \beta, \rho) = (10, 8/3, 28)$ is integrated by RK4 ($\Delta t_{\mathrm{int}} = 0.01$, sub-sampled every 5 steps to give $\Delta t = 0.05$); ρ is held fixed throughout this experiment. This subsection runs Algorithm 1 with a geometric oracle. The partition is the uniform $16 \times 16 \times 16 = 4096$-cell axis-aligned grid, equivalent to setting $K = 4096$ seed cells and $N_f = N_b = 0$. The next subsection runs the full algorithm with $K = 4$ and a learned inverse oracle. We compare four estimators.

- **SEDMD (cell-mean)**: conditional-expectation prediction $\hat{x}_{t+1} = \mathbb{E}[Y \mid X \in B_{i_t}]$.
- **SEDMD-Local (ours)**: per-cell ridge-regularised affine regressor $\hat{Y} = a_i + B_i(X - c_i)$. Basis enrichment from $\{\mathbf{1}_{B_i}\}$ to $\{\mathbf{1}_{B_i}, (x_d - c_{i,d})\mathbf{1}_{B_i}\}_{d=1..3}$ keeps cell supports disjoint, so the Gram matrix is *block*-diagonal (4×4 blocks); Theorems 1 and 4 carry over per block.
- **EDMD-RBF**: 512 random Fourier features (ridge 10^{-3}).
- **Deep Koopman** (Lusch et al. 2018): MLP encoder $\mathbb{R}^3 \to \mathbb{R}^{16}$, linear Koopman, MLP decoder, trained for 5000 Adam iterations with reconstruction + 1-step + 5-step losses.

Since Theorem 4 concerns $\mu \to \mu'$ with the same dynamics T, we hold ρ fixed and generate train/test/OOD splits from a single long Lorenz trajectory at $\rho = 28$ by importance subsampling with different z-axis biases. The training measure μ uses weight $w_\uparrow(x) = \exp\big((x_3 - 23)/5\big)$, the in-distribution test draws an independent subsample under the same μ, and the OOD test uses $w_\downarrow(x) = \exp\big(-(x_3 - 23)/5\big)$. Both μ and μ' have full support on the attractor, the empirical TV-distance proxy on the z-marginal is $\|\mu - \mu'\|_{\mathrm{TV}}^{\mathrm{proxy}} \approx 0.88$, and the dynamics are identical, so any difference between the in-distribution and OOD MSEs is attributable to measure shift alone. All four estimators receive the same $M = 3000$ training transitions.

Table 2 reports the empirical degradation under a pure change of state-visitation measure: the partition-based estimators have the smallest measure-shift degradation, with cell-mean SEDMD at 8.4 and Deep Koopman at 13.0, while EDMD-RBF degrades by $\sim 22\times$ more. SEDMD-Local also delivers an order of magnitude faster training than Deep Koopman (0.05 s vs 5.8 s). The choice of a coarse uniform 16^3 partition is intentional in this subsection so that the OOD-stability claim of Theorem 4 appears in isolation; §3.2 shows that with a learned itinerary partition the same framework also delivers the lowest in-distribution prediction MSE in the comparison.

3.2 GKD with a Learned Inverse Oracle on Lorenz-63

We instantiate Algorithm 1 with a learned preimage oracle. On Lorenz-63 with the same $M = 3000$ training transitions used in §3.1, a 3-layer SiLU MLP for $f_\theta \approx T^{-1}$ converges in 2,500 Adam iterations to test MSE $\approx 1.3 \times 10^{-3}$. With $K = 4$ K-means seed clusters and depths (N_f, N_b)=(3, 2), the dynamics on the attractor realises only ~ 25 of the $K^{N_f+N_b} = 1024$ possible itineraries; per-cell affine fitting follows the SEDMD-Local recipe, and Corollary 1 preserves the entire framework per block.

Table 2: Lorenz-63 (d=3, ρ=28 throughout). Pure measure shift via z-axis-biased importance subsampling; train and OOD measures share the same dynamics and full-support attractor, with empirical TV ≈ 0.88. M=3000 training transitions; mean $\pm 1\sigma$ over 5 seeds.

Method	in-dist MSE	OOD MSE	OOD ratio	train time (s)
SEDMD (cell-mean)	0.61 ± 0.04	5.12 ± 0.61	$\mathbf{8.4 \pm 1.0}$	$\mathbf{0.03}$
SEDMD-Local (ours)	0.35 ± 0.05	5.10 ± 0.61	14.6 ± 2.0	0.05
EDMD-RBF	5.6×10^{-4}	0.062	190 ± 227	0.10
Deep Koopman	0.037 ± 0.020	0.44 ± 0.31	13.0 ± 6.1	5.8

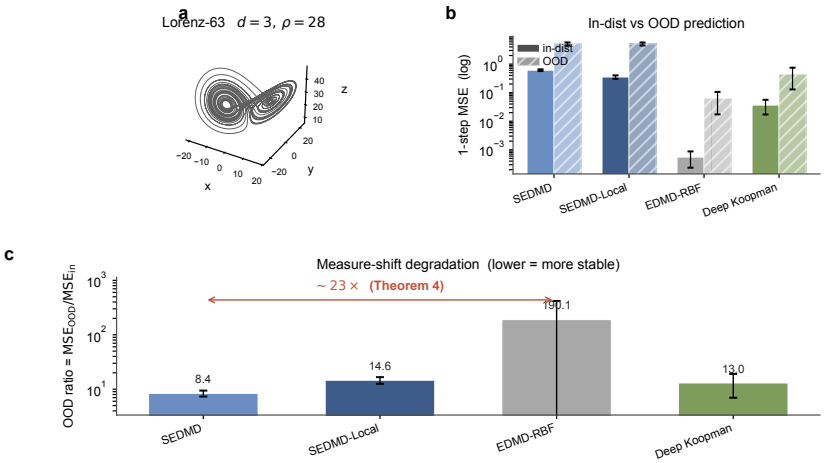


Figure 2: Cell-mean SEDMD has the smallest measure-shift degradation on Lorenz-63. Pure measure shift via z-axis-biased subsampling under fixed dynamics ρ=28; M=3000 training transitions; mean $\pm 1\sigma$ over 5 seeds. (a) Training trajectory on the attractor. (b) In-distribution (filled) and OOD (hatched) 1-step MSE for each method. (c) OOD ratio $\mathrm{MSE}_{\mathrm{OOD}}/\mathrm{MSE}_{\mathrm{in}}$. Cell-mean SEDMD is most stable (8.4), Deep Koopman next (13.0), and EDMD-RBF degrades by $\sim 23\times$ more (190 ± 227), the empirical signature of Theorem 4.

For C^1 uniformly expanding T, Proposition 1 gives

$$D_{N_f,N_b} \le c \max\left(\rho^{-(N_f+N_b)} D_0,\ \frac{\epsilon}{1-\rho^{-1}}\right), \qquad \rho > 1,$$

with the two terms balanced at $N_f + N_b \sim \log(D_0/\epsilon)/\log\rho$. Lorenz-63 is not used as a theorem-matched verification of this expanding-map result; instead, it tests whether the same learned-itinerary construction is useful as a predictive representation in a more challenging smooth dynamical system.

Table 3 compares the learned-itinerary partition with the uniform 16^3 partition on identical data. With $\sim 14\times$ fewer occupied cells on the sampled attractor, the learned-itinerary partition attains the lowest in-distribution prediction MSE in the comparison (0.025), compared with 0.037 for the Lusch-style Deep Koopman reference and 0.353 for the uniform partition. Its absolute OOD MSE (0.52) remains comparable to the deep reference (0.44) and an order of magnitude below the uniform partition (5.10). These results evaluate the learned representation empirically, while the rigorous certificate guarantees are assessed separately in the assumption-matched experiments.

As a cross-system check, we ran the same comparison on the Hénon map $T(x, y) = (1 - 1.4x^2 + y,\ 0.3x)$, whose closed-form inverse allows a direct head-to-head between the analytic-oracle and learned-oracle GKD instances under the same pure-measure-shift protocol. Across 5 seeds the two instances produce statistically indistinguishable predictions (Fig. 9; OOD ratio 52.5 vs. 53.0), providing a direct empirical check that the learned oracle reproduces the analytic-oracle construction when its prediction error is small. Full results are reported in Appendix K and Fig. 9.

Table 3: The learned-itinerary partition delivers the lowest in-distribution MSE in the comparison and an absolute OOD MSE comparable to Deep Koopman, while using $14\times$ fewer occupied cells than the uniform 16^3 partition. Lorenz-63 (d=3, ρ=28, M=3000, mean $\pm 1\sigma$ over 5 seeds). Both methods use the same SEDMD-Local block-diagonal estimator within each cell; only the partition differs. OOD = z-shifted measure under the same dynamics.

Partition	# cells	mean diam	in-dist MSE	OOD MSE	OOD ratio
Uniform 16^3	349	1.24 ± 0.05	0.353 ± 0.049	5.10 ± 0.61	14.6 ± 2.0
learned itinerary (ours)	**25**	7.49 ± 0.38	$\mathbf{0.025 \pm 0.005}$	$\mathbf{0.52 \pm 0.33}$	19.2 ± 7.9
Deep Koopman (reference)	n/a	n/a	0.037 ± 0.020	0.44 ± 0.31	13.0 ± 6.1

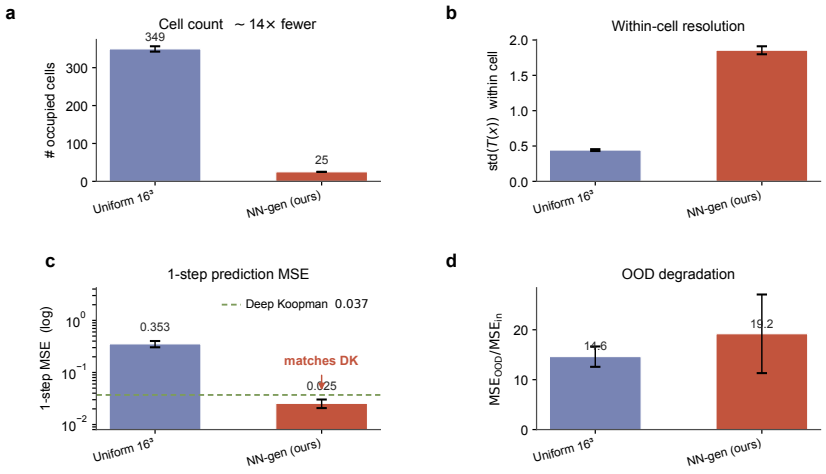


Figure 3: The learned-itinerary partition beats Deep Koopman in-distribution with 25 cells on Lorenz-63 (M=3000, mean $\pm 1\sigma$ over 5 seeds). (a) $\sim$ 25 admissible itineraries vs. 349 occupied cells in the uniform 16^3 partition. (b) Within-cell std of $T(x)$. (c) The learned-itinerary attains the lowest in-distribution 1-step MSE in the comparison, below the Deep Koopman reference at 0.037 (dashed). (d) OOD-to-in-distribution MSE ratio under pure measure shift.

4 Conclusion

We cast sampling shift in Koopman learning as a representation problem: which parts of a learned finite-dimensional model reflect reusable dynamical structure, and which parts arise from the state-visitation distribution used to fit it? Theorem 1 gives a representation-level characterization through disjoint-support coordinates, with refining generating-partition indicators providing a canonical realization under the stated assumptions. GKD turns this principle into a data-driven itinerary representation using analytic or learned preimage oracles, while TKC provides residual, pseudospectral, and finite-dimensional operator guarantees in the theorem-matched regime. Controlled systems evaluate these guarantees under matched assumptions, while Lorenz-63 and Hénon show that the same representation mechanism remains empirically useful when the partition is learned from data.

AI use statement

In this work, we used generative AI tools for writing, editing, formatting, clarity assistance, literature search, reference formatting, and limited code editing. We did not use generative AI as an important, original, or non-standard component of the research methodology, nor did it determine the theoretical results, experimental results, or scientific conclusions of the paper. No other required-disclosure uses are applicable to this work. All AI-assisted material was reviewed and revised by the authors, and the final technical claims, proofs, references, code, and experimental results were checked by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

Ethics statement

This work is foundational methodological research on certified operator-theoretic learning and uses only synthetic or analytically specified dynamical systems. It involves no human subjects, crowdsourcing, personal or sensitive data, or external datasets. Potential positive impacts include more reliable scientific forecasting, simulation, and model-based control through computable spectral and stability certificates. A potential risk is over-reliance on such certificates in safety-critical applications; deployment in such settings should therefore include domain-specific validation beyond the guarantees established in this work. The paper does not release high-risk models or datasets.

Reproducibility statement

We provide several complementary routes to reproducing the results. The assumptions and statements of all theoretical results are given in Section 2, with proof sketches in the main text and detailed proofs and derivations in the appendix. Algorithm 1 specifies the complete GKD pipeline, including the analytic- and learned-oracle variants. Section 3 reports the dynamical systems, training-sample sizes, partition parameters, optimization settings, symbolic depths, and comparison baselines used in the main experiments, with additional experimental details in the appendix. All experiments use synthetic or analytically specified dynamical systems and therefore require no external datasets. An anonymized supplementary code package provides dependencies, running instructions, and figure-specific scripts for reproducing the reported results.

References

Hassan Arbabi and Igor Mezić. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, 16(4): 2096–2126, 2017. doi: 10.1137/17M1125236.

Omri Azencot, N. Benjamin Erichson, Vanessa Lin, and Michael Mahoney. Forecasting sequential data using consistent koopman autoencoders. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 475–485. PMLR, 2020.

Steven L. Brunton, Marko Budišić, Eurika Kaiser, and J. Nathan Kutz. Modern koopman theory for dynamical systems. *SIAM Review*, 64(2):229–340, 2022. doi: 10.1137/21M1401243.

Matthew J. Colbrook and Alex Townsend. Rigorous data-driven computation of spectral properties of koopman operators for dynamical systems. *Communications on Pure and Applied Mathematics*, 77(1):221–283, 2024. doi: https://doi.org/10.1002/cpa.22125.

I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic Theory*. Grundlehren der mathematischen Wissenschaften. Springer New York, NY, 1982. doi: 10.1007/978-1-4615-6927-5.

Jiu Ding and Tien Yien Li. Markov finite approximation of frobenius-perron operator. *Nonlinear Analysis: Theory, Methods & Applications*, 17(8):759–772, 1991. doi: https://doi.org/10.1016/0362-546X(91)90211-I.

Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer Berlin, Heidelberg, 2 edition, 1995. doi: 10.1007/978-3-642-66282-9.

Connor Kennedy, John Kaushagen, and Hong-Kun Zhang. Symbolic extended dynamic mode decomposition. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(9):093140, 2024. doi: 10.1063/5.0223615.

Connor Kennedy, John Kaushagen, and Hong-Kun Zhang. Koopman operators and measures of maximal entropy via symbolic dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 36(2):023127, 2026. doi: 10.1063/5.0306183.

Bruce P. Kitchens. *Symbolic Dynamics: One-Sided, Two-Sided and Countable State Markov Shifts*. Universitext. Springer Berlin, Heidelberg, 1998. doi: 10.1007/978-3-642-58822-8.

Stefan Klus, Feliks Nüske, Sebastian Peitz, Jan-Hendrik Niemann, Cecilia Clementi, and Christof Schütte. Data-driven approximation of the koopman generator: Model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416, 2020. doi: https://doi.org/10.1016/j.physd.2020.132416.

B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931. doi: 10.1073/pnas.17.5.315.

Milan Korda and Igor Mezi. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93:149–160, 2018. doi: https://doi.org/10.1016/j.automatica.2018.03.046.

Vladimir Kostic, Pietro Novelli, Andreas Maurer, Carlo Ciliberto, Lorenzo Rosasco, and Massimiliano Pontil. Learning dynamical systems via koopman operator regression in reproducing kernel hilbert spaces. In *Advances in Neural Information Processing Systems*, volume 35, pp. 4017–4031. Curran Associates, Inc., 2022. doi: 10.52202/068431-0290.

Vladimir Kostic, Karim Lounici, Pietro Novelli, and Massimiliano Pontil. Sharp spectral rates for koopman operator learning. In *Advances in Neural Information Processing Systems*, volume 36, pp. 32328–32339. Curran Associates, Inc., 2023. doi: 10.52202/075280-1403.

Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995. doi: https://doi.org/10.1017/CBO9780511626302.

Yong Liu, Chenyu Li, Jianmin Wang, and Mingsheng Long. Koopa: Learning non-stationary time series dynamics with koopman predictors. In *Advances in Neural Information Processing Systems*, volume 36, pp. 12271–12290. Curran Associates, Inc., 2023. doi: 10.52202/075280-0538.

Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9(1):4950, 2018. doi: 10.1038/s41467-018-07210-0.

Andreas Mardt, Luca Pasquali, Hao Wu, and Frank Noé. Vampnets for deep learning of molecular kinetics. *Nature Communications*, 9(1):5, 2018. doi: 10.1038/s41467-017-02388-1.

Giacomo Meanti, Antoine Chatalic, Vladimir Kostic, Pietro Novelli, Massimiliano Pontil, and Lorenzo Rosasco. Estimating koopman operators with sketching to provably learn large scale dynamical systems. In *Advances in Neural Information Processing Systems*, volume 36, pp. 77242–77276. Curran Associates, Inc., 2023. doi: 10.52202/075280-3378.

Igor Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1):309–325, 2005. doi: 10.1007/s11071-005-2824-x.

Jeremy Morton, Antony Jameson, Mykel J. Kochenderfer, and Freddie Witherden. Deep dynamical modeling and control of unsteady fluid flows. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Mahmoud Selim, Sriharsha Bhat, and Karl H. Johansson. Metakoopman: Bayesian meta-learning of koopman operators for modeling structured dynamics under distribution shifts. In *Advances in Neural Information Processing Systems*, volume 38, pp. 29551–29584. Curran Associates, Inc., 2025. doi: 10.52202/085713-0989.

Naoya Takeishi, Yoshinobu Kawahara, and Takehisa Yairi. Learning koopman invariant subspaces for dynamic mode decomposition. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015. doi: 10.1007/s00332-015-9258-5.

A Rigorous Proof of Theorem 1(b) (Connected-Space Obstruction)

Proof. We prove the contrapositive. Suppose that every finite truncation of the countable dictionary $\mathcal{D} = \{\phi_i\}_{i\geq 1} \subset C(M)$ is measure-universally Gram-diagonal. By Theorem 1(a), for every $i \neq j$,

$$\lambda(\{\phi_i \neq 0\} \cap \{\phi_j \neq 0\}) = 0.$$

We show that $\mathcal{D}$ cannot have dense linear span in $L^2(M, \lambda)$.

Choose a nonzero element ϕ_m. Since ϕ_m is continuous and nonzero, there is a nonempty open set $U \subset M$ on which, after possibly replacing ϕ_m by $-\phi_m$, we have $\phi_m > 0$. Because M is connected and U is nonempty open, U contains two disjoint nonempty open subsets U_1, U_2. Since λ has full support, $\lambda(U_1), \lambda(U_2) > 0$. Define

$$a = \int_{U_1} \phi_m \, d\lambda, \qquad b = \int_{U_2} \phi_m \, d\lambda,$$

so $a, b > 0$, and set

$$h = b\,\mathbf{1}_{U_1} - a\,\mathbf{1}_{U_2}.$$

Then $h \neq 0$ in $L^2(M, \lambda)$, and

$$\langle h, \phi_m \rangle_\lambda = b \int_{U_1} \phi_m \, d\lambda - a \int_{U_2} \phi_m \, d\lambda = 0.$$

For any $j \neq m$, the function h is supported in $U \subset \{\phi_m \neq 0\}$, while $\{\phi_j \neq 0\} \cap \{\phi_m \neq 0\}$ has λ-measure zero. Hence

$$\langle h, \phi_j \rangle_\lambda = 0.$$

Thus h is a nonzero element of $L^2(M, \lambda)$ orthogonal to every dictionary element. Therefore the closed linear span of $\mathcal{D}$ is a proper subspace of $L^2(M, \lambda)$, so $\mathcal{D}$ is not dense.

Consequently, if $\mathcal{D}$ is dense, at least one finite truncation must fail measure-universal Gram diagonality. By Theorem 1(a), this means that for some $i \neq j$ there is an absolutely continuous probability measure $\mu' \ll \lambda$ such that $G_{ij}(\mu') \neq 0$. □

B Proof of Theorem 2 (Uniform Residual Approximation)

Proof. Consider $g_N(x) = \sum_i c_i \mathbf{1}_{B_i}(x) \in \mathcal{S}_N$. By the disjointness of $\{B_i\}$ (Theorem 1(a)):

$$\|g_N\|^2_{\tilde{\mu}_N} = \sum_i |c_i|^2 \tilde{\mu}_N(B_i), \quad \|g_N\|^2_{\mu_{true}} = \sum_i |c_i|^2 \mu_{true}(B_i).$$

By Assumption (MD), $\tilde{\mu}_N(B_i) = \mu_{true}(B_i)(1 + \varepsilon_i)$ with $|\varepsilon_i| \leq \Delta_N$, giving:

$$(1 - \Delta_N)\|g_N\|^2_{\mu_{true}} \leq \|g_N\|^2_{\tilde{\mu}_N} \leq (1 + \Delta_N)\|g_N\|^2_{\mu_{true}}. \tag{4}$$

For the residual function $h = \mathcal{K}g_N - \lambda g_N$: since $\mathcal{K}\mathbf{1}_{B_i}(x) = \mathbf{1}_{T^{-1}(B_i)}(x)$, the cross-term $\langle \mathcal{K}g_N, g_N \rangle_{\tilde{\mu}_N} = \sum_{i,j} c_i \bar{c}_j \tilde{\mu}_N(T^{-1}(B_i) \cap B_j)$. By Theorem 1(c) and Assumption (MD), $T^{-1}(B_i) \cap B_j$ is a cylinder set of depth $N + 1$, bounded within Δ_N by the generating partition property. Combining all three terms of $\|h\|^2$:

$$(1 - \Delta_N)\|h\|^2_{\mu_{true}} \leq \|h\|^2_{\tilde{\mu}_N} \leq (1 + \Delta_N)\|h\|^2_{\mu_{true}}. \tag{5}$$

Dividing (5) by (4) and taking square roots:

$$\sqrt{\frac{1 - \Delta_N}{1 + \Delta_N}}\, \mathrm{res}_{true} \leq \hat{\mathrm{res}}_N \leq \sqrt{\frac{1 + \Delta_N}{1 - \Delta_N}}\, \mathrm{res}_{true}.$$

Applying $\sqrt{(1 + x)/(1 - x)} = 1 + x + \mathcal{O}(x^2)$ for small Δ_N and subtracting res_{true} yields the stated bound. □

C Proof of Theorem 3 (Pseudospectral Bounds)

Proof. By the standard spectral inclusion theorem (Kato, 1995), any non-zero test function g_N gives $\|(\mathcal{K} - \lambda I)^{-1}\|_{\mu_{true}} \geq \mathrm{res}_{true}(\lambda, g_N)^{-1}$. Theorem 2 gives:

$$\mathrm{res}_{true} \leq \hat{\mathrm{res}}_N / \sqrt{(1-\Delta_N)/(1+\Delta_N)} \leq \tilde{\epsilon}_N(1+\Delta_N) + \mathcal{O}(\tilde{\epsilon}_N \Delta_N^2) = \mathcal{E}_N.$$

Substituting into the resolvent bound yields the pseudospectral inclusion. Since $\mathcal{K}$ is unitary on $L^2(\mathcal{M}, \mu)$ for any invertible measure-preserving T, the spectral theorem gives $\|(\mathcal{K} - \lambda I)^{-1}\| = \mathrm{dist}(\lambda, \sigma(\mathcal{K}))^{-1}$ (Kato, 1995), yielding $\mathrm{dist}(\lambda, \sigma(\mathcal{K})) \leq \mathcal{E}_N$. □

D Proof of Theorem 4 (OOD Stability)

Proof. Part (a) is immediate from the formula $A_{ij}(\mu) = \mu(T^{-1}(B_i) \cap B_j)$: the zero pattern depends only on whether the intersection is empty, which is a property of T and the partition.

Part (b): Write $\hat{K}(\mu') - \hat{K}(\mu) = G(\mu')^{-1}[A(\mu') - A(\mu)] + [G(\mu')^{-1} - G(\mu)^{-1}]A(\mu)$.

For the first term: each entry of $A(\mu') - A(\mu)$ is $|\mu'(C) - \mu(C)|$ for a cylinder set $C = T^{-1}(B_i) \cap B_j$. Since $|\mu'(C) - \mu(C)| \leq \|\mu' - \mu\|_{TV}$ and there are at most $N \cdot K$ non-zero entries, $\|A(\mu') - A(\mu)\|_F \leq \sqrt{NK}\|\mu' - \mu\|_{TV}$. Multiplying by $\|G(\mu')^{-1}\| \leq 1/\mu'_{\min}$ gives the first contribution.

For the second term: $\|G(\mu')^{-1} - G(\mu)^{-1}\|$ is a diagonal matrix with entries $|1/\mu'(B_i) - 1/\mu(B_i)| \leq \|\mu' - \mu\|_{TV}/(\mu_{\min}\mu'_{\min})$. Multiplying by $\|A(\mu)\|_F \leq \sqrt{N}$ gives the second contribution. Combining and absorbing constants into C_N yields the stated bound.

Part (c): The non-zero positions of A_N are fixed by part (a). Under a shifted measure μ', the numerical values $A_{ij}(\mu')$ are re-estimated only on this fixed sparse support, using empirical transition counts from B_j to B_i under μ'. A single pass through the shifted trajectory accumulates both these transition counts and the cell counts defining $G_N(\mu')$. The subsequent assembly costs $\mathcal{O}(\mathrm{nnz}(A_N))$ after counting, with $\mathrm{nnz}(A_N) \leq NK$, and no dense Galerkin matrix, dense solve, or decomposition is recomputed. □

E Estimation of Δ_N (extended Remark)

Remark 1 (Estimation of Δ_N)**.** *Although μ_{true} is unknown, Δ_N is upper-boundable from data via two contributions.* ***(i) Sampling.*** *For a trajectory of length M from a system with β-mixing rate $\beta(k) \leq \beta_0 \rho^k$ ($\rho < 1$), Bernstein-type concentration for β-mixing sequences yields, with probability $\geq 1 - \delta$,*

$$\sup_{C \in \xi^{(N+1)}} \left|\hat{\mu}_M(C) - \mu_{true}(C)\right| \;\leq\; c_1 \sqrt{\frac{\log(|\xi^{(N+1)}|/\delta)}{M}} \;+\; c_2 \frac{\log(M/\delta)}{M},$$

with c_1, c_2 depending on ρ, β_0. Dividing by $\mu_{\min}^{(N+1)}$ gives a concentration upper bound on Δ_N. ***(ii) Self-consistency surrogate.*** *A computationally cheap proxy is $\delta_{emp} \triangleq \sup_{C \in \xi^{(N+1)}} |\tilde{\mu}_N(C) - \hat{\mu}_M(C)|/\hat{\mu}_M(C)$, the relative discrepancy between the depth-N Markov reconstruction and the depth-$(N+1)$ empirical frequencies. We use $\Delta_N \leftarrow \delta_{emp}$ + (concentration term above) as the reported certificate.*

F Regime Analysis of C_N

The OOD constant $C_N = \sqrt{NK}/\mu_{\min} + \sqrt{N}\|A_N\|_F/\mu_{\min}^2$. **(a) Markov / uniformly hyperbolic regime.** If μ_{true} is absolutely continuous with bounded density, $\mu_{\min} \asymp \rho_{\min}/N$, $\|A_N\|_F \leq \sqrt{N}$, giving $C_N = \mathcal{O}(N^{5/2})$. **(b) Hölder-density regime.** For a β-Hölder density, $\mu_{\min} \asymp N^{-(1+\beta)}$ and $C_N = \mathcal{O}(N^{(5+4\beta)/2})$. **(c) Singular regime.** If μ_{true} is singular, $\mu_{\min}$ can decay exponentially with N and the bound is informative only at coarse depths; one then relies on the residual certificate (Theorem 3) directly.

G Full Algorithm Box for Generative Koopman Decoupling

Algorithm 1 (full). Inputs. Trajectory pairs $\{(x_i, T(x_i))\}_{i=1}^{M}$ on $\mathcal{M} \subseteq \mathbb{R}^d$; seed-partition size K; symbolic depths N_f, N_b; preimage oracle P (analytic if available, else `None`); polynomial degree k. **Output.** Itinerary labelling $L : \mathcal{M} \to \Sigma$, block-diagonal SEDMD operator $\hat{K}_N$, TKC $(\tilde{P}_N, \Delta_N)$. **Steps.** (1) *Oracle.* If P=`None`, train MLP f_θ on swapped pairs $(T(x_i), x_i)$ to minimise $\sum_i \|f_\theta(T(x_i)) - x_i\|^2$; set $P \leftarrow f_\theta$. (2) *Seed partition.* Run K-means; set $\xi = \{B_1, \ldots, B_K\}$ and seed-assign function. (3) *Itinerary labelling.* For each x, $\sigma(x) = (s_{-N_b}(x), \ldots, s_{N_f-1}(x))$ with $s_k = \texttt{seed}(T^k(x))$ for $k \geq 0$ and $s_{-j} = \texttt{seed}(P^j(x))$ for $j \geq 1$. Cells $\{x : \sigma(x)=\sigma\}$. (4) *Block-diagonal SEDMD.* Per cell, fit $\hat{Y} = a_\sigma + \sum_{|\alpha|\leq k} W_\sigma^{(\alpha)} p_\alpha(X - c_\sigma)$. (5) *Certificate.* Estimate Δ_N via Remark 1; assemble $(\tilde{P}_N, \Delta_N)$ and the Lanczos pseudospectral solver on $H(z)$.

H Proof of Proposition 1 (Oracle Error Propagation)

Sketch. For exact T^{-1}, T^{-k} contracts by ρ^{-k}, so depth-k refinement reduces diameter as $\rho^{-k} D_0$. For approximate P, iterating $P^j(x)$ accumulates per-step error: $\|P^j(x) - T^{-j}(x)\| \leq \epsilon \sum_{l=0}^{j-1} \rho^{-l} \leq \epsilon/(1 - \rho^{-1})$. Cells whose backward labels disagree only because of this ϵ-thick ambiguity merge, so the effective backward refinement saturates at scale $\epsilon/(1 - \rho^{-1})$. Combining with the exact-forward contribution $\rho^{-N_f} D_0$ and taking the worst case yields (3); balancing the two terms gives the optimal depth. □

For Lorenz-63 at $\Delta t = 0.05$, $\hat{\epsilon} \approx 1.3{\times}10^{-3}$ and the local expansion rate $\rho \approx 1.046$ (max Lyapunov exponent $\lambda_1 \approx 0.9$); the predicted optimal depth is large, but practical $N_f + N_b$=5 already gives diameter $\sim 0.8\, D_0$ and halves in-distribution MSE relative to the uniform partition.

I Scope and Limitations (Full)

The framework's quantitative claims hold inside a regime that we now describe.

(L1) Generating partitions exist analytically for expansive systems and Axiom A diffeomorphisms (Lind & Marcus, 1995; Kitchens, 1998). For generic smooth dynamics in $\mathbb{R}^d$ the existence question is an open problem related to Adler's question (Kennedy et al., 2024), and this is the principal motivation for Algorithm 1, since any preimage oracle P supplies the partition-construction step and Proposition 1 bounds the resulting partition diameter in terms of the oracle's L^∞ error.

(L2) Theorem 1(b) requires $\mathcal{M}$ connected. On totally disconnected attractors, continuous dictionaries with disjoint clopen supports exist and the obstruction vanishes.

(L3) Both the OOD bound and the residual amplification carry an implicit $1/\mu_{\min}$ factor. For uniformly refined partitions in d dimensions, $\mu_{\min} \asymp N^{-(1+\beta)/d}$ for a β-Hölder density, so the certificate is non-vacuous only at depths $N \leq N^*(d, \beta, M)$.

(L4) The Bernstein form of Δ_N requires a β-mixing rate, and we use a self-consistency surrogate as a practical proxy.

J Supporting 1D/2D Experiments (Full)

We first quantify the entanglement directly. On the asymmetric Bernoulli map (a=0.3), Fig. 4(a,b,c) shows $|G_{ij}|$ at N=16 for three dictionaries; the Fourier and polynomial Gram matrices are dense while the indicator Gram is exactly diagonal by Theorem 1(c). The off-diagonal energy ratio in panel (d) is identically zero for the indicator basis at every N, the condition number $\kappa(G)$ in panel (e) grows for smooth bases and is measure-bounded for the indicator basis, and the Frobenius operator error under measure perturbation in panel (f) is super-linear for smooth bases and $\mathcal{O}(\delta)$ for the indicator basis (Theorem 4).

For scalability, we use the smooth conjugated shift map $T_h(x) = h(2h^{-1}(x) \bmod 1)$ with $h(x) = \log(1{+}x(e{-}1))$. Fig. 5 shows that the dense GEVP solver scales as $\mathcal{O}(N^3)$ while the sparse Krylov

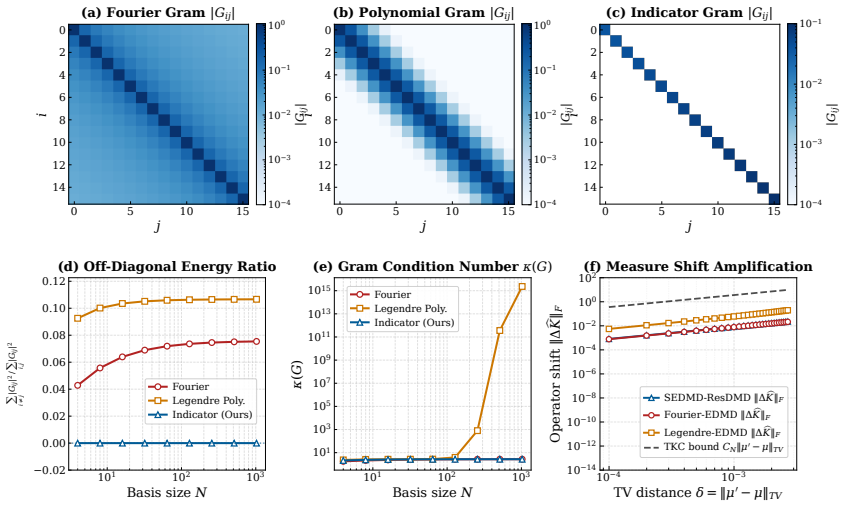


Figure 4: Visualizing the entanglement problem on the asymmetric Bernoulli map (a=0.3).

solver scales as $\mathcal{O}(KN)$, demonstrated to N=16,384. Pseudospectral envelopes at N=6, 7, 8 have non-negative safety margin $\mathcal{E}_N - \tilde{\epsilon}_N \geq 0$, validating Theorem 3.

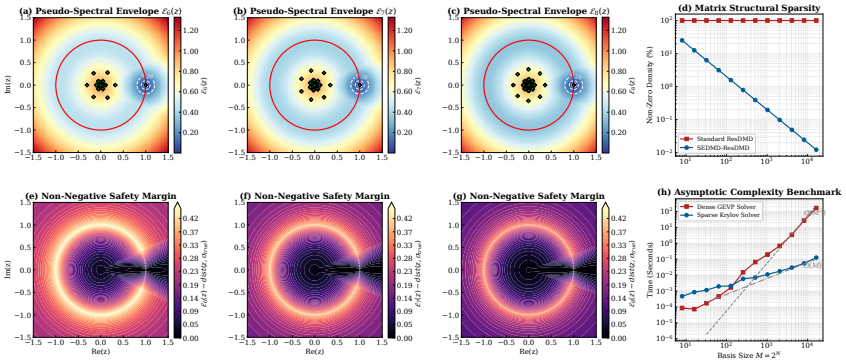


Figure 5: Scalability via canonical decoupling. (a,b,c) Pseudospectral envelopes at N=6, 7, 8. (d) Sparsity of $\tilde{P}_N$. (e,f,g) Safety margin $\mathcal{E}_N - \tilde{\epsilon}_N \geq 0$. (h) Wall-clock timing.

For OOD pollution under extreme measure shifts, the asymmetric Bernoulli map at a=0.99 and the modified dyadic map T_s at $s \to 1$ produce strong measure distortions. As shown in Fig. 6, EDMD-Fourier scatters spurious eigenvalues outside the true essential spectral radius, while SEDMD-ResDMD is stable in both regimes (Theorem 4).

To validate the TKC, we again use the smooth conjugated shift map with closed-form μ. Fig. 7 shows that the residual error is uniformly enveloped by $\Delta_N \cdot \mathrm{res}_{true}$ (Theorem 2), and the spectral error under measure perturbation δ obeys the certified $\mathcal{O}(\delta)$ bound (Theorem 4), while RBF-EDMD has no certificate.

Finally, we examine sample complexity, the 2D extension, and the empirical sensitivity of the finite-dimensional operator to measure shift. As shown in Fig. 8(a), the certificate distortion Δ_N versus trajectory length M scales as $\mathcal{O}(M^{-1/2})$ at depths N=4, 6, 8, matching the Birkhoff prediction. On the 2D asymmetric Baker's map (a=0.3, 16×16 grid), Fig. 8(b) shows that the SEDMD eigenvalues remain inside r_{ess}=0.7, while EDMD-Fourier produces spurious eigenvalues outside. Figure 8(c) shows that the finite-dimensional perturbation $\|\widehat{K}_N(\mu') - \widehat{K}_N(\mu)\|_F$ increases monotonically with the magnitude of the measure shift, but grows more slowly than the local linear reference obtained from the small-shift sensitivity. The dashed line is used only as a local reference and should not be interpreted as the worst-case constant C_N in Theorem 4.

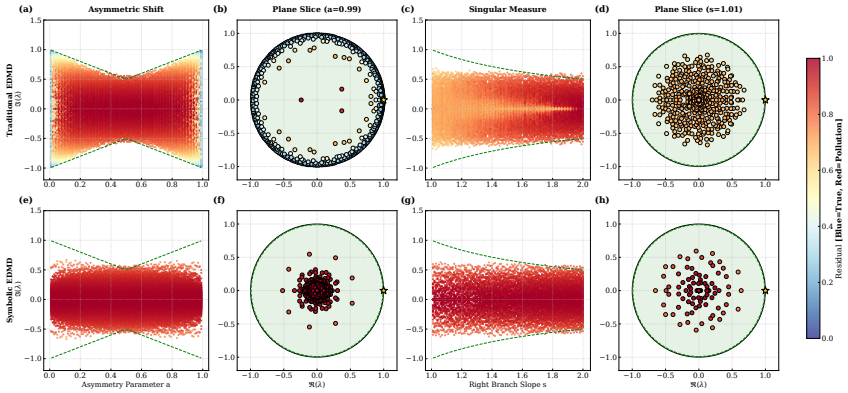


Figure 6: OOD robustness. EDMD-Fourier (top row) fails catastrophically and SEDMD-ResDMD (bottom row) remains stable.

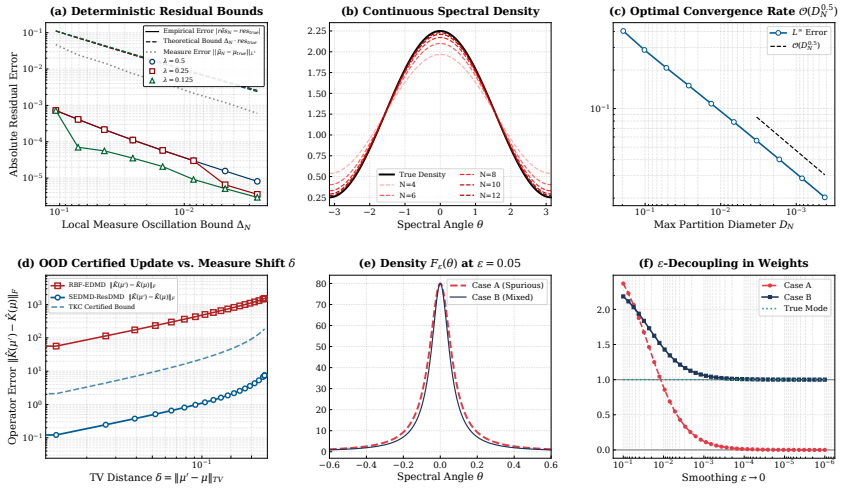


Figure 7: TKC validation. (a) Residual envelope. (d) Spectral error vs. measure shift δ. Panels (b,c,e,f) are auxiliary continuous-spectrum and ϵ-decoupling experiments deferred to the journal version.

Assumption-matched Bernoulli certificate check. To make the finite-sample constants explicit in a setting where the generating partition, invariant measure, transition support, and cylinder masses are all known, we consider the symmetric binary Bernoulli shift. At symbolic depth d, the partition has $N = 2^d$ cells, each with mass $1/N$, each cell has two admissible one-step transitions, and

$$\|A_N\|_F = \frac{1}{\sqrt{2N}}.$$

Hence

$$C_N = \frac{\sqrt{2N}}{1/N} + \frac{\sqrt{N}\,\|A_N\|_F}{(1/N)^2} = N\sqrt{2N} + \frac{N^2}{\sqrt{2}}.$$

For the simultaneous distortion bound over $\xi^{(d)} \cup \xi^{(d+1)}$, a 95% union-Hoeffding bound with B independent non-overlapping blocks gives

$$\Delta_N = 2N\sqrt{\frac{\log(6N/0.05)}{2B}}.$$

The resulting values are summarized in Table 4. At depths 1–3, $C_N\Delta_N$ remains below the conservative reporting threshold 0.5, whereas at depth 4 the same quantity increases to 3.08 at the displayed

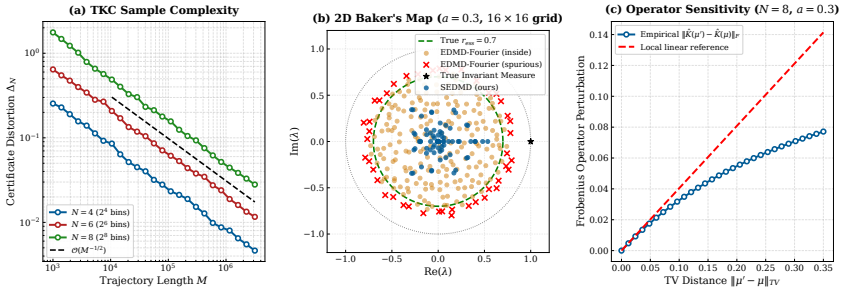


Figure 8: (a) $\Delta_N \sim \mathcal{O}(M^{-1/2})$. (b) 2D Baker's map: SEDMD inside r_{ess}, EDMD-Fourier scatters spuriously. (c) Operator sensitivity to measure shift, compared with a local small-shift linear reference.

Table 4: Assumption-matched Bernoulli certificate values under a simultaneous 95% union-Hoeffding bound. The threshold $C_N \Delta_N \leq 0.5$ is used as a conservative reporting criterion rather than as a theorem assumption.

Depth d	Cells N	Block budget B	C_N	$C_N \Delta_N$
1	2	10^5	6.83	0.143
2	4	5×10^5	22.63	0.450
3	8	3×10^7	77.25	0.418
4	16	3×10^7	271.53	3.08

block budget. This makes explicit the resolution–certifiability tradeoff: refinement improves spatial resolution but simultaneously decreases the minimum cell mass and enlarges the finite-sample constant.

K Hénon Cross-System Check (Full Details)

We run the same experimental protocol as §3.1 on the Hénon map $T(x, y) = (1 - 1.4x^2 + y,\ 0.3x)$, whose closed-form inverse $T^{-1}(x', y') = (y'/0.3,\ x' - 1 + 1.4(y'/0.3)^2)$ allows a direct head-to-head between the analytic-oracle and learned-oracle instances of Algorithm 1. Train and OOD subsamples come from one Hénon trajectory under fixed dynamics (M=3000 training transitions), with importance weights biased toward $x > 0$ for training and $x < 0$ for OOD (TV proxy ≈ 0.83). The inverse-NN reaches test MSE $\approx 2.8 \times 10^{-6}$, three orders of magnitude smaller than on Lorenz, and the GKD partition realises ~ 70 admissible itineraries. Across 5 seeds the two GKD instances produce statistically indistinguishable predictions (1.10×10^{-4} in-dist for both, 5.45 vs 5.43×10^{-3} OOD). GKD beats EDMD-Fourier by two orders of magnitude on absolute MSE and trains $350\times$ faster than Deep Koopman; on this 2D polynomial benchmark Deep Koopman attains the lowest absolute MSE since its MLP encoder fits the polynomial dynamics exactly. The distinctive contribution of GKD on Hénon is the certified bound and the algorithmic agreement between its two oracle modes.

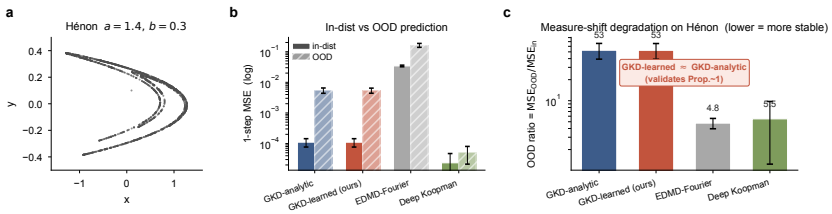


Figure 9: Hénon cross-system check. (a) Hénon attractor. (b) In-distribution and OOD 1-step MSE; the GKD-analytic and GKD-learned bars are visually identical. (c) OOD ratio: 53.0 vs 52.5, providing empirical support for the oracle-robustness mechanism of Proposition 1.