FROM CROSS MAPPING TO STABLE DYNAMICAL FACTORS: LEARNING ALIGNED DYNAMICAL STRUCTURE IN NONLINEAR SYSTEMS

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning aligned dynamical structure from paired nonlinear time series calls for a population correspondence alongside predictive skill. Cross-map prediction leaves this structural target unspecified. We define an aligned stable dynamical factor: a map that aligns paired states, commutes with dynamics, transports transmitted tangent directions, and controls metric-relative conditioning. Atlas GCCM–BCD and Neural GCCM estimate this relation with hard local and global differentiable surrogates, respectively. Matched-budget evidence isolates lower state and tangent incompatibility from structured losses, without uniform factor-error or categorical-direction gains; geometry and rank differences remain unresolved. The target uses *latent dynamical manifolds* (LDMs), combining reconstructed state, induced dynamics, a metric, and invariant measure. We prove reconstruction transport of an assumed factor and conditional regularized-target consistency for idealized globally smooth estimators under dependent sampling, plug-in consistency, risk separation, and approximate global optimization; hard-chart Atlas and no-factor consistency remain outside the result. An ancestry reading is external and conditional on ordered generative dynamics, component-local sensing, causal sufficiency, and absence of observational equivalence; otherwise the relation describes aligned dynamical information containment. Synthetic studies examine factor recovery, matched-independent negative controls, adverse feedback, synchronization, robustness, controlled observation-map dimension and conditioning, representation diagnostics, and targeted external baselines.

1 INTRODUCTION

Convergent cross mapping (CCM) uses delay reconstruction and cross-map prediction skill as evidence that one history contains information about another (Sugihara et al., 2012). Prediction skill and library-size convergence do not specify the population correspondence when each series preserves only part of hidden dynamics.

We make that target explicit as a population aligned stable dynamical factor. Each history induces an observable quotient, represented as $\mathfrak{L}_i = (\mathcal{M}_i, \widetilde{F}_i, g_i, \nu_i)$ with reconstructed state, dynamics, metric, and invariant measure. The factor aligns paired states, commutes with dynamics, transports transmitted tangent directions, and supports metric-relative conditioning diagnostics.

We retain $X \to_G Y$ for aligned dynamical information containment. Restricted ancestry requires the external Convention 4.2; direct edges, intervention effects, confounder robustness, and synchronized-system orientation are outside the factor relation alone.

Atlas GCCM–BCD uses hard local charts and validation-stopped block-coordinate updates; Neural GCCM uses a global differentiable parameterization. Their distinct frozen surrogates estimate the same relation class.

Theory transports an assumed factor through delay reconstruction and gives regularized-target consistency for idealized globally smooth estimators under uniform laws, risk separation, and approximate global optimization. Section 6 specifies estimator applicability.

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.

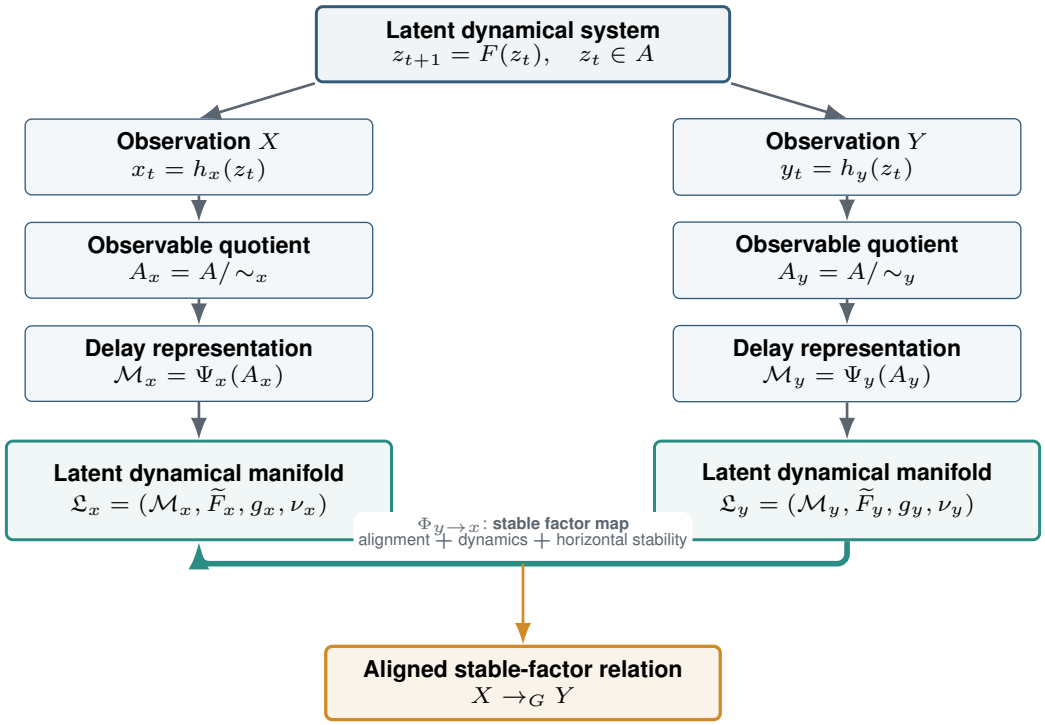


Figure 1: Overall framework. Each history induces an observable quotient and delay representation. An LDM augments the reconstructed manifold with dynamics, a chosen metric, and measure. The mathematical relation $X \to_G Y$ is witnessed by the reverse aligned factor $\Phi_{y\to x}$; ancestry is an external interpretation under Assumption 4.1 and Convention 4.2.

Synthetic studies characterize recovery, negative controls, feedback, synchronization, robustness, observation maps, representation diagnostics, and targeted baselines. Matched-budget results isolate state and tangent compatibility gains; factor-error, geometry, and rank advantages remain unresolved, and categorical outcomes are unchanged.

Our contributions are:

1. **Population target.** Define an aligned stable-factor relation beyond cross-map prediction skill.
2. **Structured estimators.** Introduce Atlas GCCM–BCD and Neural GCCM as hard local and global differentiable estimators of that relation class.
3. **Conditional theory.** Derive reconstruction transport and regularized-target consistency for assumed factors and idealized globally smooth estimators.
4. **Falsifiable evaluation.** Characterize recovery, negative controls, synchronization, robustness, representations, targeted baselines, and matched-budget structural effects.

2 Related Work

Reconstruction and dynamical dependence. Delay embeddings support state reconstruction (Takens, 1981; Sauer et al., 1991), while stable embeddings distinguish injectivity from conditioning (Eftekhari et al., 2016). CCM uses cross-map prediction and library-size convergence as evidence (Sugihara et al., 2012); GCCM specifies a population factor target. Topological Causality measures local expansion (Harnack et al., 2017), whereas GCCM regularizes horizontal derivatives of a

dimension-reducing factor. Generalized synchronization can make both observational directions recoverable under one-way forcing (Kocarev & Parlitz, 1996).

Dynamical representations. State-space, neural-ODE, predictive-state, and Koopman models describe individual systems (Karl et al., 2017; Chen et al., 2018; Littman et al., 2001; Lusch et al., 2018); L2CDS, conjugate mappings, and DSA address cross-system correspondence or similarity (Kim et al., 2020; Bramburger et al., 2021; Ostrow et al., 2023).

Recent work scales operator learning by sketching (Meanti et al., 2023), learns probabilistic or spectral representations (Bevanda et al., 2025; Xu et al., 2025), and compares systems through Koopman spectra (Zhang et al., 2025a). Structured recurrent, continuous-time, and disentangled models target forecasting (Zhang et al., 2025b; Bai & Ding, 2025; Fotiadis et al., 2023). Our target is instead a directed, potentially non-invertible paired-state map with derivative compatibility and metric-relative conditioning (Appendix H). A shared mapping-valued benchmark remains future work.

Temporal causal discovery. Structural models and conditional independence target intervention semantics or graph equivalence classes (Pearl, 2009; Spirtes et al., 2000); Granger causality and transfer entropy quantify predictive dependence (Granger, 1969; Schreiber, 2000).

Recent graph methods address amortized inference (Löwe et al., 2022), latent contexts and confounders (Günther et al., 2023; Yehezkel Rohekar et al., 2023), and irregular sampling or subsampling (Cheng et al., 2024; Wu et al., 2024). CADYT models continuous-time causal dynamics (Tagliapietra et al., 2026), while an industrial benchmark provides expert-constructed graph evidence (Mogensen et al., 2024). Graphs and causal orderings complement our mapping-valued relation between paired reconstructed systems; these estimation targets are not interchangeable.

Manifold learning. Low-dimensional coordinates and diffusion geometry (Tenenbaum et al., 2000; Roweis & Saul, 2000; Coifman & Lafon, 2006), with local tangent estimation (Aamari & Levrard, 2019), support our representation; LDMs additionally specify dynamics, metric, and measure.

3 Problem Formulation

3.1 Latent and observable systems

Let A be a compact invariant state set, let $F : A \to A$ be an invertible discrete-time evolution map, and let μ be an ergodic F-invariant probability measure. We observe paired trajectories

$$z_{t+1} = F(z_t), \qquad x_t = h_x(z_t), \qquad y_t = h_y(z_t). \tag{1}$$

An observable need not distinguish the full latent state. For $i \in \{x, y\}$, define history equivalence

$$z \sim_i z' \quad \Longleftrightarrow \quad h_i(F^k z) = h_i(F^k z') \quad \text{for every admissible } k. \tag{2}$$

The observable state space is $A_i = A/\!\sim_i$, with quotient projection $\varpi_i : A \to A_i$. Shift invariance of the equivalence relation induces dynamics $F_i : A_i \to A_i$ and a marginal invariant measure $\mu_i = (\varpi_i)_{\#}\mu$. Appendix B gives the construction and its regularity assumptions.

3.2 Delay reconstruction

For delay dimension E_i and lag τ_i, define

$$\Psi_i(a) = \big(h_i(a), h_i(F_i^{-\tau_i} a), \ldots, h_i(F_i^{-(E_i-1)\tau_i} a)\big). \tag{3}$$

When Ψ_i is an embedding, it reconstructs the observable system as

$$\mathcal{M}_i = \Psi_i(A_i), \qquad \widetilde{F}_i = \Psi_i \circ F_i \circ \Psi_i^{-1}, \qquad \nu_i = (\Psi_i)_{\#}\mu_i. \tag{4}$$

Let ν_{xy} denote the stationary joint law of simultaneous reconstructed states (q_x, q_y). Given a finite paired trajectory $\{(x_t, y_t)\}_{t=1}^{L}$, our goal is to decide whether there exists a stable map from $\mathcal{M}_y$ to $\mathcal{M}_x$ and to estimate that map, its derivative, and its dynamical residuals.

4 ALIGNED STABLE FACTORS AND CONDITIONAL ANCESTRY INTERPRETATION

Assumption 4.1 (Structural conditions for a component-local dynamical model)**.** The structural model satisfies: **(S1)** ordered generative dynamics, e.g. $u_{t+1} = f(u_t)$ and $v_{t+1} = g(u_t, v_t)$ for a one-way pair; **(S2)** component-local sensing, $h_x(u, v) = \bar{h}_x(u)$ and $h_y(u, v) = \bar{h}_y(v)$; **(S3)** causal sufficiency at the analyzed component level; and **(S4)** no generalized synchronization, invertible conjugacy, observational equivalence, or response collapse that makes both directions recoverable. These are externally supplied structural conditions, not conclusions inferred from the two observed series.

Interpretation Convention 4.2 (Conditional ancestry interpretation)**.** The estimator targets an aligned stable factor. When independent domain knowledge supports Assumption 4.1, the relation may be described as compatible with restricted ancestry. This model-side convention is not inferred from the two series, is neither a theorem premise nor a proof of factor existence, and enters neither training nor decision gates.

Latent dynamical manifold. Let (A_i, F_i, μ_i) be a smooth observable dynamical system and let Ψ_i be a delay embedding. Its latent dynamical manifold is

$$\boxed{\mathfrak{L}_i = (\mathcal{M}_i, \widetilde{F}_i, g_i, \nu_i),} \tag{5}$$

where $\mathcal{M}_i$, $\widetilde{F}_i$, and ν_i are defined in equation 4, and $g_i = (\Psi_i^{-1})^* g_{A_i}$ transports a selected reference metric g_{A_i}.

The metric is part of the specified population object, not uniquely identified from scalar observations. Transporting a fixed metric makes the relation invariant to smooth coordinate changes, but changing the reference metric can change singular values, margins, and condition numbers. These are therefore metric-relative conditioning quantities. Empirically, we use standardized delay-space Euclidean neighborhoods and local SVD tangent coordinates (Appendix A.3).

4.1 STABLE DYNAMICAL FACTOR MAPPINGS

Let $d_i = \dim \mathcal{M}_i$ and suppose $d_x \le d_y$. For a C^1 map $\Phi : \mathcal{M}_y \to \mathcal{M}_x$, define the vertical and horizontal tangent spaces

$$V_q = \ker D\Phi_q, \qquad H_q = V_q^{\perp_{g_y}}, \qquad T_q\mathcal{M}_y = H_q \oplus V_q. \tag{6}$$

Definition 4.3 (Aligned stable-factor relation)**.** We write $X \to_G Y$ if there exists a C^1 map $\Phi_{y\to x} : \mathcal{M}_y \to \mathcal{M}_x$ that is an (m, M)-stable dynamical factor map for finite $0 < m \le M < \infty$, meaning:

1. **Dynamics compatibility:**
$$\Phi_{y\to x} \circ \widetilde{F}_y = \widetilde{F}_x \circ \Phi_{y\to x} \qquad \nu_y\text{-a.e.} \tag{7}$$
2. **Paired-state alignment and non-degeneracy:**
$$q_x = \Phi_{y\to x}(q_y) \quad \nu_{xy}\text{-a.s.}, \qquad \operatorname{rank} D\Phi_{y\to x}(q) = d_x \quad \nu_y\text{-a.e.} \tag{8}$$
The graph condition implies $(\Phi_{y\to x})_{\#}\nu_y = \nu_x$ and surjectivity on the support of ν_x.
3. **Horizontal geometric stability:** for every $v \in H_q$,
$$m\|v\|_{g_y} \le \|D\Phi_{y\to x}(q)v\|_{g_x} \le M\|v\|_{g_y} \qquad \nu_y\text{-a.e.} \tag{9}$$

Definition scope. The symbol $X \to_G Y$ denotes aligned information containment; ancestry is the external Convention 4.2. A mixed Y sensor that reconstructs independent (u, v) can admit $(u, v) \mapsto u$ without influence, violating (S2); Appendix B.3 gives the counterexample.

When $d_x < d_y$, vertical directions encode response-specific state that must be collapsed, so global distance preservation is neither required nor possible. When $d_x = d_y$, the condition becomes local bi-Lipschitz stability, and conjugacy or synchronization can make both directions admissible.

4.2 SCORES AND IDENTIFICATION BOUNDARIES

For a candidate factor, define the root dynamics residual and geometric distortion

$$\mathcal{R}_{\mathrm{dyn}}(\Phi)^{1/2} = \left[\int d_{g_x}^2\Big(\Phi(\widetilde{F}_y q), \widetilde{F}_x(\Phi(q))\Big)\, d\nu_y(q)\right]^{1/2}, \tag{10}$$

$$\Delta_2^2(\Phi) = \int \sum_{j=1}^{d_x} \big(\log \sigma_j(D\Phi_q|_{H_q})\big)^2 d\nu_y(q). \tag{11}$$

The condition number $\kappa_G = M/m$ measures metric-relative conditioning, not coupling strength or causal identification. Proposition D.4 gives multiplicative stability bounds for surjective derivatives with the same intermediate metric. Direct edges additionally require conditional factorization; Convention 4.2 governs ancestry, with synchronization and hidden drivers as boundaries.

5 STABLE DYNAMICAL FACTOR LEARNING

5.1 REFERENCE OBJECTIVE AND FROZEN ESTIMATOR SURROGATES

For a candidate aligned relation $X \to_G Y$, stable dynamical factor learning estimates a structured map from synchronized finite trajectories. For a hypothesis class $\mathcal{H}$, we use the following idealized reference functional:

$$\begin{aligned}\widehat{\Phi}_{\mathcal{H},L} \in \arg\min_{\Phi\in\mathcal{H}} \widehat{\mathcal{J}}_{L,\boldsymbol{\lambda}}(\Phi), \qquad \widehat{\mathcal{J}}_{L,\boldsymbol{\lambda}}(\Phi) &= \widehat{\mathcal{R}}_{\mathrm{pair}}(\Phi) + \lambda_{\mathrm{dyn}}\widehat{\mathcal{R}}_{\mathrm{dyn}}(\Phi) + \lambda_{\mathrm{tan}}\widehat{\mathcal{R}}_{\mathrm{tan}}(\Phi) \\ &\quad + \lambda_{\mathrm{geo}}\widehat{\mathcal{P}}_{\mathrm{geo}}(\Phi) + \lambda_{\mathrm{reg}}\Omega(\Phi),\end{aligned} \tag{12}$$

where

$$\widehat{\mathcal{R}}_{\mathrm{pair}} = \frac{1}{L'}\sum_t d_{g_x}^2\big(\widehat{\Phi}(q_{y,t}), q_{x,t}\big), \tag{13}$$

$$\widehat{\mathcal{R}}_{\mathrm{dyn}} = \frac{1}{L'-1}\sum_t d_{g_x}^2\Big(\widehat{\Phi}(q_{y,t+1}), \widehat{F}_x(\widehat{\Phi}(q_{y,t}))\Big), \tag{14}$$

$$\widehat{\mathcal{R}}_{\mathrm{tan}} = \frac{1}{L'-1}\sum_t \left\|\widehat{B}_{t+1}\widehat{A}_{y,t} - \widehat{A}_{x,t}\widehat{B}_t\right\|_F^2, \tag{15}$$

$$\widehat{\mathcal{P}}_{\mathrm{geo}} = \frac{1}{L'}\sum_t \sum_{j=1}^{d_x} \left[(m - \widehat{\sigma}_{j,t})_+^2 + (\widehat{\sigma}_{j,t} - M)_+^2\right]. \tag{16}$$

Here $\boldsymbol{\lambda} = (\lambda_{\mathrm{dyn}}, \lambda_{\mathrm{tan}}, \lambda_{\mathrm{geo}}, \lambda_{\mathrm{reg}})$ is fixed. Theorem 6.2 concerns the population counterpart of this displayed functional, whose target is its regularized minimizer; it coincides with an unregularized paired-state factor only under an additional identification condition or a separate vanishing-regularization regime. The frozen Atlas and Neural implementations use documented finite-sample surrogates and therefore need not have the same fixed-weight minimizer. Under exact paired alignment on an invariant population support, dynamics compatibility follows by applying alignment at the next time point. Pair, dynamics, tangent, and geometry losses encode state correspondence, finite-sample commutation, derivative compatibility, and metric-relative conditioning; their sum is a multi-objective compromise rather than independent causal identification. The tangent identity is

$$D\Phi_{\widetilde{F}_y(q)} D\widetilde{F}_y(q) = D\widetilde{F}_x(\Phi(q)) D\Phi_q. \tag{17}$$

The local atlas class $\mathcal{H}_{\mathrm{atlas}}$ and differentiable neural class $\mathcal{H}_{\mathrm{NN}}$ share the aligned relation and diagnostics; objective reductions and fixed-weight minimizers are estimator-specific. Convention 4.2 remains external.

5.2 ATLAS GCCM–BCD

Atlas GCCM–BCD is our official nonparametric estimator. It constructs synchronized delay vectors, estimates local PCA charts and tangent dynamics $A_{i,t} \approx D\widetilde{F}_i(q_{i,t})$, and represents a factor chart as

$$\xi_x \approx B_t \xi_y + b_t, \qquad B_t \in \mathbb{R}^{d_x \times d_y}. \tag{18}$$

A lower-dimensional domain is rejected before fitting as full row rank is impossible. The global C^1 reference relation is estimated by a finite hard local atlas (G2): queries use the nearest source-chart center without partition-of-unity gluing or a global C^1 guarantee. Blocks couple along temporal dynamics and tangent-transition edges. Starting from the historical one-pass solution, monotone Gauss–Seidel sweeps update (B_t, b_t); a nested chronological validation block selects the sweep and stops after two relative improvements below 10^{-4}, with $T_{\max} = 40$. Direction labels and test diagnostics are not read. This hard-chart estimator is outside Theorem 6.2's globally smooth class; Appendix E.4 records the boundary.

5.3 Neural Stable Factor Estimator

Neural GCCM provides a flexible parametric estimator for the same stable-factor relation class through a distinct frozen surrogate. For standardized delay states, it uses

$$\widehat{\Phi}_\theta(q_y) = W_2\,\mathrm{GELU}(\mathrm{LayerNorm}(W_1 q_y + b_1)) + b_2, \tag{19}$$

and minimizes

$$\mathcal{L}(\theta) = \mathcal{L}_{\mathrm{pair}} + \lambda_{\mathrm{dyn}}\mathcal{L}_{\mathrm{dyn}} + \lambda_{\mathrm{tan}}\mathcal{L}_{\mathrm{tan}} + \lambda_{\mathrm{geo}}\mathcal{L}_{\mathrm{geo}} + \lambda_{\mathrm{reg}}\mathcal{L}_{\mathrm{reg}}, \tag{20}$$

The implementation averages pair and dynamics squared errors over samples and coordinates and uses a squared, scale-normalized Frobenius tangent residual; its geometry term averages the summed singular-value barrier. These reductions differ from equation 13–equation 15, so fixed weights define a surrogate-specific regularized target $\Phi^{\dagger,\mathrm{NN}}_{\boldsymbol{\lambda}}$ rather than necessarily $\Phi^{\dagger}_{\boldsymbol{\lambda}}$. Autodiff is restricted to the intrinsic tangent spaces,

$$B_\theta(q) = \widehat{U}_x^\top D\widehat{\Phi}_\theta(q)\widehat{U}_y, \qquad m \le \sigma_j(B_\theta(q)) \le M, \tag{21}$$

not regularized through the full ambient Jacobian. The estimated tangent bases and dynamics are frozen nuisance quantities during training. The differentiable parameterization is compatible with the idealized smooth class; Theorem 6.2 requires the surrogate's uniform-law, plug-in, separation, and approximate-global-minimization premises, which remain unproved. Appendix K gives implementation details.

5.4 Shared diagnostics and finite-sample screen

Algorithm 1: Atlas-GCCM–BCD. For each candidate direction: (1) construct aligned delay clouds and chronological train/validation/test blocks; (2) estimate fixed charts, neighbors, tangent bases, and local dynamics on training data; (3) reject dimension-infeasible directions; (4) initialize factor blocks with Atlas GCCM–1pass; (5) perform monotone block-coordinate sweeps; (6) stop using only the blocked validation structural objective; (7) evaluate held-out diagnostics once; and (8) apply the frozen finite-sample decision protocol. One accepted direction supports a detected aligned factor; two indicate feedback or observational equivalence; neither means no detected factor or insufficient observation. An ancestry reading separately requires Convention 4.2.

The coordinate-NRMSE, cross-map, vector-RMS state, and tangent gates are 0.8, 0.35, 0.8, and 0.9. The historical evaluator serialized the state residual as per-coordinate RMSE; R3.3 therefore derives vector RMS as $\sqrt{D}$ times that frozen value and reapplies the unchanged state threshold without refitting. The empirical non-degeneracy screen requires a one-sided 95% lower bound for the held-out 5% minimum-singular-value quantile, using 1,000 circular moving-block replicates of length 32, to exceed zero. This fitted-representation screen replaces the historical 0.01 cutoff but is neither a calibrated factor-absence test nor a Type-I guarantee and does not prove population full rank; Appendix G states its full scope.

6 Theoretical Analysis

Appendices D and E give the assumptions and proofs. The theory concerns smooth observable quotients; Lorenz experiments are controlled out-of-class stress tests unless a stratified extension is invoked.

Theorem 6.1 (Transport of aligned stable factors under valid reconstruction)**.** *Let A_x, A_y be compact smooth observable quotient manifolds of a common latent system, with stable delay embeddings Ψ_x, Ψ_y. Suppose there is an observationally aligned C^1 surjective submersion $\pi_{y\to x} : A_y \to A_x$ such that*

$$\pi_{y\to x} \circ F_y = F_x \circ \pi_{y\to x}, \qquad \varpi_x = \pi_{y\to x} \circ \varpi_y \quad \mu\text{-a.e.}, \tag{22}$$

and the horizontal singular values of $D\pi_{y\to x}$ lie in $[m_\pi, M_\pi]$. Then

$$\Phi_* = \Psi_x \circ \pi_{y\to x} \circ \Psi_y^{-1} \tag{23}$$

is an aligned stable dynamical factor and witnesses $X \to_G Y$. Relative to the transported reference metrics its stability constants are $[m_\pi, M_\pi]$. If ambient delay metrics obey $c_i\|v\| \le \|D\Psi_i v\|_2 \le C_i\|v\|$, valid conservative constants are

$$m_* \ge \frac{c_x}{C_y} m_\pi, \qquad M_* \le \frac{C_x}{c_y} M_\pi. \tag{24}$$

Conversely, every aligned stable factor between valid reconstructions induces an aligned factor between the observable quotient systems.

This result transports an assumed aligned factor across reconstruction coordinates; Assumption 4.1 and the ancestry Convention 4.2 require external support. Measure-theoretically on the invariant support, up to μ-null sets, existence means the Y history partition refines the X partition; smoothness, rank, and metric-relative bounds provide stable geometric conditioning.

Theorem 6.2 (Conditional consistency for a regularized population target)**.** *Fix $\boldsymbol{\lambda}$ and let $\Phi^\dagger_{\boldsymbol{\lambda}} = \arg\min_{\Phi\in\mathcal{H}^s_{m,M}} \mathcal{J}_{\boldsymbol{\lambda}}(\Phi)$ be the unique regularized population minimizer in a class of globally defined C^s candidate maps, $s \ge 2$, precompact in the topology supporting the state-level uniform law. Assume (B1)–(B8) in Appendix E, including risk separation around this minimizer. If the plug-in chart and tangent-dynamics estimators are consistent and*

$$\sup_{\Phi\in\mathcal{H}^s_{m,M}} |\widehat{\mathcal{J}}_{L,\boldsymbol{\lambda}}(\Phi) - \mathcal{J}_{\boldsymbol{\lambda}}(\Phi)| \xrightarrow{p} 0, \tag{25}$$

then every $o_p(1)$-approximate empirical minimizer has the state-level consistency

$$\|\widehat{\Phi}_L - \Phi^\dagger_{\boldsymbol{\lambda}}\|_{L^2(\nu_y)} \xrightarrow{p} 0. \tag{26}$$

If the derivative-topology separation clause of (B8) holds, the class is also precompact in C^1, and the derivative plug-in and uniform-law conditions hold, the stronger derivative consistency equation 50 follows. Consequently, paired and state-dynamics functionals converge under state-level continuity; tangent and horizontal singular-value functionals additionally require the derivative conclusion. Held-out estimates require the same dependent-sample law of large numbers.

Remark 6.3 (Estimator-class scope). The theorem covers the displayed idealized functional. Fixed weights target its $\Phi^\dagger_{\boldsymbol{\lambda}}$, while Corollary E.2 gives a distinct route to the unregularized Φ_*. Frozen Neural has a surrogate-specific minimizer and requires separate verification of the theorem premises; hard-chart Atlas is excluded. No-factor rejection, global optimization, and sensor locality remain unproved.

Under generalized synchronization, the two directional model classes may each have a well-defined, even unique, stable factor. Orientation fails because both directions are observationally admissible, not necessarily because either directional risk has multiple minimizers.

7 Synthetic Experiments

Experiment scope. These tests evaluate factor recovery, structural compatibility, and identification boundaries. Lorenz known-factor/noise and Experiment 6 response observations include driver state, placing them outside (S2) and the ancestry reading in Convention 4.2. Experiments 1, 2, 3, and 5 use 30 seeds and $L \in \{250, 500, 1000, 2000, 4000, 8000\}$; Experiment 4 uses $L \in \{2000, 4000, 8000\}$ and seven coupling ratios. Atlas GCCM–BCD is official; one-pass is a retained solver ablation.

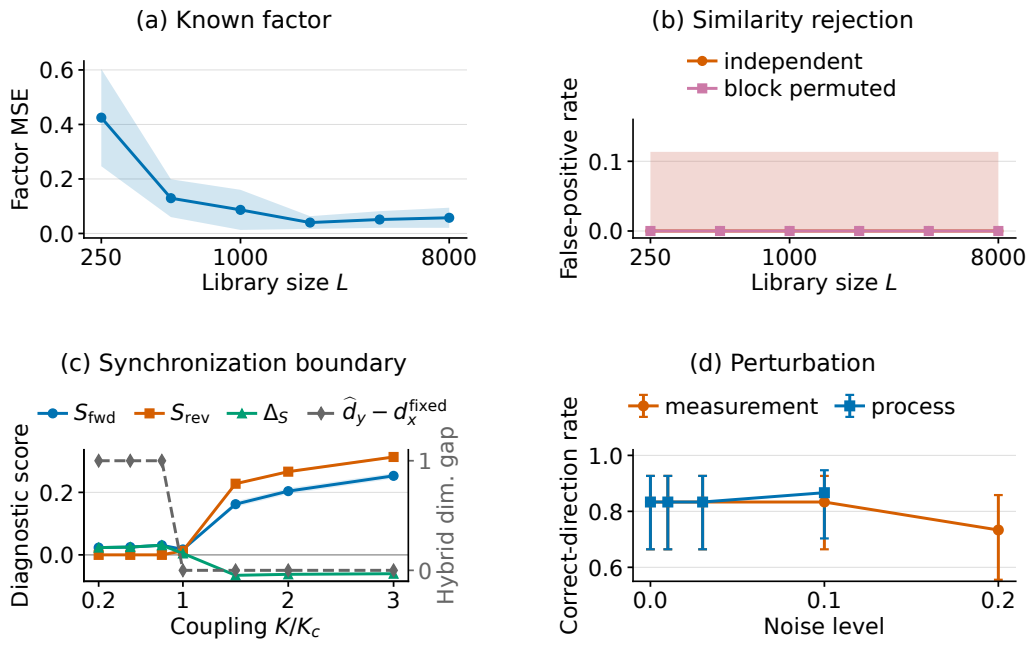


Figure 4: Atlas GCCM–BCD across four diagnostic roles: (a) known-factor recovery; (b) rejection of unaligned dynamical similarity; (c) directional admissibility and the hybrid configured/estimated dimension gap across synchronization; and (d) measurement/process robustness. Bands are 95% Student-t intervals; binary rates use Wilson intervals.

7.1 Protocol, interface, and metrics

All tasks share simulation, ground-truth, fitting, and evaluation interfaces. The maximum delay horizon and Theiler window separate blocked splits; hyperparameters exclude direction labels. Summaries use means, medians, standard deviations, 95% Student-t intervals for continuous seed means, and Wilson intervals for binary proportions.

We report labeled factor MSE/NRMSE, state RMSE per coordinate $R^{\mathrm{pc}}_{\mathrm{state}}$ and as vector RMS $R^{\mathrm{vec}}_{\mathrm{state}} = \sqrt{D} R^{\mathrm{pc}}_{\mathrm{state}}$, normalized tangent residual, log-singular distortion, rank, cross-map convergence, direction accuracy, and L_{crit}. Evaluation-only replay applies $R^{\mathrm{vec}}_{\mathrm{state}}$ to the frozen 0.8 gate and Atlas score. Experiment 4 reports bidirectional scores, signed margin, and the hybrid configured/estimated dimension gap $\widehat{d}_y - d_x^{\mathrm{fixed}}$; it has no ambiguity threshold and is excluded from direction accuracy. Standard scalar CCM reports coordinate error; controlled-observation CCM can report factor NRMSE. Residuals and direction are joint; $\rho(L)$ is never the sole endpoint.

Experiment 1: known factor recovery. For the Lorenz skew product with $\pi(u, \theta, r) = u$, Atlas GCCM–BCD at $L = 8000$ accepts the forward factor in $25/30$ seeds and the dimension-infeasible reverse in $0/30$. Mean per-coordinate factor MSE is 0.0576 (95% CI $[0.0207, 0.0945]$); state residual is 0.2196 per coordinate and 0.5809 as vector RMS, tangent residual is 0.4380, and the frozen minimum-singular-value diagnostic is 0.0583. The cell-level direction result is $25/30$ (Wilson 95% CI $[0.6644, 0.9266]$); across the full library grid, accuracy is 0.7833, and corrected BCD and one-pass labels differ in $16/180$ cells.

Experiment 2: alignment versus dynamical similarity. Two uncoupled Lorenz copies share parameters and marginal spectra but have a product joint law. At $L = 8000$, Atlas GCCM–BCD gives $0/30$ false positives for simultaneous-independent and block-permuted pairing (Wilson/exact upper $0.1135/0.1157$), mean correlations $-0.0049/0.0049$, and NRMSEs $1.1183/1.1166$. Matched abstract dynamics do not replace alignment.

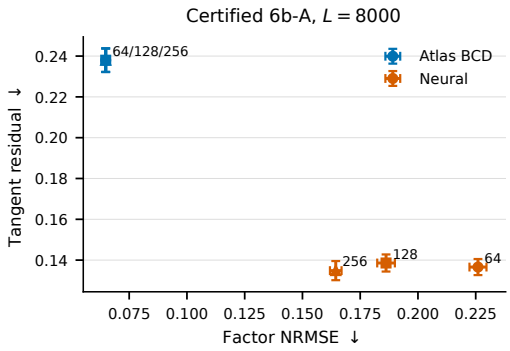


Figure 5: Experiment 6b-A at $L = 8000$ (30 seeds; 95% Student-t intervals). The two fixed estimators show a trade-off: Atlas lowers factor NRMSE and Neural lowers tangent residual.

Experiment 3: sample-size-dependent feedback recovery across orbit-sampling protocols. At $L = 250$ both protocols recover $0/30$; at 8000, one/eight-orbit recovery is $27/30$ versus $25/30$; and at 2000 it is $29/30$ versus $13/30$. Orbit count also changes samples per orbit, so this is not an isolated support effect. Full-grid BCD accuracy is 0.6333 (one-pass 0.7139; 70 differing labels).

Experiment 4: factor admissibility under synchronization. Generalized synchronization can make both directions admissible (Kocarev & Parlitz, 1996). In the Kuramoto sweep (Kuramoto, 1984), $S_{\mathrm{fwd}}/S_{\mathrm{rev}} = 0.0308/0$ and $\Delta_S = 0.0308$ at $K/K_c = 0.8$; at locking the hybrid configured/estimated dimension gap $\widehat{d}_y - d_x^{\mathrm{fixed}}$ falls $1 \to 0$ and Δ_S to 0.0045. Here only $\widehat{d}_y$ is estimated and $d_x^{\mathrm{fixed}} = 1$ is configured. The margin is -0.0655 at 1.5; scores at 3 are $0.2529/0.3133$ as admissibility becomes bidirectional. The ungated contrast and gap diagnose boundary behavior without assuming monotonicity.

Experiment 5: perturbation robustness. Measurement and process noise are applied separately. At $L = 8000$, Atlas is correct in $25/30$ seeds at zero and at measurement/process $0.01, 0.03$, $25/30$ at measurement 0.10, $22/30$ at measurement 0.20, and $26/30$ at process 0.10. The range is 22–$26/30$ (Wilson lower endpoint 0.5555 at $22/30$).

Experiment 6: controlled observation-map dimension and conditioning audit. All $180/180$ draws passed while state-retaining features vary $D \in \{64, 128, 256\}$ and $\kappa \in \{1, 5, 10\}$ at $D = 128$; Atlas is $30/30$ per $L = 8000$ condition. Paired-grid Atlas-minus-Neural differences are -0.1437 (factor NRMSE), -0.01295 (state pc), -0.23891 (state vector), and $+0.0825$ (tangent); geometry/rank depend on condition, and failed Experiment 6a remains separate.

Targeted external baselines and matched budgets. RSB-v2 compares five methods on 10 seeds, eight conditions, and $L = 2000$, retaining failures and synchronization outcomes (Appendix L). With matched initialization and 72 fixed epochs, full versus pair-only training lowers state and tangent incompatibility; factor-error, geometry, and rank advantages remain unresolved, and categorical outcomes are unchanged. These are metric-specific effects (Appendix G.11).

8 Conclusion

Aligned stable factors provide a population target for directed dynamical structure learning, estimated by hard local Atlas GCCM–BCD and differentiable Neural GCCM. Theory gives reconstruction transport and regularized-target consistency within the stated smooth-estimator scope. Matched budgets isolate state and tangent compatibility gains; factor-error, geometry, and rank advantages remain unresolved, with unchanged categorical outcomes. Restricted ancestry requires Convention 4.2; synchronization marks an identification boundary. Future work covers no-factor testing, a globally smooth Atlas, equal-dimension benchmarks, full bootstrap, and real data.

AI USE STATEMENT

Generative AI tools assisted literature organization, code generation, debugging, testing, experiment orchestration. The human authors selected the research question and final formulation, verified all AI-assisted mathematics, code, results, and citations, and take responsibility for the final content.

REPRODUCIBILITY STATEMENT

The anonymous supplementary archive supports protocol inspection, verification of frozen numerical claims, reconstruction of reported tables and figures, and lightweight smoke tests.

REFERENCES

Eddie Aamari and Clément Levrard. Nonasymptotic rates for manifold, tangent space and curvature estimation. *The Annals of Statistics*, 47(1):177–204, 2019.

Hanru Bai and Weiyang Ding. KoNODE: Koopman-driven neural ordinary differential equations with evolving parameters for time series analysis. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 2484–2519. PMLR, 2025. URL https://proceedings.mlr.press/v267/bai25b.html.

Petar Bevanda, Max Beier, Alexandre Capone, Stefan Georg Sosnowski, Sandra Hirche, and Armin Lederer. Koopman-equivariant Gaussian processes. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 3151–3159. PMLR, 2025. URL https://proceedings.mlr.press/v258/bevanda25a.html.

Jason J. Bramburger, Steven L. Brunton, and J. Nathan Kutz. Deep learning of conjugate mappings. *Physica D: Nonlinear Phenomena*, 427:133008, 2021. doi: 10.1016/j.physd.2021.133008. URL https://doi.org/10.1016/j.physd.2021.133008.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Yuxiao Cheng, Lianglong Li, Tingxiong Xiao, Zongren Li, Jinli Suo, Kunlun He, and Qionghai Dai. CUTS+: High-dimensional causal discovery from irregular time-series. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 11525–11533, 2024. doi: 10.1609/aaai.v38i10.29034. URL https://ojs.aaai.org/index.php/AAAI/article/view/29034.

Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

Armin Eftekhari, Han Lun Yap, Michael B. Wakin, and Christopher J. Rozell. Stabilizing embedology: Geometry-preserving delay-coordinate maps. *arXiv preprint arXiv:1609.06347*, 2016.

Stathi Fotiadis, Mario Lino Valencia, Shunlong Hu, Stef Garasto, Chris D Cantwell, and Anil Anthony Bharath. Disentangled generative models for robust prediction of system dynamics. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 10222–10248. PMLR, 2023. URL https://proceedings.mlr.press/v202/fotiadis23a.html.

Clive W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

Wiebke Günther, Urmi Ninad, and Jakob Runge. Causal discovery for time series from multiple datasets with latent contexts. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 766–776. PMLR, 2023. URL https://proceedings.mlr.press/v216/gunther23a.html.

Daniel Harnack, Erik Laminski, and Klaus Richard Pawelzik. Topological causality in dynamical systems. *Physical Review Letters*, 119(9):098301, 2017.

Maximilian Karl, Maximilian Soelch, Justin Bayer, and Patrick van der Smagt. Deep variational bayes filters: Unsupervised learning of state space models. In *International Conference on Learning Representations*, 2017.

Nam Hee Kim, Zhaoming Xie, and Michiel van de Panne. Learning to correspond dynamical systems. In *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, volume 120 of *Proceedings of Machine Learning Research*, pp. 105–117. PMLR, 2020. URL https://proceedings.mlr.press/v120/kim20a.html.

Ljupco Kocarev and Ulrich Parlitz. Generalized synchronization, predictability, and equivalence of unidirectionally coupled dynamical systems. *Physical Review Letters*, 76(11):1816–1819, 1996. doi: 10.1103/PhysRevLett.76.1816.

Yoshiki Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer, 1984.

Michael L. Littman, Richard S. Sutton, and Satinder Singh. Predictive representations of state. In *Advances in Neural Information Processing Systems*, volume 14, 2001.

Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2): 130–141, 1963.

Sindy Löwe, David Madras, Richard Zemel, and Max Welling. Amortized causal discovery: Learning to infer causal graphs from time-series data. In *Proceedings of the First Conference on Causal Learning and Reasoning*, volume 177 of *Proceedings of Machine Learning Research*, pp. 509–525. PMLR, 2022. URL https://proceedings.mlr.press/v177/lowe22a.html.

Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9:4950, 2018.

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. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/f3d1e34a15c0af0954ae36a7f811c754-Paper-Conference.pdf.

Søren Wengel Mogensen, Karin Rathsman, and Per Nilsson. Causal discovery in a complex industrial system: A time series benchmark. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, volume 236 of *Proceedings of Machine Learning Research*, pp. 1218–1236. PMLR, 2024. URL https://proceedings.mlr.press/v236/mogensen24a.html.

Mitchell Ostrow, Adam Eisen, Leo Kozachkov, and Ila Fiete. Beyond geometry: Comparing the temporal structure of computation in neural circuits with dynamical similarity analysis. In *Advances in Neural Information Processing Systems*, volume 36, pp. 33824–33837. Curran Associates, Inc., 2023. doi: 10.52202/075280-1467. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/6ac807c9b296964409b277369e55621a-Abstract-Conference.html.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2 edition, 2009.

Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

Jakob Runge. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*, volume 124 of *Proceedings of Machine Learning Research*, pp. 1388–1397. PMLR, 2020. URL https://proceedings.mlr.press/v124/runge20a.html.

Tim Sauer, James A. Yorke, and Martin Casdagli. Embedology. *Journal of Statistical Physics*, 65 (3–4):579–616, 1991.

Thomas Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. MIT Press, 2 edition, 2000.

George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *Science*, 338(6106):496–500, 2012.

Nicholas Tagliapietra, Katharina Ensinger, Christoph Zimmer, and Osman Mian. Causal structure learning for dynamical systems with theoretical score analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 36740–36748, 2026. doi: 10.1609/aaai.v40i43.40999. URL https://ojs.aaai.org/index.php/AAAI/article/view/40999.

Floris Takens. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pp. 366–381. Springer, 1981.

Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

Anpeng Wu, Haoxuan Li, Kun Kuang, Keli Zhang, and Fei Wu. Learning causal relations from subsampled time series with two time-slices. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 53635–53655. PMLR, 2024. URL https://proceedings.mlr.press/v235/wu24p.html.

Yuanchao Xu, Kaidi Shao, Nikos K. Logothetis, and Zhongwei Shen. ResKoopNet: Learning Koopman representations for complex dynamics with spectral residuals. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 69647–69674. PMLR, 2025. URL https://proceedings.mlr.press/v267/xu25y.html.

Raanan Yehezkel Rohekar, Shami Nisimov, Yaniv Gurwicz, and Gal Novik. From temporal to contemporaneous iterative causal discovery in the presence of latent confounders. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 39939–39950. PMLR, 2023. URL https://proceedings.mlr.press/v202/yehezkel-rohekar23a.html.

Shimin Zhang, Ziyuan Ye, Yinsong Yan, Zeyang Song, Yujie Wu, and Jibin Wu. KoopSTD: Reliable similarity analysis between dynamical systems via approximating Koopman spectrum with timescale decoupling. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 76563–76582. PMLR, 2025a. URL https://proceedings.mlr.press/v267/zhang25cp.html.

Yitian Zhang, Liheng Ma, Antonios Valkanas, Boris N. Oreshkin, and Mark Coates. SKOLR: Structured Koopman operator linear RNN for time-series forecasting. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 75734–75756. PMLR, 2025b. URL https://proceedings.mlr.press/v267/zhang25be.html.