From Covariance to Response: What Is Identifiable, and What Should We Measure?

Anonymous authors
Paper under double-blind review

Abstract

Equilibrium measurements describe how variables fluctuate together, whereas a perturbation experiment asks how a change propagates from one variable to another. When can covariance and a list of decay rates answer this response question, and what should be measured when they cannot? We study stable Ornstein–Uhlenbeck networks with a known acyclic graph and unknown diagonal diffusion. For complete two-layer networks, we establish a sharp identification boundary using the rate sum and show how the full anonymous spectrum can resolve the remaining rate assignments. For a sparse family, an odd-cycle criterion identifies the missing information and guides targeted temporal measurements. Matching bounds on stars quantify the value of target self-lags; multiroot results connect response risk to graph geometry and amplitude calibration. On the star, with unknown independent sensor variances, tracking source and target attains the four-coordinate risk order. A local risk calculation predicts when temporal inversion improves on a prior-center prediction. All sixteen known-amplitude comparisons in independent synthetic experiments support the predicted directions, alongside matched-readout and likelihood-based comparisons. These results connect response identification to the choice and budget of temporal measurements.

1 Introduction

In protein-signaling studies, measuring a cell can destroy it. One may therefore have many independent measurements near equilibrium but few trajectories of the same cell. Graphical continuous Lyapunov models use the resulting cross-sectional covariance to study directed dynamics (Varando & Hansen, 2020; Dettling et al., 2023). The measured variables have names: the covariance records, for example, which two proteins fluctuate together. But a planned perturbation asks a different question: if one named protein is briefly displaced, how much will a specified downstream protein change after a chosen delay? This is a query about transient propagation. Optogenetic Ras activation followed by a downstream Erk readout makes the source–target timing question concrete (Toettcher et al., 2013); our estimand is an idealized response to an initial-state pulse.

Equilibrium and temporal assays reveal different features of the dynamics. Graph restrictions can identify drift from equilibrium covariance when the diffusion matrix is known (Dettling et al., 2023); with unknown diagonal diffusion, even the absolute time scale is lost, and recent work asks which edge *signs* survive (van Seeventer & Salehkaleybar, 2026). Under suitable recoverability conditions, temporal analysis can recover eigenvalues from partially observed deterministic affine trajectories (Cheng & Tang, 2022), while eigenvalue timescales are also used to interpret fitted dynamical models (Qian et al., 2024). Consider two assays of the same stationary system: a multivariable snapshot panel and time-resolved recordings with only one reporter at a time. Such separately observed channels are considered by Browning et al. (2025). A reporter can contain several decay modes; their rates need not label the variables in the snapshot panel. We ask whether the named covariance and the recovered rate list jointly determine a preselected source–target response. The assays must share the same dynamics and operating state; we study the information retained in these two summaries.

In stable linear Ornstein–Uhlenbeck (OU) networks with a known acyclic graph and unknown positive diagonal diffusion, we connect response identification to the sampling cost of selected temporal coordinates.

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.

Our main contributions are threefold:

1. **Sharp criteria for identifying a named response.** We establish a root–leaf count boundary for complete two-layer networks given covariance and the rate sum, and an odd-cycle criterion for a sparse family. We also characterize when the full anonymous spectrum resolves rate assignments that their sum leaves ambiguous.
2. **Targeted measurements that complete the missing information.** We characterize which root measurements remove response ambiguity and show how target self-lags reveal rate information more strongly than oriented cross-lags under weak coupling. With unknown independent sensor variances, tracking only source and target attains the same risk order as tracking all four coordinates in the star model.
3. **Matching statistical limits and finite-budget evidence.** We derive matching minimax bounds for snapshot and temporal observation schemes on stars, and connect graph geometry to response risk in multi-root families, including estimated common covariance amplitude. A local risk calculation predicts when temporal inversion improves on the prior center. Same-record comparisons, matched readout budgets, and full-likelihood static decoding test response accuracy; independent experiments test the predicted budget transition.

2 Related Work

Equilibrium identification in dynamical graphs. Graphical continuous Lyapunov models connect equilibrium covariance to directed drift (Varando & Hansen, 2020), with identification results under fixed diffusion (Dettling et al., 2023) and estimation methods (Dettling et al., 2024). With unknown diagonal diffusion, van Seeventer & Salehkaleybar (2026) study edge signs and Améndola et al. (2025) study equivalence of acyclic graph models. We condition on a known graph and quantify identification and estimation of a named finite-horizon response from supplementary information. Non-Gaussian cumulants (Recke & Hansen, 2026) and selected steady-state interventions (Salehkaleybar, 2026) offer alternative routes to identification, up to a common time scale under their respective assumptions. Our temporal observations supply time-scale information for a specified transient response.

Spectral and partial-observation ambiguity. Qian et al. (2024) show how partial observation can distort mechanistic interpretation, including spectral timescales, in fitted neural dynamical models. Cheng & Tang (2022) establish conditions for recovering eigenvalues from a partially observed deterministic affine trajectory. In a two-state OU example, Browning et al. (2025, Section 3.1) show that separately observed marginal processes can preserve eigenvalue information while leaving cross-covariance undetermined. These works motivate our joint-information question: can named equilibrium covariance complete the anonymous rate information? We give exact criteria in known two-layer graphs and quantify the cost of additional temporal coordinates for one response.

Temporal measurements and targeted outputs. Initial-state response is linked to stationary correlations by generalized fluctuation–response theory (Marconi et al., 2008). Gilson et al. (2018) use the OU Green function to construct dynamic communicability and noise-weighted network flow. Lagged observations also support OU estimation: Singh et al. (2018) derive likelihood-based drift estimates from discrete observations, and Gilson et al. (2016) fit zero- and nonzero-lag covariances for directed connectivity. Minimal-output-set work studies local structural parameter identifiability in ordinary differential equations (Anguelova et al., 2012; Joubert et al., 2018). Task-dependent information acquisition appears in the exploration bounds and design algorithm of Wagenmaker et al. (2021) for controlled linear systems. Our matching risk bounds instead quantify the value of selected coordinates recorded from an autonomous stationary system. Identification likewise underpins Gaussian-process parameter interpretation (Qaqish & Li, 2025), longitudinal effects (Ying, 2025), and temporal representations (Li et al., 2025); here the requested quantity is one observed source–target response.

3 Setting

Let $X_t \in \mathbb{R}^n$ satisfy

$$dX_t = AX_t\,dt + Q^{1/2}dW_t, \qquad A\Sigma + \Sigma A^\top + Q = 0. \tag{1}$$

The known graph specifies the allowed off-diagonal entries of A. It is acyclic, all diagonal entries of A are negative, and the unknown Q is positive diagonal. Thus A is stable and Σ is its stationary covariance. Write $T = -\operatorname{tr} A$ for the sum of decay rates and $\mathcal{R} = \{-\lambda : \lambda \in \operatorname{spec}(A)\}$ for their unordered multiset, with multiplicities. These positive rates are the negatives of the drift eigenvalues.

For a preselected source k, target i, and horizon $H > 0$, the estimand is

$$g_{k\to i}(H) = e_i^\top e^{AH} e_k. \tag{2}$$

This is the mean response to a unit *initial-state pulse*. With future-first lag covariance $\Gamma(h) = \mathbb{E}[X_{t+h}X_t^\top]$, the OU solution gives

$$\Gamma(h) = e^{Ah}\Sigma, \qquad g_{k\to i}(H) = e_i^\top \Gamma(H)\Sigma^{-1} e_k. \tag{3}$$

Full lag covariance determines every response. We ask which responses remain determined by rate summaries or selected lag entries. Let $\mathcal{F}_G(\Sigma, \mathcal{I})$ collect admissible drifts with graph G, covariance Σ, and supplied rate information $\mathcal{I} \in \{T, \mathcal{R}\}$. The response is identified exactly when

$$\left|\left\{e_i^\top e^{AH} e_k : A \in \mathcal{F}_G(\Sigma, \mathcal{I})\right\}\right| = 1. \tag{4}$$

Finite-sample analysis asks how accurately this quantity can be estimated from snapshots and specified temporal coordinates.

The role of temporal information is visible already from scale. If (A, Q) produces Σ, then (cA, cQ) produces the same Σ for every $c > 0$, whereas the response at H becomes the old response at cH. The rate sum removes this scale ambiguity. The compatible covariance set also applies to independent finite-variance Lévy innovations; our likelihood and risk analysis uses Gaussian OU samples. The extension is proved in Appendix A.2. We write $f \asymp g$ for two-sided bounds up to positive constants depending only on the stated parameter class.

4 Method

We first determine whether compatible dynamics agree on the response, then select temporal coordinates that remove ambiguity. Moment inversions and matching risk bounds connect this measurement choice to response accuracy. Two-layer graphs establish the structural criteria; stars isolate the statistical value of each measurement, and multiroot families quantify the additional cost of graph geometry.

4.1 Identifying Responses from Covariance and Rates

Start with a two-layer graph containing every root-to-leaf edge. With r roots and m leaves, its drift has the form

$$A = \begin{pmatrix} -\operatorname{diag}(a_1, \dots, a_r) & 0 \\ B & -\operatorname{diag}(d_1, \dots, d_m) \end{pmatrix}. \tag{5}$$

In root–leaf formulas, $k \to i$ denotes root R_k to leaf L_i. Write $s_k = \operatorname{Var}(R_k)$ and $x_{ik} = \operatorname{Cov}(L_i, R_k)$. The Lyapunov equation gives

$$\begin{aligned} B_{ik} &= \frac{(a_k + d_i)x_{ik}}{s_k}, \qquad g_{k\to i}(H) = \frac{x_{ik}}{s_k} G_H(a_k, d_i), \\ \phi_H(a, d) &= \int_0^H e^{-au - d(H-u)}\, du, \qquad G_H(a, d) = (a + d)\phi_H(a, d). \end{aligned} \tag{6}$$

Recovering the named rates therefore recovers edges and responses. For leaf pair $i < j$, write $c_{ijk} = x_{ik}x_{jk}/s_k$ and $\Delta_{ij} = \Sigma_{r+i,r+j} - \sum_k c_{ijk}$. Its Lyapunov equation is

$$\Delta_{ij}(d_i + d_j) - 2\sum_{k=1}^{r} c_{ijk} a_k = 0. \tag{7}$$

Collecting these linear constraints in K_Σ yields the boundary.

Theorem 1 (Trace information boundary)**.** *For $m \geq 3$, on a generic subset of positive-rate, positive-diffusion complete two-layer models,*

$$\operatorname{rank} K_\Sigma = \min\left\{ \binom{m}{2}, r + m - 1 \right\}. \tag{8}$$

Consequently (Σ, T) determines the drift and every named response if and only if $r \leq (m-1)(m-2)/2$. Above that threshold, each preselected root-to-leaf response at a fixed $H > 0$ is generically nonidentified over the full positive parameter space.

The threshold counts independent constraints from *pairs of leaves*. With four leaves, the rate sum resolves three roots; a fourth leaves continuous ambiguity. Genericity refers to the full parameter space, allowing exceptional fixed-T slices. The proof is in Appendices A.3–A.4.

Identification from the complete anonymous spectrum. For complete two-layer models with positive diffusion, pairwise distinct rates, and $m \geq 3$, $(\Sigma, \mathcal{R})$ generically identifies the whole drift for any r. Covariance constraints can therefore assign an anonymous rate to its variable even beyond the trace boundary. A one-root, two-leaf model admits a counterexample. The proof checks finitely many rate assignments (Appendix A.10).

On a three-leaf star, covariance ratios recover relative rates and T supplies their scale. But $O(\epsilon)$ root–leaf covariances give informative residuals of only $O(\epsilon^2)$, against $O(N^{-1/2})$ noise (inverse in Appendix A.11).

Graph topology matters beyond this complete family. Consider a universal root connected to all leaves and extra roots each connected to a pair of leaves. Remove the selected pairs from the complete undirected graph on leaves and call the remaining graph H_{pair}. Its unsigned incidence matrix has one null direction per bipartite connected component. Let S be the set of distinct selected pairs, p the number of extra roots, and $b(H_{\text{pair}})$ the number of bipartite components, including isolated vertices. The unsigned incidence matrix U_H has row $e_i^\top + e_j^\top$ for each pair (i, j) of H.

Theorem 2 (Topology and targeted completion)**.** *For pair-root models with $m \geq 3$ leaves, positive rates and diffusion, and nonzero specified edges,*

$$\dim \mathcal{F}_G(\Sigma, T) = p - |S| + b(H_{\text{pair}}). \tag{9}$$

Thus the whole drift is identified exactly when the selected pairs are distinct and every component of H_{pair} has an odd cycle. If the pairs are distinct and $\Delta_e \neq 0$ for every selected pair, choose $b(H_{\text{pair}})$ extra roots R such that

$$\operatorname{im} N_H = \ker U_{H_{\text{pair}}}, \qquad \operatorname{rank}(U_R N_H) = b(H_{\text{pair}}), \tag{10}$$

where U_R contains the selected-pair incidence rows. Such a choice always exists. Replacing the supplied trace, stationary covariance plus root autocovariances at any fixed lag $h > 0$ for the universal root and these selected extra roots identify the whole drift globally, without T or the full spectrum. This is the minimum number of continuous scalar temporal quantities for whole-drift identification under this interface. A task-specific response can need fewer.

Here $\Delta_e \neq 0$ excludes signed-edge cancellation. Reading selected rates from $C_k(h) = \Sigma_{kk} e^{-a_k h}$ leaves a linear system whose nullspace the selected roots must resolve. With four leaves, selecting adjacent pairs leaves an odd cycle and requires one root lag; selecting disjoint pairs leaves a four-cycle and requires two. The same node and edge counts therefore give different answers. After measuring the universal root and roots R, its response curve to leaf i is identified exactly when

$$e_i^\top N_H \in \operatorname{rowspan}(U_R N_H). \tag{11}$$

For a response from an extra root on pair e, also require $U_e N_H \in \operatorname{rowspan}(U_R N_H)$. Thus a response can be identified before the whole drift: the measured rows need only remove ambiguity affecting its endpoints. Proofs are in Appendices A.5–A.7.

4.2 Graph-Dependent Measurement and Estimation

Take a regular leaf-pair graph H (equal degree at every leaf) with full-column-rank U_H, one universal root, and one extra root for each pair outside H. Define $\mathcal{C}_{H,\epsilon}$ by unit stationary variances, allowed

root–leaf covariances $x_{ik} = \epsilon$, and

$$a_0 = 1, \quad d_i \in [2-\delta, 2+\delta], \quad \sum_i d_i = 2m, \quad a_{(i,j)} = \frac{3}{5}(d_i + d_j) - 1. \tag{12}$$

Here $\delta > 0$ is fixed and sufficiently small, and $0 < \epsilon \le c_0/m$. The stationary equations determine the edges and positive diagonal diffusion. All models in this class have the same trace; only leaf-pair covariances on H vary. This construction isolates the difficulty contributed by graph geometry.

Let $R_{i,H}(N)$ be the minimax expected error in $g_{0\to i}(1)$ divided by ϵ, from N independent stationary vectors. The graph, trace, variances, a_0, and root–leaf covariances are known. Write $\sigma_H = \min_{v\perp\mathbf{1}, \|v\|_2=1} \|U_H v\|_2$. For $N \ge C\epsilon^{-4}(|E(H)| + \sigma_H^{-2})$,

$$R_{i,H}(N) \asymp \frac{\kappa_i(H)}{\epsilon^2\sqrt{N}}, \qquad \kappa_i(H) = \sup_{\substack{\mathbf{1}^\top v = 0 \\ \|U_H v\|_2 \le 1}} |v_i|. \tag{13}$$

For an odd cycle $H = C_m$, the same class gives an all-budget law:

$$R_{0,C_m}(N) \asymp 1 \wedge \frac{\sqrt{m}}{\epsilon^2\sqrt{N}}, \qquad N \asymp m\epsilon^{-4} \quad \text{for fixed small error.} \tag{14}$$

The factor $\kappa_i(H)$ measures how much the target rate can change while the observed pair sums barely move. Odd cycles identify every rate, but their growing sensitivity makes longer cycles statistically more expensive. Both bounds are proved in Appendix A.5.

The static rate has a finite-budget consequence: cycle inversion can lose to a class-center prediction. A target-leaf self-lag offers a different route. On the same odd-cycle class, its leading term is $e^{-d_0 h}$; the parent-path correction is of order $m\epsilon^2$. A center-corrected estimator uses the average product $\widehat{\rho}_h = K^{-1}\sum_{j=1}^K L_0^{(j)}(0)L_0^{(j)}(h)$. For $I_h = [e^{-(2+\delta)h}, e^{-(2-\delta)h}]$ and interval projection Π_{I_h}, set

$$\begin{aligned} b_{m,\epsilon}(h) &= \epsilon^2\left[3\phi_h(1,2) + (m-3)\tfrac{17}{5}\phi_h(7/5,2)\right], \\ \widehat{d}_0 &= -h^{-1}\log\Pi_{I_h}(\widehat{\rho}_h - b_{m,\epsilon}(h)), \qquad \widehat{g} = \epsilon G_1(1,\widehat{d}_0). \end{aligned} \tag{15}$$

Subtracting the parent-path term at the class center gives

$$\sup_{\theta\in\mathcal{C}_{C_m,\epsilon}} \frac{\mathbb{E}_\theta|\widehat{g} - g_{0\to 0}(1)|}{\epsilon} \le C_h\big(K^{-1/2} + m\epsilon^2\delta\big). \tag{16}$$

The two terms bound sampling error and center-correction error. Thus the lag bypasses graph inversion, at the cost of a controlled approximation error (proof in Appendix A.5).

Proposition 3 (Learning the common covariance amplitude)**.** *In the odd-cycle model* (12)*, retain the known unit variances, $a_0 = 1$, and the known root-rate linkage, but replace every allowed root–leaf covariance by an unknown common $\eta \in [\epsilon, 2\epsilon]$. For fixed $0 < \delta \le 1/2$ and $0 < \epsilon \le 1/(4m)$, the static minimax law* (14) *still holds. For M independent simultaneous source–target calibration pairs and K independent target self-lag pairs, set $\widehat{\eta} = \Pi_{[\epsilon,2\epsilon]} M^{-1}\sum_{j=1}^M R_0^{(j)}L_0^{(j)}$. Substitution into both the amplitude and center correction yields*

$$\sup \frac{\mathbb{E}|\widehat{g} - g|}{\epsilon} \le C_h\left[1 \wedge \left\{(M\epsilon^2)^{-1/2} + K^{-1/2} + m\epsilon^2\delta\right\}\right]. \tag{17}$$

The static graph penalty survives estimated calibration. Temporal acquisition has a separate amplitude cost: extra lag pairs help most after calibration uncertainty is controlled. Appendix C provides the estimator and full proof.

For several targets in any two-layer DAG, one fixed-lag self-covariance per requested leaf and distinct parent, alongside stationary covariance, identifies all incoming response curves. Parent probes are shared across targets; this sufficient design also allows repeated rates and signed-edge cancellation (proof in Appendix A.9).

There is also a narrower design for one active source–target edge. Once a_k is measured, the two lag orientations give

$$\frac{\Gamma_{r+i,k}(h)}{x_{ik}} = e^{-a_k h} + 2a_k\phi_h(a_k, d_i), \qquad \frac{\Gamma_{k,r+i}(h)}{x_{ik}} = e^{-a_k h}. \tag{18}$$

The first expression decreases strictly with $d_i > 0$, identifying the target rate and response curve; the second contains no d_i. This determines which orientation to record (proof in Appendix A.8).

4.3 Statistical Value of Temporal Coordinates

Identification alone gives no finite-sample precision. Under weak coupling, exchanging two named leaf rates preserves the spectrum and changes the response, while barely changing covariance. We first quantify the cost of distinguishing these assignments.

Assigning supplied rates. Consider a one-root, three-leaf star with root rate in $[.9, 1.1]$ and one leaf rate in each of $[1.8, 2.2]$, $[2.8, 3.2]$, $[3.8, 4.2]$, assigned to leaves in an unknown order. Variances lie in $[.9, 1.1]$ and root–leaf covariances in $[.8\epsilon, 1.2\epsilon]$; the graph, ϵ, and exact unordered spectrum are supplied. Let $R_\epsilon^{\mathcal{R}}(N, M)$ be normalized minimax response risk from N full snapshots and M independent oriented pairs $(R_0, L_{1,1})$.

Theorem 4 (The cost of assigning supplied rates)**.** *On this class, for sufficiently small ϵ, there are $c_0, c_1 > 0$ such that*

$$N\epsilon^2 \le c_1 \text{ or } N\epsilon^4 + M\epsilon^2 \le c_1 \quad \Longrightarrow \quad R_\epsilon^{\mathcal{R}}(N, M) \ge c_0. \tag{19}$$

For every fixed error target $\eta > 0$, a suitable C_η suffices: $N \ge C_\eta \epsilon^{-4}$ attains risk at most η using snapshots alone, and $N, M \ge C_\eta \epsilon^{-2}$ do so using both records.

The root band supplies its label; observations still identify leaf assignments and amplitudes (proof in Appendix B.3). We next replace supplied rates by measured temporal information.

Parameter class and observations. To compare measurements, use a one-root, three-leaf star. Write $s = \mathrm{Var}(R)$, $v_i = \mathrm{Var}(L_i)$, and $x_i = \mathrm{Cov}(R, L_i) = \epsilon u_i$. The class Θ_ϵ is defined by

$$\tfrac{1}{2} \le a, d_i \le \tfrac{13}{2}, \quad \tfrac{1}{2} \le s, v_i \le 2, \quad 1 \le u_i \le 2, \quad a + \sum_{i=1}^{3} d_i = 10. \tag{20}$$

Set $B_i = (a + d_i)x_i/s$, $Q_{00} = 2as$, and $Q_{ii} = 2d_i v_i - 2B_i x_i > 0$. Positivity holds throughout the box for $0 < \epsilon \le .01$; ϵ is known.

Independently prepare N stationary vectors, M oriented pairs, and M_0 root pairs, all with noiseless endpoints:

$$\mathcal{D}_{N,M,M_0} = \left(\{X_0^{(n)}\}_{n=1}^N, \{(R_0, L_{1,1})^{(j)}\}_{j=1}^M, \{(R_0, R_1)^{(\ell)}\}_{\ell=1}^{M_0} \right). \tag{21}$$

Here $L_{1,1}$ is leaf 1 at time 1. Define normalized minimax risk by

$$R_\epsilon(N, M, M_0) = \inf_{\widehat{g}} \sup_{\theta \in \Theta_\epsilon} \frac{\mathbb{E}_\theta |\widehat{g}(\mathcal{D}_{N,M,M_0}) - g_{0\to 1}(1)|}{\epsilon}. \tag{22}$$

Normalization measures precision relative to the response's coupling scale. The infimum includes every estimator using the stated data.

Theorem 5 (Risk from snapshots, oriented pairs, and root pairs)**.** *For Θ_ϵ and the independent records in* (21)*, uniformly for sufficiently small ϵ,*

$$R_\epsilon(N, M, M_0) \asymp 1 \wedge \max\left\{ (N\epsilon^2)^{-1/2}, (N\epsilon^4 + M\epsilon^2)^{-1/2}, (N\epsilon^4 + M_0)^{-1/2} \right\}. \tag{23}$$

The parameter class fixes the rate sum. The static-only upper bound uses it; an attaining temporal decoder ignores it. The lower bounds allow the exact sum to be supplied to every estimator.

The terms isolate amplitude, target-rate, and root-rate uncertainty. Amplitude has an $O(\epsilon)$ signal; static rate residuals have an $O(\epsilon^2)$ signal. Oriented pairs strengthen target-rate information to $O(\epsilon)$, and root self-lags resolve the remaining rate ambiguity at order one. Clipped moment inversion attains the bound (proof in Appendix B.4). At fixed small normalized error, the resulting sample requirements are

$$N_{\text{static}} \asymp \epsilon^{-4}, \qquad (N, M, M_0)_{\text{targeted}} \asymp (\epsilon^{-2}, \epsilon^{-2}, 1). \tag{24}$$

Counts refer to independent records; preparation and synchronization have separate physical costs.

Full bivariate time pairs. Return to Θ_ϵ. Retain N snapshots and replace the restricted temporal batches by K independent records $(R_0, L_{1,0}, R_1, L_{1,1})$. The new coordinates supply the target's own self-lag. Define R_ϵ^{bi} using the same loss.

Theorem 6 (Risk from full bivariate time pairs)**.** *On Θ_ϵ, with known ϵ and independent noiseless records as above,*

$$R_\epsilon^{\text{bi}}(N, K) \asymp 1 \wedge \max\left\{[\epsilon^2(N+K)]^{-1/2}, [N\epsilon^4 + K]^{-1/2}\right\}. \tag{25}$$

This holds uniformly for sufficiently small ϵ, $N \geq 1$, and $K \geq 0$. The lower bound allows a free exact trace; an attaining temporal estimator needs no supplied spectrum.

The target self-lag contains a strong rate signal. With $F_1(a, d) = e^{-d} + G_1(a, d)$, its sensitivity contrasts with the oriented lag as follows:

$$\begin{aligned} C_1(1) &= v_1 e^{-d_1} + \frac{x_1^2}{s} G_1(a, d_1), & \partial_{d_1} C_1(1) &= -v_1 e^{-d_1} + O(\epsilon^2), \\ \Gamma_{L_1,R}(1) &= x_1 F_1(a, d_1), & \partial_{d_1} \Gamma_{L_1,R}(1) &= O(\epsilon). \end{aligned} \tag{26}$$

Once amplitude is estimated from $N \asymp \epsilon^{-2}$ snapshots, full pairs require constant-order K for a fixed nonzero error target, whereas restricted pairs require $M \asymp \epsilon^{-2}$. The constant may be large; both total readout budgets remain of order ϵ^{-2}. The full proof is in Appendix B.7.

With $a = -\log(C_R(1)/s)$, the decoder inverts

$$d \longmapsto C_1(1; d) = v_1 e^{-d} + \frac{x_1^2}{s}(a + d)\phi_1(a, d) \tag{27}$$

which is strictly decreasing for small ϵ. Projected empirical moments give $\widehat{g} = (\widehat{x}_1/\widehat{s}) G_1(\widehat{a}, \widehat{d}_1)$; estimating x_1 still costs the first term of (25). Known bounded, independent endpoint noise preserves these risk orders after correction of same-time moments; the next result treats unknown sensor variances on a single trajectory.

Theorem 7 (Two-coordinate tracking with unknown sensor variances)**.** *On Θ_ϵ, observe $Y_t = X_t + \xi_t$ for $t = 0, \ldots, L$, where X_t is one stationary trajectory and $\xi_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \operatorname{diag}\omega)$ are independent of it. The unknown sensor variances obey $0 \leq \omega_j \leq \omega_{\max} < \infty$, with known upper bound. Let $R_\epsilon^{\text{traj},q}(L)$ denote normalized minimax response risk over (θ, ω) when observing either all four coordinates ($q = 4$) or only (R, L_1) ($q = 2$). For known small ϵ and $L \geq 3$,*

$$R_\epsilon^{\text{traj},4}(L) \asymp R_\epsilon^{\text{traj},2}(L) \asymp 1 \wedge (L\epsilon^2)^{-1/2}. \tag{28}$$

Positive-lag covariances retain the latent temporal signal without requiring sensor-variance correction. Define $\gamma_R(h) = \mathbb{E}(Y_{0,R} Y_{h,R})$. Then

$$a = \log \frac{\gamma_R(1)}{\gamma_R(2)}, \qquad s = \frac{\gamma_R(1)^2}{\gamma_R(2)}. \tag{29}$$

Two target self-lags similarly determine its rate and variance; same-time cross-coordinate covariance estimates coupling. All these statistics use only the source and target. Tracking the other two leaves therefore does not improve the optimal risk order in this class, even with unknown diagonal sensor noise. The proof and two-coordinate reduction are in Appendix B.8.

4.4 Choosing Coordinates and Sampling Times

For normalized error at most a fixed small $\delta > 0$, the restricted interface requires, and suffices up to class constants,

$$\min\{N\epsilon^2,\ N\epsilon^4 + M\epsilon^2,\ N\epsilon^4 + M_0\} \gtrsim \delta^{-2}. \tag{30}$$

Full pairs replace these terms by $\epsilon^2(N + K)$ and $N\epsilon^4 + K$. The smallest term identifies the unmet measurement need. A sufficient staged design estimates amplitude from snapshots, then adds full pairs to resolve rates. Full pairs also estimate amplitude, as the $N + K$ term shows. Other leaves need not be tracked; lag and parameter-class constants set the numerical budget.

With known root variance, the inverse self-lag estimator has asymptotic variance $(e^{2ah}+1)/(Kh^2)$ at an interior rate. Shortening the lag can improve a fast root-rate estimate while weakening information about a slower target. To select a lag for the response, fix the static moments and invert the two self-lags jointly. The source self-lag determines a; the target self-lag is strictly decreasing in d_1 throughout the feasible star domain. At an interior parameter, the resulting estimator obeys

$$\sqrt{K}(\widehat{g} - g) \Rightarrow \mathcal{N}(0, \mathcal{V}_h), \qquad \mathcal{V}_h = \nabla g^\top J_h^{-1} \Omega_h J_h^{-\top} \nabla g, \tag{31}$$

where $J_h = D_{(a,d_1)}(\Gamma_{RR}(h), \Gamma_{L_1 L_1}(h))$, Ω_h is the covariance of the two lag products, and the gradient is with respect to (a, d_1). This criterion weights uncertainty in both rates by their effects on the requested response. Its local, fixed-static-moment interpretation and proof are given in Appendix B.9.

When does temporal inversion improve on the prior center? On the calibrated odd-cycle class, fix m, h, set $\delta = 2w$, and use the decoder in (15). For $d_i = 2 + wD_i$, with $D_i = U_i - m^{-1}\sum_j U_j$ and independent $U_j \sim U[-1, 1]$, define $J_m(s) = \mathbb{E}|\Pi_{[-2,2]}(D_0 + sZ) - D_0|$ for independent standard normal Z. Let $g_c = \epsilon G_1(1, 2)$ and $q = |\partial_d G_1(1, 2)|$. Averaging over the prior and observation noise, as $w, \epsilon \to 0$ and $Kw^2 \to \kappa > 0$,

$$\frac{\mathbb{E}|g_c - g|}{\epsilon w q} \to c_m = \frac{1}{2} - \frac{1}{3m}, \qquad \frac{\mathbb{E}|\widehat{g}_K - g|}{\epsilon w q} \to J_m(\sigma_h/\sqrt{\kappa}), \quad \sigma_h^2 = \frac{e^{4h}+1}{h^2}.$$

The unique root $J_m(s_m^*) = c_m$ gives the local budget prediction

$$K_m^*(w, h) = \frac{e^{4h}+1}{h^2 w^2 (s_m^*)^2}. \tag{32}$$

Above this budget, timing precision outweighs the center's prior precision in the local experiment. Proposition 28 proves the limit and gives a Gaussian integral for computing the prediction before evaluating new observations.

5 Experiments

Our identification theory guides which temporal measurements to add. Without a supplied full spectrum, can they estimate the response accurately, and at what budget? Synthetic OU systems provide the exact unit-pulse response g at $H = 1$ for evaluation. Loss is $|\widehat{g} - g|/\epsilon$ for multiroot experiments and $|\widehat{g} - g|/|g|$ for stars; protocols are in Appendix D.

Can target timing outperform full snapshots? For $m = 3, 5, 7, 9$, static estimators receive K full state vectors; temporal inversion receives K pairs recording the target twice, .5 time units apart. All know the prior, variances, root–leaf covariances, and universal root rate. The static comparator uses the response posterior median under the full Gaussian likelihood and exact prior (32,768 Sobol points). At $K = 8192, 32768$, with 2,000 systems per cell, temporal inversion reduces mean error by 20.7–67.0%; all eight contrasts pass Holm correction. Thus, recording target persistence outperforms full snapshots in this family, despite fewer scalar readouts per record.

Does the gain survive equal readout budgets? On stars, static estimators receive snapshots and exact rate sum $T = 10$; targeted estimators receive snapshots, source-before-target pairs, and source autocorrelation pairs, without T. With equal scalar-endpoint counts, targeted moments lower mean error in all 16 cells (300 systems each; twelve positive adjusted intervals). Eight-start likelihood fits on 100 fresh systems per $\epsilon = .1, .05, .01$ give static errors $.1724, .2108, .2771$, versus targeted errors $.0626, .0684, .0553$. The temporal gain therefore survives both matched readout counts and likelihood-based estimation.

Why do target self-lags help? At matched temporal endpoints, K full source–target time pairs compete with K source-before-target plus K source-autocorrelation pairs. Both receive $N = \lceil 16/\epsilon^2 \rceil$ snapshots and no rates; $\epsilon \in \{.1, .05, .025, .01\}$ and $K \in \{16, 64, 256\}$. Full pairs lower mean error in 11/12 cells for moments (3,600 systems; seven positive adjusted intervals) and local likelihood (2,400 systems; six). A same-record ablation holds static moments and the root-rate estimate fixed, replacing cross-lag inversion with target-self-lag inversion: 11/12 means improve, with seven positive adjusted intervals. Using target persistence helps even with identical records, supporting its predicted informational role. The smallest-budget intervals include zero.

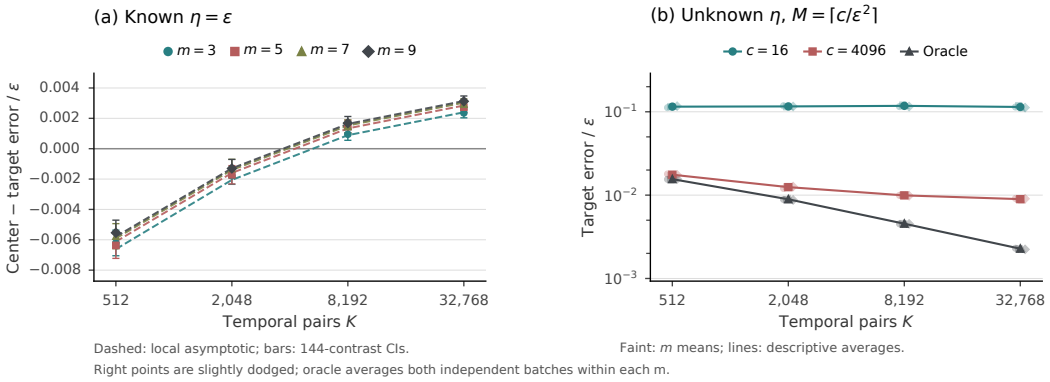


Figure 1: When temporal measurements help. Left: center-minus-inversion error with known amplitude; simultaneous 95% intervals and dashed local predictions. Positive values favor inversion. Right: mean inversion errors across four graph sizes, with estimated or known amplitude. Calibration budget M and temporal budget K control different sources of error.

How many temporal records are needed? We compare inversion with the response at the parameter-class center, freezing (32) before a new 48,000-system experiment. Predicted crossover budgets are $4878, 3914, 3586, 3419$ for $m = 3, 5, 7, 9$. All sixteen known-amplitude comparisons support the predicted directions: the center wins at $K = 512, 2048$, inversion at $8192, 32768$ (Figure 1, left). Each prediction lies within the observed transition bracket. The center's low-budget advantage shows the estimation noise that temporal inversion must overcome.

Why does amplitude calibration matter? The remaining 32 cells estimate $\eta \sim U[\epsilon, 2\epsilon]$ from M calibration pairs shared by both estimators. At $M\epsilon^2 = 4096$ and $K = 32768$, all four graphs favor temporal inversion, with loss .00876–.00913. With the smaller calibration budget $M\epsilon^2 = 16$, its loss remains .11181–.11893 (Figure 1, right). At this temporal budget, amplitude calibration remains the limiting source of error. Each cell has 1,000 systems; simultaneous intervals use one Bonferroni family of 144 paired contrasts, including oracle diagnostics.

Which sampling interval should we use? In the fast-root sensitivity stratum, $a \in [5.5, 6]$ and $d_1 \in [.8, 1]$. At $K = 1024$ noiseless pairs, shortening h from .5 to $1/6.5$ improves source-rate MAE from .605 to .406, but increases relative response error from .0741 to .1612. Sampling intervals should therefore be evaluated for the requested response, as in (31), since improving one decay-rate estimate can worsen response accuracy.

6 Conclusion

Our graph criteria determine when covariance and rate summaries identify a named response and which temporal measurements remove ambiguity. Matching risk bounds quantify their sampling cost; experiments show when they improve response accuracy. Together, these results guide which coordinates to track, how many records to collect, and how to balance amplitude calibration with temporal precision.

Limitations. The structural results assume a known two-layer graph and diagonal diffusion. The matching statistical bounds concern the stated star classes and multiroot families with specified covariance structure and linked rates; the common covariance amplitude can be learned. Experiments use synthetic OU systems, and finite-budget gains depend on coupling, lag choice, and the observation scheme. Measurement counts omit system-specific preparation costs. General multiroot trajectories, uncertain graphs, and nonstationary dynamics remain open extensions.

Ethics statement

This work develops mathematical identification and sampling results and evaluates them on synthetic stochastic systems. No human-subject, animal, or identifiable personal data were collected. Biological examples motivate the response questions.

Reproducibility statement

The main text specifies the observation schemes, parameter classes, and estimation targets. Appendix A gives the structural identification proofs, Appendix B the statistical proofs, and Appendix C the calibration and budget-transition analysis. Appendix D collects simulation protocols, estimator settings, budget matching, and uncertainty calculations. The experimental descriptions include parameter ranges, sample counts, decoder settings, and multiple-comparison procedures needed to interpret the reported tests.

AI use statement

Generative AI tools were used to improve the clarity and grammar of the manuscript. The authors reviewed the edited text and take responsibility for the final content.

References

Carlos Améndola, Tobias Boege, Benjamin Hollering, and Pratik Misra. Structural identifiability of graphical continuous Lyapunov models. *arXiv preprint arXiv:2510.04985*, 2025. URL https://arxiv.org/abs/2510.04985.

Milena Anguelova, Johan Karlsson, and Mats Jirstrand. Minimal output sets for identifiability. *Mathematical Biosciences*, 239(1):139–153, 2012. doi: 10.1016/j.mbs.2012.04.005.

Alexander P. Browning, Michael J. Chappell, Hamid Rahkooy, Torkel E. Loman, and Ruth E. Baker. Exact identifiability analysis for a class of partially observed near-linear stochastic differential equation models. *arXiv preprint arXiv:2503.19241*, 2025. URL https://arxiv.org/abs/2503.19241.

Jiahui Cheng and Sui Tang. Estimate the spectrum of affine dynamical systems from partial observations of a single trajectory data. *Inverse Problems*, 38(1):015004, 2022. doi: 10.1088/1361-6420/ac37fb.

Philipp Dettling, Roser Homs, Carlos Améndola, Mathias Drton, and Niels Richard Hansen. Identifiability in continuous Lyapunov models. *SIAM Journal on Matrix Analysis and Applications*, 44 (4):1799–1821, 2023. doi: 10.1137/22M1520311.

Philipp Dettling, Mathias Drton, and Mladen Kolar. On the Lasso for graphical continuous Lyapunov models. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, volume 236 of *Proceedings of Machine Learning Research*, pp. 514–550. PMLR, 2024. URL https://proceedings.mlr.press/v236/dettling24a.html.

Matthieu Gilson, Ruben Moreno-Bote, Adrián Ponce-Alvarez, Petra Ritter, and Gustavo Deco. Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. *PLOS Computational Biology*, 12(3):e1004762, 2016. doi: 10.1371/journal.pcbi.1004762.

Matthieu Gilson, Nikos E. Kouvaris, Gustavo Deco, and Gorka Zamora-López. Framework based on communicability and flow to analyze complex network dynamics. *Physical Review E*, 97(5): 052301, 2018. doi: 10.1103/PhysRevE.97.052301.

D. Joubert, J. D. Stigter, and J. Molenaar. Determining minimal output sets that ensure structural identifiability. *PLOS ONE*, 13(11):e0207334, 2018. doi: 10.1371/journal.pone.0207334.

Zijian Li, Yifan Shen, Kaitao Zheng, Ruichu Cai, Xiangchen Song, Mingming Gong, Guangyi Chen, and Kun Zhang. On the identification of temporal causal representation with instantaneous dependence. In *The Thirteenth International Conference on Learning Representations*, volume 2025, pp. 35203–35242, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/5726513facc85f802be4a25e77fb9765-Abstract-Conference.html.

Umberto Marini Bettolo Marconi, Andrea Puglisi, Lamberto Rondoni, and Angelo Vulpiani. Fluctuation-dissipation: Response theory in statistical physics. *Physics Reports*, 461(4–6):111–195, 2008. doi: 10.1016/j.physrep.2008.02.002.

Ameer Qaqish and Didong Li. Identifiability for Gaussian processes with holomorphic kernels. In *The Thirteenth International Conference on Learning Representations*, volume 2025, pp. 7311–7339, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/15f4cefb0e143c7ad9d40e879b0a9d0c-Abstract-Conference.html.

William Qian, Jacob A. Zavatone-Veth, Benjamin S. Ruben, and Cengiz Pehlevan. Partial observation can induce mechanistic mismatches in data-constrained models of neural dynamics. In *Advances in Neural Information Processing Systems*, volume 37, pp. 67467–67510, 2024. doi: 10.52202/079017-2154. URL https://doi.org/10.52202/079017-2154.

Cecilie Olesen Recke and Niels Richard Hansen. Identifiability and estimation in continuous Lyapunov models. *arXiv preprint arXiv:2603.17142*, 2026. URL https://arxiv.org/abs/2603.17142.

Saber Salehkaleybar. One intervention per component is enough: Towards identifiability in linear stochastic dynamics from steady state. *arXiv preprint arXiv:2609.19955*, 2026. URL https://arxiv.org/abs/2609.19955.

Rajesh Singh, Dipanjan Ghosh, and R. Adhikari. Fast Bayesian inference of the multivariate Ornstein–Uhlenbeck process. *Physical Review E*, 98(1):012136, 2018. doi: 10.1103/PhysRevE.98.012136.

Jared E. Toettcher, Orion D. Weiner, and Wendell A. Lim. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell*, 155(6):1422–1434, 2013. doi: 10.1016/j.cell.2013.11.004.

Gijs van Seeventer and Saber Salehkaleybar. Sign identifiability of causal effects in stationary stochastic dynamical systems. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pp. 6091–6124, 2026. URL https://proceedings.mlr.press/v337/seeventer26a.html.

Gherardo Varando and Niels Richard Hansen. Graphical continuous Lyapunov models. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*, volume 124 of *Proceedings of Machine Learning Research*, pp. 989–998, 2020. URL https://proceedings.mlr.press/v124/varando20a.html.

Andrew J. Wagenmaker, Max Simchowitz, and Kevin Jamieson. Task-optimal exploration in linear dynamical systems. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 10641–10652. PMLR, 2021. URL https://proceedings.mlr.press/v139/wagenmaker21a.html.

Andrew Ying. Causal identification for complex functional longitudinal studies. In *The Thirteenth International Conference on Learning Representations*, volume 2025, pp. 70816–70838, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/b078eaadc927d7f7ae67371c1363195b-Abstract-Conference.html.

A Structural identification proofs

A.1 Compatible fibers and reconstruction

We use the stable DAG model and initial-state response of Section 3. Diffusion is positive diagonal; $\Sigma \succ 0$ is named, and supplied rates are either their sum T or their unordered multiset, including multiplicities. For the two-layer drift (5), write $s_k = \Sigma_{kk}$, $v_i = \Sigma_{r+i,r+i}$, $x_{ik} = \Sigma_{r+i,k}$, and $z_{ij} = \Sigma_{r+i,r+j}$. Distinct roots are independent at stationarity, with $s_k = q_k/(2a_k) > 0$. The root–leaf and leaf–leaf Lyapunov entries are

$$B_{ik}s_k = (a_k + d_i)x_{ik}, \tag{33}$$

$$(d_i + d_j)z_{ij} = \sum_k (B_{ik}x_{jk} + B_{jk}x_{ik}) \quad (i \neq j). \tag{34}$$

Substitution gives the constraint (7), with $c_{ijk} = x_{ik}x_{jk}/s_k$ and $\Delta_{ij} = z_{ij} - \sum_k c_{ijk}$. Let K_Σ collect these rows and let $\theta = (a_1, \dots, a_r, d_1, \dots, d_m)^\top$. The remaining diffusion entries are

$$q_{r+i} = 2d_i v_i - 2\sum_k B_{ik}x_{ik}, \qquad q_k = 2a_k s_k. \tag{35}$$

Thus each positive θ satisfying $K_\Sigma\theta = 0$ determines unique B, Q. Conversely, these formulas satisfy every Lyapunov entry; stability makes Σ the stationary covariance. At fixed (Σ, T) the exact rate fiber is therefore

$$\{\theta > 0 : K_\Sigma\theta = 0,\ \mathbf{1}^\top\theta = T,\ q_j(\theta) > 0 \text{ for all } j\}. \tag{36}$$

Missing edges have $x_{ik} = 0$ and are reconstructed as zero. Nonzero specified edges retain their signs because $a_k + d_i > 0$. The strict inequalities leave a relatively open neighborhood in the affine fiber. The trace row is independent of K_Σ, since it evaluates positively on the true $\theta \in \ker K_\Sigma$.

For an active edge,

$$g'_{k\to i}(0) = B_{ik} = (x_{ik}/s_k)(a_k + d_i). \tag{37}$$

Its slope is constant on the fiber exactly when $e_{a_k}^\top + e_{d_i}^\top$ belongs to rowspan$[K_\Sigma; \mathbf{1}^\top]$. Fixed-horizon responses require the additional argument below.

For a general DAG in topological order, let $J = \{1, \dots, i-1\}$ and $A_{ii} = -\rho_i$. The (i, J) Lyapunov block yields

$$A_{iJ} = (\rho_i\Sigma_{iJ} - \Sigma_{iJ}A_{JJ}^\top)\Sigma_{JJ}^{-1}. \tag{38}$$

Every row is recursively linear in the rates; forbidden entries impose homogeneous linear equations. This fixed-covariance linear parametrization is consistent with graphical OU feasibility (van Seeventer & Salehkaleybar, 2026).

A.2 Finite-variance innovations

Proposition 8 (Second-order innovation universality)**.** *Let A be stable and let M_t be a two-sided, mean-zero, square-integrable Lévy process with independent coordinates and covariance rate $Q \succ 0$ diagonal. The stationary solution of $dX_t = AX_t\,dt + dM_t$ satisfies*

$$\Sigma = \int_0^\infty e^{Au}Qe^{A^\top u}\,du, \quad A\Sigma + \Sigma A^\top + Q = 0, \quad \Gamma(h) = e^{Ah}\Sigma \quad (h \geq 0). \tag{39}$$

When the innovation law may vary in this class, the compatible drifts given only covariance and supplied functions of A are exactly those of Gaussian OU. Hence the corresponding population identification results apply; Gaussian likelihood and sampling bounds are separate.

Proof. Stability defines the L^2 convolution $X_t = \int_{-\infty}^t e^{A(t-s)}\,dM_s$. Orthogonality of disjoint increments gives the covariance integral. Integrating its derivative gives the Lyapunov equation. In $X_{t+h} = e^{Ah}X_t + \int_t^{t+h} e^{A(t+h-s)}\,dM_s$, the second term is orthogonal to X_t, proving the lag identity; an initial-state pulse has mean response $e^{Ah}e_k$. Conversely, every stable A with positive diagonal $Q = -A\Sigma - \Sigma A^\top$ is realized both by Brownian noise and by independent rate-one compound-Poisson coordinates with equiprobable jumps $\pm\sqrt{q_i}$. The compatible drift sets therefore coincide. □

A.3 Trace boundary: rank witness and response ambiguity

We prove Theorem 1. Put $E = \binom{m}{2}$ and let B_m be the unsigned incidence matrix of K_m. For $m \geq 3$ it has rank m: equations $h_i + h_j = 0$ on a triangle force its entries to zero, and incident edges force all remaining entries to zero. Since the true positive rate vector lies in $\ker K_\Sigma$,

$$\operatorname{rank} K_\Sigma \leq \min(E, r + m - 1). \tag{40}$$

Choose a triangle with a tree edge attaching each remaining leaf. This spanning graph F has m edges and invertible unsigned incidence matrix B_F, by propagation from its odd cycle. Let $S = E(K_m)\backslash F$. Temporarily connect root 1 positively to every leaf and each of the next $\min(r-1, E-m)$ roots positively to a distinct pair in S. Give any further roots any positive pair of edges; set other edges to zero. Choose positive rates and diffusion. Every leaf pair shares root 1, so

$$\Delta_{ij} = \frac{2\sum_k a_k c_{ijk}}{d_i + d_j} > 0. \tag{41}$$

The leaf columns are $\operatorname{diag}(\Delta)B_m$. Each specially assigned extra-root column is supported on its own selected row. In a vanishing combination of these columns, restriction to F gives $\operatorname{diag}(\Delta_F)B_F h = 0$, hence $h = 0$; the selected rows then set every extra-root coefficient to zero. Their rank is $m + \min(r-1, E-m) = \min(E, r+m-1)$. Root 1's column is dependent by $K_\Sigma \theta = 0$ and $a_1 > 0$.

The witnessing minor is rational in rates, edges, and diffusion, with nonzero denominators on the positive domain. It remains nonzero after the temporarily zero edges receive sufficiently small nonzero perturbations. The resulting nonzero polynomial condition defines a Zariski-open dense subset of the fully connected model. By (36), the independent trace row yields uniqueness exactly when $r + m - 1 \leq E$, or $r \leq (m-1)(m-2)/2$.

For the converse for a specified edge $k \to i$, repeat the witness above the threshold with root k isolated. The other $r-1$ roots can achieve row rank E, since $r - 1 \geq E - m + 1$. With $T_{-k} = \mathbf{1}^\top \theta_{-k}$, define

$$h_{a_k} = 1, \qquad h_{-k} = -\theta_{-k}/T_{-k}. \tag{42}$$

Then $K_\Sigma h = \mathbf{1}^\top h = 0$, whereas $h_{a_k} + h_{d_i} = 1 - d_i/T_{-k} > 0$. Thus appending the slope row raises the rank of $[K_\Sigma; \mathbf{1}^\top]$. This minor stays nonzero when small nonzero edges are added to the isolated root. At the perturbed witness $x_{ik} \neq 0$, so (37) varies along the strictly feasible fiber. The same nonzero-minor argument proves generic nonidentification for each specified edge slope.

A.4 A fixed response horizon

For an active root–leaf edge the only directed path is that edge, so

$$g_{k\to i}(H) = \frac{x_{ik}}{s_k}(a_k + d_i)\phi_H(a_k, d_i). \tag{43}$$

Use the fully connected witness just constructed, with $h \in \ker K_\Sigma \cap \mathbf{1}^\perp$ and $x_{ik}(h_{a_k} + h_{d_i}) \neq 0$. Scaling (A, Q) to $(\tau A, \tau Q)$ preserves covariance and support. At this model h is still tangent to the fixed-$(\Sigma, \tau T)$ fiber. Writing $a = a_k, d = d_i$,

$$\begin{aligned} D_h g(H; \tau\theta) = \frac{x_{ik}}{s_k}\big[&(h_{a_k} + h_{d_i})\phi_H(\tau a, \tau d) \\ &+ \tau(a+d)(h_{a_k}\partial_a + h_{d_i}\partial_d)\phi_H(\tau a, \tau d)\big] \\ \longrightarrow\ &(x_{ik}/s_k)(h_{a_k} + h_{d_i})H \neq 0 \quad (\tau \downarrow 0). \end{aligned} \tag{44}$$

Here $\phi_H \to H$ and its rate derivatives stay bounded. For every fixed $H > 0$, some strictly feasible witness therefore has a varying response. Constancy would require a rank drop in $[K_\Sigma; \mathbf{1}^\top; \nabla_\theta g(H)]$. Its entries are real analytic; the witness supplies a nonidentically zero maximal minor. Its proper real-analytic zero set contains the exceptional models. This proves fixed-H generic nonidentification over the full parameter space, without asserting it on every preassigned trace slice or at all horizons simultaneously.

A.5 Pair-root topology and covariance inversion

In the pair-root family of Theorem 2, root 0 is universal, p additional roots select pairs S (possibly with repetitions), and $H = K_m \setminus S$. On every unselected pair, only root 0 is common, so $\Delta_{ij} = 2a_0c_{ij0}/(d_i + d_j) \neq 0$. On selected pairs Δ_{ij} may vanish through signed cancellation.

Each extra-root column of K_Σ is supported on its selected pair, with a nonzero coefficient. Delete the universal-root column, which is dependent by $K_\Sigma\theta = 0$. Ordering selected rows first leaves the matrix

$$\begin{pmatrix} D_S U_S & P \\ D_H U_H & 0 \end{pmatrix}, \qquad D = \operatorname{diag}(\Delta). \tag{45}$$

The extra-root block P has rank $|S|$, including with repeated pairs. Its columns eliminate every selected-row leaf entry, while D_H is invertible. Hence

$$\operatorname{rank} K_\Sigma = |S| + \operatorname{rank} U_H. \tag{46}$$

On a connected component, $U_H h = 0$ means $h_i = -h_j$ across every edge. A bipartite component, including an isolated vertex, contributes one free scalar; an odd cycle forces all entries in its component to zero. Consequently

$$\operatorname{rank} U_H = m - b(H), \qquad \operatorname{rank} K_\Sigma = |S| + m - b(H). \tag{47}$$

The independent trace row leaves dimension $p - |S| + b(H)$ in the $p + 1 + m$ rates. This is zero precisely when $p = |S|$ and $b(H) = 0$. Otherwise strict feasibility permits a continuum of distinct drifts with identical (Σ, T) and exact signed support. Distinct drifts differ in the derivative at zero of some entry of e^{At}, though not necessarily in a preselected response.

When the criterion holds, define

$$w_e = \frac{2c_{e0}}{\Delta_e}, \qquad U_H v = w, \quad d = a_0 v, \qquad v = U_H^+ w, \tag{48}$$

$$\beta_k = \frac{\Delta_{e_k}(v_{i_k} + v_{j_k}) - 2c_{e_k,0}}{2c_{e_k,k}}, \tag{49}$$

$$a_0 = \frac{T}{1 + \sum_i v_i + \sum_k \beta_k}, \qquad a_k = a_0\beta_k. \tag{50}$$

The denominators are nonzero at the true model, and the last is positive. These equations and (35) reconstruct A, Q. For a fixed graph the ratio perturbation obeys the sharp bound

$$\|\widehat{v} - v\|_2 \leq \frac{\|\widehat{w} - w\|_2}{\sigma_{\min}(U_H)}. \tag{51}$$

Equality is attained in a minimum left-singular direction. Conversion from covariance error to ratio error additionally needs denominator margins. For a cycle, $U^\top U = 2I + \operatorname{Adj}(C_\ell)$ has eigenvalues $2 + 2\cos(2\pi j/\ell)$; its minimum is zero for even ℓ, and

$$\sigma_{\min}(U_{C_\ell}) = 2\sin(\pi/(2\ell)) \sim \pi/\ell \qquad (\ell \text{ odd}). \tag{52}$$

A.5.1 Regular finite-sample and local guarantees

Proposition 9 (Pair-root plug-in stability)**.** *Fix dimensions and an identifiable pair-root graph. Assume a compact parameter set with covariance eigenvalues in* $[\lambda_-, \lambda_+]$, $s_k \geq s_* > 0$, $|\Delta_e| \geq \delta_* > 0$ *for unselected pairs,* $|c_{e_k,k}| \geq c_* > 0$ *for selected pairs, and trace denominator* $1 + \sum v_i + \sum \beta_k \geq D_* > 0$. *From* N *iid centered Gaussian snapshots and exact* T*, a guarded plug-in estimator satisfies, at any fixed horizon,*

$$\sup_\theta \mathbb{E}_\theta|\widehat{g} - g| \leq \frac{C}{\sqrt{N}}\big(1 + \sigma_{\min}(U_H)^{-1}\big). \tag{53}$$

C *depends on the dimensions, compact box, horizon and stated margins, but not otherwise on the incidence matrix.*

Proof. Gaussian fourth moments give $\mathbb{E}\|\widehat{\Sigma} - \Sigma\|_F \leq C/\sqrt{N}$. Covariance-to-ratio maps are uniformly smooth under the margins. Put $\sigma = \sigma_{\min}(U_H)$. On $\|\widehat{\Sigma} - \Sigma\|_F \leq c\min(1, \sigma)$, equations

(48)–(50) give errors bounded by $C(1+\sigma^{-1})\|\widehat{\Sigma}-\Sigma\|_F$ in the rates, edges, and fixed-horizon response; the trace denominator stays above $D_*/2$. Guard empirical denominators below half their margins and rates outside an enlarged compact box, output zero on a guard, and clip every response to a known bounded interval. No guard fires on the good event. Markov's inequality bounds its complement by $C(1+\sigma^{-1})/\sqrt{N}$, proving the claim. □

Proposition 10 (Local task difficulty)**.** *At an interior model satisfying the preceding margins, use free rates (subject to the exact trace), root and leaf variances, and allowed root–leaf covariances as coordinates θ. For $u=\nabla g(\theta_0)\neq 0$, define*

$$I_{ij}=\tfrac{1}{2}\operatorname{tr}(\Sigma_0^{-1}\partial_i\Sigma_0\Sigma_0^{-1}\partial_j\Sigma_0),\qquad V_g^2=u^\top I^{-1}u. \tag{54}$$

Then $I\succ 0$. If the center is known to the estimator, minimax absolute risk on $\|\theta-\theta_0\|_I\le\rho/\sqrt{N}$ satisfies, for fixed $\rho>0$ and sufficiently large N,

$$c_\rho V_g/\sqrt{N}\le\mathcal{R}_N^{\mathrm{loc}}\le C_\rho V_g/\sqrt{N}. \tag{55}$$

This is local at a fixed model, not uniform in weak coupling or graph size.

Proof. The smooth inverse (48)–(50) is a local left inverse of $\theta\mapsto\Sigma_\theta$, so its Jacobian is injective. Thus $b^\top Ib=\frac{1}{2}\|\Sigma_0^{-1/2}(\sum_j b_j\partial_j\Sigma_0)\Sigma_0^{-1/2}\|_F^2>0$ for $b\neq 0$. The Gaussian score is

$$S_j(X;\theta_0)=\tfrac{1}{2}\{X^\top\Sigma_0^{-1}(\partial_j\Sigma_0)\Sigma_0^{-1}X-\operatorname{tr}(\Sigma_0^{-1}\partial_j\Sigma_0)\}. \tag{56}$$

For $\Delta=\theta-\theta_0$, its mean is $I\Delta+O(\|\Delta\|^2)$ and covariance $I+O(\|\Delta\|)$. The estimator $g(\theta_0)+u^\top I^{-1}N^{-1}\sum_n S(X_n;\theta_0)$ therefore has bias $O(N^{-1})$ and variance $(V_g^2+o(1))/N$, uniformly on the local ellipsoid, proving the upper bound. For the lower bound take $v=I^{-1}u/V_g$ and $\theta_\pm=\theta_0\pm tv/\sqrt{N}$, with $0<t<\min(\rho,1/4)$. Their N-sample KL is $2t^2+o(1)$ and response separation is $2tV_g/\sqrt{N}+o(N^{-1/2})$. Pinsker and the two-point absolute-loss inequality prove the claim. □

A.5.2 Calibrated graph risk: shared construction

For (12), fix a sufficiently small numerical $\delta>0$, put $\alpha=1,\beta=7/5,d_*=2$, and let the q-regular graph H have full-column-rank unsigned incidence U_H. Every nonedge of H has one dedicated root, with

$$\beta_{(i,j)}=\frac{\alpha+\beta}{2d_*}(d_i+d_j)-\alpha,\qquad d_i\in[2-\delta,2+\delta],\quad\sum_i d_i=2m. \tag{57}$$

Unit variances and allowed root–leaf covariances ϵ are fixed and known. The dedicated-pair graph is $(m-1-q)$-regular; hence its root-rate sum is fixed, as is $T=1+2m+(7/5)|E(K_m)\setminus E(H)|$. The only varying covariance entries are

$$z_{ij}=\epsilon^2\left(1+\frac{2\alpha}{d_i+d_j}\right),\qquad (i,j)\in H, \tag{58}$$

$$z_{ij}=2\epsilon^2+\frac{2\epsilon^2(\alpha+\beta_{(i,j)})}{d_i+d_j}=2\epsilon^2+\frac{\epsilon^2(\alpha+\beta)}{d_*},\qquad (i,j)\notin H. \tag{59}$$

Reconstruct B,Q by (33) and (35). The root–leaf covariance block has Frobenius norm at most $Cm\epsilon$, and the leaf off-diagonal block has operator norm at most $Cm\epsilon^2$. Thus $\epsilon\le c_0/m$ gives covariance eigenvalues in a fixed positive interval. Each leaf has at most $m-2$ bounded-rate parents, so its diffusion is $2d_i-O(m\epsilon^2)>0$ after decreasing c_0. This verifies feasibility uniformly in the graph and m.

We use two common estimates. The pair-sum inverse is

$$s_e=d_i+d_j=f_\epsilon(z_e),\qquad f_\epsilon(z)=\frac{2\alpha\epsilon^2}{z-\epsilon^2}. \tag{60}$$

Choose a fixed smooth cutoff χ, equal to one on the compact range of z_e/ϵ^2 and supported inside $(1,\infty)$. Replacing the inverse globally by $4+\chi(z/\epsilon^2)\{2\alpha/(z/\epsilon^2-1)-4\}$ (defined as 4 off the cutoff support) makes it bounded, agrees at every true covariance, and gives $|f_\epsilon'|\le C\epsilon^{-2}$, $|f_\epsilon''|\le C\epsilon^{-4}$

on every sample. For empirical products $\widehat{z}_e = N^{-1}\sum_n L_i^{(n)} L_j^{(n)}$ and any deterministic weights on distinct leaf pairs,

$$\mathrm{Var}\left(\sum_e w_e(\widehat{z}_e - z_e)\right) \le \frac{C}{N}\|w\|_2^2, \qquad \mathbb{E}(\widehat{z}_e - z_e)^2 \le C/N. \tag{61}$$

Indeed $\sum_e w_e L_i L_j = L^\top A_w L$, where the two symmetric edge entries are $w_e/2$, so $\|A_w\|_F^2 = \|w\|_2^2/2$. Gaussian fourth moments give variance $2\,\mathrm{tr}(A_w\Sigma_{LL}A_w\Sigma_{LL}) \le C\|w\|_2^2$; independence gives the factor N^{-1}. We also use the dimension-free Gaussian KL bound $D_{\mathrm{KL}}(P_\Sigma\|P_{\Sigma'}) \le C\|\Sigma - \Sigma'\|_F^2$ under these eigenvalue bounds. It follows by diagonalizing $\Sigma'^{-1/2}\Sigma\Sigma'^{-1/2}$ and bounding $x - 1 - \log x$ on its fixed positive spectral interval. For two models separated by response Δ, the maximum absolute risk is at least $\Delta(1 - \mathrm{TV})/4$, by classifying according to the nearer response.

Proposition 11 (Regular-graph risk)**.** *For the preceding class and information, let B have orthonormal columns spanning $\mathbf{1}^\perp$, $M = U_H B$, and $\sigma = \sigma_{\min}(M) > 0$. Then*

$$\kappa_i(H) = \|e_i^\top B M^+\|_2 = \sup_{\mathbf{1}^\top h = 0,\ \|U_H h\|_2 \le 1} |h_i|. \tag{62}$$

The matching risk law (13) *holds for $N \ge C\epsilon^{-4}(|E(H)| + \sigma^{-2})$. Constants are graph-independent and may depend on the fixed rate box.*

Proof. For the upper bound, put $\widehat{s}_e = f_\epsilon(\widehat{z}_e)$ and project $2 + e_i^\top B M^+(\widehat{s} - 4\mathbf{1})$ onto $[2-\delta, 2+\delta]$. Since $s = 4\mathbf{1} + MB^\top(d - 2\mathbf{1})$, the linear Taylor term has coefficient norm at most $C\kappa_i/\epsilon^2$. Equation (61) bounds its expected absolute value by $C\kappa_i/(\epsilon^2\sqrt{N})$. The expected absolute remainder is at most $C\|e_i^\top B M^+\|_1/(N\epsilon^4) \le C\sqrt{|E(H)|}\kappa_i/(N\epsilon^4)$, which the sample-size condition absorbs into the linear bound. Projection and Lipschitz continuity of $G_1(1, d_i)$ finish the upper bound.

For the lower bound put $v = B^\top e_i$ and $h = BM^+(M^+)^\top v/\kappa_i$. Then $\mathbf{1}^\top h = 0$, $\|U_H h\|_2 = 1$, $h_i = \kappa_i$, and $\|h\|_\infty \le \sigma^{-1}$. Take $d(t) = 2\mathbf{1} + th$ with $t = c/(\epsilon^2\sqrt{N})$. The sample-size condition keeps this path inside the box. Dedicated-pair covariances and the trace remain fixed, while $\|\Sigma_t - \Sigma_0\|_F \le C\epsilon^2 t$. The N-sample KL is at most $CN\epsilon^4 t^2$. Since $\partial_d G_1(1,2) = 5e^{-2} - 2e^{-1} \ne 0$, choosing the fixed box sufficiently small gives normalized response separation at least $ct\kappa_i$. A small constant in t makes TV at most $1/2$ by Pinsker, and the two-point bound proves the result. □

Theorem 12 (All-budget odd-cycle risk)**.** *For $H = C_m$ with odd $m \ge 3$, denote the above class by $\mathcal{C}_{m,\epsilon} = \mathcal{C}_{C_m,\epsilon}$. The all-budget law* (14) *holds uniformly in m, N, ϵ for $\epsilon \le c_0/m$, even when the graph, trace, variances, universal-root rate, and allowed root–leaf covariances are supplied.*

Proof. Let $U = U_{C_m}$ and $h_j = \sqrt{2/m}\cos(2\pi(m-1)j/(2m))$, $j = 0, \dots, m-1$. Then $\mathbf{1}^\top h = 0$, $\|h\|_2 = 1$, $h_0 = \|h\|_\infty = \sqrt{2/m}$, and $\|Uh\|_2 = \sigma_m = 2\sin(\pi/(2m))$. For $d(t) = 2\mathbf{1} + th$, $|t| \le t_0\sqrt{m}$, the common construction is feasible and

$$\|\Sigma_t - \Sigma_0\|_F \le C\epsilon^2|t|\|Uh\|_2 = C\epsilon^2\sigma_m|t|. \tag{63}$$

Thus KL is at most $CN\epsilon^4\sigma_m^2 t^2$. Taking $t = c\min\{\sqrt{m}, (N\epsilon^4\sigma_m^2)^{-1/2}\}$ keeps TV below $1/2$ and gives normalized response gap $c|t|/\sqrt{m}$ by the derivative used above. Since $\sigma_m \asymp m^{-1}$, the lower bound is $c(1 \wedge \sqrt{m}/(\epsilon^2\sqrt{N}))$.

For the upper bound project the first coordinate of $U^{-1}(f_\epsilon(\widehat{z}_e))_{e\in C_m}$ to the rate box. Odd-cycle recursion shows that every entry in a row of U^{-1} has absolute value $1/2$; its ℓ_2 and ℓ_1 norms are $\sqrt{m}/2$ and $m/2$. Taylor expansion and (61) bound expected rate error by $C\{\sqrt{m}/(\epsilon^2\sqrt{N}) + m/(N\epsilon^4)\}$. For $N \ge m\epsilon^{-4}$ the second term is dominated by the first; projection and the response map preserve this bound. For smaller N, $\widehat{g} = 0$ has uniformly bounded normalized risk. □

Proposition 13 (Odd-cycle target self-lag)**.** *On $\mathcal{C}_{m,\epsilon}$, take K independent stationary target-leaf pairs at fixed $h > 0$. The center-corrected estimator* (15) *satisfies* (16)*, with C_h independent of m, K, ϵ.*

Proof. Let P_0 be the $m - 2$ parents of leaf 0. The leaf row of $e^{Ah}\Sigma$ gives the exact self-lag

$$\rho_h(d) = e^{-d_0 h} + \epsilon^2 \sum_{k\in P_0} (a_k(d) + d_0)\phi_h(a_k(d), d_0). \tag{64}$$

At $d^c = 2\mathbf{1}$ the sum is $3\phi_h(1,2) + (m-3)(17/5)\phi_h(7/5,2)$. All rates lie in fixed positive compact intervals, and the dedicated rates are affine in the incident leaf rates. Bounded first derivatives therefore give

$$|\rho_h(d) - e^{-d_0 h} - b_{m,\epsilon}(h)| \le C_h m\epsilon^2 \|d - d^c\|_\infty \le C_h m\epsilon^2 \delta. \tag{65}$$

A unit-variance Gaussian pair has product variance $1+\rho_h^2 \le 2$, so $\mathbb{E}|\widehat{\rho}_h - \rho_h| \le \sqrt{2/K}$. Projection toward the true $e^{-d_0 h}$ is nonexpansive, and $-h^{-1}\log$ is Lipschitz on I_h. Applying the Lipschitz response map proves the bound. The center approximation can persist at fixed (m,ϵ,δ) as $K \to \infty$; the sampling term uses Gaussian fourth moments, not merely finite variance. □

Corollary 14 (Unknown pair-root support)**.** *The same odd-cycle risk order holds over relabelings within the known root and leaf sets when their allowed root–leaf support and the identity of the universal root are not supplied, but the class form, m, and ϵ remain known.*

Proof. The known-support class gives the lower bound. For the upper bound threshold empirical root–leaf covariances at $\epsilon/2$, separating their population values 0 and ϵ. Gaussian products have uniformly bounded subexponential norms; Bernstein and a union bound over $O(m^3)$ entries give support error probability $Cm^3 e^{-cN\epsilon^2}$. On correct recovery, the universal root is the unique degree-m root for $m \ge 5$ and the only root for $m = 3$. Reject supports outside the stipulated class and apply the preceding estimator otherwise, always clipping its response to the known range. On the correct-support event the output equals the known-support estimator on the same data, so its risk is bounded by that estimator's unconditional risk plus the bounded loss on support error. For $N \ge m\epsilon^{-4}$ and $\epsilon \le c_0/m$, the exponential term is at most $C\sqrt{m}/(\epsilon^2\sqrt{N})$; for smaller N use the constant estimator. The separation gap is part of this class. □

A.5.3 Discrete-time independent-innovation analogue

Theorem 15. *For every odd $m \ge 3$ there is a stable two-layer Gaussian VAR(1) class $\mathcal{D}_{m,\epsilon}$ with pair-root support, positive diagonal innovation covariance, unit stationary variances, known universal-root coefficient $\alpha = 3/5$, and allowed root–leaf covariances ϵ, for which the one-step universal-root-to-leaf-0 response satisfies*

$$\inf_{\widehat{g}} \sup_{\mathcal{D}_{m,\epsilon}} \frac{\mathbb{E}|\widehat{g} - g|}{\epsilon} \asymp 1 \wedge \frac{\sqrt{m}}{\epsilon^2\sqrt{N}} \quad (0 < \epsilon \le c_0/m). \tag{66}$$

Observations are N iid full stationary vectors; graph, α, variances, and root–leaf covariances may be known, but no trace or spectrum is supplied.

Proof. Write $X_{t+1} = FX_t + \eta_t$, with root coefficients $a_k \in (0,1)$, leaf coefficients $d_i \in (0,1)$, and edges b_{ik}. Independent innovations give

$$x_{ik} = a_k b_{ik} s_k + a_k d_i x_{ik}, \qquad b_{ik} = \frac{1 - a_k d_i}{a_k s_k} x_{ik}. \tag{67}$$

Substituting into $(1 - d_i d_j) z_{ij} = \sum_k (b_{ik} b_{jk} s_k + b_{ik} d_j x_{jk} + b_{jk} d_i x_{ik})$ yields

$$z_{ij} = \sum_k \frac{x_{ik} x_{jk}}{s_k} \frac{a_k^{-2} - d_i d_j}{1 - d_i d_j}. \tag{68}$$

Set variances to one, allowed $x_{ik} = \epsilon$, $\alpha = 3/5$, baseline dedicated coefficient $\beta = 7/10$, and baseline leaf coefficient $d_* = 1/2$. For $d_i \in [d_* - \delta, d_* + \delta]$ set

$$a_{(i,j)}^{-2} = C_*(1 - d_i d_j) - \alpha^{-2} + 2 d_i d_j, \qquad C_* = \frac{\alpha^{-2} + \beta^{-2} - 2d_*^2}{1 - d_*^2}. \tag{69}$$

For sufficiently small fixed δ, its right side stays above one and all coefficients stay in compact subsets of $(0,1)$. Dedicated-pair covariances are exactly $\epsilon^2 C_*$, whereas

$$z_e = \epsilon^2 \frac{\alpha^{-2} - p_e}{1 - p_e}, \qquad p_e = d_i d_j, \quad e \in C_m. \tag{70}$$

All other entries are fixed. The same block-norm bounds give uniformly positive covariance for $\epsilon \le c_0/m$. Root innovation variances are $1-a_k^2$ and leaf innovation variances are $1-d_i^2-\sum_k(b_{ik}^2+2d_ib_{ik}\epsilon) = 1-d_i^2-O(m\epsilon^2) > 0$. The displayed stationary equations set every off-diagonal innovation covariance to zero.

For the lower bound use the previous odd-cycle vector h and $d_i(t) = d_* \exp(th_i)$, $|t| \le t_0\sqrt{m}$. Then $p_e(t) = d_*^2 \exp(t(Uh)_e)$ and $\|\Sigma_t - \Sigma_0\|_F \le C\epsilon^2\sigma_m|t|$. The N-sample KL is at most $CN\epsilon^4\sigma_m^2t^2$; $g_t = \epsilon(1-\alpha d_0(t))/\alpha$ has normalized separation at least $c|t|/\sqrt{m}$. The same choice of t and two-point bound prove the lower rate. For the upper bound,

$$p_e = \frac{\alpha^{-2} - z_e/\epsilon^2}{1 - z_e/\epsilon^2}, \qquad U(\log d_i)_i = (\log p_e)_e. \tag{71}$$

The map $z_e \mapsto \log p_e$ has a bounded smooth extension whose first two derivatives are $O(\epsilon^{-2})$ and $O(\epsilon^{-4})$. Apply it to empirical covariances, multiply by U^{-1}, project $\log d_0$ to its compact interval, and substitute into g. The row norms and (61) give error $C\{\sqrt{m}/(\epsilon^2\sqrt{N}) + m/(N\epsilon^4)\}$. Switch to a constant estimator for $N < m\epsilon^{-4}$. □

This is a separate innovation model: sampling OU at lag h instead produces $Q_h = \int_0^h e^{Au}Qe^{A^\top u}\,du$, generally nondiagonal.

A.6 Minimal scalar temporal probes

Theorem 16 (Slope probes)**.** *In a strictly feasible pair-root model with all specified edges nonzero, put $f = p - |S| + b(H)$. Given (Σ, T), exactly f appropriately selected coordinate autocovariance slopes suffice to identify the whole drift. Given only Σ, the universal-root slope and f further slopes suffice. Fewer continuous scalar measurements cannot locally identify the whole drift. Replacing these slopes by autocovariances at one sufficiently small positive lag preserves local identification.*

Proof. Define the one-sided slopes by

$$\ell_u = \frac{d}{dh}\operatorname{Cov}(X_u(t+h), X_u(t))\Big|_{h=0+} = (A\Sigma)_{uu}. \tag{72}$$

The root and leaf expressions are

$$\ell_k = -s_k a_k, \tag{73}$$

$$\ell_{r+i} = \sum_k (x_{ik}^2/s_k)a_k - \nu_i d_i, \qquad \nu_i = v_i - \sum_k x_{ik}^2/s_k > 0. \tag{74}$$

The last inequality is the Schur complement of the diagonal root covariance block. Hence $\ell = L_\Sigma\theta$, with

$$L_\Sigma = \begin{pmatrix} -\operatorname{diag}(s_k) & 0 \\ (x_{ik}^2/s_k)_{ik} & -\operatorname{diag}(\nu_i) \end{pmatrix} \tag{75}$$

invertible. Let N_0, N_T span $\ker K_\Sigma$ and $\ker K_\Sigma \cap \mathbf{1}^\perp$, of dimensions $f+1, f$. The matrices $L_\Sigma N_0, L_\Sigma N_T$ have full column rank; selecting independent rows gives injective linear maps on the corresponding affine fibers. The universal-root row is nonzero on $\ker K_\Sigma$, since it evaluates to $-s_0a_0$ on the true rate vector, and can therefore be included in a row basis. Reconstruction then gives A, Q.

Each strictly feasible fiber contains a relatively open set of its stated dimension. Invariance of dimension rules out a continuous injection of this set into fewer real coordinates. This lower count is for the whole drift. For finite lag, with covariance fixed,

$$c_u(h;\theta) = e_u^\top e^{A(\theta)h}\Sigma e_u = \Sigma_{uu} + h(L_\Sigma\theta)_u + O(h^2). \tag{76}$$

The remainder is continuously differentiable in θ, since $A(\theta)$ is linear. On fiber coordinates the selected map has Jacobian $hL_{\text{selected}}N + O(h^2)$, nonsingular for sufficiently small positive h. The inverse function theorem gives the claimed local, rather than global, finite-lag conclusion. □

A.7 Root-only completion and task certificates

We prove the finite-lag part of Theorem 2 and the task test (11). Here selected pairs are distinct and $\Delta_e \neq 0$ for every selected pair; every specified edge is nonzero. The unselected equations are

$$U_H d = a_0 w, \qquad w_e = 2c_{e0}/\Delta_e. \tag{77}$$

Their denominators are nonzero because only root 0 is shared. Let $L = U_H^+ w$ and let N_H be a basis of $\ker U_H$. Consistency gives the full leaf-rate fiber

$$d = a_0 L + N_H t, \qquad t \in \mathbb{R}^{b(H)}. \tag{78}$$

For the extra root k on pair e_k, the remaining equation gives

$$\begin{aligned} a_k &= a_0 \beta_k + \lambda_k U_{e_k} N_H t, \\ \beta_k &= \frac{\Delta_{e_k} U_{e_k} L - 2c_{e_k 0}}{2c_{e_k k}}, \qquad \lambda_k = \frac{\Delta_{e_k}}{2c_{e_k k}} \neq 0. \end{aligned} \tag{79}$$

Strict positivity of rates and the linear diffusion entries leaves a relatively open convex cone in (a_0, t) of dimension $b(H) + 1$.

At any fixed $h > 0$, including coincident rates,

$$C_k(h) = s_k e^{-a_k h}, \qquad a_k = -h^{-1} \log\{C_k(h)/s_k\}. \tag{80}$$

After root 0 is measured, additional roots R determine t exactly when $U_R N_H$ has full column rank. Such $b(H)$ rows exist: if $U_S N_H v = 0$, then stacking these with $U_H N_H v = 0$ gives $U_{K_m} N_H v = 0$. The triangle argument makes $\ker U_{K_m} = \{0\}$, whence $v = 0$. The recovered rates reconstruct A, Q globally. Invariance of dimension on the open $(b(H) + 1)$-dimensional fiber proves the minimum scalar count without trace or spectrum.

For the task test, remaining changes satisfy $\Delta t \in \ker(U_R N_H)$. For an active edge,

$$\mathcal{L}\{g_{k \to i}\}(z) = \frac{x_{ik}}{s_k} \frac{a_k + d_i}{(z + a_k)(z + d_i)}. \tag{81}$$

If $k = 0$, its rate is fixed, so the whole curve is constant exactly when d_i is constant, equivalent to $e_i^\top N_H \in \text{rowspan}(U_R N_H)$. For an extra root, the rational function determines $\{a_k, d_i\}$, since the numerator is nonzero and constant in z. Along a connected affine fiber an unordered pair of fixed values cannot vary continuously. Thus both rates must be fixed, giving the additional condition $U_{e_k} N_H \in \text{rowspan}(U_R N_H)$ by (79). Failure of either condition gives a local strictly feasible change of the curve. The noncancellation hypothesis is needed here: $\Delta_e = 0$ can make an extra-root rate insensitive to leaf ambiguity. Positive edges exclude this cancellation; arbitrary signed edges do not. The criteria concern the entire curve, not equality at one horizon.

Perturbation bound. Let P index the $b(H) + 1$ probed rates and U the remaining rates. The full-column-rank augmented system $[K_\Sigma; E_P]$ implies $K_{\Sigma,U}$ has full column rank. With $\widehat{K} = K_{\widehat{\Sigma}}$ and logarithmic root estimates $\widehat{a}_P$, put $\widehat{\theta}_P = \widehat{a}_P$ and

$$\widehat{\theta}_U = \arg\min_z \|\widehat{K}_U z + \widehat{K}_P \widehat{a}_P\|_2. \tag{82}$$

If $\gamma = \sigma_{\min}(K_{\Sigma,U})$ and $\varepsilon_K = \|\widehat{K} - K_\Sigma\|_2 < \gamma$, then

$$\|\widehat{\theta}_U - \theta_U\|_2 \leq \frac{\varepsilon_K \|\theta\|_2 + \|\widehat{K}_P\|_2 \|\widehat{a}_P - a_P\|_2}{\gamma - \varepsilon_K}. \tag{83}$$

Indeed, the residual at θ_U is $(\widehat{K} - K_\Sigma)\theta + \widehat{K}_P(\widehat{a}_P - a_P)$; apply $\widehat{K}_U^+$ and Weyl's bound. For $G_k = C_k(h) > 0$, relative errors in $\widehat{G}_k, \widehat{s}_k$ at most $1/2$ give

$$|\widehat{a}_k - a_k| \leq \frac{2}{h} \left(\frac{|\widehat{G}_k - G_k|}{G_k} + \frac{|\widehat{s}_k - s_k|}{s_k} \right), \tag{84}$$

using $|\log(1 + u)| \leq 2|u|$. For N iid noiseless Gaussian snapshots and M independent pairs per probed root, covariance-product errors have root-mean-square orders $N^{-1/2}$ and $M^{-1/2}$. The latter also holds with independent mean-zero endpoint noise of bounded variance, independent of the process: the lag product stays unbiased and has bounded variance. On fixed compact interior sets with s_k, G_k, γ bounded below, the rational map $\Sigma \mapsto K_\Sigma$ and the displayed bounds give $O_P(N^{-1/2} + M^{-1/2})$ error in rates, drift and smooth responses. This assertion is for the unconstrained interior fit with these margins.

A.8 Oriented source–target observations

Proposition 17 (Oriented-lag sufficiency)**.** *For any active root–leaf edge $k \to i$ in a two-layer DAG, let $s = \Sigma_{kk} > 0$, $x = \Sigma_{r+i,k} \neq 0$, $a = a_k$, and $d = d_i$. At any fixed $h > 0$, $(s, x, \Gamma_{kk}(h), \Gamma_{r+i,k}(h))$ determines its entire response curve without a trace or spectrum. The reverse cross-lag generally does not determine d.*

Proof. The root and leaf diagonal entries of e^{Ah} are e^{-ah} and e^{-dh}, and

$$(e^{Ah})_{r+i,k} = B_{ik}\phi_h(a,d), \qquad \phi_h(a,d) = \int_0^h e^{-au-d(h-u)}\,du. \tag{85}$$

Root independence and $\Gamma(h) = e^{Ah}\Sigma$ give

$$\Gamma_{kk}(h) = se^{-ah}, \qquad \Gamma_{r+i,k}(h) = xe^{-dh} + B_{ik}s\phi_h(a,d). \tag{86}$$

Since $B_{ik}s = (a+d)x$ and $(d-a)\phi_h = e^{-ah} - e^{-dh}$, including at $a = d$ by continuity,

$$\frac{\Gamma_{r+i,k}(h)}{x} = e^{-ah} + 2a\phi_h(a,d), \qquad \partial_d\left(\frac{\Gamma_{r+i,k}(h)}{x}\right) = -2a\int_0^h (h-u)e^{-au-d(h-u)}\,du < 0. \tag{87}$$

The root logarithm determines a. The last quotient decreases from $2 - e^{-ah}$ to e^{-ah} as d goes from zero to infinity, so the true observation has a unique positive inverse d. Then $B_{ik} = (a+d)x/s$ gives $B_{ik}\phi_t(a,d)$ for all $t > 0$. The reverse lag is $\Gamma_{k,r+i}(h) = xe^{-ah}$, independent of d. Only the same-time moments s, x, not the full covariance, are needed. □

A.9 Self-lags for multiple targets

Proposition 18 (Multi-target completion)**.** *In an arbitrary stable two-layer DAG with positive diagonal diffusion, let I be a set of target leaves and $P(I)$ their distinct root parents. Named stationary covariance and, at any fixed $h > 0$, one self-lag for each node in $I \cup P(I)$ determine every incoming root-to-target response curve. No distinct-rate assumption is needed.*

Proof. For each parent, $C_k(h) = s_k e^{-a_k h}$ determines a_k. For target i put $c_{ik} = x_{ik}^2/s_k$ and $\nu_i = v_i - \sum_{k\in P(i)} c_{ik} > 0$, the root-block Schur complement. The leaf row of $e^{Ah}\Sigma$ gives

$$\begin{aligned} C_i(h) &= v_i e^{-d_i h} + \sum_{k\in P(i)} B_{ik}x_{ik}\phi_h(a_k,d_i) \\ &= \nu_i e^{-d_i h} + \sum_{k\in P(i)} c_{ik}\{e^{-a_k h} + 2a_k\phi_h(a_k,d_i)\}. \end{aligned} \tag{88}$$

Here we used the edge-recovery and kernel identities above. Holding the measured parent rates fixed,

$$\partial_{d_i} C_i(h) = -h\nu_i e^{-d_i h} - 2\sum_{k\in P(i)} a_k c_{ik}\int_0^h (h-u)e^{-a_k u - d_i(h-u)}\,du < 0. \tag{89}$$

Thus each observed self-lag uniquely determines its positive target rate and hence all incoming coefficients and curves. Parent rates are shared across targets, giving the sufficient count $|P(I)| + |I|$. Squared covariances and $\nu_i > 0$ make the argument valid with signed edges, cancellation, or repeated rates. □

Parent-rate measurements cannot simply be removed. In a one-root, one-leaf model take $s = v = 1$, $(a,d) = (1,2)$, and sufficiently small $x > 0$, so covariance and diffusion are positive. The nonzero target derivative gives, by the implicit-function theorem, a local curve $d(a)$ preserving both Σ and $C_L(1)$. Its slope at $a = 1$ is $O(x^2)$, whereas

$$\partial_a G_1(a,d)|_{(1,2)} = e^{-1} - 4e^{-2} \neq 0. \tag{90}$$

Thus $xG_1(a, d(a))$ varies along this curve for small x. This establishes insufficiency of target self-lag alone, not necessity of every scalar in the multi-target design.

A.10 Full anonymous spectrum

Theorem 19 (Anonymous rate assignments)**.** *In the complete two-layer model with pairwise distinct positive rates and positive diagonal diffusion, $(\Sigma, \mathcal{R})$ generically identifies A for every $r \geq 1, m \geq 3$. With one root and two leaves, two different named response curves can share these observations.*

Proof. Let $\theta = (a_1, \ldots, a_r, d_1, \ldots, d_m)$ be the true assignment. Any compatible permutation θ' must satisfy

$$\sum_k c_{ijk} \left[a_k \frac{d'_i + d'_j}{d_i + d_j} - a'_k \right] = 0 \quad \text{for every } i < j. \tag{91}$$

Fix an incorrect permutation. If all these equations were identities in the model parameters, shrink every edge except those from an arbitrary root k to an arbitrary pair (i, j). Positive rates and diffusion can be retained, and only c_{ijk} survives in that row. The supposed identities would force

$$\frac{d'_i + d'_j}{d_i + d_j} = \frac{a'_k}{a_k} \quad \text{for every } k, i < j. \tag{92}$$

All roots therefore share a scale factor γ. Full column rank of U_{K_m} gives $d' = \gamma d$ as well. Permutation preserves the total rate sum, so $\gamma = 1$; distinctness then makes the permutation correct, a contradiction. For each wrong permutation, at least one residual is a nonzero rational function. Clearing denominators and excluding the finite union of its proper algebraic zero sets leaves only the true assignment. Reconstruction gives unique A, Q.

For the obstruction take one root, $a = 1$, leaf rates $(2, 3)$, unit variances, $x_1 = x_2 = \varepsilon > 0$, and $z_{12} = 7\varepsilon^2/5$. Edges are $(3\varepsilon, 4\varepsilon)$ and diffusion diagonal is $(2, 4 - 6\varepsilon^2, 6 - 8\varepsilon^2)$. Swapping the leaf rates gives edges $(4\varepsilon, 3\varepsilon)$ and diffusion $(2, 6 - 8\varepsilon^2, 4 - 6\varepsilon^2)$. For sufficiently small ε both are strictly feasible, with the same named covariance and spectrum $\{1, 2, 3\}$. Their first-leaf slopes 3ε and 4ε differ, hence so do their curves for small positive horizons. □

A.11 Worked three-leaf inverse

For one root and three leaves with nonzero x_i, the leaf-pair equation gives

$$w_{ij} = \left(\frac{s z_{ij}}{x_i x_j} - 1 \right)^{-1} = \frac{d_i + d_j}{2a}. \tag{93}$$

Solving the three pair sums and then the trace gives

$$\begin{aligned} \ell_i = d_i/a = w_{ij} + w_{ik} - w_{jk}, \qquad \{i, j, k\} = \{1, 2, 3\}, \\ a = \frac{T}{1 + \sum_i \ell_i}, \qquad d_i = a\ell_i, \qquad g_{0 \to i}(H) = (x_i/s) G_H(a, d_i). \end{aligned} \tag{94}$$

For $x_i = O(\epsilon)$ and bounded positive rates, the informative residual $z_{ij} - x_i x_j / s$ is $O(\epsilon^2)$; this identifies the quantity whose inversion amplifies sampling error.

A.12 Chains: naming only the rates needed by the task

In a directed chain with nonzero edges $b_i = A_{i,i-1}$, $i \geq 2$, and diagonal rates $r_i > 0$, the root-column equation is

$$(r_i + r_1)\Sigma_{i1} = b_i \Sigma_{i-1,1}. \tag{95}$$

All these covariances are nonzero. Multiplication telescopes to

$$\prod_{i=2}^{n} b_i = \frac{\Sigma_{n1}}{\Sigma_{11}} \prod_{i=2}^{n} (r_i + r_1). \tag{96}$$

The lower-bidiagonal resolvent therefore gives

$$\mathcal{L}\{g_{1 \to n}\}(z) = \frac{\Sigma_{n1}}{\Sigma_{11}} \frac{\prod_{i=2}^{n} (r_1 + r_i)}{\prod_{i=1}^{n} (z + r_i)}. \tag{97}$$

Thus the source rate, unordered spectrum, and one covariance ratio suffice for the end-to-end curve, without naming every intermediate rate.

For $n \geq 3$ and pairwise distinct rates, the source can itself be named generically. On the first three nodes write $(s, x, y, z, w) = (\Sigma_{11}, \Sigma_{21}, \Sigma_{31}, \Sigma_{22}, \Sigma_{32})$ and rates (a, b, c). Eliminating the two edges from the Lyapunov equations gives

$$a(yz/x + xy/s) + b(xy/s - w) + c(yz/x - w) = 0. \tag{98}$$

This excludes generic wrong spectral triples. A witness has rates $(1, 2, 4)$, edges $(1, 1)$, and diffusion $2I$, yielding $(s, x, y, z, w) = (1, 1/3, 1/15, 2/3, 11/90)$. The coefficient vector is $(14, -9, 1)/90$: it annihilates $(1, 2, 4)$ but none of its five other orderings. Choose any additional positive rates algebraically independent over $\mathbb{Q}$; no wrong triple using them satisfies the same equation. For each candidate triple the residual is rational in the true parameters, so this witness excludes its being an identity. Removing their finitely many proper algebraic zero sets proves the claim.

The remaining rates can also be recovered generically. Set $v_i = \Sigma_{i1}/\Sigma_{i-1,1}$ and $b_i = (r_i + r_1)v_i$. The node-2 covariance equation $(r_i + r_2)\Sigma_{i2} = b_i\Sigma_{i-1,2} + b_2\Sigma_{i1}$ gives

$$r_i = \frac{r_1 v_i \Sigma_{i-1,2} + b_2\Sigma_{i1} - r_2\Sigma_{i2}}{\Sigma_{i2} - v_i\Sigma_{i-1,2}} \quad (i \geq 3). \tag{99}$$

For $D_j = (r_1 - r_2)\Sigma_{j2} + b_2\Sigma_{j1}$, the two covariance recursions imply

$$\Sigma_{i2} - v_i\Sigma_{i-1,2} = \frac{b_i D_{i-1}}{(r_i + r_1)(r_i + r_2)}, \qquad D_j = \frac{b_j}{r_j + r_2} D_{j-1} \quad (j \geq 3). \tag{100}$$

The witness has $D_2 = -1/3$, and nonzero edges propagate this nondegeneracy. Hence all recursive denominators are generically nonzero at every fixed chain length. The end-to-end task needs only the weaker source-rate assignment established above; its covariance ratio can nevertheless become small in long chains.

B Statistical proofs

Common star model and identities. Throughout this section, stationary means are zero. A one-root star has independent Brownian drivers and dynamics

$$dR_t = -aR_t\,dt + \sqrt{q_0}\,dW_{0,t}, \qquad dL_{i,t} = (-d_iL_{i,t} + B_iR_t)\,dt + \sqrt{q_i}\,dW_{i,t}. \tag{101}$$

Write $s = \mathrm{Var}(R)$, $v_i = \mathrm{Var}(L_i)$ and $x_i = \mathrm{Cov}(R, L_i)$. The Lyapunov equations determine $q_0 = 2as$, $B_i = (a + d_i)x_i/s$, $q_i = 2d_iv_i - 2B_ix_i$, and

$$z_{ij} := \mathrm{Cov}(L_i, L_j) = \frac{x_ix_j}{s}\left(1 + \frac{2a}{d_i + d_j}\right), \qquad i \neq j. \tag{102}$$

Indeed, $(d_i + d_j)z_{ij} = B_ix_j + B_jx_i$. Using ϕ_h from (6), define

$$\begin{aligned} F_h(a, d) &= e^{-dh} + (a + d)\phi_h(a, d) = e^{-ah} + 2a\phi_h(a, d), \\ G_h(a, d) &= (a + d)\phi_h(a, d) = F_h(a, d) - e^{-dh}, \\ \mathrm{Cov}(L_{i,t+h}, R_t) &= x_iF_h(a, d_i), \qquad \mathrm{Cov}(R_{t+h}, R_t) = se^{-ah}. \end{aligned} \tag{103}$$

The equality defining F_h follows from $(d - a)\phi_h = e^{-ah} - e^{-dh}$, including $a = d$ by continuity. The horizon-one response is

$$g_\theta = g_{0\to1}(1) = \frac{x_1}{s}G_1(a, d_1). \tag{104}$$

On a compact positive rate box $a, d \in [a_-, a_+]$,

$$-\partial_dF_h(a, d) = 2a\int_0^h (h - u)e^{-au-d(h-u)}\,du \geq a_-h^2e^{-a_+h} > 0. \tag{105}$$

All derivatives used below are bounded on the relevant compact boxes. Unless noise or a trajectory is explicitly specified, records are independent stationary replicates with noiseless endpoints, and different batches are mutually independent. Endpoints within one record need not be independent. Risk is expected absolute response error divided by ϵ; on the positive-amplitude classes below, $c\epsilon \leq g_\theta \leq C\epsilon$, so this is equivalent to relative absolute risk up to class constants. We use $0^{-1/2} = +\infty$.

Moment errors and Gaussian comparison. For centered jointly Gaussian U, V, Isserlis' identity gives $\mathrm{Var}(UV) = \mathrm{Var}(U)\,\mathrm{Var}(V) + \mathrm{Cov}(U,V)^2$. Bounded observation covariances therefore give expected empirical second-moment error $C/\sqrt{n}$ from n iid records. Projections onto known intervals are nonexpansive. Bounded clipped responses supply the truncation at one in subsequent risk bounds. For $E = C_1 - C_0$, set $U = C_0^{-1/2} E C_0^{-1/2}$. If $C_0 \succeq \lambda_0 I$ and $\|U\|_{\mathrm{op}} \le 1/2$, then

$$\begin{aligned} D_{\mathrm{KL}}\{N(0,C_1)\|N(0,C_0)\} &= \tfrac{1}{2}\sum_j \{\lambda_j(U) - \log(1+\lambda_j(U))\} \\ &\le \frac{\|E\|_F^2}{2\lambda_0^2}. \end{aligned} \tag{106}$$

This uses $u - \log(1+u) \le u^2$ for $|u| \le 1/2$; the bound is $2\|E\|_F^2$ when $\lambda_0 = 1/2$. Independent observations add KL. For two observation laws P_0, P_1 and response gap Δ_g,

$$\inf_{\widehat{g}} \max_{j=0,1} \mathbb{E}_j|\widehat{g} - g_j| \ge \frac{\Delta_g}{4}\{1 - \mathrm{TV}(P_0, P_1)\}. \tag{107}$$

To prove this, classify by the nearer response: a mistake costs at least $\Delta_g/2$, and optimal equal-prior testing error is $(1-\mathrm{TV})/2$. Pinsker gives $\mathrm{TV} \le \sqrt{D_{\mathrm{KL}}/2}$. Thus a path with normalized response gap at least bt and joint KL at most CIt^2, on a fixed interval $0 \le t \le t_0$, yields risk at least $c(1 \wedge I^{-1/2})$ by taking $t = \min\{t_0, (8CI)^{-1/2}\}$.

B.1 Weak covariance witnesses

Take three leaves, $a = s = v_i = 1$, $x = (t,t,\eta)$ with $0 < t, \eta \le .1$, and compare rates $(2,3,4)$ with $(3,2,4)$. Both models share the graph and exact anonymous spectrum. In this case

$$B_i = (1+d_i)x_i, \quad q_0 = 2, \quad q_i = 2d_i - 2(1+d_i)x_i^2, \quad \Sigma_{ij} = x_i x_j \left(1 + \frac{2}{d_i + d_j}\right) \ (i \ne j). \tag{108}$$

Their first-leaf response gap is

$$\delta = 3t(e^{-1} - e^{-2}) - 2t(e^{-1} - e^{-3}) = te^{-3}(e-1)(e-2) > 0. \tag{109}$$

Theorem 20 (Weak-witness sample barrier)**.** *For these two models, every response estimator using N stationary snapshots, the graph and exact spectrum satisfies*

$$\max_{M \in \{P, P'\}} \mathbb{E}_M|\widehat{g} - g_M| \ge \frac{\delta}{4}\left(1 - \frac{\sqrt{N}\,t\eta}{14.7}\right)_+. \tag{110}$$

Selecting the more responsive of leaves 1 and 2 after a common root pulse has worst-case expected regret at least twice this bound. For the general known binary construction below, uniformly conditioned covariances make the reciprocal squared witness energy the testing scale.

Proof. Equation (108) gives positive diffusion. Covariance diagonals are one; absolute off-diagonal row sums are at most .3 for the root and less than .13 for leaves. Gershgorin gives minimum eigenvalues at least .7. Only entries $(1,3), (3,1)$ differ by $t\eta/21$, and $(2,3), (3,2)$ by its negative, so

$$\|\Sigma - \Sigma'\|_F^2 = 4t^2\eta^2/21^2. \tag{111}$$

The whitened difference has norm at most $2t\eta/(21 \cdot .7) < 1/2$. Consequently

$$D_{\mathrm{KL}}(P\|P') \le \frac{2t^2\eta^2}{21^2(.7)^2}, \tag{112}$$

and

$$\mathrm{TV}(P^N, P'^N) \le \frac{\sqrt{N}\,t\eta}{21(.7)}. \tag{113}$$

Apply (107). Swapping rates exchanges the two preferred leaves; choosing the wrong leaf incurs regret exactly δ. The same testing bound therefore proves the regret claim. □

Several witnesses and an attaining binary test. Keep a root of rate $a > 0$ and variance one, two leaves p, q with $x_p = x_q = t$ and distinct rates, and other root–leaf covariances x_j. Swap d_p, d_q and adjust edges and diffusion by the Lyapunov identities, assuming both resulting models are admissible. Define

$$\mu_j = 2atx_j \left(\frac{1}{d_p + d_j} - \frac{1}{d_q + d_j} \right), \qquad j \notin \{p, q\}. \tag{114}$$

Only the p, j and q, j covariances and transposes change, by μ_j and $-\mu_j$, respectively; hence $\|\Sigma - \Sigma'\|_F^2 = 4\|\mu\|_2^2$. If both minimum eigenvalues are at least λ_0 and $2\|\mu\|_2/\lambda_0 \le 1/2$,

$$\mathrm{TV}(P^N, P'^N) \le \sqrt{N}\, \|\mu\|_2/\lambda_0. \tag{115}$$

The target gap at horizon H is $t|G_H(a, d_p) - G_H(a, d_q)|$; response and ranking bounds follow as above. For a matching upper bound, also assume both covariance operator norms are at most L. With $Y = X_p - X_q$ and Z the vector of witness leaves, the means of YZ are μ and $-\mu$, and

$$\mathrm{Cov}(YZ) = \mathrm{Var}(Y)\Sigma_{ZZ} + \mu\mu^\top, \qquad \|\,\mathrm{Cov}(YZ)\|_{\mathrm{op}} \le 2L^2 + \|\mu\|_2^2. \tag{116}$$

The sign of $\mu^\top N^{-1} \sum_j Y^{(j)} Z^{(j)}$ has error at most $(2L^2 + \|\mu\|_2^2)/(N\|\mu\|_2^2)$ by Chebyshev. Thus $N \asymp \|\mu\|_2^{-2}$ gives fixed nontrivial binary precision under uniform conditioning. The test knows both candidates; without an upper covariance bound, increasing witness variance can invalidate this sample scale.

Local testing limit. For fixed t and $\eta_N = c/\sqrt{N}$, both covariances approach Σ_*. Their difference is $\eta_N E_0$, with E_0 equal to $t/21$ on $(1, 3), (3, 1)$ and $-t/21$ on $(2, 3), (3, 2)$. The vector $(0, 1, -1)$ in the first three coordinates of Σ_* has eigenvalue $1 - 7t^2/5$, while leaf 3 is independent with variance one. It follows that

$$J(t) = \tfrac{1}{2}\,\mathrm{tr}[(\Sigma_*^{-1}E_0)^2] = \frac{2t^2}{21^2(1 - 7t^2/5)}. \tag{117}$$

Expanding the Gaussian log-likelihood ratio to second order gives limiting means $\mp c^2 J(t)/2$ and variance $c^2 J(t)$ under the two models: the first-order centered quadratic forms obey the CLT, the second-order term has vanishing variance, and the third-order remainder is $O(N\eta_N^3) = o(1)$ in expectation, by uniform Gaussian moments. Thresholding at zero gives equal-prior optimal error

$$\Phi\Big(-c\sqrt{J(t)}/2\Big). \tag{118}$$

The response gap (109) stays fixed although every $\eta_N > 0$ is on the identifiable side of Theorem 19.

B.2 Static versus oriented-pair testing

Set $t = \eta = \epsilon$ in the preceding pair. An oriented observation records only (R_{t_0}, L_{1,t_0+h}); its endpoints may have independent centered Gaussian noise of known variance $\omega \ge 0$.

Theorem 21 (Quartic-to-quadratic testing separation)**.** *For fixed $h > 0$, these known alternatives share their graph, exact spectrum and entire root-only temporal law. Their response gap is $e^{-3}(e-1)(e-2)\epsilon$, and their per-record divergences satisfy*

$$D_{\mathrm{KL}}^{\mathrm{static}} = \epsilon^4/21^2 + o(\epsilon^4), \qquad D_{\mathrm{KL}}^{\mathrm{oriented}} = \frac{\kappa_h^2 \epsilon^2}{2(1+\omega)^2} + o(\epsilon^2), \quad \kappa_h = e^{-h}(1 - e^{-h})^2. \tag{119}$$

For fixed classification error in $(0, 1/2)$, static-only and oriented-only sample orders are respectively ϵ^{-4} and ϵ^{-2}; independent batches have information scale $N\epsilon^4 + M\epsilon^2$.

Proof. The models are admissible for $\epsilon \le .1$, their covariances converge to I_4, and

$$\|\Sigma_\epsilon - \Sigma'_\epsilon\|_F^2 = 4\epsilon^4/21^2. \tag{120}$$

The Gaussian KL expansion therefore gives, for the whitened difference U_ϵ,

$$D_{\mathrm{KL}}(\Sigma_\epsilon \| \Sigma'_\epsilon) = \tfrac{1}{4}\|U_\epsilon\|_F^2 + o(\epsilon^4) = \epsilon^4/21^2 + o(\epsilon^4). \tag{121}$$

For oriented pairs the candidate covariances are

$$V_d = \begin{pmatrix} v & \epsilon F_h(1,d) \\ \epsilon F_h(1,d) & v \end{pmatrix}, \qquad v = 1+\omega, \quad d \in \{2,3\}. \tag{122}$$

Since $F_h(1,2) = 3e^{-h} - 2e^{-2h}$ and $F_h(1,3) = 2e^{-h} - e^{-3h}$, their difference is $\kappa_h > 0$ and

$$D_{\mathrm{KL}}(V_2 \| V_3) = \epsilon^2 \kappa_h^2/(2v^2) + o(\epsilon^2). \tag{123}$$

Both reverse divergences have the same leading terms. Pinsker and additivity prove necessity; common spectra and root-only observations add no divergence. For sufficiency, $(X_1 - X_2)X_3$ has candidate means $\pm\epsilon^2/21$ and variance $2-14\epsilon^2/5+\epsilon^4/21^2$. Its sample-average sign has error $O((N\epsilon^4)^{-1})$. The oriented cross-product has means separated by $\epsilon\kappa_h$ and variance $v^2+O(\epsilon^2)$, so midpoint thresholding gives $O((M\epsilon^2)^{-1})$ error. These are tests of two known models. A full bivariate pair additionally contains a target self-lag whose difference tends to $e^{-2h} - e^{-3h} > 0$; it distinguishes these alternatives with order-one pairs. □

B.3 Assigning supplied rates with unknown amplitudes

Consider the three-leaf class with $a \in [.9, 1.1]$, one leaf rate in each of $[1.8, 2.2]$, $[2.8, 3.2]$, $[3.8, 4.2]$ in unknown order, $s, v_i \in [.9, 1.1]$, and $x_i/\epsilon \in [.8, 1.2]$. The Lyapunov identities above give positive diffusion for uniformly small ϵ. The graph, ϵ and exact unordered spectrum are supplied; the separated root band makes a its smallest member. Let $R_\epsilon^{\mathcal{R}}(N, M)$ be normalized minimax response risk from N static four-vectors and M oriented pairs at lag one.

Proof of Theorem 4. For the static upper bound, define $R_{ij} = sz_{ij}/(x_i x_j)$ and $w_{ij} = (R_{ij} - 1)^{-1}$. For distinct i, j, k,

$$\lambda_i = w_{ij} + w_{ik} - w_{jk} = d_i/a, \qquad a = \frac{T}{1 + \sum_i \lambda_i}, \qquad d_i = a\lambda_i. \tag{124}$$

Only the supplied spectrum's sum T is used. Replace moments by their sample second moments and project every intermediate quantity to its known compact range, as in (131) with this class's intervals. Since $x_i \asymp \epsilon$, $z_{ij} = O(\epsilon^2)$ and $R_{ij} - 1$ is bounded away from zero, subtraction and Lipschitzness give

$$\frac{|\widehat{g}_{\mathrm{stat}} - g|}{\epsilon} \le C\epsilon^{-2} \|\widehat{\Sigma} - \Sigma\|_{\max}. \tag{125}$$

The moment bound yields risk $C/(\epsilon^2\sqrt{N})$.

For the temporal upper bound, the remaining three supplied rates are leaf candidates. By (105), compactness and separated bands give a uniform gap between their $F_1(a, d)$ values. Estimate s, x_1 from snapshots and $y = x_1 F_1(a, d_1)$ from oriented pairs, select the candidate minimizing $|\widehat{y} - \widehat{x}_1 F_1(a, d)|$, and substitute it into the response, clipping s, x_1 to their class intervals. A wrong assignment requires $|\widehat{y} - y| + C|\widehat{x}_1 - x_1| \ge c\epsilon$. Chebyshev, the bounded response range, and the amplitude moment error give

$$\sup \frac{\mathbb{E}|\widehat{g}_{\mathrm{time}} - g|}{\epsilon} \le C\left\{\frac{1}{\epsilon\sqrt{N}} + \frac{1}{N\epsilon^2} + \frac{1}{M\epsilon^2}\right\}. \tag{126}$$

These two bounds prove the stated fixed-precision upper corners.

For the first lower bound use the preceding rate swap with $x_i = \epsilon$. It has common spectrum, gap $c\epsilon$ and joint KL $C(N\epsilon^4 + M\epsilon^2)$, so (107) applies. For the second keep that rate swap and $s = v_i = 1$, $x_2 = x_3 = \epsilon$, but use $x_1 = .8\epsilon$ at $d_1 = 2$ and $x_1' = .8\rho\epsilon$ at $d_1' = 3$, where, with $q = e^{-1}$,

$$\rho = \frac{F_1(1,2)}{F_1(1,3)} = \frac{3-2q}{2-q^2} \in (1, 1.25). \tag{127}$$

Both amplitudes are allowed. The oriented marginal variances and cross lag agree exactly, hence its entire Gaussian law is identical for every M. Static covariance differences are $O(\epsilon)$ in the root–target entry and $O(\epsilon^2)$ in leaf pairs, giving KL $CN\epsilon^2$. Using $G_1(1,2) = 3q(1-q)$ and $G_1(1,3) = 2q(1-q^2)$, the gap is

$$|g - g'| = .8\epsilon \frac{q^2(1-q)(2-q)}{2-q^2} > 0. \tag{128}$$

Equation (107) proves the second necessary condition. This argument uses the amplitude interval's width; the upper bound uses the root's separated spectral band. Neither assumption follows from anonymous spectral information alone. □

B.4 Minimax risk with an exact rate sum

For the remaining star results, unless stated otherwise, use three leaves and the compact class $\Theta_\epsilon^{\mathrm{tf}}$:

$$\begin{aligned} &\tfrac{1}{2} \le a, d_i \le \tfrac{13}{2}, \quad \tfrac{1}{2} \le s, v_i \le 2, \quad x_i = \epsilon u_i, \quad 1 \le u_i \le 2, \\ &a + \sum_i d_i = T = 10, \quad q_0 = 2as, \quad B_i = (a + d_i)x_i/s, \quad q_i = 2d_i v_i - 2B_i x_i > 0. \end{aligned} \tag{129}$$

For $\epsilon \le 1/100$, $q_i \ge 1/2 - 208\epsilon^2 > 0$ uniformly. The graph, ϵ and exact trace are known, not individual rates or the full spectrum. Observe $N \ge 1$ static four-vectors, $M \ge 0$ oriented pairs $(R_t, L_{1,t+1})$, and $M_0 \ge 0$ root pairs (R_t, R_{t+1}). They are independent batches of fresh stationary replicates, and

$$R_\epsilon(N, M, M_0) = \inf_{\widehat{g}} \sup_{\theta \in \Theta_\epsilon^{\mathrm{tf}}} \frac{\mathbb{E}_\theta |\widehat{g} - g_\theta|}{\epsilon}. \tag{130}$$

Proof of Theorem 5. Static upper bound. Let $\widehat{S}$ be the snapshot second-moment matrix. Set $\widehat{s} = \Pi_{[1/2,2]}\widehat{S}_{00}$, $\widehat{x}_i = \Pi_{[\epsilon,2\epsilon]}\widehat{S}_{0i}$ and $\widehat{z}_{ij} = \widehat{S}_{ij}$. For $I_R = [1 + 1/13, 14]$, $I_\lambda = [1/13, 13]$ and $I_a = [1/2, 13/2]$, define

$$\begin{aligned} &\widehat{R}_{ij} = \Pi_{I_R}\frac{\widehat{s}\widehat{z}_{ij}}{\widehat{x}_i\widehat{x}_j}, \quad \widehat{w}_{ij} = (\widehat{R}_{ij} - 1)^{-1}, \qquad \widehat{\lambda}_i = \Pi_{I_\lambda}(\widehat{w}_{ij} + \widehat{w}_{ik} - \widehat{w}_{jk}), \\ &\widehat{a} = \Pi_{I_a}\frac{T}{1 + \sum_i \widehat{\lambda}_i}, \quad \widehat{d}_1 = \Pi_{I_a}(\widehat{a}\widehat{\lambda}_1), \quad \widehat{g}_{\text{stat}} = \frac{\widehat{x}_1}{\widehat{s}} G_1(\widehat{a}, \widehat{d}_1), \qquad \{i, j, k\} = \{1, 2, 3\}. \end{aligned} \tag{131}$$

The true values lie in all projection intervals, and (124) gives the correct population inverse. All denominators are protected on every sample. Since $z_{ij} = O(\epsilon^2)$,

$$|\widehat{R}_{ij} - R_{ij}| \le C\left\{\frac{|\widehat{z}_{ij} - z_{ij}|}{\epsilon^2} + \frac{|\widehat{x}_i - x_i| + |\widehat{x}_j - x_j|}{\epsilon} + |\widehat{s} - s|\right\}.$$

The remaining compact-domain maps are Lipschitz, so the common moment bound and bounded normalized responses give

$$\sup_\theta \frac{\mathbb{E}_\theta |\widehat{g}_{\text{stat}} - g_\theta|}{\epsilon} \le C\{1 \wedge (N\epsilon^4)^{-1/2}\}. \tag{132}$$

Targeted upper bound. For $N, M, M_0 \ge 1$, let $\widehat{s}_W$ and $\widehat{c}_W$ be the root-pair first-endpoint second moment and lag, and $\widehat{y}$ the oriented cross-product mean. Use static $\widehat{s}, \widehat{x}_1$ above and put

$$\begin{aligned} &\widehat{a}_W = -\log \Pi_{[e^{-13/2}, e^{-1/2}]}\frac{\widehat{c}_W}{\Pi_{[1/2,2]}\widehat{s}_W}, \qquad \widehat{f} = \widehat{y}/\widehat{x}_1, \\ &\widehat{d}_1 = \arg\min_{d \in I_a} |F_1(\widehat{a}_W, d) - \widehat{f}|, \qquad \widehat{g}_{\text{targ}} = (\widehat{x}_1/\widehat{s})G_1(\widehat{a}_W, \widehat{d}_1). \end{aligned} \tag{133}$$

The logarithm is Lipschitz on its fixed interval and $\mathbb{E}|\widehat{a}_W - a| \le C/\sqrt{M_0}$. Equation (105), with $\kappa = e^{-13/2}/2$, makes the minimizer unique, including when the observed value is outside the image. Projection onto that image gives

$$|\widehat{d}_1 - d_1| \le \kappa^{-1}\{|\widehat{f} - F_1(a, d_1)| + L_a|\widehat{a}_W - a|\}, \qquad |\widehat{f} - F_1(a, d_1)| \le C\epsilon^{-1}(|\widehat{y} - y| + |\widehat{x}_1 - x_1|),$$

where $L_a = \sup |\partial_a F_1|$. Thus

$$\sup_\theta \frac{\mathbb{E}_\theta |\widehat{g}_{\text{targ}} - g_\theta|}{\epsilon} \le C\left[1 \wedge \left\{(N\epsilon^2)^{-1/2} + (M\epsilon^2)^{-1/2} + M_0^{-1/2}\right\}\right]. \tag{134}$$

No supplied rate is used.

Three lower paths. Set $s = v_i = 1$ and $x_i = \epsilon$ except where stated, and write $\zeta = e^{-1}$. All paths below lie in the same trace-fixed class on a sufficiently short fixed interval $[0, t_0]$. Static and oriented covariance baselines tend to identity, and the root-pair minimum eigenvalue at the base is $1 - e^{-1} > 1/2$. Shrinking t_0 and ϵ makes (106) applicable uniformly to every path.

Path A keeps $a = 1$ and changes

$$(d_1, d_2, d_3) = (2 + t, 3 - t, 4). \tag{135}$$

The root-pair law and z_{12} are fixed; the changing static entries are $z_{13} = \epsilon^2\{1 + 2/(6 + t)\}$ and $z_{23} = \epsilon^2\{1 + 2/(7 - t)\}$ and their transposes. The oriented covariance is $\epsilon F_1(1, 2 + t)$. Their bounded derivatives yield per-record bounds

$$D_{\mathrm{KL}}^{\mathrm{stat}}(t\|0) \le C_A\epsilon^4 t^2, \qquad D_{\mathrm{KL}}^{\mathrm{orient}}(t\|0) \le C_A\epsilon^2 t^2, \qquad D_{\mathrm{KL}}^{\mathrm{root}}(t\|0) = 0. \tag{136}$$

Since $\partial_d G_1(1, 2) = \zeta(5\zeta - 2) \neq 0$, its normalized response gap is at least $b_A t$. The common two-point argument with $I = N\epsilon^4 + M\epsilon^2$ gives the target-rate lower term.

Path B uses the same rates but

$$x_1(t) = \epsilon r(t), \qquad r(t) = \frac{F_1(1, 2)}{F_1(1, 2 + t)}, \qquad x_2 = x_3 = \epsilon. \tag{137}$$

Monotonicity gives $1 \le r(t) \le 2$ on a short interval. Both oriented variances and $x_1(t)F_1(1, 2 + t)$ remain fixed, as does the root-pair law. The changing root–target entry is $O(\epsilon t)$; leaf-pair changes are $O(\epsilon^2 t)$ because r, r' and all rate denominators are uniformly bounded. Thus

$$D_{\mathrm{KL}}^{\mathrm{stat}}(t\|0) \le C_B\epsilon^2 t^2, \qquad D_{\mathrm{KL}}^{\mathrm{orient}} = D_{\mathrm{KL}}^{\mathrm{root}} = 0. \tag{138}$$

Direct differentiation gives $r'(0) = 2(1 - 2\zeta)/(3 - 2\zeta)$ and $g'(0)/\epsilon = \zeta^2(1 + 2\zeta)/(3 - 2\zeta) > 0$. Its gap is therefore at least $b_B\epsilon t$, and $I = N\epsilon^2$ proves the amplitude lower term regardless of temporal sample counts.

Path C sets $a(t) = 1 + t$ and uses the implicit-function theorem and (105) to define

$$F_1(1 + t, d_1(t)) = F_1(1, 2), \quad d_1(0) = 2, \quad d_2(t) = 3, \quad d_3(t) = 7 - a(t) - d_1(t). \tag{139}$$

All static variances and root–leaf covariances stay fixed, and the entire oriented-pair law is unchanged. At the base,

$$\partial_a F_1 = \zeta(1 - 4\zeta), \quad \partial_d F_1 = -2\zeta(1 - 2\zeta), \quad d_1'(0) = \frac{1 - 4\zeta}{2(1 - 2\zeta)} \neq 0. \tag{140}$$

As $G_1 = F_1 - e^{-d}$, $g'(0)/\epsilon = \zeta^2 d_1'(0) \neq 0$. The branch has bounded derivative (for example $\sup|d_1'| \le 15e^{13/2}$), so each leaf-pair covariance derivative is $O(\epsilon^2)$ by (102); the root-pair off-diagonal entry changes from ζ to ζe^{-t}. Consequently

$$D_{\mathrm{KL}}^{\mathrm{stat}}(t\|0) \le C_C\epsilon^4 t^2, \qquad D_{\mathrm{KL}}^{\mathrm{orient}}(t\|0) = 0, \qquad D_{\mathrm{KL}}^{\mathrm{root}}(t\|0) \le 4\zeta^2 t^2. \tag{141}$$

Its gap is at least $b_C\epsilon t$, giving the third lower term with $I = N\epsilon^4 + M_0$. Taking the maximum of the three lower bounds is valid since each path lies in the same parameter class.

Matching all budgets. Put $A = N\epsilon^4$, $B = M\epsilon^2$ and $C_0 = M_0$. If $B \le A$ or $C_0 \le A$, the static bound is within a constant of the corresponding $(A + B)^{-1/2}$ or $(A + C_0)^{-1/2}$ lower factor. Otherwise $B, C_0 > A$, and the targeted bound matches the maximum because $B^{-1/2} \le \sqrt{2}(A + B)^{-1/2}$ and similarly for C_0. These comparisons survive truncation at one and include zero temporal budgets through the static branch. □

For fixed small normalized error the three information quantities in (23) must, and up to target-dependent constants need only, exceed fixed thresholds. Thus static-only sampling costs ϵ^{-4}, while the targeted corner costs $(N, M, M_0) \asymp (\epsilon^{-2}, \epsilon^{-2}, 1)$. Fixed M_0 ensures only fixed precision when $N\epsilon^4 = o(1)$. Exact T is an external scale observation: replacing (A, Q) by (cA, cQ) preserves covariance, but at the base the response becomes $3\epsilon(e^{-c} - e^{-2c})$, with nonzero derivative at $c = 1$. A plug-in estimate $\widehat{T}$ adds $C\mathbb{E}|\widehat{T} - T|$ to the normalized static upper bound; its acquisition cost belongs to the enlarged experiment below.

B.5 A noisy rate sum

Replace $T = 10$ in (129) by $T \in [9, 11]$, keeping all other bounds. The estimator receives $N \geq 1$ snapshots, $J \geq 0$ external readings $Y_j = T + N(0, \sigma_T^2)$, and $M_0 \geq 0$ root pairs, in independent batches; $0 < \sigma_T^2 < \infty$ is fixed and known. There are no oriented pairs, and T is unknown.

Theorem 22 (Risk with a noisy rate sum)**.** *The normalized minimax risk on this enlarged class obeys, uniformly over sample counts and sufficiently small ϵ,*

$$R_\epsilon^{\text{anchor}}(N, J, M_0) \asymp 1 \wedge \max\{(N\epsilon^4)^{-1/2}, (J/\sigma_T^2 + M_0)^{-1/2}\}. \tag{142}$$

Constants may depend on the fixed box and σ_T^2.

Proof. Path A keeps both the trace-reading and root-pair laws fixed, so its KL is $CN\epsilon^4 t^2$ and its response gap is $c\epsilon t$. For the other lower term, scale the base drift and diffusion by $1 + t$. The whole static law stays fixed, $T_t = 10(1 + t) \in [9, 11]$ for small $t \geq 0$, and one external reading has divergence

$$D_{\text{KL}}\{N(T_t, \sigma_T^2) \| N(10, \sigma_T^2)\} = 50t^2/\sigma_T^2. \tag{143}$$

The root-pair KL is $O(t^2)$, while the response derivative at zero is $3\epsilon(-e^{-1} + 2e^{-2}) \neq 0$. The common two-point argument gives the two claimed terms, including $J = M_0 = 0$ by a fixed scale change.

For the upper bound, the ratio stage of (131) does not need T and gives

$$\max_i \mathbb{E}|\widehat{\lambda}_i - d_i/a| \leq C(N\epsilon^4)^{-1/2}. \tag{144}$$

If $J \geq M_0$ and $J > 0$, set $\widehat{T} = \Pi_{[9,11]}\overline{Y}$ and $\widehat{a} = \Pi_{I_a}\{\widehat{T}/(1 + \sum_i \widehat{\lambda}_i)\}$; its error is at most $C\{(N\epsilon^4)^{-1/2} + \sigma_T/\sqrt{J}\}$. If $M_0 > J$, use the root-pair inverse in (133), with error $C/\sqrt{M_0}$. Then set $\widehat{d}_1 = \Pi_{I_a}(\widehat{a}\widehat{\lambda}_1)$ and substitute into $(\widehat{x}_1/\widehat{s})G_1(\widehat{a}, \widehat{d}_1)$. Lipschitzness, bounded responses and fixed σ_T^2 prove the upper bound, since $\max\{J, M_0\} \geq (J + M_0)/2$. When both vanish use zero. □

If $\sigma_T^2 = 0$, one reading supplies the exact trace instead; that case is not covered by the positive-noise rate above.

B.6 A fixed number of leaves

For fixed $m \geq 3$, let $\Theta_{\epsilon,m}$ have the same coordinate bounds and Lyapunov parameterization, with m leaves and known $T_m = m + 7$. Observe $N \geq 1$ full $(m + 1)$-vectors, M oriented pairs and M_0 root pairs as before.

Theorem 23 (Fixed-leaf-count extension)**.** *For fixed m, all sample counts and sufficiently small ϵ,*

$$R_{\epsilon,m}(N, M, M_0) \asymp_m 1 \wedge \max\{(N\epsilon^2)^{-1/2}, (N\epsilon^4 + M\epsilon^2)^{-1/2}, (N\epsilon^4 + M_0)^{-1/2}\}. \tag{145}$$

Proof. For each leaf, any two other leaves give $d_i/a = w_{ij} + w_{ik} - w_{jk}$; use the first three leaves for their ratios and leaves 1 and 2 for each additional leaf. Then $a = T_m/(1 + \sum_i d_i/a)$. The clipped static proof gives $C_m[1 \wedge (N\epsilon^4)^{-1/2}]$; the targeted decoder is unchanged. The same budget split combines them.

Embed paths A–C at $a = 1$, $(d_1, d_2, d_3) = (2, 3, 4)$, $d_4 = \cdots = d_m = 1$, $s = v_i = u_i = 1$, holding the extra rates fixed. Their traces remain $m+7$ and their nonzero response derivatives are unchanged. In A and C every new covariance derivative is $O(\epsilon^2)$; in B the root–target change is $O(\epsilon t)$ and all new leaf-pair changes are $O(\epsilon^2 t)$. With finitely many entries, joint KL is bounded respectively by $C_m(N\epsilon^4 + M\epsilon^2)t^2$, $C_m N\epsilon^2 t^2$, and $C_m(N\epsilon^4 + M_0)t^2$. Positive covariance and diffusion hold on fixed short paths for small ϵ. The three two-point bounds finish the proof. Constants and the allowed coupling range may depend on m; this is not a growing-dimension assertion. □

B.7 Full bivariate time pairs

On $\Theta_\epsilon^{\mathrm{tf}}$, replace the restricted temporal batches by $K \geq 0$ independent records $(R_t, L_{1,t}, R_{t+1}, L_{1,t+1})$, independent also of $N \geq 1$ snapshots. Let $R_\epsilon^{\mathrm{bi}}(N, K)$ denote normalized minimax response risk.

Proof of Theorem 6. Write $x = x_1, v = v_1, d = d_1$. Besides s, v, x and the oriented lag, each full pair supplies

$$C_R = se^{-a}, \qquad C_L = ve^{-d} + (x^2/s)G_1(a, d). \tag{146}$$

The leaf transition $L_{t+1} = e^{-d}L_t + B_1\phi_1(a, d)R_t +$ independent innovation proves the second identity. Uniformly on the box, $\partial_d C_L = -ve^{-d} + O(\epsilon^2)$ is bounded above by a negative constant for sufficiently small ϵ.

For the first lower term keep the base rates $(1, 2, 3, 4)$ and variances one, and change x_1 from ϵ to $\epsilon(1 + t)$. The normalized response gap is exactly $tG_1(1, 2) > 0$. Every changed static or full-pair covariance entry is $O(\epsilon t)$; covariance eigenvalues are uniformly positive. Equation (106) gives joint KL $C(N + K)\epsilon^2 t^2$. For the second term use path A: the static KL is $CN\epsilon^4 t^2$, the full-pair covariance changes by $O(t)$ because of C_L, and its KL is CKt^2. The normalized gap remains at least ct. Applying (107) to these two paths yields both lower factors.

For the upper bound, retain the static decoder. When $K \geq 1$, the full-pair empirical moments estimate s, v, x, C_R, C_L with expected error $C/\sqrt{K}$, without requiring independence within a record. Project s, v, x to their known intervals and invert the clipped root lag to obtain $\widehat{a}$. Recover d from

$$\widehat{C}_L = \widehat{v}e^{-\widehat{d}} + (\widehat{x}^2/\widehat{s})G_1(\widehat{a}, \widehat{d}), \qquad \widehat{d} \in I_a, \tag{147}$$

projecting $\widehat{C}_L$ to the fitted endpoint range first. The uniform derivative margin applies even to projected parameters; this inverse and the root inverse have error $C/\sqrt{K}$ on average. Substituting into the response gives full-pair-only risk $C[1 \wedge (K\epsilon^2)^{-1/2}]$.

A hybrid uses static $\widehat{x}, \widehat{s}$ and full-pair $\widehat{a}, \widehat{v}, \widehat{C}_L$. Its normalized risk is $C[1 \wedge \{(N\epsilon^2)^{-1/2} + K^{-1/2}\}]$: static error in the $O(\epsilon^2)$ correction contributes at most $C\epsilon N^{-1/2}$ to rate error and is absorbed. Choose the static decoder if $K \leq N\epsilon^4$, the hybrid if $N\epsilon^4 < K \leq N$, and the full-pair-only decoder if $K > N$. In each regime this bound is within a constant of (25). □

Scalar-endpoint costs are $4N + 2M + 2M_0$ for restricted batches and $4N + 4K$ for full pairs. Both attain fixed normalized precision at order ϵ^{-2} total readouts; once amplitude is learned, full pairs need only order-one temporal records for rates. These are information counts, not physical preparation or synchronization costs.

B.8 One trajectory with unknown sensor variances

On $\Theta_\epsilon^{\mathrm{tf}}$ observe a stationary trajectory at integer times $0, \ldots, L$, $L \geq 3$, with

$$Y_t = X_t + \xi_t, \qquad \xi_t \overset{\mathrm{iid}}{\sim} N(0, \operatorname{diag}\omega), \qquad \omega \in [0, \omega_{\max}]^4. \tag{148}$$

Errors are independent of the process and across coordinates and times. The finite bound $\omega_{\max}$ and ϵ are known, but individual sensor variances are not. No separate static or pair batch is supplied. For $\mathcal{Y}_4 = (Y_t)_{t=0}^L$ and $\mathcal{Y}_2 = (Y_{t,R}, Y_{t,L_1})_{t=0}^L$, define

$$R_k = R_\epsilon^{\mathrm{traj,k}}(L) := \inf_{\widehat{g}_k} \sup_{\theta \in \Theta_\epsilon^{\mathrm{tf}},\, \omega \in [0,\omega_{\max}]^4} \frac{\mathbb{E}_{\theta,\omega}|\widehat{g}_k(\mathcal{Y}_k) - g_\theta|}{\epsilon}, \qquad k \in \{2, 4\}. \tag{149}$$

Both infima allow the same known class and trace; unused leaf-noise parameters do not affect the two-coordinate marginal. Write $R_\epsilon^{\mathrm{traj}}(L) = R_4$ and $r_L = 1 \wedge (L\epsilon^2)^{-1/2}$.

Proof of Theorem 7. We establish, uniformly for $L \geq 3$ and sufficiently small ϵ,

$$cr_L \leq R_4 \leq R_2 \leq Cr_L. \tag{150}$$

Coordinate projection gives $R_4 \le R_2$: composing a two-coordinate estimator with projection preserves its risk for each (θ, ω). It remains to prove the two-coordinate upper and four-coordinate lower bounds on the common nuisance class.

Population inverse. Put $x = x_1, v = v_1, d = d_1$. Independence of sensor errors removes them from positive-lag covariances and from the same-time cross-coordinate moment. For $h = 1, 2$,

$$\gamma_R(h) = se^{-ah}, \qquad \gamma_L(h) = ve^{-dh} + (x^2/s)G_h(a,d), \qquad \mathbb{E}(Y_{0,R}Y_{0,L_1}) = x. \tag{151}$$

Consequently

$$a = \log \frac{\gamma_R(1)}{\gamma_R(2)}, \qquad s = \frac{\gamma_R(1)^2}{\gamma_R(2)}. \tag{152}$$

The denominator is at least $e^{-13}/2$. Define $F_{a,s,x}(d,v) = (ve^{-d} + x^2 G_1(a,d)/s, ve^{-2d} + x^2 G_2(a,d)/s)$. At $x = 0$ its inverse is $d = \log(C_1/C_2)$, $v = C_1^2/C_2$, Lipschitz on a positive rectangle containing the compact image. Hence for $z = (d,v)$, z' in the box, $\|F_{a,s,0}(z) - F_{a,s,0}(z')\| \ge c_0\|z - z'\|$. Bounded derivatives and $x^2/s \le 8\epsilon^2$ imply

$$\|F_{a,s,x}(z) - F_{a,s,x}(z')\| \ge (c_0 - C_0\epsilon^2)\|z - z'\| \ge (c_0/2)\|z - z'\|. \tag{153}$$

Choose $\epsilon_0 \le \min\{1/100, \sqrt{c_0/(2C_0)}\}$. The inverse bound is global on the compact box and uniform in projected a, s, x, not merely a pointwise Jacobian argument.

Dependent empirical moments. The available statistics are

$$\widehat{\gamma}_U(h) = \frac{1}{L+1-h}\sum_{t=0}^{L-h} Y_{t,U}Y_{t+h,U}, \quad U \in \{R, L_1\},\ h = 1,2, \qquad \widehat{x} = \frac{1}{L+1}\sum_{t=0}^{L} Y_{t,R}Y_{t,L_1}.$$

They need not be mutually independent. The star transition obeys $\|e^{Ak}\| \le C(1+k)e^{-k/2} \le C'e^{-k/4}$ uniformly, and $\mathrm{Cov}(X_0, X_k) = \Sigma e^{A^\top k}$. Sensor noise affects only lag zero. Isserlis therefore bounds covariance between quadratic summands separated by $k > 2$ by $Ce^{-k/4}$; the finitely many overlapping cases have bounded covariance. Summation gives variance C/L and expected absolute error $C/\sqrt{L}$ for all five moments, uniformly also over the unobserved leaves' nuisance parameters.

Clipped estimator. Project the root lags to $I_h = [e^{-13h/2}/2, 2e^{-h/2}]$, then apply (152), projecting the outputs to their rate and variance intervals. Project $\widehat{x}$ to $[\epsilon, 2\epsilon]$. The inverses are Lipschitz on $I_1 \times I_2$, even for negative raw lags before projection, so these three estimates have error $C/\sqrt{L}$. Choose a measurable minimum-distance fit $(\widehat{d}, \widehat{v})$ of the target-lag vector to $F_{\widehat{a},\widehat{s},\widehat{x}}(d,v)$ over the compact box. Comparison with the feasible true (d,v) and (153) bounds its error by a constant times the empirical target-lag error plus $\epsilon^2(|\widehat{a} - a| + |\widehat{s} - s|) + \epsilon|\widehat{x} - x|$. It too has expected error $C/\sqrt{L}$. Substitution gives

$$\frac{\mathbb{E}|\widehat{g} - g|}{\epsilon} \le C\left\{\epsilon^{-1}\mathbb{E}|\widehat{x} - x| + \mathbb{E}|\widehat{a} - a| + \mathbb{E}|\widehat{d} - d| + \mathbb{E}|\widehat{s} - s|\right\} \le \frac{C'}{\epsilon\sqrt{L}}. \tag{154}$$

Use this estimator when $L\epsilon^2 \ge 1$, and zero otherwise, proving $R_2 \le Cr_L$ with no sensor-variance estimates or rate oracle.

Full-record lower bound. Fix rates $(1,2,3,4)$, variances one, $x_i = 3\epsilon/2$, and vary only $x_1 = \epsilon(3/2 + t)$, $|t| \le 1/4$. The trace remains ten and diffusion remains positive. The normalized response gap is exactly $|t|G_1(1,2)$. Only $B_1 = 3x_1$ and $q_1 = 4 - 6x_1^2$ vary in the drift and diffusion. The noiseless sampled process is a Gaussian Markov chain with transition $E = e^A$ and innovation covariance

$$V = \Sigma - E\Sigma E^\top = \int_0^1 e^{Au}Qe^{A^\top u}\,du, \qquad \Sigma = \int_0^\infty e^{Au}Qe^{A^\top u}\,du.$$

Compactness, uniform stability and $Q \succeq (1/2 - 208\epsilon_0^2)I$ bound the eigenvalues of both V, Σ above and away from zero. Smoothness gives changes $O(\epsilon|t|)$ in E, V, Σ. Initial Gaussian KL is $C\epsilon^2t^2$; conditional transition KL is at most $C\epsilon^2t^2(1 + \|X_j\|^2)$, including the changing mean EX_j. The chain rule and stationarity bound the full latent-record KL by $C(L+1)\epsilon^2t^2$. For any common admissible ω, adding sensor noise is a parameter-independent Markov kernel, so the same KL bound holds for the noisy four-coordinate record by data processing. With $t = c_1r_L$, $c_1 \le 1/4$, the bound stays small because $(L+1)\epsilon^2r_L^2 \le 4/3$ for $L \ge 3$. Pinsker and (107) give $R_4 \ge cr_L$, including the supremum over unknown noise. This proves both results. □

The result concerns minimax order at the same sampling times; its constants need not coincide between records. Unknown cross-coordinate or temporally correlated sensor errors are not covered.

B.9 Lag sensitivity and endpoint noise

Proposition 24 (Scalar lag variance and known endpoint noise)**.** *For a scalar stationary OU root with known variance $s > 0$, fixed $h > 0$ and $0 < a_{\min} < a < a_{\max} < \infty$, observe K iid root pairs. Let $\widehat{C}_h = K^{-1}\sum_j R_{j,0}R_{j,h}$, $\ell = e^{-a_{\max}h}$, $u = e^{-a_{\min}h}$, and $\widehat{a}_h = f(\widehat{C}_h/s)$ with $f(z) = -h^{-1}\log \Pi_{[\ell,u]}z$. Then, with $V_h = (e^{2ah}+1)/h^2$,*

$$\sqrt{K}(\widehat{a}_h - a) \Rightarrow N(0, V_h), \qquad K\mathbb{E}(\widehat{a}_h - a)^2 \to V_h, \qquad K\,\mathrm{Var}(\widehat{a}_h) \to V_h. \tag{155}$$

Separately, the risk equivalences in Theorems 5 and 6 remain valid, up to constants, after adding centered Gaussian endpoint errors of known, uniformly bounded covariance, independent across times, records and latent paths. Known same-time error covariance is subtracted from second moments.

Proof. For $\rho = e^{-ah}$, Isserlis gives $\mathrm{Var}(R_0R_h) = s^2(1+\rho^2)$. The iid CLT and $f'(\rho) = -(h\rho)^{-1}$ give the normal limit because $\rho \in (\ell, u)$. To justify moment convergence, put $Z_j = R_{j,0}R_{j,h}/s$, $\sigma^2 = 1+\rho^2$ and $\mu_4 = \mathbb{E}(Z_j - \rho)^4 < \infty$. Independence yields

$$\mathbb{E}|\overline{Z}_K - \rho|^4 = \mu_4/K^3 + 3(K-1)\sigma^4/K^3 = O(K^{-2}).$$

The fixed clipped map is globally $(h\ell)^{-1}$-Lipschitz, so $U_K = \sqrt{K}(\widehat{a}_h - a)$ satisfies $\sup_K \mathbb{E}|U_K|^4 \le (h\ell)^{-4}(\mu_4 + 3\sigma^4)$. Uniform integrability of U_K, U_K^2 gives $\sqrt{K}(\mathbb{E}\widehat{a}_h - a) \to 0$ and $\mathbb{E}U_K^2 \to V_h$, hence also $K\,\mathrm{Var}(\widehat{a}_h) = \mathbb{E}U_K^2 - (\mathbb{E}U_K)^2 \to V_h$. At a clipping boundary the normal and variance limits can differ.

For the noise assertion, independent centered errors leave lagged moments unchanged; same-time moments become $\Sigma + \Omega_t$. Subtracting known Ω_t gives unbiased moments with error $C/\sqrt{n}$ by bounded Gaussian fourth moments. Substitution into the original clipped decoders preserves their upper-bound arguments. Adding parameter-independent noise is a Markov transformation, so data processing preserves the noiseless minimax lower bounds. The result uses known bounded noise, not unknown or serially correlated errors. □

Proposition 25 (Task-weighted lag uncertainty)**.** *Consider a centered Gaussian OU star channel with independent drivers, known static $s, v > 0$, $x \neq 0$ and $sv > x^2$. Put $c = x^2/s < v$ and let (a, d) be an interior feasible point*

$$\mathcal{F} = \{a, d > 0 :\ q_L = 2[dv - (a+d)c] > 0\}.$$

For fixed lag $h > 0$ and horizon $H > 0$, the self-lag map

$$m_h(a,d) = (m_R, m_L)^\top = \left(se^{-ah},\ ve^{-dh} + cG_h(a,d)\right)^\top$$

is globally injective on $\mathcal{F}$, and its Jacobian satisfies $\det J_h \ge h^2 s(v-c)e^{-(a+d)h} > 0$. Observe K iid stationary full pairs (R_0, L_0, R_h, L_h) and average their two lag products to obtain $\widehat{m}_h$. For a locally consistent inverse $(\widehat{a}, \widehat{d})$ solving $m_h(\widehat{a}, \widehat{d}) = \widehat{m}_h$ with probability tending to one, let $g = (x/s)G_H(a,d)$ and $\widehat{g} = g(\widehat{a}, \widehat{d})$. Then

$$\sqrt{K}(\widehat{g} - g) \Rightarrow N(0, V_{g,h}), \qquad V_{g,h} = \nabla g^\top J_h^{-1}\Omega_h J_h^{-\top}\nabla g, \tag{156}$$

with derivatives in (a, d) holding s, v, x fixed, and

$$\Omega_h = \begin{pmatrix} s^2 + m_R^2 & x^2 + \Gamma_{RL}(h)\Gamma_{LR}(h) \\ x^2 + \Gamma_{RL}(h)\Gamma_{LR}(h) & v^2 + m_L^2 \end{pmatrix}, \qquad \Gamma_{RL}(h) = xe^{-ah}, \qquad \Gamma_{LR}(h) = x\{e^{-ah} + 2a\phi_h(a,d)\}. \tag{157}$$

Here $\Gamma_{UV}(h) = \mathrm{Cov}(U_h, V_0)$ is future-first. All parameters, static moments and lags remain fixed in this limit. With independent centered Gaussian errors of known variances ω_R, ω_L across coordinates, times and pairs, independent of the process, use noisy products but retain the exact latent s, v, x. The same limit holds with

$$\Omega_h^{(\omega)} = \begin{pmatrix} (s+\omega_R)^2 + m_R^2 & x^2 + \Gamma_{RL}(h)\Gamma_{LR}(h) \\ x^2 + \Gamma_{RL}(h)\Gamma_{LR}(h) & (v+\omega_L)^2 + m_L^2 \end{pmatrix}. \tag{158}$$

Proof. The drift $A = \left(\begin{smallmatrix} -a & 0 \\ b & -d \end{smallmatrix}\right)$ with $b = (a+d)x/s$ and diffusion variances $2as$ and q_L realizes the positive-definite specified static covariance by the Lyapunov equation. Its transition is $e^{Ah} = \left(\begin{smallmatrix} e^{-ah} & 0 \\ b\phi_h & e^{-dh} \end{smallmatrix}\right)$; multiplication by the static covariance gives the stated four lag entries, using $(d-a)\phi_h = e^{-ah} - e^{-dh}$. Differentiation under the integral gives

$$\nabla G_t(a,d) = \int_0^t e^{-au-d(t-u)} \begin{pmatrix} 1-(a+d)u \\ 1-(a+d)(t-u) \end{pmatrix} du, \qquad \nabla g = (x/s)\nabla G_H,$$

smooth also at $a = d$. Thus

$$J_h = \begin{pmatrix} -hse^{-ah} & 0 \\ c\partial_a G_h & -hve^{-dh} + c\partial_d G_h \end{pmatrix}. \tag{159}$$

More precisely,

$$\begin{aligned} m_L &= (v-c)e^{-dh} + ce^{-ah} + 2ac\phi_h, \\ \partial_d m_L &= -h(v-c)e^{-dh} - 2ac\int_0^h (h-u)e^{-au-d(h-u)}\,du < 0. \end{aligned}$$

Multiplying by $-hse^{-ah}$ proves the determinant bound, uniform away from zero on any compact neighborhood inside $\mathcal{F}$. Equality of moment vectors forces equal a because se^{-ah} is strictly decreasing. For this fixed a, feasible d form the interval $d > ac/(v-c)$, on which m_L is strictly decreasing, proving global injectivity without a small-coupling condition. The inverse function theorem gives a smooth inverse near the true moments; the law of large numbers puts $\widehat{m}_h$ there with probability tending to one. Any measurable extension used only outside this neighborhood leaves the distributional limit unchanged.

For $Z_j = (R_{j,0}R_{j,h}, L_{j,0}L_{j,h})^\top$, Isserlis gives $\mathrm{Var}(R_0R_h) = s^2 + m_R^2$, $\mathrm{Var}(L_0L_h) = v^2 + m_L^2$, and

$$\begin{aligned} \mathrm{Cov}(R_0R_h, L_0L_h) &= \mathrm{Cov}(R_0, L_0)\,\mathrm{Cov}(R_h, L_h) \\ &\quad + \mathrm{Cov}(R_0, L_h)\,\mathrm{Cov}(R_h, L_0) \\ &= x^2 + \Gamma_{LR}(h)\Gamma_{RL}(h). \end{aligned}$$

Thus $\mathrm{Cov}(Z_j) = \Omega_h$, and the iid multivariate CLT followed by the delta method for $g \circ m_h^{-1}$ proves (156). Independent Gaussian sensor errors change only the endpoint variances in this calculation, to $s + \omega_R, v + \omega_L$; all cross-coordinate and positive-lag covariances, the mean map and its derivatives stay unchanged. This proves the noisy version by the same argument. □

The classical delta method here weights uncertainty by response sensitivity and retains the covariance of the two same-pair products. Its preferred lag need not minimize the root-only constant (155). This proposition supplies the variance of a pointwise limiting normal law, not finite-budget MSE or variance convergence; the latter were established separately for the fixed-clipped scalar inverse. Empirical comparisons with estimated static moments are reported in Appendix D.5.

C Calibration and finite-budget measurement design

C.1 Shared model and admissibility

Let $m \geq 3$ be odd, $S = E(K_m)\backslash E(C_m)$, and index leaves by $0, \ldots, m-1$. A universal root $R_\star$ feeds every leaf; each dedicated root $R_{\{i,j\}}$, $\{i,j\} \in S$, feeds its two endpoints. For fixed $0 < \delta \leq 1/2$ and $0 < \epsilon \leq 1/(4m)$, set

$$a_\star = 1, \qquad d_i \in [2-\delta, 2+\delta], \qquad \sum_i d_i = 2m, \qquad a_{\{i,j\}} = \frac{3}{5}(d_i + d_j) - 1, \qquad \eta \in [\epsilon, 2\epsilon]. \tag{160}$$

All stationary variances are one. Distinct roots are uncorrelated; allowed root–leaf covariances equal η and forbidden ones vanish. With $P(i)$ the parents of leaf i, define the triangular drift A by root and

leaf diagonals $-a_k, -d_i$, allowed entries B_{ik}, and zero remaining entries, and take diagonal diffusion Q with

$$B_{ik} = (a_k + d_i)\eta, \qquad Q_{kk} = 2a_k \quad (k \text{ a root}), \qquad Q_{ii} = 2d_i - 2\eta^2 \sum_{k \in P(i)} (a_k + d_i). \tag{161}$$

The leaf-pair Lyapunov equation is $(d_i + d_j)z_{ij} = \eta^2 \sum_{k \in P(i) \cap P(j)} (2a_k + d_i + d_j)$, so

$$z_{ij} = \begin{cases} \eta^2\{1 + 2/(d_i + d_j)\}, & \{i, j\} \in E(C_m), \\ (16/5)\eta^2, & \{i, j\} \in S. \end{cases} \tag{162}$$

Since S is $(m-3)$-regular, the sum of positive rates is

$$T_m = 1 + 2m + \frac{7}{5}|S|. \tag{163}$$

The graph, variances, $a_\star$, and root-rate linkage are known; neither individual dedicated-root rates nor the full spectrum is supplied. Supplying T_m adds no information. Observations are centered, noiseless Gaussian stationary preparations, independent across records. Write $\mathcal{U}_{m,\epsilon}$ for this class (with δ fixed) and define the time-one response by

$$g_{\eta,d} = \eta G_1(1, d_0), \qquad G_h(a, d) = (a+d)\phi_h(a, d), \qquad \phi_h(a, d) = \int_0^h e^{-au-d(h-u)}\, du. \tag{164}$$

Admissibility and response margin. Orient the root–leaf block X with leaves as rows. The diagonal, cycle off-diagonal, and other off-diagonal entries of $XX^\top$ are $(m-2)\eta^2, \eta^2, 2\eta^2$, respectively. Its constant row sum is $3(m-2)\eta^2$. If D is the leaf off-diagonal block, then

$$\|X\|_{\text{op}} = \sqrt{3(m-2)}\eta \le \frac{\sqrt{6}}{8}, \qquad \|D\|_{\text{op}} \le \frac{16}{5}(m-1)\eta^2 \le \frac{8}{45},$$
$$\|\Sigma - I\|_{\text{op}} \le \frac{\sqrt{6}}{8} + \frac{8}{45} < \frac{1}{2}.$$

Here $\eta \le 1/(2m)$, $(m-2)/m^2 \le 1/8$, and $(m-1)/m^2 \le 2/9$. Also $a_k \in [4/5, 2]$, $d_i \in [3/2, 5/2]$, and each leaf has $m-2$ parents, giving $Q_{ii} \ge 3 - 9(m-2)\eta^2 \ge 87/32$ and $Q_{kk} \ge 8/5$. Thus $Q > 0$, Σ has eigenvalues in $[1/2, 3/2]$, and the stable triangular drift and displayed covariance satisfy every Lyapunov entry. These are uniform bounds with coupling constant $1/4$. For $r = d - 1 \in [1/2, 3/2]$, the exponential series gives $2e^r - (r^2 + 2r + 2) \ge r^3/3$, hence

$$-\partial_d G_1(1, d) = \frac{e^{-(1+r)}}{r^2}\{2e^r - (r^2 + 2r + 2)\} \ge \frac{e^{-5/2}}{6} =: c_G > 0. \tag{165}$$

C.2 Unknown common amplitude: static and temporal risk

Theorem 26 (Static calibration)**.** *For $N \ge 1$ independent full named stationary vectors, put $q_N = \sqrt{m}/(\epsilon^2\sqrt{N})$. Then*

$$\inf_{\widehat{g}} \sup_{(\eta,d) \in \mathcal{U}_{m,\epsilon}} \frac{\mathbb{E}_{\eta,d}|\widehat{g} - g_{\eta,d}|}{\epsilon} \asymp 1 \wedge q_N. \tag{166}$$

Constants are independent of m, ϵ, N and may depend on fixed $\delta > 0$; the estimator knows the class but not η, d.

Proof. Let U be the unsigned edge-by-vertex incidence matrix of C_m. The recursion $d_{i+1} = s_i - d_i$ and odd closure make U invertible, with every entry of $\ell^\top = e_0^\top U^{-1}$ equal to $\pm 1/2$:

$$\|\ell\|_2 = \sqrt{m}/2, \qquad \|\ell\|_1 = m/2. \tag{167}$$

The population inverse is

$$s_e = d_i + d_j = f(z_e; \eta), \qquad f(z; u) = \frac{2u^2}{z - u^2}, \qquad d = U^{-1}s. \tag{168}$$

To control empirical denominators on every sample, choose smooth $0 \le \chi \le 1$, equal to one on $[7/5, 5/3]$ and supported in $(5/4, 7/4)$. Set $\psi(v) = 4 + \chi(v)\{2/(v-1) - 4\}$ on that open interval and $\psi(v) = 4$ elsewhere; let $\widetilde{f}(z;u) = \psi(z/u^2)$. Since $z_e/\eta^2 \in [1 + 1/(2+\delta), 1 + 1/(2-\delta)] \subset [7/5, 5/3]$, this agrees with the inverse at every population covariance. Globally for $z \in \mathbb{R}$, $u \in [\epsilon, 2\epsilon]$,

$$|\widetilde{f}| \le C, \qquad |\partial_z \widetilde{f}| \le C\epsilon^{-2}, \qquad |\partial_{zz} \widetilde{f}| \le C\epsilon^{-4}, \qquad |\partial_u \widetilde{f}| \le C\epsilon^{-1}. \tag{169}$$

The last inequality follows from $\partial_u \widetilde{f} = -2v\psi'(v)/u$ and boundedness of $v\psi'(v)$.

For $N \ge m\epsilon^{-4}$, split the sample into $n_1 = \lfloor N/2 \rfloor$, $n_2 = N - n_1$, both at least $N/3$. With Π_I denoting interval projection, use the first batch for

$$\widehat{\eta} = \Pi_{[\epsilon, 2\epsilon]} \left(n_1^{-1} \sum_{j=1}^{n_1} R_\star^{(j)} L_0^{(j)} \right) \tag{170}$$

and second-batch averages $\widehat{z}_{\{i,j\}} = n_2^{-1} \sum_{b=1}^{n_2} L_i^{(b)} L_j^{(b)}$. Define

$$\widehat{d}_0 = \Pi_{[2-\delta, 2+\delta]} \left\{ \sum_e \ell_e \widetilde{f}(\widehat{z}_e; \widehat{\eta}) \right\}, \qquad \widehat{g} = \widehat{\eta} G_1(1, \widehat{d}_0). \tag{171}$$

Write $\Delta_\eta = \widehat{\eta} - \eta$, $\Delta_e = \widehat{z}_e - z_e$. Gaussian fourth moments and projection give

$$\mathbb{E}|\Delta_\eta| \le C n_1^{-1/2}, \qquad \mathbb{E}\Delta_e^2 \le C/n_2, \qquad \mathrm{Var}\left(\sum_e v_e \Delta_e \right) \le C\|v\|_2^2/n_2 \tag{172}$$

for deterministic v. Indeed $\sum_e v_e L_i L_j = L^\top A_v L$, where $(A_v)_{ij} = (A_v)_{ji} = v_e/2$ on cycle edges. Its variance is $2\,\mathrm{tr}(A_v \Sigma_{LL} A_v \Sigma_{LL}) \le 2\|\Sigma_{LL}\|_{\mathrm{op}}^2 \|A_v\|_F^2 \le C\|v\|_2^2$; averaging independent vectors divides this by n_2.

The exact pre-projection error is

$$\sum_e \ell_e \{\widetilde{f}(\widehat{z}_e; \widehat{\eta}) - \widetilde{f}(z_e; \eta)\} = C_\eta + L + R, \tag{173}$$

where $C_\eta = \sum_e \ell_e \{\widetilde{f}(\widehat{z}_e; \widehat{\eta}) - \widetilde{f}(\widehat{z}_e; \eta)\}$, $L = \sum_e \ell_e \partial_z \widetilde{f}(z_e; \eta) \Delta_e$, and Taylor's theorem gives $|R| \le C\epsilon^{-4} \sum_e |\ell_e| \Delta_e^2$. Thus $|C_\eta| \le Cm|\Delta_\eta|/\epsilon$ on every sample, $\mathbb{E}|L| \le C\sqrt{m}/(\epsilon^2\sqrt{n_2})$ by (172), and $\mathbb{E}|R| \le Cm/(n_2\epsilon^4)$. Nonexpansiveness of projection yields

$$\mathbb{E}|\widehat{d}_0 - d_0| \le C \left\{ \frac{m}{\epsilon\sqrt{n_1}} + \frac{\sqrt{m}}{\epsilon^2\sqrt{n_2}} + \frac{m}{n_2\epsilon^4} \right\}. \tag{174}$$

Boundedness and smoothness of G_1, and $\eta/\epsilon \le 2$, imply

$$\begin{aligned} \frac{\mathbb{E}|\widehat{g} - g|}{\epsilon} &\le C\{\mathbb{E}|\Delta_\eta|/\epsilon + \mathbb{E}|\widehat{d}_0 - d_0|\} \\ &\le C \left\{ \frac{1}{\epsilon\sqrt{N}} + \frac{m}{\epsilon\sqrt{N}} + q_N + q_N^2 \right\}. \end{aligned} \tag{175}$$

Here $q_N \le 1$, $m/(\epsilon\sqrt{N}) = \sqrt{m}\epsilon q_N \le q_N/(4\sqrt{m})$, and $1/(\epsilon\sqrt{N}) = (\epsilon/\sqrt{m}) q_N$. This proves the upper bound without excluding any bad event. For $N < m\epsilon^{-4}$, $\widehat{g} = 0$ has bounded normalized risk.

For the lower bound fix and, if desired, reveal $\eta = \epsilon$. Let $v_j = \sqrt{2/m}\cos(2\pi(m-1)j/(2m))$. Then $\mathbf{1}^\top v = 0$, $\|v\|_2 = 1$, $v_0 = \|v\|_\infty = \sqrt{2/m}$, and $\|Uv\|_2 = \sigma_m = 2\sin(\pi/(2m)) \asymp m^{-1}$. The path $d(t) = 2\mathbf{1} + tv$, $0 \le t \le \delta\sqrt{m/2}$, preserves all supplied quantities. Only cycle covariances change, so (162) and the covariance eigenvalue bounds imply

$$\|\Sigma_t - \Sigma_0\|_F \le C\epsilon^2 t \sigma_m, \qquad D_{\mathrm{KL}}(P_t^{\otimes N} \| P_0^{\otimes N}) \le CN\epsilon^4 t^2 \sigma_m^2.$$

The latter follows from the Gaussian KL formula and $x - 1 - \log x \le C(x-1)^2$ on $[1/3, 3]$. The normalized response gap is at least $ct/\sqrt{m}$ by (165). Choose $t = c_1 \min\{\sqrt{m}, (N\epsilon^4\sigma_m^2)^{-1/2}\}$, with $c_1 > 0$ small depending only on δ; the path is admissible and Pinsker gives total variation at most $1/2$. Classifying by the nearer of two targets separated by Δ gives maximum absolute risk at least $\Delta(1 - \mathrm{TV})/4$. Consequently the normalized risk is at least $c \min\{1, (\epsilon^2\sigma_m\sqrt{mN})^{-1}\} \asymp 1 \wedge q_N$. $\square$

Proposition 27 (Calibrated target self-lag)**.** *Fix $h > 0$. Observe $M \geq 1$ independent simultaneous pairs $(R_\star, L_0)$ and, independently, $K \geq 1$ target self-lag pairs $(L_0(0), L_0(h))$, independent across preparations. Define*

$$\begin{aligned} \widehat{\eta} &= \Pi_{[\epsilon,2\epsilon]}\left(M^{-1}\sum_{j=1}^{M} R_\star^{(j)} L_0^{(j)}\right), \qquad \widehat{\rho}_h = K^{-1}\sum_{j=1}^{K} L_0^{(j)}(0) L_0^{(j)}(h), \\ A_m(h) &= 3\phi_h(1,2) + (m-3)\frac{17}{5}\phi_h(7/5,2), \qquad I_h = [e^{-(2+\delta)h}, e^{-(2-\delta)h}], \\ \widehat{d}_0 &= -h^{-1}\log \Pi_{I_h}\{\widehat{\rho}_h - \widehat{\eta}^2 A_m(h)\}, \qquad \widehat{g} = \widehat{\eta} G_1(1, \widehat{d}_0). \end{aligned} \tag{176}$$

Uniformly in m, ϵ, M, K,

$$\sup_{\mathcal{U}_{m,\epsilon}} \frac{\mathbb{E}|\widehat{g} - g_{\eta,d}|}{\epsilon} \leq C_h\left[1 \wedge \left\{(\epsilon\sqrt{M})^{-1} + K^{-1/2} + m\epsilon^2\delta\right\}\right]. \tag{177}$$

Proof. The leaf row of e^{Ah} has diagonal $e^{-d_0 h}$ and parent entries $B_{0k}\phi_h(a_k, d_0)$, giving the exact self-covariance

$$\rho_h = e^{-d_0 h} + \eta^2 B_m(d; h), \qquad B_m(d; h) = \sum_{k \in P(0)} G_h(a_k(d), d_0). \tag{178}$$

There are $m - 2$ summands, $B_m(2\mathbf{1}; h) = A_m(h)$, and $|a_k - 7/5| \leq 6\delta/5$ for dedicated roots. Smoothness on the fixed rate intervals proves

$$|A_m(h)| \leq C_h m, \qquad |B_m(d; h) - A_m(h)| \leq C_h m\delta. \tag{179}$$

Gaussian product variance $1 + \rho^2 \leq 2$ gives $\mathbb{E}|\widehat{\eta} - \eta| \leq C/\sqrt{M}$ and $\mathbb{E}|\widehat{\rho}_h - \rho_h| \leq \sqrt{2/K}$. The corrected error is exactly

$$\widehat{\rho}_h - \widehat{\eta}^2 A_m - e^{-d_0 h} = \widehat{\rho}_h - \rho_h + \eta^2(B_m - A_m) + (\eta^2 - \widehat{\eta}^2)A_m.$$

Since $|\widehat{\eta}^2 - \eta^2| \leq 4\epsilon|\widehat{\eta} - \eta|$, projection and the Lipschitz logarithm on I_h yield

$$\mathbb{E}|\widehat{d}_0 - d_0| \leq C_h\left\{K^{-1/2} + m\epsilon^2\delta + m\epsilon/\sqrt{M}\right\}. \tag{180}$$

The response inequality in (175) adds $(\epsilon\sqrt{M})^{-1}$ for the amplitude itself. The center-correction term $m\epsilon/\sqrt{M} = (m\epsilon^2)/(\epsilon\sqrt{M})$ is absorbed because $m\epsilon^2 \leq 1/(16m)$. The projected response has bounded normalized error on every sample, proving the minimum with one. □

C.3 Local temporal-budget transition

For this local comparison the amplitude is known: $\eta = \epsilon$. Fix odd $m \geq 3$ and $h > 0$, draw independent $U_i \sim U[-1, 1]$, and set $D_i = U_i - m^{-1}\sum_j U_j$, $d_i = 2 + wD_i$. For $0 < w \leq 1/4$ this is the shared construction with radius $\delta = 2w$; $\sum_i d_i = 2m$ and $|D_i| \leq 2$. Let $g = \epsilon G_1(1, 2 + wD_0)$, $g_c = \epsilon G_1(1, 2)$, and let $\widehat{g}_K$ be (176) with $\widehat{\eta} = \epsilon$, no calibration sample, and $I_h = [e^{-(2+2w)h}, e^{-(2-2w)h}]$.

Proposition 28 (Prior width and temporal budget)**.** *Let $w, \epsilon \downarrow 0$ and $Kw^2 \to \kappa \in (0, \infty)$. With expectations over both prior and observations, define*

$$q = |\partial_d G_1(1, 2)| = |5e^{-2} - 2e^{-1}| > 0, \qquad \sigma_h^2 = (e^{4h} + 1)/h^2, \qquad c_m = \frac{1}{2} - \frac{1}{3m}, \tag{181}$$

and, for standard normal Z independent of D_0,

$$J_m(s) = \mathbb{E}|\Pi_{[-2,2]}(D_0 + sZ) - D_0|. \tag{182}$$

Then

$$\frac{\mathbb{E}|g_c - g|}{\epsilon w q} \to c_m, \qquad \frac{\mathbb{E}|\widehat{g}_K - g|}{\epsilon w q} \to J_m(\sigma_h/\sqrt{\kappa}). \tag{183}$$

The optimal Bayes risk from N full named static vectors has the same normalized limit c_m if Nw^2 is bounded, even with the prior supplied. Moreover J_m increases continuously and strictly from zero to two, so its unique root $J_m(s_m^*) = c_m$ gives the local transition $K_m^*(w, h) = \sigma_h^2/[w^2(s_m^*)^2]$.

Proof. Put $b = \epsilon^2 A_m(h)$. Equations (178)– (179) give $\rho_h = e^{-d_0 h} + b + r(D)$ with $\sup_D |r(D)| \le C_{m,h}\epsilon^2 w$. Each Gaussian lag product has variance $1 + \rho_h^2$ and uniformly bounded third centered absolute moment. Since $\rho_h \to e^{-2h} < 1$ uniformly in D, the conditional triangular-array central limit theorem gives

$$\sqrt{K}(\widehat{\rho}_h - \rho_h) \Rightarrow \sqrt{1 + e^{-4h}}Z,$$
$$\frac{\widehat{\rho}_h - b - e^{-2h}}{w} \Rightarrow -he^{-2h}D_0 + \kappa^{-1/2}\sqrt{1 + e^{-4h}}Z.$$

The uniform conditional limit makes Z independent of D; the second convergence uses $r(D)/w \to 0$. Projection endpoints are $e^{-2h} \mp 2whe^{-2h} + O(w^2)$. Smoothness of the logarithm and continuity of interval projection therefore give the joint limit

$$\left(D_0, \frac{\widehat{d}_0 - 2}{w}\right) \Rightarrow \left(D_0, \Pi_{[-2,2]}(D_0 + \sigma_h Z/\sqrt{\kappa})\right),$$

where the Gaussian sign is immaterial. Both scaled rates lie in $[-2, 2]$, so a uniform Taylor expansion of $G_1(1, 2 + wv)$ and bounded loss give (183). To compute $\mathbb{E}|D_0|$, write $D_0 = b_0 U_0 - V$, where $b_0 = (m-1)/m$ and $V = m^{-1}\sum_{j=1}^{m-1} U_j$. Since $|V| \le b_0$,

$$\mathbb{E}(|b_0 U_0 - V| \mid V) = b_0/2 + V^2/(2b_0), \qquad \mathbb{E}V^2 = (m-1)/(3m^2),$$

which gives $\mathbb{E}|D_0| = c_m$.

For static data, only cycle covariances vary with D, each by $O(\epsilon^2 w)$ at fixed m. The same Gaussian KL bound yields $\sup_D D_{\mathrm{KL}}(P_D^{\otimes N} \| P_0^{\otimes N}) \le C_m N \epsilon^4 w^2 \to 0$. Thus the actual joint law $\pi(dD)P_D^{\otimes N}$ and the independent-data law $\pi(dD)P_0^{\otimes N}$ have vanishing total variation by Pinsker. Clipping actions to the attainable response interval bounds loss divided by $\epsilon w q$ uniformly and cannot increase absolute loss; their optimal normalized Bayes risks therefore differ by $o(1)$. Under the independent-data law g_c is a prior median, since D_0 is symmetric and $G_1(1, \cdot)$ is strictly monotone on the support. Its risk limit is c_m.

For $v \in (-2, 2)$, integrating the positive and negative Gaussian tails gives

$$\begin{aligned} \mathbb{E}_Z|\Pi_{[-2,2]}(v + sZ) - v| &= F(s, 2 - v) + F(s, 2 + v), \\ F(s,t) &= s\{\varphi(0) - \varphi(t/s)\} + t\Phi(-t/s), \quad t > 0, \end{aligned} \tag{184}$$

where φ, Φ are the standard normal density and cdf. Here $\partial_s F(s,t) = \varphi(0) - \varphi(t/s) > 0$; conditional loss tends to zero as $s \downarrow 0$ and to $(2 - v)/2 + (2 + v)/2 = 2$ as $s \to \infty$. The prior has $|D_0| \le 2(m-1)/m < 2$ and the loss is bounded by four. Averaging preserves continuity, strict increase, and these limits. Since $0 < c_m < 2$, the root exists uniquely, and $J_m(\sigma_h/\sqrt{\kappa}) < c_m$ exactly when $\kappa > \sigma_h^2/(s_m^*)^2$. □

The transition is a fixed-(m, h) local-limit comparison for this decoder; at finite w, ϵ its budget threshold is an approximation.

C.4 Unknown global time scale

Retain known $\eta = \epsilon$ and fixed $\delta > 0$, and let (A_d, Q_d) be the shared model. Define $\mathcal{V}_{m,\epsilon}$ by

$$(A, Q) = (\tau A_d, \tau Q_d), \qquad \tau \in [9/10, 11/10], \qquad g_{\tau,d} = \epsilon G_1(\tau, \tau d_0). \tag{185}$$

The baseline sum $\sum_i d_i = 2m$ fixes the meaning of τ. Stationary covariance is unchanged; the actual root rate τ, trace τT_m, and full spectrum are not supplied.

Corollary 29 (Estimated time scale)**.** *For $N \ge 1$ full stationary vectors and $M \ge 0$ universal-root lag-one pairs, all independent across records and batches,*

$$\inf_{\widehat{g}} \sup_{\mathcal{V}_{m,\epsilon}} \frac{\mathbb{E}|\widehat{g} - g_{\tau,d}|}{\epsilon} \asymp 1 \wedge \max\{q_N, M^{-1/2}\}, \qquad q_N = \frac{\sqrt{m}}{\epsilon^2\sqrt{N}}. \tag{186}$$

Here $M^{-1/2} = +\infty$ at $M = 0$; constants may depend on δ but are uniform in m, ϵ, N, M on the shared admissible ranges.

Proof. Fixing $\tau = 1$ gives the static lower-bound path above. Root pairs have constant correlation e^{-1}, hence add zero KL and leave the lower bound $c(1 \wedge q_N)$ unchanged. Conversely, fix $d = 2\mathbf{1}$ and compare $\tau = 1$ and $1 + t$. Static laws coincide, while a root pair has covariance $V_\tau = \left(\begin{smallmatrix} 1 & e^{-\tau} \\ e^{-\tau} & 1 \end{smallmatrix}\right)$ with uniformly positive eigenvalues, so the M-pair KL is at most CMt^2. Also

$$G_1(\tau, 2\tau) = 3(e^{-\tau} - e^{-2\tau}), \qquad \frac{d}{d\tau} G_1(\tau, 2\tau)\Big|_{\tau=1} = 3(-e^{-1} + 2e^{-2}) \neq 0.$$

Taking $t = c(1 \wedge M^{-1/2})$ (fixed small t at $M = 0$) and using the two-point argument gives $c(1 \wedge M^{-1/2})$. The maximum of these two lower bounds is the claimed rate.

For the upper bound use (171) with known amplitude, all N vectors, and no split when $N \geq m\epsilon^{-4}$. The same Taylor bound, without calibration error, gives $\mathbb{E}|\widehat{d}_0 - d_0| \leq C(q_N + q_N^2) \leq Cq_N$. For smaller N use $\widehat{d}_0 = 2$. For $M \geq 1$, set

$$\widehat{\tau} = -\log \Pi_{[e^{-11/10}, e^{-9/10}]} \left(M^{-1} \sum_{j=1}^{M} R_\star^{(j)}(0) R_\star^{(j)}(1) \right).$$

Product variance and the clipped logarithm give $\mathbb{E}|\widehat{\tau} - \tau| \leq CM^{-1/2}$; use $\widehat{\tau} = 1$ at $M = 0$. Compact-rectangle Lipschitz continuity of $(\tau, d) \mapsto G_1(\tau, \tau d)$ bounds the normalized risk of $\epsilon G_1(\widehat{\tau}, \widehat{\tau}\widehat{d}_0)$ by $C\{(1 \wedge q_N) + (1 \wedge M^{-1/2})\} \leq 2C[1 \wedge \max\{q_N, M^{-1/2}\}]$. □

Thus static observations determine relative rates and root self-lags calibrate time; without root pairs a positive risk floor remains.

D Experimental protocols and results

The experiments ask whether temporal measurements improve response estimation beyond a strong static decoder, at what budget, and through which coordinates. Unless stated otherwise, systems are stable two-layer OU models with positive diagonal diffusion, the response horizon is one, and observations are independent stationary Gaussian replicates. Exact Bartlett–Wishart sufficient statistics reproduce snapshot laws without storing the nominal number of vectors. Only the trajectory check uses correlated observations. Paired contrasts compare errors on the same system, not necessarily the same records; cells and batches are not pooled.

We distinguish absolute error, relative error $|\widehat{g} - g|/g$ for positive-response star comparisons, and coupling-normalized error $|\widehat{g} - g|/\epsilon$ for multiroot, anchor, and trajectory tests. Positive contrasts mean the stated comparator has larger error. A resolved contrast has its specified adjusted interval excluding zero. Endpoint counts price scalar readings, not preparation, synchronization, elapsed time, or acquisition of a supplied trace. Refits, stratification, loss renormalization, and response-lag calculations are post-hoc diagnostics specified after aggregate results were observed. Raw results and experiment manifests retain seeds, generating parameters, observations or sufficient statistics, estimates, losses, software versions, and source hashes. Named star artifacts include `compare_no_spectrum_star.py`, `cloud_static_targeted_diagnostics.py`, `cloud_star_diagnostics.py`, and the matched-sensor, full-bivariate, and self-lag experiment scripts.

D.1 Multiroot comparison with a full static likelihood

For $m \in \{3, 5, 7, 9\}$, set $\epsilon = .25/m$ and $d_i = 2 + u_i - \bar{u}$, with independent $u_i \sim U[-.17, .17]$. The universal root rate is one; dedicated rates are $a_{ij} = .6(d_i + d_j) - 1 \in [.992, 1.808]$. The target is the universal-root response to leaf zero. All arms know the generating prior, unit marginal variances, fixed root–leaf covariance block X, and rate support $[1.66, 2.34]$, but no anonymous spectrum. Static arms receive K full records; the temporal arm receives K target pairs at $h = .5$, independently conditional on the system. A static record contains $r + m$ scalars, a temporal record two: equal record counts favor static sensing in scalar count.

With known zero mean and regression X, generate $KS \sim \text{Wishart}_m(K, V(d))$ for residuals $L - XR$, where

$$V_{ii} = 1 - (m-2)\epsilon^2, \qquad V_{ij} = \begin{cases} 2\epsilon^2/(d_i + d_j), & (i,j) \in C_m, \\ 6\epsilon^2/5, & (i,j) \notin C_m, \end{cases} \quad i \neq j. \tag{187}$$

The full static log-likelihood and absolute-response Bayes action are

$$\ell(d; S) = -\frac{K}{2}[\log\det V(d) + \text{tr}\{V(d)^{-1}S\}] + c(S), \tag{188}$$

$$\widehat{g}_{\text{B}}(S) \in \text{Median}_{d\sim\pi_m(\cdot|S)}\{\epsilon G_1(1, d_0)\}, \qquad \pi_m(dd \mid S) \propto e^{\ell(d;S)}\pi_m(dd). \tag{189}$$

Here π_m is the induced centered-uniform prior, not a uniform law on the constrained box. All entries of S, including those with constant population covariance, enter the likelihood. The posterior median uses 32,768 scrambled Sobol prior points and likelihood weights, without linearizing either model or prior. The linear comparator uses the same S and prior covariance $(.17^2/3)(I - \mathbf{1}\mathbf{1}^\top/m)$. Target-lag inversion uses the fixed center correction and the same projection support. For pair correlation ρ_h, the exact product mean is $[(1+\rho_h)Z_+ - (1-\rho_h)Z_-]/(2K)$, $Z_\pm \overset{\text{iid}}{\sim} \chi^2_K$.

Table 1 uses 2,000 fresh systems per cell (seed base 2,026,092,617). One-sided paired t tests with Holm adjustment over eight cells have maximum adjusted $p = 1.25 \times 10^{-24}$; 10,000-resample bootstrap intervals in the table are unadjusted. The independent 65,536-point audit on the first 256 systems per cell changes normalized predictions by at most 2.898×10^{-5} (tolerance 10^{-4}); main integration ESS exceeds 99.65%. Scalar/batched log-likelihood differences agree within 1.79×10^{-11}, and covariance implementations within 1.74×10^{-18} on 44 center, boundary, and random points. All 16,000 system records are retained.

m	K	static linear	static nonlinear	target lag	paired 95% CI
3	8,192	3.888	3.888	3.084	[.651,.957]
5	8,192	4.334	4.334	3.027	[1.149,1.464]
7	8,192	4.524	4.523	3.040	[1.328,1.640]
9	8,192	4.690	4.689	3.050	[1.480,1.801]
3	32,768	3.876	3.876	1.502	[2.246,2.504]
5	32,768	4.347	4.346	1.513	[2.700,2.968]
7	32,768	4.586	4.586	1.532	[2.917,3.192]
9	32,768	4.591	4.592	1.515	[2.944,3.209]

Table 1: Full static likelihood versus target lag. All entries, including the unadjusted paired 95% interval for nonlinear static minus lag, are in units of 10^{-3} for $|\widehat{g} - g|/\epsilon$; 2,000 systems per row. All eight one-sided tests pass Holm adjustment. Linear and nonlinear mean errors differ by less than 8×10^{-7} per cell.

The earlier center/linear comparison used a separate final batch (seed base 8,870,000), after two pilots aligned projection support and required beating both baselines. Its intersection–union cell tests passed Holm adjustment across eight cells (largest unadjusted $p = 2.48 \times 10^{-23}$; 2,000 bootstrap resamples). At the secondary $K = 2048$ budget, however, lag error .00582–.00599 exceeded the center's .00380–.00464. The next independent experiment tests this budget dependence.

D.2 Independent budget and amplitude-calibration validation

The frozen batch has 1,000 systems in each of 48 cells: $m = 3, 5, 7, 9$, $K = 512, 2048, 8192, 32768$, and three amplitude regimes. It uses the prior, $\epsilon = 1/(4m)$, $h = .5$, and horizon above; coupled roots follow (160)–(162). Covariance and diffusion positivity are checked without rejected-model redraws. Before test observations, 32,768 independently seeded scrambled Sobol points evaluated the local prediction of Proposition 28. Its thresholds are $4878, 3914, 3586, 3419$ for the four graph sizes.

Sixteen cells supply $\eta = \epsilon$. The other 32 draw $\eta \sim U[\epsilon, 2\epsilon]$ independently of rates and provide $M = \lceil c/\epsilon^2 \rceil$ independent simultaneous source–target calibration pairs, with $c = 16$ or 4096. Both decoders use the same $\widehat{\eta} = \Pi_{[\epsilon,2\epsilon]}\widehat{\text{Cov}}(R_\star, L_0)$: the center returns $\widehat{\eta}G_1(1,2)$; the lag decoder substitutes $\widehat{\eta}$ in both its center correction and final response, projecting the target rate onto $[1.66, 2.34]$.

Calibration uses its exact two-dimensional Wishart law; temporal means use the chi-square law above. The center receives M pairs and lag receives $M + K$, costing $2M$ and $2M + 2K$ scalar readings. Oracle-amplitude versions on these same unknown-amplitude systems are separate diagnostics, never replacements for estimated-amplitude estimates.

All 48 rows appear in Tables 2–4. Loss is $|\widehat{g} - g|/\epsilon$. One Bonferroni family contains 144 two-sided paired Student-t contrasts: one gain in each known-amplitude cell and four in each unknown-amplitude cell (estimated gain, oracle gain, and the center and lag calibration penalties). Directional claims use these approximate simultaneous 95% intervals, not the additional unadjusted bootstrap intervals.

All sixteen known-amplitude directions are resolved: the center wins at $K = 512, 2048$ and lag inversion at $8192, 32768$. Each observed transition bracket $(2048, 8192)$ contains the frozen prediction. With $c = 4096$, the two smaller budgets still favor the center; at 8192 all mean gains are positive but two intervals remain unresolved; at 32768 all four favor lag. With $c = 16$, five of sixteen estimated gains are resolved and eleven unresolved. At $K = 32768$, lag error is .11181–.11893, versus oracle error .00224–.00234: amplitude calibration, not timing precision, dominates.

m	K	M	Center	Target lag	Gap [simultaneous 95% CI]
3	512	0	0.003909	0.010117	-0.006209 $[-0.007052, -0.005366]$
3	2,048	0	0.003916	0.005703	-0.001787 $[-0.002348, -0.001227]$
3	8,192	0	0.003963	0.003011	0.000952 $[0.000553, 0.001350]$
3	32,768	0	0.003841	0.001483	0.002358 $[0.002034, 0.002683]$
5	512	0	0.004306	0.010675	-0.006369 $[-0.007226, -0.005511]$
5	2,048	0	0.004412	0.006131	-0.001718 $[-0.002322, -0.001115]$
5	8,192	0	0.004449	0.002984	0.001465 $[0.001048, 0.001882]$
5	32,768	0	0.004195	0.001460	0.002735 $[0.002395, 0.003075]$
7	512	0	0.004499	0.010261	-0.005762 $[-0.006586, -0.004937]$
7	2,048	0	0.004551	0.005854	-0.001303 $[-0.001907, -0.000699]$
7	8,192	0	0.004513	0.003043	0.001471 $[0.001046, 0.001895]$
7	32,768	0	0.004507	0.001569	0.002937 $[0.002589, 0.003286]$
9	512	0	0.004793	0.010332	-0.005539 $[-0.006371, -0.004707]$
9	2,048	0	0.004634	0.005937	-0.001303 $[-0.001898, -0.000707]$
9	8,192	0	0.004771	0.003084	0.001687 $[0.001252, 0.002122]$
9	32,768	0	0.004665	0.001542	0.003123 $[0.002773, 0.003473]$

Table 2: All sixteen known-amplitude cells, $\eta = \epsilon$, $M = 0$. Mean loss is $|\widehat{g} - g|/\epsilon$ over 1,000 systems per row. Gap is center minus lag; intervals use the common 144-contrast Bonferroni family. The center uses no timing records; lag uses $2K$ scalar readings.

m	K	M	Center	Target lag	Gap [simultaneous 95% CI]
3	512	2,304	0.111486	0.112112	-0.000626 $[-0.002820, 0.001569]$
3	2,048	2,304	0.118905	0.118080	0.000825 $[-0.000651, 0.002301]$
3	8,192	2,304	0.116876	0.115519	0.001358 $[0.000329, 0.002386]$
3	32,768	2,304	0.116139	0.114076	0.002063 $[0.001198, 0.002928]$
5	512	6,400	0.116955	0.115749	0.001207 $[-0.000959, 0.003372]$
5	2,048	6,400	0.115892	0.114525	0.001368 $[-0.000130, 0.002865]$
5	8,192	6,400	0.121937	0.120600	0.001337 $[0.000213, 0.002462]$
5	32,768	6,400	0.120373	0.118933	0.001440 $[0.000465, 0.002415]$
7	512	12,544	0.115588	0.115981	-0.000393 $[-0.002591, 0.001804]$
7	2,048	12,544	0.116771	0.115614	0.001157 $[-0.000414, 0.002727]$
7	8,192	12,544	0.120839	0.119508	0.001331 $[0.000203, 0.002460]$
7	32,768	12,544	0.112396	0.111810	0.000585 $[-0.000408, 0.001579]$
9	512	20,736	0.115824	0.116850	-0.001026 $[-0.003193, 0.001141]$
9	2,048	20,736	0.117216	0.115752	0.001464 $[-0.000067, 0.002994]$
9	8,192	20,736	0.115988	0.115471	0.000517 $[-0.000579, 0.001613]$
9	32,768	20,736	0.113080	0.112216	0.000864 $[-0.000127, 0.001855]$

Table 3: All sixteen unknown-amplitude cells at calibration scale $c = 16$. Both decoders use the same estimated amplitude. $M = \lceil c/\epsilon^2 \rceil$; center and lag costs are $2M$ and $2M + 2K$ scalar readings. Loss, gap, system count, and simultaneous intervals follow Table 2.

m	K	M	Center	Target lag	Gap [simultaneous 95% CI]
3	512	589,824	0.010685	0.017937	-0.007252 $[-0.008790, -0.005715]$
3	2,048	589,824	0.009948	0.012045	-0.002097 $[-0.003212, -0.000982]$
3	8,192	589,824	0.010803	0.009934	0.000870 $[-0.000003, 0.001743]$
3	32,768	589,824	0.010406	0.009134	0.001272 $[0.000512, 0.002032]$
5	512	1,638,400	0.010371	0.017170	-0.006799 $[-0.008369, -0.005229]$
5	2,048	1,638,400	0.010955	0.012626	-0.001671 $[-0.002854, -0.000489]$
5	8,192	1,638,400	0.010725	0.009998	0.000727 $[-0.000179, 0.001633]$
5	32,768	1,638,400	0.011017	0.008873	0.002144 $[0.001371, 0.002916]$
7	512	3,211,264	0.010768	0.017721	-0.006953 $[-0.008570, -0.005337]$
7	2,048	3,211,264	0.010730	0.012536	-0.001805 $[-0.003020, -0.000590]$
7	8,192	3,211,264	0.011211	0.009963	0.001248 $[0.000296, 0.002200]$
7	32,768	3,211,264	0.011006	0.008757	0.002248 $[0.001448, 0.003048]$
9	512	5,308,416	0.010724	0.017501	-0.006777 $[-0.008394, -0.005160]$
9	2,048	5,308,416	0.011341	0.012683	-0.001342 $[-0.002543, -0.000141]$
9	8,192	5,308,416	0.010930	0.009831	0.001100 $[0.000162, 0.002037]$
9	32,768	5,308,416	0.011101	0.009047	0.002054 $[0.001229, 0.002878]$

Table 4: All sixteen unknown-amplitude cells at $c = 4096$, with the same shared-calibration protocol, loss, and 144-contrast interval family. These large M values are calibration records, separate from temporal budget K.

D.3 Star acquisition without a supplied spectrum

For the following star comparisons, draw $a \sim U[.75, 1.5]$, $d_1 \sim U[1.3, 3]$, $d_2 \sim U[1.8, 3.8]$, and $d_3 = 10 - a - d_1 - d_2$; independently draw $s, v_i \sim U[.8, 1.2]$ and $u_i \sim U[1, 2]$, with $x_i = \epsilon u_i$. Diffusion positivity is checked, rejecting invalid generating systems at the largest coupling. Fits and clipped inversions use physical rate bounds $[.5, 6.5]$, variance bounds $[.5, 2]$, and amplitude bounds $[1, 2]$; likelihood fits also enforce positive diffusion. For $\epsilon > .01$, positivity is checked individually beyond the theorem's uniform small-coupling range.

The static arm gets exact $T = 10$ and N_A four-vectors. The targeted arm, without trace or spectrum, gets $N = M$ four-vectors, M oriented $(R_0, L_{1,1})$ pairs, and $M_0 = 512$ root pairs. Set M to the next even integer at least C/ϵ^2 and $N_A = N + (M + M_0)/2$: both arms use $4N_A = 4N + 2M + 2M_0$ endpoints. The frozen moment grid has 300 fresh systems in each $\epsilon \in \{.1, .05, .025, .01\}$, $C \in \{4, 16, 64, 256\}$ cell. Targeted error is lower in all sixteen cells; twelve paired intervals are positive (200,000 bootstrap resamples, Bonferroni family sixteen). All four $C = 4$ cells are unresolved.

A separate likelihood batch has 100 fresh systems for each $\epsilon = .1, .05, .01$, with $C = 64$. Static fitting enforces the trace in ten parameters; targeted fitting has eleven parameters and omits it. Thus feasible sets differ as specified, although scalar budgets match. Eight SLSQP starts with a 1,000-iteration cap select the lowest objective among converged, feasible, physically repaired fits. All 300 selected fits per arm qualify. Table 5 retains both moment and likelihood comparators. Targeted moments beat static moments by $.1045, .0758, .0693$, with familywise intervals $[.0697, .1417]$, $[.0346, .1223]$, $[.0357, .1093]$; against static likelihood the reductions are $.1076, .1388, .2203$, also resolved. The family covers all twelve cross-arm decoder comparisons, not only those favoring the proposed estimator.

ϵ	static moment	static likelihood	targeted moment	targeted likelihood
.100	.1694	.1724	.0648	.0626
.050	.1479	.2108	.0721	.0684
.010	.1261	.2771	.0568	.0553

Table 5: Mean relative response errors on 100 fresh systems per coupling; static alone receives exact trace. Both static comparators are retained, including the more accurate moment decoder.

D.3.1 Static–targeted optimization diagnostic

All 300 original cases reconstruct within 10^{-10}, including truth, moments, and selected likelihood scores; all 4,800 original starts converge in class. Latent-system seeds are shared across coupling values, so cells remain separate. The first sixteen indices per cell replay all eight starts successfully. Sixteen stronger starts then use tolerance 10^{-11} and the same 1,000-iteration cap, parameterization, repair, constraints, and convergence eligibility. Original candidates remain eligible; selection uses same-arm likelihood only, never truth or response error. Multiply normalized objectives by N_A (static) or $N + M + M_0$ (targeted) before computing total negative-log-likelihood (NLL) gaps.

	static			targeted		
ϵ	moment	original MLE	stronger MLE	moment	original MLE	stronger MLE
.100	.2070	.2385	.2511	.0638	.0639	.0637
.050	.1646	.2391	.2445	.0865	.0792	.0790
.010	.1492	.3094	.2945	.0757	.0734	.0747

Table 6: Fixed sixteen-system-per-cell static–targeted diagnostic. Mean relative errors before and after stronger fitting, selected by likelihood only. The primary 100-system result remains unchanged.

Stronger fitting preserves the targeted advantage over both static comparators in all three subset cells; static moments remain better than stronger static likelihood in all three. Maximum per-case NLL improvements are 1.288 (static) and .323 (targeted). Across all 100 original systems per coupling, mean found-candidate response ranges divided by ϵ, at total NLL gap $\leq .5$, are $.0701, .1713, .4543$ versus $.0027, .0071, .0459$. At gap ≤ 2, they are $.0841, .1940, .5177$ versus $.0046, .0084, .0621$. These descriptive attained-fit spreads are not confidence intervals. Selected-fit boundary counts are 78,69,59 versus 78,66,61, using log-box distance at most 10^{-4}.

D.4 Which temporal records carry the gain?

D.4.1 Full-pair acquisition package

A fresh 4,800-system grid repeats the sixteen (ϵ, C) cells and the static/restricted budgets above. A third arm instead receives N_A full records $(R_0, L_{1,0}, R_1, L_{1,1})$, without trace or additional snapshots, again costing $4N_A$ endpoints. It estimates s, v, x from both slices, the root rate from C_R/s, and the target rate from its self-lag, then returns $(x/s)G_1(a, d_1)$. Full-pair mean relative error is lower in all sixteen cells; fourteen restricted-minus-full intervals are positive (200,000 resamples, Bonferroni sixteen). The unresolved cells are $(.025, 4)$ and $(.01, 4)$, with intervals $[-.0262, .0429]$ and $[-.0177, .0453]$. This comparison changes the entire moment allocation. The next comparison isolates the target self-lag on identical records.

D.4.2 Same-record self-lag ablation

For each of twelve (ϵ, K) cells, with the same four couplings and $K = 16, 64, 256$, 300 new systems supply $N = \lceil 16/\epsilon^2 \rceil$ static four-vectors and K full pairs. Both decoders use exactly the same records, static moments, root estimate, and clipping; only target-rate inversion changes from the oriented cross-lag to the target self-lag. Neither receives trace or spectrum. The grid and a diagnostic repeating it with the true root rate form one Bonferroni family of 24 contrasts (50,000 paired bootstrap resamples). The primary self-lag decoder lowers the mean in eleven cells, with seven resolved gains; the oracle-root version gives twelve lower means and eight resolved gains. All $K = 16$ primary intervals contain zero. At $(.1, 16)$ the primary errors are $.2656$ versus $.2726$, difference $-.0069$ with interval $[-.0537, .0391]$: self-lag inversion therefore has no resolved gain in this cell.

D.4.3 Matched sensors and likelihood

Both arms receive the same $N = \lceil 16/\epsilon^2 \rceil$ static records. Restricted sensing adds K oriented pairs and K root pairs; full sensing adds K full pairs. Each costs $4K$ temporal readings, with batches independent conditional on the same system; neither gets trace or spectrum. Moment decoding uses 300 systems per cell (3,600 total). Likelihood uses the first 200 per cell (2,400 total), the same eleven-parameter box and positive-diffusion constraint for both arms, four SLSQP starts, and 250 iterations

per start. Only the observation likelihood changes; all selected original fits converge. Each decoder has its own twelve-comparison Bonferroni family with 100,000 paired bootstrap resamples. Table 7 retains every cell for both decoders.

		moment, 300 systems/cell		likelihood, 200 systems/cell	
ϵ	K	restricted / full	difference interval	restricted / full	difference interval
.100	16	.2505 / .2417	[-.0425,.0599]	.2167 / .2036	[-.0420,.0681]
.100	64	.2148 / .2142	[-.0396,.0401]	.1934 / .1920	[-.0455,.0473]
.100	256	.1736 / .1403	[.0076,.0593]	.1750 / .1326	[.0108,.0747]
.050	16	.2827 / .2712	[-.0466,.0692]	.2619 / .2580	[-.0637,.0710]
.050	64	.2480 / .1900	[.0220,.0939]	.2314 / .1775	[.0116,.0961]
.050	256	.2073 / .1300	[.0517,.1035]	.1919 / .1180	[.0444,.1048]
.025	16	.2744 / .2756	[-.0556,.0531]	.2605 / .2705	[-.0691,.0475]
.025	64	.2633 / .1932	[.0316,.1086]	.2513 / .1792	[.0273,.1194]
.025	256	.2379 / .1470	[.0599,.1214]	.2312 / .1503	[.0432,.1180]
.010	16	.2866 / .2730	[-.0384,.0666]	.2642 / .2636	[-.0543,.0554]
.010	64	.2543 / .1993	[.0148,.0946]	.2270 / .2002	[-.0213,.0731]
.010	256	.2339 / .1429	[.0642,.1175]	.2141 / .1464	[.0339,.1000]

Table 7: All matched-sensor cells. Mean relative errors and familywise 95% intervals for restricted minus full error. The two decoders use separate twelve-contrast bootstrap families.

Both decoders have eleven lower full-pair means, but only seven moment and six likelihood gains are resolved; none at $K = 16$ is resolved. Post-hoc replacement of relative loss by $|\widehat{g} - g|/\epsilon$ uses the same saved predictions, with a combined 24-contrast family and 100,000 resamples. Moment gains are positive in twelve cells (seven resolved), likelihood gains in ten (six resolved). The negative likelihood differences at $(.025, 16)$ and $(.01, 16)$ are $-.00601$ and $-.00280$, with intervals $[-.06905, .05768]$ and $[-.06424, .05908]$.

D.4.4 Likelihood geometry and operating regime

All 3,600 moment and 2,400 likelihood rows reconstruct within 10^{-10}. On the fixed first 24 systems per cell, all 288 four-start replays match responses, objectives, and flags. Twelve additional starts use 1,000 iterations and tolerance 10^{-11} in the same box; all four original candidates remain eligible under likelihood-only selection. Total NLL divisors are $N + K$ for full and $N + 2K$ for restricted sensing.

Full pairs have lower mean response error in the same ten subset cells before and after stronger fitting. The two exceptions remain $(.1, 64)$ and $(.025, 16)$; better likelihood can worsen response loss. Mean near-best response spread is smaller for full pairs in all twelve cells at gaps .5 and 2. At $(.01, 256)$, restricted/full spreads divided by ϵ are $.2878/.0913$ and $.3077/.1044$, respectively. Among 4,608 retained candidates per arm, any-bound counts are 4,086 restricted and 4,011 full; rate-bound counts are 3,414 and 3,240 (normalized log-distance 10^{-4}). All 19,200 original starts and 4,800 selected original fits converge.

All original rows are split within cell at fixed generator midpoints $a = 1.125$, $d_1 = 2.15$, and population target correlation $\rho_L = e^{-2.15}$, with coupling included in ρ_L. One Bonferroni family of 144 paired t intervals covers two methods, twelve cells, three stratifiers, and two groups; group sizes are 119–181 (moment) and 76–124 (likelihood), with none omitted. High/low target-correlation groups give 6/3 resolved moment gains and 5/1 likelihood gains out of twelve cells each; no adjusted interval in the family is strictly negative. At $(.01, 256)$, high-correlation moment errors are $.2240/.1192$ (difference interval $[.0589, .1508]$); low-correlation errors are $.2437/.1665$ ($[.0273, .1271]$). Target-inversion boundary counts are 134/2 and 138/16 out of 150 per group. The stronger target-persistence stratum has more resolved gains, while low-budget comparisons remain uncertain.

D.5 Lag choice, noisy anchors, and dependent observations

For the existing lag study, evaluate all 3,600 generating systems (300 per slow/fast-root, noise-$0/.05$, $K = 64, 256, 1024$ condition). At $h \in \{1, .5, 1/6.5\}$, define $J_h = D_{(a,d_1)}(\Gamma_{RR}(h), \Gamma_{L_1L_1}(h))$ and let Ω_h be the covariance of the two products from the same pair, including their cross-covariance.

With stationary moments fixed, $\mathcal{V}_h = \nabla g^\top J_h^{-1} \Omega_h J_h^{-\top} \nabla g$; the local relative absolute-error approximation is $\sqrt{2\mathcal{V}_h/(\pi K)}/g$. The derivation belongs to Appendix B.9, Proposition 25, not this empirical protocol. Analytic derivatives agree with centered differences within 2.33×10^{-11}.

Root	noise	K	local choice	error $h = 1$	error $h = .5$	error $h = 1/6.5$
slow	0	64	1	.1254	.1066	.1678
slow	.05	64	1	.1255	.1067	.1880
slow	0	256	1	.0629	.0630	.1140
slow	.05	256	1	.0703	.0669	.1084
slow	0	1024	1	.0432	.0443	.0727
slow	.05	1024	1	.0436	.0453	.0729
fast	0	64	.5	.2083	.2089	.3500
fast	.05	64	.5	.2148	.2191	.3341
fast	0	256	.5	.1441	.1283	.2476
fast	.05	256	.5	.1492	.1414	.2648
fast	0	1024	.5	.0962	.0741	.1612
fast	.05	1024	.5	.0983	.0807	.1596

Table 8: All twelve existing lag conditions, 300 systems per row. Entries are mean relative response errors; noise is known endpoint-noise variance. The local choice minimizes the mean conditional approximation at generating parameters, not a fitted-data selection rule.

The local choice matches seven of twelve empirical minima, including all four at $K = 1024$. For fast roots, no noise, and $K = 1024$, local errors at $h = .5, 1/6.5$ are $.0738, .1619$, versus observed $.0741, .1612$. At $h = 1$, however, $.3775$ substantially exceeds the observed $.0962$: weak root covariance defeats the local linearization, whereas the actual decoder estimates static moments and clips the inverse. Shorter lags can improve root-rate estimation while worsening the response; the conditional calculation does not fully predict finite-budget performance.

Noisy anchors. Eight cells cross the four star couplings with $N = \lceil C/\epsilon^4 \rceil$, $C = 1, 16$, and 200 systems each. Time rescaling by $U(.95, 1.05)$ varies the trace while preserving covariance; models outside the rate box are redrawn before observation. All arms share static data and clipped ratio inversion, adding either nested $J = 16, 256$ readings $T + \mathcal{N}(0, 1)$ or nested $M_0 = 16, 256$ root pairs at lag one; exact trace is an oracle only. All eight root-pair improvements are resolved but no noisy-trace improvement is (10,000 bootstrap resamples, Bonferroni sixteen). At $(.01, 16)$, gains are .1183 with $[.0603, .1817]$ for root pairs, versus .00081 with $[-.00175, .00345]$ for trace readings. The two channels have different acquisition costs.

A separate frozen trace-limited diagnostic uses 300 new systems per $\epsilon = .1, .05$, $N = \lceil 4096/\epsilon^4 \rceil$, trace in $[9, 11]$, and reading standard deviation four. Errors for $J = 4$, $J = 256$, exact trace, and center $T = 10$ are respectively $.04234, .01269, .00816, .02651$ at .1 and $.04180, .01556, .01060, .02920$ at .05. The two adjusted improvement intervals are $[.02499, .03457]$ and $[.02183, .03081]$. Four noisy readings are worse than the center in both cells; nominal snapshot budgets are 4.096×10^7 and 6.5536×10^8.

Trajectories. A separate batch uses 120 fresh systems in each of eight strata: slow/fast roots, $\Delta = 1, .25$, and unknown sensor variance $0, .05$, with $\epsilon = .02$. Each stationary trajectory has 32,768 observations and nested prefixes of 2,048 and 8,192. Six Bonferroni-adjusted intervals for error at 2,048 minus 32,768 exclude zero. For fast roots at $\Delta = 1$, intervals $[-.054, .518]$ and $[-.181, .364]$ are unresolved; at $\Delta = .25$, they are $[.121, .298]$ and $[.169, .359]$. Equal observation counts cover different physical durations.

D.6 Structural checks and additional comparators

Graph identification and measurement checks. Population spectral reconstruction uses twenty systems per graph/stress condition: $(r, m) = (1, 3), (2, 3), (2, 4), (4, 4)$ have one compatible assignment, whereas the two-leaf star has two; reconstructed covariances agree with a Lyapunov solver to floating-point precision. All 4,000 rank checks (twenty configurations, 200 systems each) match the

predicted rank. Pair-root checks cover 7,000 systems, $m = 4, 5, 6, 7, 8$, seven p choices, and 200 systems per configuration: all match the rank formula, including 258 cases with distinct additional pairs but a missing required odd cycle. A separate 3,500-system test recovers all parameters in its 2,333 identified cases, with maximum error 2.1×10^{-13}.

Root-probe counts match augmented ranks for all 3,000 positive systems on 600 graphs with three to eight leaves. In 100 witnesses consisting of an unselected triangle on leaves 1,2,3 and isolated leaf 4, observing root 0 alone leaves covariance equal within 2.3×10^{-16} and its rate identical. Responses to leaf 1 agree within 5.6×10^{-17}, but those to leaf 4 differ by at least 3.7×10^{-6} (median .001745). A 400-model multiroot audit gives maximum lag, response, and off-diagonal Lyapunov residuals 2.8×10^{-16}, 1.6×10^{-16}, and 2.8×10^{-17}.

The odd-cycle stress grid uses $m = 3, 5, 7, 9$, $\epsilon = .25/m$, 2,000 draws per cell, and 24 graph–budget cells. All covariance and diffusion checks pass; trace and dedicated-pair covariance deviations are at most 2.2×10^{-14} and 1.8×10^{-18}. At $N = m\epsilon^{-4}$, cycle-inversion error .0131–.0139 exceeds center error .0039–.0046. At $N = 16384 m\epsilon^{-4}$ it falls to .00144–.00150 and beat the center in all four graphs (largest nominal $N = 2.48 \times 10^{11}$). Fisher checks have full rank for $m = 3, 5, 7$ and one null direction for $m = 4$. The triangular prism, two triangles, Petersen graph, and K_5 each have 1,000 construction checks: restricted singular values are positive, covariance/diffusion tests pass, and trace deviations are at most 2.2×10^{-14}. VAR(1) checks at $m = 3, 5, 7, 9, 13$ use 1,000 draws each, with minimum stationary eigenvalue above .87, minimum innovation variance above .46, and off-diagonal innovation residual at most 1.4×10^{-17}.

Additional comparator studies. The same-input spectrum-supplied check covers 1,800 systems: three graphs $(1, 2), (1, 3), (2, 3)$, ordinary/weak-near-collision conditions, $N = 128, 512, 2048$, and 100 per cell. Moment inversion and constrained Gaussian MLE share data, graph, and exact unordered spectrum. Sixteen of eighteen unadjusted paired intervals contain zero and one favors each method; at weak $(2, 3), N = 2048$, both recover the named assignment in 0/100 cases. On the same 600 three-leaf samples, a post-pilot trace-only decoder has error .03310 at ordinary $N = 2048$, versus .01792 for spectrum-supplied MLE, an unequal-information comparison.

Separate four-leaf tests span strong/ordinary/weak coupling and $(N, M) = (8192, 256), (32768, 1024)$. Without covariance, trace, or spectrum oracles, joint likelihood improves on moments in strong and ordinary regimes; weak-regime differences remain unresolved. The moment estimate has negative implied diffusion in 14/100 strong, smaller-budget systems. At matched temporal endpoints, source–target versus full-vector MLE resolves only one of six differences; source–target optimization is slower. A trace-aware comparison at $N = 128, 512, 2048, 8192$ resolves only $N = 512$ in favor of likelihood. Conversely, the spectrum-supplied acquisition grid ($\epsilon = .1, .05, .025$, two budgets, 300 systems/cell) resolves all six targeted-likelihood gains, while retaining the two nonconverged paired cases among 1,800. These comparisons use separate sampling protocols and are evaluated within their respective batches.