TEMPOGEOM: OBSERVATION DESIGN FOR LEARNING SWITCHING SPEEDS IN TEMPORAL GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Predicting whether an edge is present does not determine how rapidly it forms and disappears. We study observation design for jointly learning switching clocks and unknown, finite-dimensional time-varying affinities from discretely observed temporal graphs. Affinity perturbations can vanish at observation times while changing between-observation dynamics, confounding the small clock-dependent lag under fast switching. TEMPOGEOM uses a two-scale efficient-Fisher analysis to retain these hidden nuisance effects. In an explicit shared-affinity family, reordering a fixed multiset of observation gaps changes local log-clock information from $\Theta(\lambda^{-4})$ to $\Theta(\lambda^{-2})$, where λ scales switching rates, without changing observation count or horizon. We characterize how the observation graph determines the number of clock directions at each rate. Because both rates vanish, we introduce separated affinity anchors followed by probes on the switching timescale. Under quantitative affinity-conditioning, graph-rank, and uniform regularity assumptions, this design preserves a positive information floor over bounded relative clock uncertainty; a fixed positive minimum gap precludes such a floor for arbitrarily fast clocks. For finite clocks, we combine nuisance-aware receding-horizon design with centered graph regularization. Matched-estimator comparisons in polynomial and Fourier synthetic studies yield clock-RMSE reductions of 8.3% and 11.5% at the slowest scale, respectively, with no consistent advantage at faster scales. Broader pipeline gains also reflect regularization, whose benefits can reverse under increased heterogeneity. Together, these results separate the roles of observation timing, nuisance conditioning, and regularization in learning switching speeds.

1 INTRODUCTION

Two node pairs can each be connected half the time even if one switches once an hour and the other once a second. Predicting edge presence therefore differs from learning switching speed. We separate *affinity*, which determines the instantaneous equilibrium preference for an edge, from a *clock* that scales its formation and dissolution rates together. We observe selected node pairs at known times; intermediate switches remain hidden. Inference must fit clocks jointly with the coefficients of finite-dimensional affinity trajectories.

Node-level clock parameters are shared across incident pairs. The *observation graph* records which pairs are measured, including observations of absent edges. We ask: *How do observation timing and observation-graph structure determine the information available about switching speeds when affinities are unknown?* We quantify this using *efficient Fisher information*: the local clock information remaining after accounting for fitted affinity uncertainty.

Unknown affinities create a subtle confounding mechanism. Two trajectories can agree at every observation time yet differ between observations. Under fast switching, edge probabilities approach their instantaneous equilibrium, leaving a small clock-dependent lag that hidden affinity changes can mimic. TEMPOGEOM retains both endpoint and lag effects before accounting for affinity uncertainty. Taking the leading fast-clock limit first can otherwise discard nuisance directions that absorb the clock signal (Theorem 1).

Prior work models edge durations through survival representations (Çelikkanat et al., 2024) and uses ℓ_1 latent distances to obtain exact cumulative hazards (Lai et al., 2026). Observation timing

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.

for parameter learning is also established: Fisher design accommodates nuisance parameters (Kyriakopoulos and Wolf, 2015), and Bayesian designs adapt measurements to learn Markov transition rates (Barendregt et al., 2023). Recent adaptive design refines learned policies online (Hedman et al., 2025), optimizes policies with nested particle methods (Iqbal et al., 2024), and combines contrastive acquisition with tempered inference (Iollo et al., 2024). Our theoretical focus is endpoint-invisible affinity variation at the clock-signal scale and its interaction with timing and graph structure.

Our results distinguish slowing information decay from preserving a nonvanishing information floor. Reordering the same gaps improves the decay rate but does not prevent information loss. To obtain a positive floor, we pair separated *anchor observations* that constrain affinities with follow-up *probes* on the switching timescale. Observation count and horizon remain fixed, but probe gaps shrink as clocks accelerate; a fixed positive minimum gap precludes this guarantee.

For finite clocks, we instantiate this timing principle with nuisance-coupled local covariance-trace planning and centered graph regularization. Covariance-trace design (Alexanderian et al., 2016) and ridge shrinkage are established ingredients; our construction combines them with finite-clock transition derivatives and feasible anchor/probe replanning. Matched-estimator comparisons separate acquisition from estimation gains, while fixed-schedule controls test the value of optimization beyond timing structure (Sections 4.3–5).

Our contributions are:

- **Hidden-affinity information analysis.** We derive a two-scale efficient-Fisher expansion that retains affinity directions invisible at observation endpoints and identifies when they absorb the leading clock signal (Theorem 1).
- **Sharp timing and graph rates.** In an explicit shared-affinity family, we prove matching local $\Theta(\lambda^{-4})$ and $\Theta(\lambda^{-2})$ information rates under the same gap multiset, where λ scales switching rates, and determine how observation-graph structure controls the number of clock directions at each rate (Corollary 2; Theorem 4).
- **Conditional design and finite-clock method.** We prove anchor-and-probe information floors under quantitative assumptions and identify the fixed-resolution obstruction. We instantiate the design principle with a nuisance-coupled finite-clock planner and centered regularization, separating their effects through matched-estimator experiments (Theorem 5; Proposition 7; Sections 4.3–5).

2 Model, observations, and scope

For an observed dyad $e = (i, j)$, let $S_e(t) \in \{0, 1\}$ have rates

$$a_e = e^{c_e+\eta_e/2}, \quad b_e = e^{c_e-\eta_e/2}, \quad p_e = \mathrm{sigmoid}(\eta_e), \quad q_e = a_e + b_e = 2e^{c_e}\cosh(\eta_e/2). \quad (1)$$

Affinity η determines the frozen-rate equilibrium p; c is a log-geometric-mean switching clock. Conditional on deterministic trajectories, dyads are independent nonexplosive chains. Each observed transition contributes $\log P(u, t)_{S(u),S(t)}$, integrating out hidden switches and conditioning on the initial state. Missing observations create longer gaps, not disconnected states. An observation-graph edge records that a dyad is measured, not that it is connected; absent-edge observations are included. Its unsigned incidence matrix $B \in \mathbb{R}^{m\times n}$ has row $e_i + e_j$. We parameterize dyad log clocks by $c = \log\lambda + Bv$, so each dyad combines its two node-level parameters, and measure information for v, with λ a fixed regime multiplier.

For edge-presence probability y, the exact transition equation is

$$y' = q(p - y), \qquad y(t) = e^{-H}y(u) + \int_u^t e^{-\int_s^t q}q(s)p(s)ds, \quad H = \int_u^t q. \quad (2)$$

The information results concern finite-dimensional affinity families in (1). The phase construction below and the polynomial/Fourier experiments have different nuisance spaces. No identification result transfers between them without checking its assumptions.

3 INFORMATION PRESERVED BY OBSERVATION DESIGN

3.1 ACCOUNTING FOR HIDDEN AFFINITY CHANGES

Let $q_{e,\lambda} = \lambda q_{e,0}$, where $q_{e,0} = \exp(X_e^\top v) h_e(t;\theta) > 0$ and $p_e = p_e(t;\theta)$. Taking $h_e = 2\cosh(\eta_e/2)$ recovers (1). At record endpoints define

$$\begin{aligned} A_r &= p_e'(t_r^-)/q_{e,0}(t_r), \quad J_0 = \partial_\theta p, \quad J_1 = -\partial_\theta A, \\ G_r &= A_r X_e(t_r)^\top, \qquad W = \operatorname{diag}\{[p_r(1-p_r)]^{-1}\}. \end{aligned} \tag{3}$$

Assume fixed positive gaps, q_0 bounded below, endpoint p bounded away from zero and one, and bounded endpoint-neighborhood time derivatives of p through order three and q_0 through order two, together with their first parameter derivatives. Stable left facets suffice at piecewise-smooth trajectory points. Uniformly over starting states, with $\varepsilon = \lambda^{-1}$,

$$y = p - \varepsilon A + O(\varepsilon^2), \quad \partial_v y = \varepsilon G + O(\varepsilon^2), \quad \partial_\theta y = J_0 + \varepsilon J_1 + O(\varepsilon^2). \tag{4}$$

Theorem 1 (Two-scale nuisance information)**.** *Let K_0 span $\ker J_0$ and choose L_0 so that $[L_0, K_0]$ is invertible and $J_0 L_0$ has full column rank. Suppose*

$$U = [W^{1/2} J_0 L_0,\ W^{1/2} J_1 K_0] \tag{5}$$

has full column rank. With Π_U its orthogonal projector and $\mathcal{E}_* = (I - \Pi_U) W^{1/2} G$,

$$I_{v,\mathrm{eff}} = \lambda^{-2} \mathcal{E}_*^\top \mathcal{E}_* + O(\lambda^{-3}). \tag{6}$$

Directions with $\mathcal{E}_* h \neq 0$ have $\Theta(\lambda^{-2})$ information; directions in its kernel have $O(\lambda^{-4})$ information.

Rescaling nuisance columns by $[L_0, \varepsilon^{-1} K_0]$ preserves their span and gives the nonsingular limit (5). This proves the projector limit without inverting a singular limiting information block. An endpoint-invisible parameter is therefore not a removable gauge: its lag derivative can absorb a clock score. If J_0 is full rank, the result reduces to ordinary endpoint projection. If U is rank deficient, higher scales can matter. A smooth interior-bump construction in Appendix C even defeats every endpoint derivative; it is a pointwise local Fisher statement, not a uniform minimax impossibility result. Full proofs are in Appendices B and K.

Spatial identifiability is distinct. Under homogeneous rates, efficient clock information is bounded below by $B^\top \operatorname{diag}(\widetilde{w}_e) B$ with positive per-dyad weights (Theorem 14, Appendix J). Full column rank holds exactly when every component is nonbipartite. An exact $Bh = 0$ gauge survives every temporal intervention, while positive eigenvalues may still be arbitrarily small.

3.2 EQUAL GAP BUDGETS, DIFFERENT INFORMATION RATES

Define

$$f(t) = -0.3 + 0.06(t-3)^2, \qquad \phi(t) = \frac{f'(t)\sin(2\pi t)}{2\pi} + \frac{(t-3)^2 \sin^2(2\pi t)}{8\pi^2}, \tag{7}$$

and fit all three node vectors in the shared-affinity family

$$c_e = \log\lambda + (Bv)_e, \qquad \eta_e(t) = f(t) + (B\alpha)_e + \phi(t)(B\gamma)_e. \tag{8}$$

Compare $T_A = (0.5, 1.5, 2.5, 3.5, 4.5, 5.5)$ with $T_S = (1, 2, 3, 3.5, 4.5, 5.5)$, starting at zero. Both have six observations, horizon 5.5, and gap multiset $\{0.5, 1, 1, 1, 1, 1\}$. At these endpoints $\phi = 0$, but $\phi' = f' \cos(2\pi t)$ changes phase.

Corollary 2 (Equal-gap-multiset separation)**.** *For full-column-rank B, at $v = \alpha = \gamma = 0$,*

$$I^A_{v,\mathrm{eff}} = \kappa_A \lambda^{-4} B^\top B + O(\lambda^{-5}), \qquad I^S_{v,\mathrm{eff}} = \kappa_S \lambda^{-2} B^\top B + O(\lambda^{-3}), \tag{9}$$

where $\kappa_A \approx 0.577926$ and $\kappa_S \approx 0.0112345$ are strictly positive.

Put $q_0 = 2\cosh(f/2)$, $j = \operatorname{sigmoid}(f)(1 - \operatorname{sigmoid}(f))$, $d = t - 3$, $b = 0.12d$, and

$$u_0 = \sqrt{j}, \quad u_1 = \sqrt{j}\, b/q_0, \quad z_2 = \sqrt{j}\, d^2/q_0^2. \tag{10}$$

Then $\kappa_A = \|(I - \Pi_{[u_0,u_1]}) z_2\|^2$ on T_A, whereas $\kappa_S = \|(I - \Pi_{[u_0, \cos(2\pi t)\odot u_1]}) u_1\|^2$ on T_S. For aligned observations, subtracting the invisible-affinity derivative cancels the first clock term, leaving $-\lambda^{-2} j d^2 B / q_0^2$. Its residual is nonzero because z_2, u_0 are even, u_1 is odd, and d^2/q_0^2 is not constant. Staggered endpoints impose incompatible signs on any first-order cancellation. Lemma 10 and Appendix D.3 give the matching lower orders, rather than only an upper bound or a fitted slope.

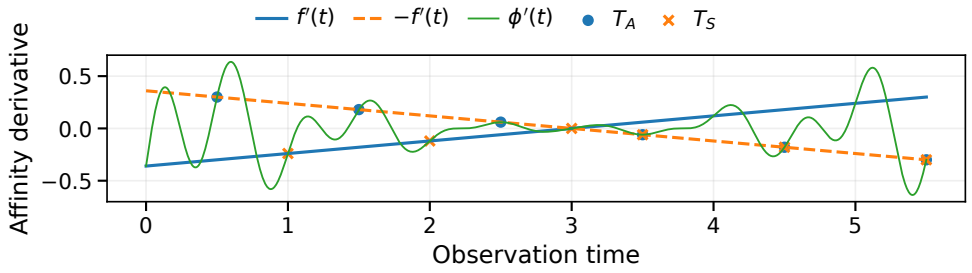

Figure 1: The derivative, not only the endpoint value, determines the lag nuisance. At T_A, $\phi' = -f'$; T_S includes both signs. The basis itself is zero at both schedules. These designs share six observations, horizon 5.5, and the same gap multiset.

3.3 How the observation graph constrains clock information

Nonzero invisible affinity can restore some quadratic directions without changing any endpoint value. Let $D_g = \mathrm{diag}(B\gamma)$, $\Omega = \mathrm{diag}(\omega_e)$, where $\omega_e = \sum_r j_{er} b_r^2 / q_{0,er}^2$, $j_{er} = \mathrm{sigmoid}(f_r + (B\alpha)_e)(1 - \mathrm{sigmoid}(f_r + (B\alpha)_e))$, and $q_{0,er} = 2e^{(Bv)_e}\cosh((f_r + (B\alpha)_e)/2)$.

Proposition 3 (Exact aligned leading matrix)**.** *For full-column-rank B on T_A,*

$$\begin{aligned} I^A_{v,\mathrm{eff}} &= \lambda^{-2} L_A + O(\lambda^{-3}), \\ L_A &= B^\top D_g [\Omega - \Omega B (B^\top \Omega B)^{-1} B^\top \Omega] D_g B. \end{aligned} \tag{11}$$

Its kernel is $\{h : D_g B h \in \mathrm{col}(B)\}$. On K_n, $n \geq 6$, generic γ gives rank $n-1$.

The endpoint-offset score is even over paired endpoints, and the lag scores are odd. Profiling the odd block yields (11). Its weights are positive; hence the following rank restriction is independent of v, α.

Theorem 4 (Observation-graph rank budget and matching spectrum)**.** *Let B have full column rank and let its observation graph have k components. Then*

$$r := \mathrm{rank}\, L_A = \mathrm{rank}[B, D_g B] - n \leq \min\{n - k, m - n\}. \tag{12}$$

For each sufficiently small fixed (v, α, γ), aligned information has r eigenvalues of order λ^{-2} and $n - r$ of order λ^{-4}. Staggered information has n eigenvalues of order λ^{-2} in a neighborhood of the origin. Constants are pointwise in the fixed parameters.

Why the rank gives matching rates. The residual projector has rank $m - n$, and a common clock on each component is in its kernel. For $h \in \ker L_A$, choose the unique z_h satisfying $Bz_h = (I - D_g)Bh$. Subtracting the exact invisible-affinity derivative in direction z_h cancels the leading clock derivative. Its whitened second coefficient is

$$H_{er}(h) = -\frac{\sqrt{j_{er}}}{q_{0,er}^2}[d_r^2 + g_e(d_r^2 - 0.24)](Bh)_e, \qquad g = B\gamma. \tag{13}$$

At the origin its nuisance-projected squared norm equals $\kappa_A \|Bh\|^2$, so injectivity persists in a neighborhood for every leading-kernel direction. The second residual is even in d, whereas the leading clock residuals outside the kernel are odd. Their orthogonality excludes cancellation by rotating eigenvectors: in a kernel/complement basis the diagonal information blocks are respectively $\varepsilon^4 A_0$ and $\varepsilon^2 A_1$ to leading order, with $A_0, A_1 \succ 0$, and the off-diagonal block is $O(\varepsilon^4)$. A Schur complement gives the asserted complete spectrum. Appendix E supplies the uniform local injectivity and remainder argument. The rank is algebraic; the matching rate also requires this second-residual argument. A connected graph needs $m \geq 2n-1$ for $n-1$ quadratic directions. An odd unicyclic graph has $m = n$ and keeps *all* directions quartic locally. Adding observed dyads with unchanged per-dyad exposure cannot decrease unnormalized efficient information; merely adding realized edges is not this intervention.

Timing and nuisance sensitivity. The joint timing–clock limit is nonuniform near endpoint-rank changes. A calibrated $1/\lambda$ shift can recover the known-nuisance leading coefficient in the phase family, but adding one unknown trend reduces its best possible coefficient by over five orders of magnitude. Appendices F and H.1 give both statements and proofs. These examples motivate checking the entire fitted nuisance space rather than relying on a favorable decay exponent alone.

4 Preserving information with anchors and probes

The exact contrast $P_{11}(u,t) - P_{01}(u,t) = e^{-\int_u^t q}$ disappears over fixed positive gaps at fast clocks. To retain it, pair each of L separated macroscopic affinity anchors with a probe x_l/Λ later, with fixed $x_l > 0$. The final pair ends at the fixed horizon. The rates use nominal scale Λ and unknown departures v; a bounded common-scale error is included in v.

Theorem 5 (Anchored-probe information floor)**.** *Let B be injective and let $\eta_e(t;\theta)$ be a fixed finite-dimensional smooth affinity family independent of v. On a compact parameter set assume uniform interior equilibria and bounded positive normalized clocks and an injective stacked anchor Jacobian $\mathcal{H} = (\partial_\theta \eta_e(a_l;\theta))_{el}$, with its smallest singular value bounded below. Under the uniform smoothness conditions in Appendix H.2, the $2L$-observation experiment satisfies*

$$I_{v,\mathrm{eff}}(\Lambda) = I_* + O(\Lambda^{-1}), \qquad I_* \succeq B^\top \operatorname{diag}\left(\sum_l w_{el}\right) B \succ 0. \tag{14}$$

Here, for each frozen anchor/probe pair, F is its conditional Fisher matrix for (c_e, η_e), $j = p(1-p)$, and

$$w = F_{cc} - \frac{F_{c\eta}^2}{F_{\eta\eta} + j} \geq \frac{jF_{cc}}{F_{\eta\eta} + j} > 0. \tag{15}$$

Thus all clock eigenvalues are $\Theta(1)$ uniformly on the stated compact set.

The observable information in one pair. For local $c = (Bv)_e$ at an anchor put $p = \mathrm{sigmoid}(\eta)$, $r = 2xe^c \cosh(\eta/2)$, $z = e^{-r}$, and $y_a = p + (a-p)z$ for anchor state $a \in \{0,1\}$. The clock and affinity derivatives are

$$d_a = (p-a)rz, \qquad e_a = j(1-z) + \tfrac{1}{2}\tanh(\eta/2)d_a, \qquad F = \sum_{a=0}^{1} \frac{\mu_a}{y_a(1-y_a)} \begin{pmatrix} d_a^2 & d_a e_a \\ d_a e_a & e_a^2 \end{pmatrix}, \tag{16}$$

where $\mu_1 = p$ and $\mu_0 = 1-p$. The anchor supplies an additional affinity information j; its density is not differentiated a second time inside F. Positive semidefiniteness gives $F_{c\eta}^2 \leq F_{cc}F_{\eta\eta}$, proving the strictly positive weight. At $\eta = 0$, it reduces to the familiar two-state weight $r^2/(e^{2r}-1)$. Hence neither infinitesimal gaps at fixed clock nor gaps much longer than the local switching time are efficient probes.

Why the nuisance condition matters. Long gaps supply anchor affinity information $j\mathcal{H}^\top\mathcal{H}$ with appropriate dyad weights. Rescaled short gaps converge to a frozen two-state transition with nonzero initial-state contrast. The anchor Jacobian makes the limiting nuisance block nonsingular, excluding hidden rescaled directions. Enlarging that nuisance to separate affinity values per dyad and anchor yields (15); positive weights and injective B give the floor. Appendix H.3 proves the exact limiting matrix and uniform remainder.

For bases $[1, \phi, t-3]$, three anchors at $.75, 2.75, 5.5$ suffice: their temporal design determinant is $-3/(10\pi) - 109/(64\pi^2) \neq 0$. No affinity coefficient is supplied to inference. For more bases, the anchor-rank condition must be checked again. Arbitrary smooth nuisance functions, unbounded unknown clock ranges, and corrupted timestamps are outside this guarantee. The temporal-resolution boundary is formalized in Proposition 7.

4.1 Conditioning, temporal resolution, and statistical exposure

The anchor condition is quantitative, not merely a rank test. Write the whitened clock and nuisance score matrices as C_Λ, N_Λ, with limits C_*, N_*, in fixed normalized parameter coordinates. The following bound connects a well-conditioned anchor design to a finite-clock guarantee.

Algorithm 1 Conditioning-aware finite-pool acquisition

Require: Observation graph B, fitted affinity basis, weighted scenarios $\mathcal{P}$, candidate pool $\mathcal{D}$, threshold τ, objective J.

1: Check graph rank, fitted-nuisance coverage, horizon, count, and gap feasibility.
2: Form $\mathcal{D}_\tau = \{D \in \mathcal{D} : \sigma_{\min}(\Phi(t_1, t_3, t_5)) \geq \tau\}$; report failure if empty.
3: Evaluate the same declared J on every $D \in \mathcal{D}_\tau$.
4: Select $\widehat{D} \in \arg\max_{D \in \mathcal{D}_\tau} J(D)$; retain scores and termination metadata.
5: Report conditioning and, when available, separate finite-clock transfer diagnostics.

Proposition 6 (A finite-clock stability margin)**.** *Suppose* $\|C_\Lambda - C_*\|, \|N_\Lambda - N_*\| \leq \epsilon$, $\sigma_{\min}(N_*) \geq s > 0$, *and* $\epsilon \leq s/2$. *Put* $M = \|C_*\|$ *and* $a = \epsilon(1 + 4M/s)$. *Then*

$$\|I_{v,\mathrm{eff}}(\Lambda) - I_*\| \leq (2M + a)a. \tag{17}$$

In Theorem 5, $s \geq \sqrt{j_{\min}}\sigma_{\min}(\mathcal{H})$ *and* $\epsilon = O(\Lambda^{-1})$. *Thus a positive limiting floor gives a finite-clock floor once the right side is smaller than its smallest eigenvalue.*

The bound follows by perturbing the nuisance projector before squaring the residual scores. It also shows why a small singular value can delay the asymptotic regime. Failure of this sufficient margin does not imply nonidentifiability. Appendix L.5 makes it a pointwise finite-clock check against an explicit frozen comparison law, rather than equating a surrogate value with a certificate.

Proposition 7 (Fixed-resolution boundary)**.** *Fix the number of observations and horizon. Under the uniform long-gap smoothness assumptions of Theorem 1, any deterministic schedule with every gap at least a fixed* $\delta > 0$ *has* $I_{v,\mathrm{eff}} = O(\Lambda^{-2})$. *If affinities are constant in time, the stronger bound is* $O(\Lambda^2 e^{-2q_{\min}\Lambda\delta})$, *where* $q_{\min} > 0$ *bounds the normalized total rate below.*

Indeed the unprofiled clock score is $O(\Lambda^{-1})$; with constant affinity only the initial-state contrast remains. Nuisance projection cannot increase its norm. The constrained experiments therefore compare finite declared rate ranges.

For R independent replicated graph paths, define the local clock-field variance scale

$$V/R = \frac{\mathrm{tr}(BI_{v,\mathrm{eff}}^{-1}B^\top)}{mR}. \tag{18}$$

When $I_{v,\mathrm{eff}} \succeq \kappa B^\top B$, $V/R \leq n/(\kappa mR)$. A regular consistent local likelihood estimator has this asymptotic covariance; convergence of its scaled second moment additionally requires uniform integrability (Appendix I). This is a fixed-dimension, correctly specified local statement, not a global optimization or finite-sample guarantee. It predicts $1/R$ mean-squared error on bounded-degree graphs, while more observed dyads can reduce the field variance through n/m. The experiments test this scale and report where boundary fits or misspecification invalidate its interpretation.

4.2 From information conditions to acquisition

For linear shared affinities $\eta_e(t) = f(t) + \sum_j \psi_j(t)(B\theta_j)_e$, the anchor Jacobian factors through B and the temporal matrix $\Phi = (\psi_j(a_l))_{lj}$. This motivates the common temporal-conditioning constraint below. It is not a substitute for the full hypotheses or finite-clock remainder check. Rank-bound uses $J(D) = \min_{\xi \in \mathcal{P}} \sum_l w_l(D;\xi)$ from (15); other objectives use the *same* feasible set. The pairwise surrogate and direct finite-clock profiling are alternative objectives. Appendix L.4 defines them; Appendix L.6 explains why a larger lower bound need not rank actual information or estimation error correctly. Section 5.3 tests this distinction empirically.

4.3 Finite-clock risk and centered graph regularization

The asymptotic floor need not rank finite-clock recovery risk. We therefore add *TempoGeom-R*, a receding-horizon design within the same conditioned anchor/probe class. For a dyad's history h, let $w_s(h)$ be its posterior weights over the declared design scenarios, and let $F_s(h, D)$ combine

conditional row Fisher information using observed origin-state counts with expected information along a feasible remaining completion D. In dyad coordinates (c, α), set

$$V_s(h, D) = \{F_s(h, D) + \mathrm{diag}(0, \Sigma_\alpha^{-1})\}^{-1},$$
$$\mathcal{R}_\beta(D \mid h) = \sum_s w_s(h)\{[V_s]_{cc} + \beta \,\mathrm{tr}(G[V_s]_{\alpha\alpha})\}. \tag{19}$$

Here G is the temporal-basis Gram matrix and Σ_α is the prior's affinity covariance. Full finite-clock derivatives retain clock–affinity coupling before inversion. Algorithm 2 executes only the first pair of the best completion, updates the weights, and replans. R/C fixes $\beta = 0$; R/J uses a development-selected positive weight. This weighted A-optimal approximation (Alexanderian et al., 2016) scores open-loop completions, not all future adaptive branches. Its dyad-local inverse-Fisher proxy omits graph-wide posterior dependence and shrinkage bias; it is neither exact Bayes risk nor the MSE of the graph estimator.

For graph recovery, a complementary regularizer pools node contrasts without supplying the common clock level. With $P_\perp = I - \mathbf{1}\mathbf{1}^\top/n$, fit

$$(\widehat{v}, \widehat{\theta}) = \arg\min_{v,\theta}\{-\ell(v, \theta) + \tfrac{\kappa}{2}(\|P_\perp v\|^2 + \|\theta\|^2)\}. \tag{20}$$

Since $P_\perp \mathbf{1} = 0$, the penalty does not directly shrink the common clock level; likelihood coupling can still change its fitted value. We tune the *same* regularizer separately for every baseline and retain MLE fits to isolate acquisition gains. The information theorems apply under their stated assumptions; guarantees for the adaptive regularized estimator require separate analysis (Appendix A).

Algorithm 2 TempoGeom-R: dyad-local planning and joint graph recovery

Require: Observation graph B, feasible completions $\mathcal{D}_\tau$, scenarios and prior weights, frozen β, κ.
for each observed dyad e **do**
  Initialize count history h_e and scenario weights w_s.
  for each remaining observation pair **do**
    Enumerate completions $D \in \mathcal{D}_\tau$ consistent with the acquired prefix.
    Form $F_s(h_e, D)$ from past origin counts and expected future counts.
    Select $D^* \in \arg\min_D \mathcal{R}_\beta(D \mid h_e)$ by (19).
    Commit to its first pair; observe counts and update h_e, w_s.
  end for
end for
Fit all graph clocks and affinities jointly by (20); retain schedules and diagnostics.

5 Experiments: separating acquisition and estimator gains

5.1 Held-out data and matched comparisons

We use polynomial $[1, s, (3s^2 - 1)/2]$ and Fourier $[1, \sin(2\pi t/5.5), \cos(2\pi t/5.5)]$ affinities, $s = 2t/5.5 - 1$, with base $-.2$ and independent node coefficients and clock deviations $U[-.15, .15]$. Inference jointly fits all coefficients and $c = \log 32 + Bv$ without the true scale. Thirty held-out graph seeds are paired across families and scales $8, 32, 64$. Each acquisition uses K_6, 100 persistent paths per dyad, six post-initial observations, 9,000 observed transition intervals (not latent switches), horizon 5.5, gap floor $1/64$, and a three-pair grid with anchor singular value at least $.5$. Initial counts are parameter-independent 50/50. Adaptive methods see only their dyad's past counts; fixed schedules have no feedback. Clock RMSE is $\|B(\widehat{v} - v)\|/\sqrt{m}$, including the unknown common shift. Affinity RMSE is root mean squared logit-trajectory error over all dyads and 21 equally spaced times in $[0, 5.5]$. The reproducibility statement and Appendix A.6 describe the accompanying artifact.

Eight disjoint development seeds select R/J's risk weight and each method's precision from the same four-candidate grid; R/C fixes the clock-only weight. Selections are reported as frozen before confirmation. Each dataset is fitted with MLE and its method's selected regularizer; all select $\kappa = 133\frac{1}{3}$. Appendix A.3 specifies the 288-test Holm family for external comparisons, conditional on frozen policies. The graph seed is the sampling unit. Internal comparisons are exploratory and outside this family.

5.2 RECENT ACQUISITION METHODS AND ATTRIBUTION

Six external rows represent J/C variants of Step-DAD–grid (ICML'25), Inside-Out SMC^2–grid (ICML'24), and PASOA–grid (ICML'24) (Hedman et al., 2025; Iqbal et al., 2024; Iollo et al., 2024). Appendices O–P describe these source-informed task adaptations rather than end-to-end reproductions. Preselected acquisition policies are reused on new confirmation datasets. J denotes a joint clock–affinity target and C a clock target: R minimizes local risk, whereas the external adapters maximize information. The targets share labels, but their objectives differ.

All methods receive the same 108 scenarios and prestructured timing grid. However, the information-gain adapters use a finite-support belief (four nominal clock levels and 27 affinity configurations), whereas R uses continuous local derivatives around those scenarios. This may favor local precision after discrete clock uncertainty is small. Increasing particle count preserves the finite support; continuous-posterior comparisons require a separate evaluation. Rank-bound, fixed A/P, fixed MC-EIG and no-refinement DAD remain internal controls.

Table 1: Held-out mean clock RMSE over 30 paired graph seeds per cell; lower is better. Reg is independently development-tuned for each method. External adapters all use Reg; MLE/Reg pairs separate design from estimator changes. Every row uses the same 9,000-transition budget and timing grid. Full SDs, affinity errors and all MLE/Reg controls are in Appendix A.4.

	Polynomial			Fourier		
Design	$\lambda = 8$	32	64	$\lambda = 8$	32	64
Internal references (estimator shown)						
Rank-bound / MLE	0.1160	0.0959	0.1329	0.1250	0.0851	0.1355
Rank-bound / Reg	0.0862	0.0728	0.0936	0.0869	0.0697	0.0860
Fixed A/P / Reg	0.0861	0.0733	0.0945	0.0870	0.0705	0.0861
External task adapters: all with Reg						
Step-DAD–grid/J (ICML'25)	0.0857	0.0777	0.0950	0.0791	0.0798	0.0969
Step-DAD–grid/C (ICML'25)	0.0865	0.0736	0.0959	0.0868	0.0705	0.0871
IO-SMC^2–grid/J (ICML'24)	0.0821	0.0837	0.1064	0.0825	0.0814	0.1000
IO-SMC^2–grid/C (ICML'24)	0.0855	0.0731	0.0948	0.0866	0.0700	0.0877
PASOA–grid/J (ICML'24)	0.0911	0.0807	0.1005	0.0893	0.0799	0.0978
PASOA–grid/C (ICML'24)	0.0820	0.0735	0.0953	0.0829	0.0716	0.0902
Proposed planner (two frozen targets)						
TempoGeom-R/J / MLE	0.0976	0.0966	0.1340	0.1036	0.0868	0.1355
TempoGeom-R/C / MLE	0.0972	0.0963	0.1345	0.1038	0.0864	0.1349
TempoGeom-R/J / Reg	0.0791	0.0733	0.0941	0.0768	0.0704	0.0860
TempoGeom-R/C / Reg	0.0789	0.0728	0.0940	0.0773	0.0704	0.0853

R/J with regularization reduces mean clock RMSE by 17.3–38.5% and affinity RMSE by 16.5–41.5% relative to the *unregularized* Rank-bound pipeline. This changes both design and estimation. Holding Reg fixed, clock RMSE improves by $8.25\%/11.54\%$ at scale 8 (polynomial/Fourier), is slightly higher in three faster-scale cells, and is identical in the reported Fourier scale-64 results. Table 2 reports every paired difference against Rank-bound/Reg; its intervals are exploratory, not part of the 288-test family. With MLE held fixed, R/J likewise improves scale-8 clock RMSE by $15.9\%/17.1\%$, but not uniformly at faster scales. Its mean final probe gap grows to approximately .059 at scale 8 and returns to $1/64$ at scale 64; these are action diagnostics, not a causal ablation.

Against external MLE fits, R/J/Reg has lower clock and affinity means in all 36 comparisons per metric; 34 clock and 36 affinity differences survive the planned Holm adjustment. Against the stronger Reg fits, it has lower means in 34/36 clock and 22/36 affinity comparisons. After adjustment, nine clock differences favor R/J and none favor competitors; eight affinity differences favor R/J and four favor competitors. R/C has lower clock means in 35/36 Reg comparisons but loses affinity accuracy more often. R/J is the development-selected default. These related comparisons describe condition-specific trade-offs, rather than independent trials or evidence of equivalence.

Table 2: Matched-estimator internal comparison: R/J/Reg minus Rank-bound/Reg, in units of 10^{-3}; negative favors R/J. Mean paired differences and descriptive 95% t intervals over 30 graph seeds. These post-hoc internal comparisons are outside the frozen 288-test family. †: reported identical saved estimates; not an inferential interval or population equivalence.

Family	λ	Clock RMSE difference	Affinity RMSE difference
Polynomial	8	-7.11 $[-9.11, -5.11]$	-7.64 $[-9.27, -6.01]$
Polynomial	32	$+0.54$ $[-0.24, +1.31]$	-3.38 $[-4.97, -1.80]$
Polynomial	64	$+0.52$ $[-0.90, +1.94]$	-3.81 $[-5.33, -2.29]$
Fourier	8	-10.03 $[-13.08, -6.98]$	-5.71 $[-7.38, -4.04]$
Fourier	32	$+0.70$ $[+0.14, +1.26]$	-0.06 $[-0.34, +0.22]$
Fourier	64	$0^{\dagger}$	$0^{\dagger}$

5.3 Mechanism checks and limits

The safeguard ablation in Appendices M–N improves affinity estimation but not uniformly clock recovery. Negative surrogate/information correlations within the top score quartile motivate direct finite-clock risk rather than treating the bound as a recovery objective. Those appendices also report fixed-exposure sparse-graph, timestamp-corruption and omitted-basis studies, using schedules and estimators distinct from R/Reg. The R/Reg evaluation uses a predefined anchor/probe structure; larger or sparse graphs, unrestricted timing discovery and continuous-prior calibration remain untested.

Secondary ten-seed checks with frozen settings show lower clock error for R/J/Reg than Rank-bound/MLE at scales $16, 48$, but doubled heterogeneity can reverse regularization gains (Appendix A.4).

5.4 Auxiliary predictive evidence

A separate study fits ten endpoint-state predictors on five paired 32-node graphs per condition. TempoGeom-HR adds a bounded logit residual to the H rate head; Proxy-MLP controls its added inputs. The external adapters are ℓ_1LD, GraSSP, FreeDyG, CRAFT and TAMI (Lai et al., 2026; Çelikkanat et al., 2024; Tian et al., 2024; Yi et al., 2025; Yu et al., 2025). These are prediction, not clock-acquisition, comparisons. Appendix A.5 gives model descriptions, standard deviations, effects, and tuning settings.

HR has lower mean NLL than all five external adapters; 13/15 differences survive the planned paired-t/27-test Holm analysis. With only five graphs, this result is assumption-sensitive: a post-hoc exact two-sided sign-flip check has minimum $p = 2/2^5 = .0625$. The planned and sensitivity analyses use different null assumptions. None of HR's improvements over H survives adjustment, and the MLP has lower homogeneous-condition mean NLL. H and HR differ in learning rate and regularization, and 72/150 fits reach the 1,200-update cap (all 15 H fits and eight HR fits). These differently tuned, budget-limited comparisons assess prediction, not an isolated residual effect or clock identification, and do not establish comparable optimizer convergence.

6 Discussion

The theory distinguishes slower information decay from a nonvanishing information floor: reordering gaps can change decay rates, whereas retaining a floor requires probes on the switching timescale and sufficiently conditioned affinity anchors. At finite clocks, regularization accounts for much of the full-pipeline gain, while matched-estimator acquisition gains concentrate at the slowest tested scale. Clock–affinity trade-offs and reversals under increased heterogeneity show why timing and shrinkage must be evaluated separately. The R/Reg evaluation uses small K_6 graphs, structured timing grids, finite design priors, and frozen policies. Larger and sparse graphs, continuous-prior comparisons, and real-data evaluation are priorities for future work. The fixed-design information guarantees do not establish corresponding guarantees for the adaptive regularized estimator or predictive residual head.

AI USE STATEMENT

Generative AI tools were used only to assist with literature search and the identification of potentially relevant related work. All cited references were independently checked by the authors. Generative AI was not used to develop the theoretical contributions, derive mathematical results or proofs, design or conduct experiments, implement the proposed method, analyze results, or generate the main manuscript content. The authors are fully responsible for the accuracy and content of the paper.

ETHICS STATEMENT

The studies described here use synthetic data. Applying these methods to measured contact or relationship data requires appropriate consent, licensing, observation-coverage checks, and privacy safeguards. A latent coefficient is a statistical representation, not a causal or psychological measurement, and weakly identified clocks should not be presented as precise behavioral facts.

REPRODUCIBILITY STATEMENT

Sections 4.3–5 and Appendix A specify the finite-clock method and evaluation; Sections 3–4 and Appendices B–K provide the information results and proofs. The supplementary artifact contains the manuscript sources, numerical tables, aggregate primary-study summaries, archived audit reports, and code and saved outputs for separately scoped supporting studies. It also includes the PASOA protocol and its associated hash manifest. The primary finite-clock and matched-grid/PASOA implementations, raw observations, per-fit outputs, predictive checkpoints, and original development and freeze records are not included, so these evaluations cannot be reproduced end-to-end from the artifact. Appendix A.6 and `STUDY_PATHS.json` give study-specific coverage; `reproduce.py` lists the supported checks. Supporting-study results are separate experimental cohorts, not additional primary replications. The manuscript builds without running experiments.

REFERENCES

Alen Alexanderian, Noemi Petra, Georg Stadler, and Omar Ghattas. A fast and scalable method for A-optimal design of experiments for infinite-dimensional bayesian nonlinear inverse problems. *SIAM Journal on Scientific Computing*, 38(1):A243–A272, 2016. doi: 10.1137/140992564.

Markus Arns, Peter Buchholz, and Andriy Panchenko. On the numerical analysis of inhomogeneous continuous-time Markov chains. *INFORMS Journal on Computing*, 22(3):416–432, 2010. doi: 10.1287/ijoc.1090.0357.

J. E. Avron, M. Fraas, G. M. Graf, and P. Grech. Adiabatic theorems for generators of contracting evolutions. *Communications in Mathematical Physics*, 314:163–191, 2012. doi: 10.1007/s00220-012-1504-1. URL https://arxiv.org/abs/1106.4661.

Nicholas W. Barendregt, Emily G. Webb, and Zachary P. Kilpatrick. Adaptive Bayesian inference of Markov transition rates. *Proceedings of the Royal Society A*, 479(2270):20220453, 2023. doi: 10.1098/rspa.2022.0453. URL https://arxiv.org/abs/2205.10371.

Freddie Bickford Smith, Andreas Kirsch, Sebastian Farquhar, Yarin Gal, Adam Foster, and Tom Rainforth. Prediction-oriented Bayesian active learning. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 7331–7348. PMLR, 2023. URL https://proceedings.mlr.press/v206/bickfordsmith23a.html.

Carter T. Butts. Continuous time graph processes with known ERGM equilibria: Contextual review, extensions, and synthesis. *The Journal of Mathematical Sociology*, 48(2):129–171, 2024. doi: 10.1080/0022250X.2023.2180001.

Abdulkadir Çelikkanat, Nikolaos Nakis, and Morten Mørup. Piecewise-velocity model for learning continuous-time dynamic node representations. In *Learning on Graphs Conference*, 2022.

Abdulkadir Çelikkanat, Nikolaos Nakis, and Morten Mørup. Continuous-time graph representation with sequential survival process. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(10):11177–11185, 2024. doi: 10.1609/aaai.v38i10.28995.

Weilin Cong, Si Zhang, Jian Kang, Baichuan Yuan, Hao Wu, Xin Zhou, Hanghang Tong, and Mehrdad Mahdavi. Do we really need complicated model architectures for temporal networks? In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=ayPPc0SyLv1.

Yves Coudière, Charlie Douanla Lontsi, and Charles Pierre. Rush–larsen time-stepping methods of high order for stiff problems in cardiac electrophysiology. *Electronic Transactions on Numerical Analysis*, 52:342–357, 2020. doi: 10.1553/etna_vol52s342.

Dragoš Cvetković, Peter Rowlinson, and Slobodan K. Simić. Signless laplacians of finite graphs. *Linear Algebra and its Applications*, 423(1):155–171, 2007. doi: 10.1016/j.laa.2007.01.009.

Alexander Erreygers and Jasper De Bock. Imprecise continuous-time Markov chains: Efficient computational methods with guaranteed error bounds. In *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 145–156, 2017. URL https://proceedings.mlr.press/v62/erreygers17a.html.

Jinwoo Go and Tobin Isaac. Robust expected information gain for optimal Bayesian experimental design using ambiguity sets. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pages 728–737. PMLR, 2022. URL https://proceedings.mlr.press/v180/go22a.html.

Marcel Hedman, Desi R. Ivanova, Cong Guan, and Tom Rainforth. Step-DAD: Semi-amortized policy-based Bayesian experimental design. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 22904–22923. PMLR, 2025. URL https://proceedings.mlr.press/v267/hedman25a.html.

Peter D. Hoff, Adrian E. Raftery, and Mark S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098, 2002. doi: 10.1198/016214502388618906.

Jacopo Iollo, Christophe Heinkelé, Pierre Alliez, and Florence Forbes. PASOA: PArticle baSed Bayesian Optimal Adaptive design. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 21020–21046, 2024. URL https://proceedings.mlr.press/v235/iollo24a.html.

Sahel Iqbal, Adrien Corenflos, Simo Särkkä, and Hany Abdulsamad. Nesting particle filters for experimental design in dynamical systems. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 21047–21068. PMLR, 2024. URL https://proceedings.mlr.press/v235/iqbal24a.html.

Pavel N. Krivitsky and Mark S. Handcock. A separable model for dynamic networks. *Journal of the Royal Statistical Society: Series B*, 76(1):29–46, 2014. doi: 10.1111/rssb.12014.

Charalampos Kyriakopoulos and Verena Wolf. Optimal observation time points in stochastic chemical kinetics. In *Hybrid Systems Biology (HSB 2014)*, volume 7699 of *Lecture Notes in Computer Science*, pages 83–96. Springer, 2015. doi: 10.1007/978-3-319-27656-4_5. URL https://arxiv.org/abs/1412.5439.

Zhao-Rong Lai, Zheng-Sen Zhou, Liangda Fang, Yongsen Zheng, and Ziliang Chen. ℓ_1 latent distance based continuous-time graph representation. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=pW1Kg9CYyw.

Farimah Poursafaei, Shenyang Huang, Kellin Pelrine, and Reihaneh Rabbany. Towards better evaluation for dynamic link prediction. In *Advances in Neural Information Processing Systems*, volume 35, pages 32928–32941, 2022. URL https://arxiv.org/abs/2207.10128.

Marco Radaelli, Gabriel T. Landi, Kavan Modi, and Felix C. Binder. Fisher information of correlated stochastic processes. *New Journal of Physics*, 25(5):053037, 2023. doi: 10.1088/1367-2630/acd321.

Andrea Rotnitzky, D. R. Cox, Matteo Bottai, and James M. Robins. Likelihood-based inference with singular information matrix. *Bernoulli*, 6(2):243–284, 2000. URL https://projecteuclid.org/journals/bernoulli/volume-6/issue-2/Likelihood-based-inference-with-singular-information-matrix/bj/1081788028.full.

Yuxing Tian, Yiyan Qi, and Fan Guo. FreeDyG: Frequency enhanced continuous-time dynamic graph model for link prediction. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/21f7b745f73ce0d1f9bcea7f40b1388e-Abstract-Conference.html.

Lu Yi, Runlin Lei, Fengran Mo, Yanping Zheng, Zhewei Wei, and Yuhang Ye. Future link prediction without memory or aggregation. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-0084. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/036912a83bdbb1fd792baf6532f102d8-Abstract-Conference.html.

Le Yu, Leilei Sun, Bowen Du, and Weifeng Lv. Towards better dynamic graph learning: New architecture and unified library. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/d611019afba70d547bd595e8a4158f55-Abstract.html.

Zhongyi Yu, Jianqiu Wu, Zhenghao Wu, Shuhan Zhong, Weifeng Su, Chul-Ho Lee, and Weipeng Zhuo. TAMI: Taming heterogeneity in temporal interactions for temporal graph link prediction. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-3923. URL https://proceedings.nips.cc/paper_files/paper/2025/hash/aa963ac256590bb7ad5fc26c68229a3a-Abstract-Conference.html.

SUPPLEMENTARY MATERIAL: READING GUIDE

Finite-clock method and primary evaluation: Appendix A. Section 4.3 and Algorithm 2 define the planner and graph estimator. Appendix A (page 14) gives the finite-clock construction, development protocol, primary and secondary outcomes, and reported numerical checks. Appendix A.5 covers the auxiliary predictive study and its post-hoc small-sample sensitivity analysis. Artifact availability is described in the reproducibility statement and Appendix A.6; `STUDY_PATHS.json` maps the packaged evidence.

Primary evaluation	Page
Finite-clock planner and centered estimator	14
Development, freezing, and statistical families	15
Complete acquisition results and ablations	16
Auxiliary prediction, tuning, and uncertainty	23
Reported execution checks and artifact coverage	27

Information theory and mechanism studies. Appendices B–K supply the information proofs and timing/nuisance-sensitivity results. Appendix L defines finite-prior coverage, acquisition criteria and transfer diagnostics. Appendices M–N report criterion comparisons, fixed-budget sparse graphs, internal controls, timestamp error and omitted-basis studies. These evaluate schedules and estimators distinct from R/Reg; their graph and misspecification results do not establish performance of that pipeline.

Adapted-baseline comparisons and supporting studies. Appendices O–P report finite-grid adaptive-design comparisons, including the post-hoc PASOA study. Appendices Q–AC, with a guide on page 71, contain restricted-prior experiments, ancillary likelihood methods and forecasting adapters. These supporting studies have separate protocols and are not additional independent replications of the primary evaluation. Complete outcomes are reported for each study.

A FINITE-CLOCK PLANNING, REGULARIZED ESTIMATION, AND HELD-OUT EVALUATION

This appendix specifies the finite-clock planner, centered graph estimator, held-out acquisition study and auxiliary predictive study. Method development was informed by the supporting studies in Appendices O–P and AB. The protocol uses disjoint development and confirmation graph seeds and reports internal configuration freezing before confirmation, rather than external preregistration. Internal-comparison and sign-flip diagnostics are post-hoc analyses of the same data, separate from the primary testing families. Appendix A.6 specifies the available code, data, and audit records.

A.1 RECEDING-HORIZON FINITE-CLOCK RISK

The planner acts separately on each observed dyad, receiving only its own past counts, with the same feedback access as the external adapters. Let $x = (c, \alpha_1, \alpha_2, \alpha_3)$ denote the dyad log-clock and three affinity coefficients. For an interval $[u, t]$, scenario s, and origin state a, put $p_{s,a} = P_s(u, t)_{a1}$ and $g_{s,a} = \nabla_x p_{s,a}$. A row of n_a observed paths contributes

$$F_{s,[u,t]} = \sum_{a\in\{0,1\}} n_a \frac{g_{s,a} g_{s,a}^\top}{p_{s,a}(1 - p_{s,a})}. \tag{21}$$

The matrix is evaluated using the full transition law and derivatives, not the frozen large-clock expansion. All hidden switches are integrated out. Previously observed row counts give the past matrix. For a future completion, expected row counts start from the currently observed occupancy and propagate through P_s; expected contributions from all remaining intervals are then added. This yields $F_s(h, D)$ in (19).

The shared 108-scenario prior combines four nominal clock levels with 27 affinity configurations. Its nuisance covariance is $\Sigma_\alpha = .015 I_3$ and $G = 21^{-1} \sum_{j=0}^{20} \psi(t_j)\psi(t_j)^\top$ at $t_j = 5.5j/20$, matching the trajectory-error evaluation grid. The nuisance precision Σ_α^{-1} is used in design only; the downstream graph penalty is separately selected. A 10^{-12} clock ridge is numerical stabilization, not knowledge of a generating clock. In block form, if $A = F_s + \mathrm{diag}(0, \Sigma_\alpha^{-1})$, then

$$[A^{-1}]_{cc} = \{A_{cc} - A_{c\alpha} A_{\alpha\alpha}^{-1} A_{\alpha c}\}^{-1}. \tag{22}$$

Thus the clock term retains clock–affinity coupling; simply inverting F_{cc} would incorrectly treat affinities as known. The trace term also uses the corresponding block of the *full* inverse. It is not the inverse of the nuisance block alone.

At every decision the planner enumerates all feasible remaining completions consistent with its acquired prefix, minimizes posterior-weighted local risk, executes the first pair, and repeats. After observing counts N_{ab} it updates

$$\log w_s \leftarrow \log w_s + \sum_{a,b} N_{ab} \log P_{s,ab} - \log \sum_r w_r \exp\left(\sum_{a,b} N_{ab} \log P_{r,ab} \right), \tag{23}$$

with numerically stable normalization. Count-combinatorial factors cancel from these weights. The finite prior is a design approximation: the actual generating parameters remain continuous. R uses derivatives around the scenario locations, whereas the external information-gain adapters use their finite-support belief. This distinction is material: a finite clock-label information objective can saturate while local continuous uncertainty remains. This is a possible explanation for performance differences rather than a demonstrated cause. Increasing PASOA's particle count preserves its prior support. Comparisons with denser-support or continuous-posterior implementations require separate baseline tuning and held-out evaluation. The specified interface accepts counts and a declared prior, never the generating parameters or scale. Guarantees for this data-dependent policy and penalized estimator require analysis beyond the conditioned-grid assumptions. Committing to the two times of a pair before either observation matches every other adaptive method.

Algorithm 2 in the main text summarizes this procedure; the row-information and posterior-update formulas above specify its implementation.

This is a receding-horizon *open-loop completion* approximation, not a dynamic program integrating over every future adaptive branch. It minimizes an inverse-Fisher proxy for squared error, not exact Bayes risk, and omits shrinkage bias and graph-wide posterior dependence in its dyad-local ranking. Covariance-trace design and Gaussian approximations have established precedents (Alexanderian et al., 2016). The contribution here is their explicit finite-clock, nuisance-coupled realization in the conditioned timing class, evaluated against Rank-bound and the task-adapted information-gain methods. Optimality and nonvanishing-information guarantees for this adaptive, regularized procedure remain open to separate analysis.

A.2 Centered regularization and controlled attribution

Equation (20) changes the downstream estimator. The MLE and regularized fits share the likelihood, parameter bounds, two truth-independent starts, numerical precision checks and selection by the fitted objective. The log-clock node contrast precision and the affinity-node precision are equal to the declared κ. Regularization is implemented as additional least-squares residuals and their exact Jacobian, scaled consistently with the count deviance. At $\kappa = 0$ the objective reduces to maximum likelihood. The development audit summary reports exact numerical agreement.

The penalty leaves the unknown global clock level free: for any $a \in \mathbb{R}$, $P_\perp(v + a\mathbf{1}) = P_\perp v$. Its log-clock Hessian is $\kappa P_\perp$, with a zero eigenvalue in the constant direction. This property is algebraic, not a claim that the likelihood is shift-invariant. The estimator can reduce variance in node contrasts at the price of bias when the population is more heterogeneous than expected. Shrinking affinity coefficients toward zero means shrinking toward the declared base trajectory $-.2$, not toward a fitted truth. Centered ridge provides the standard bias–variance trade-off within this estimator.

To isolate effects, every acquisition, including all external variants and internal controls, receives both MLE and its *own development-selected* regularizer. Comparisons with external MLE measure the effect of changing the entire estimation pipeline. Comparisons with external regularized fits hold the estimator family and its tuning access fixed, isolating acquisition differences more closely. Shift invariance is a property of the penalty; likelihood coupling can still change the fitted common clock relative to MLE. Rank-bound is also fitted both ways as an internal control. For the declared node population, $\mathrm{Var}(U[-.15, .15]) = .0075$, whose inverse is $133\frac{1}{3}$, the development-selected precision. This alignment is population-specific and does not supply test labels. The doubled-heterogeneity results retain the frozen precision and assess this sensitivity to population variation.

A.3 Development, freezing, and sampling units

We first use eight development graph seeds 6100100–6100107, paired across two affinity families and scales $8, 32, 64$. Three risk weights $\beta \in \{0, .25, 1\}$ and four estimator precisions $\kappa \in \{0, 66\frac{2}{3}, 133\frac{1}{3}, 266\frac{2}{3}\}$ are evaluated, together with Rank-bound and fixed A/P. The first development pass comprises 960 reported fits. The criterion is the balanced mean of squared clock RMSE plus squared affinity RMSE, not the mean RMSE. A common precision is first selected across the five design candidates; the risk weight is selected at that precision. This chooses $\beta = 1$ and $\kappa = 133\frac{1}{3}$. We retain both the resulting R/J operating point and a separately fixed clock-only R/C operating point with $\beta = 0$; test conditions never choose between them.

A second calibration pass applies the same four precision candidates to each of the ten remaining baseline/control rows on exactly the same development graph seeds and physical random streams. This adds 1,920 fits. Each of the resulting 14 rows separately minimizes the same development criterion over its four precision candidates; all select $133\frac{1}{3}$. These 2,880 fitting outcomes are tuning evidence, not additional independent confirmation samples. The study summary reports pre-confirmation hashing of the selected risks, policies, likelihood code, prior grids and numerical checks. The corresponding record coverage is detailed in Appendix A.6.

The primary experiment uses 30 new graph seeds 31000–31029, paired across the two families and three scales. Every row gets the same six observations on each of 100 persistent paths per dyad, K_6, horizon 5.5, minimum gap $1/64$, and temporal singular-value threshold $.5$. The budget is 9,000 observed transition intervals per graph: 15 dyads times 100 persistent paths times 6 intervals. It counts observed start/end-state pairs, not latent formation or dissolution events. All 14 rows use the same 108-point design prior. Six external rows represent three source algorithms with two informa-

tion targets. They reuse the acquisition policies and PASOA settings selected in Appendices O–P, with new acquisitions and graph fits on the confirmation seeds. Those appendices specify the task adaptations.

The planner uses no additional simulator samples and shares the observation budget and design access of the external adapters, without access to the generating clock scale. It enumerates feasible completions and evaluates analytical derivatives, so its computational work and feedback differ from fixed schedules. Wall times measured under concurrent CPU workers are implementation diagnostics rather than controlled end-to-end speed comparisons. The protocol describes precomputed transition grids, Fisher caches and adaptive decision traces; their availability is specified in Appendix A.6.

Primary inference treats the graph seed as the sampling unit. The frozen multiplicity family contains $2 \times 6 \times 2 \times 2 \times 6 = 288$ two-sided paired comparisons: R/J and R/C with regularization, each external target with MLE and with its independently tuned regularizer, clock and affinity RMSE, and six family/scale cells. Holm adjustment is applied once across all 288 tests. The 95% paired t intervals are descriptive unadjusted intervals; they are not simultaneous intervals. Related conditions reuse graph random streams and are not pooled as independent replications. Results are conditional on the selected acquisition checkpoints and regularization choices, not averaged over repeating the entire training-selection procedure. Nonsignificance does not imply equivalence.

Prespecified secondary checks use ten fresh seeds 32000–32009 at scales $16, 48$, and ten seeds 33000–33009 at scale 32 with doubled node-clock deviations and affinity coefficients. The common clock is not doubled in the latter check. These tests keep every method and its selected hyperparameters frozen; they test interpolation and a limited heterogeneity shift, not general misspecification robustness. They have separate descriptive comparisons, not additional members of the primary 288-test family.

A.4 Held-out acquisition results and complete ablations

The primary summary contains 5,040 fit labels on 180 family/scale/graph-seed cases. There are 2,520 acquired-data labels and 2,179 distinct datasets: repeated schedules and paired physical random streams explain duplicates, and the two estimators always fit the same acquired dataset. No selected optimizer fails or exhausts its cap. Two retained clock-boundary fits are PASOA/J with MLE, polynomial scale 64 seed 31020 and Fourier scale 64 seed 31010; their clock RMSEs are .72649 and .73482. No regularized fit or proposed-method fit reaches a parameter boundary. Both cases are included in the summaries, with method settings fixed.

Whole-pipeline improvement versus acquisition alone. Across the six conditions, R/J with regularization reduces mean clock RMSE by 17.3–38.5% and affinity RMSE by 16.5–41.5% relative to Rank-bound/MLE. The corresponding exploratory paired intervals all exclude zero, but are not part of the primary external-comparison family. A substantial part is transferable estimator improvement: balanced descriptive mean clock RMSE is .11507 for Rank-bound/MLE, .08251 for Rank-bound with regularization and .07995 for R/J with regularization. These means first average the six conditions within each graph-seed block; the six conditions are not treated as independent replications. On the joint squared-error criterion used in development, R/J/Reg has the smallest descriptive aggregate among all retained rows: .01415, compared with .01506 for the smallest external regularized aggregate (Inside-Out/C). This exploratory aggregate summarizes joint error across conditions; it is separate from adjusted per-condition tests.

R/J with MLE alone improves clock RMSE at scale 8 from .11603 to .09756 (polynomial) and .12501 to .10359 (Fourier). At scale 32 the means increase slightly in both families; the Fourier increase has an exploratory paired interval excluding zero. At polynomial scale 64 the mean also increases slightly, and the Fourier scale-64 acquisition is identical to Rank-bound. Acquisition gains under MLE therefore depend on scale. The complete acquisition-by-estimator comparison separates these effects from regularization gains.

Matched-estimator internal controls. Table 3 reports all 24 paired contrasts with Rank-bound/Reg and fixed A/P/Reg. This post-hoc analysis reuses the confirmation fits and is separate from the frozen 288-test family. With Reg fixed, R/J reduces clock RMSE relative to Rank-bound

by 8.25% and 11.54% at scale 8. The three other nonidentical cells have slightly higher clock means; at Fourier scale 64 the audit summary reports identical saved estimates. The descriptive Fourier/32 clock interval excludes zero in favor of Rank-bound, whereas the slow-scale intervals favor R/J. These unadjusted results are exploratory, not simultaneous significance claims. The table also reports affinity comparisons and fixed A/P controls.

Table 3: Complete post-hoc matched-estimator internal audit. R/J/Reg minus each Reg control, in units of 10^{-3}; negative favors R/J. All 24 metric/condition/control contrasts are retained with unadjusted descriptive 95% paired-t intervals on 30 graph seeds. These post-hoc comparisons are separate from the primary 288-test family. † denotes the reported identical saved estimates, not an inferential interval or population equivalence.

Family	λ	Clock RMSE difference	Affinity RMSE difference
R/J/Reg minus Rank-bound/Reg			
Polynomial	8	-7.11 $[-9.11, -5.11]$	-7.64 $[-9.27, -6.01]$
Polynomial	32	$+0.54$ $[-0.24, +1.31]$	-3.38 $[-4.97, -1.80]$
Polynomial	64	$+0.52$ $[-0.90, +1.94]$	-3.81 $[-5.33, -2.29]$
Fourier	8	-10.03 $[-13.08, -6.98]$	-5.71 $[-7.38, -4.04]$
Fourier	32	$+0.70$ $[+0.14, +1.26]$	-0.06 $[-0.34, +0.22]$
Fourier	64	$0^{\dagger}$	$0^{\dagger}$
R/J/Reg minus fixed A/P/Reg			
Polynomial	8	-6.99 $[-8.81, -5.16]$	-4.08 $[-5.50, -2.66]$
Polynomial	32	-0.03 $[-0.40, +0.35]$	$+0.06$ $[-0.17, +0.29]$
Polynomial	64	-0.34 $[-0.87, +0.19]$	-0.21 $[-0.56, +0.14]$
Fourier	8	-10.20 $[-13.39, -7.01]$	-13.15 $[-16.24, -10.07]$
Fourier	32	-0.12 $[-1.25, +1.00]$	-7.06 $[-10.29, -3.83]$
Fourier	64	-0.16 $[-1.39, +1.07]$	-7.55 $[-9.93, -5.17]$

External methods with and without the same estimator improvement. R/J with regularization has lower mean clock and affinity RMSE than every external variant fitted with MLE in every primary condition (36 comparisons per metric). Of these, 34 clock and all 36 affinity contrasts are supported at 5% after the single 288-test Holm adjustment. This comparison changes both acquisition and estimation.

Against the stronger, separately calibrated regularized external fits, R/J has lower mean clock RMSE in 34/36 comparisons and lower mean affinity RMSE in 22/36. Nine adjusted clock contrasts favor R/J and none favor an external method; eight adjusted affinity contrasts favor R/J and four favor external methods. The two contrary clock mean orderings are against Inside-Out/C at scale 32, with R/J-minus-comparator differences .000247 (polynomial) and .000388 (Fourier), both adjusted $p = 1$. Both differences are reported; their nonsignificance does not establish equivalence.

R/C with regularization has lower clock means in 35/36 strengthened comparisons, with 11 adjusted contrasts favoring it and none favoring a competitor. Its sole contrary mean ordering is Inside-Out/C at Fourier scale 32, difference .000378, adjusted $p = 1$. This clock focus sacrifices affinity recovery: only 3/36 strengthened affinity mean comparisons favor R/C, with no adjusted improvements and 14 adjusted differences favoring external methods. This clock–affinity trade-off motivates reporting R/C alongside the development-selected joint-risk operating point R/J.

Timing adaptation and limits. Both proposed planners start with the minimum probe width $1/64$, because the prior is the same across all generating scales. At scale 8, R/J's mean last gap is .05924 (polynomial) and .05882 (Fourier); at scale 64 it is $1/64$ in both families. These changes use observed counts, not the true scale. They are consistent with learning a useful finite-rate probe duration, but the diagnostic is not an independent causal ablation. The planner also changes other feasible timing choices. The finite prior, local-covariance approximation, nuisance precision, and inherited baseline adaptations remain limitations.

The following tables retain all methods, both estimator choices, sample SDs, primary multiplicity outcomes and timing diagnostics. Subsequent secondary tables are descriptive and do not select a replacement configuration. At additional scales $16, 48$, R/J/Reg lowers clock error by 18.0–34.5%

relative to Rank-bound/MLE; 23/24 clock means also favor it over regularized external variants. These ten-seed checks are descriptive; scale 16 is in the design prior, but neither scale was used for recovery tuning. The selected κ equals the inverse variance of the nominal $U[-.15, .15]$ node coefficients. Doubling node deviations and affinity coefficients at scale 32 reveals sensitivity to heterogeneity: Fourier R/J error rises from $.1011/.1178$ (clock/affinity, MLE) to $.1111/.1283$ with frozen regularization. The method was not retuned on this stress result; the nominal gains depend on the node population.

Table 4: Polynomial, MLE: held-out primary mean (sample SD) RMSE over 30 paired graph seeds per cell. All 14 method/target/control rows are reported. Reg settings were separately selected per method before confirmation; MLE and Reg fit the same acquired data.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.116 (.035)	.140 (.024)	.096 (.034)	.133 (.024)	.133 (.055)	.121 (.025)
Fixed A/P	.115 (.032)	.113 (.019)	.096 (.033)	.105 (.018)	.135 (.056)	.098 (.018)
Fixed MC-EIG/J	.107 (.031)	.109 (.018)	.115 (.038)	.101 (.018)	.178 (.051)	.095 (.018)
Fixed MC-EIG/C	.116 (.035)	.140 (.024)	.096 (.034)	.133 (.024)	.133 (.055)	.121 (.025)
DAD/J (control)	.107 (.032)	.111 (.019)	.108 (.036)	.104 (.017)	.141 (.053)	.098 (.018)
DAD/C (control)	.116 (.034)	.124 (.019)	.098 (.034)	.118 (.020)	.135 (.053)	.107 (.021)
Step-DAD–grid/J (ICML'25)	.118 (.034)	.109 (.018)	.101 (.037)	.104 (.017)	.137 (.055)	.098 (.018)
Step-DAD–grid/C (ICML'25)	.116 (.034)	.125 (.019)	.098 (.034)	.118 (.020)	.135 (.053)	.107 (.021)
IO-SMC^2–grid/J (ICML'24)	.105 (.029)	.107 (.019)	.122 (.038)	.099 (.017)	.173 (.088)	.094 (.017)
IO-SMC^2–grid/C (ICML'24)	.115 (.033)	.122 (.021)	.096 (.033)	.115 (.020)	.135 (.056)	.105 (.021)
PASOA–grid/J (ICML'24)	.126 (.038)	.108 (.016)	.107 (.037)	.109 (.017)	.196 (.127)	.100 (.019)
PASOA–grid/C (ICML'24)	.106 (.030)	.122 (.022)	.097 (.035)	.117 (.023)	.140 (.053)	.106 (.020)
TempoGeom-R/J	.098 (.028)	.107 (.016)	.097 (.033)	.105 (.018)	.134 (.057)	.097 (.018)
TempoGeom-R/C	.097 (.027)	.129 (.021)	.096 (.034)	.127 (.024)	.134 (.056)	.115 (.024)

Table 5: Polynomial, Reg: held-out primary mean (sample SD) RMSE over 30 paired graph seeds per cell. All 14 method/target/control rows are reported. Reg settings were separately selected per method before confirmation; MLE and Reg fit the same acquired data.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.086 (.020)	.089 (.015)	.073 (.026)	.084 (.013)	.094 (.021)	.079 (.013)
Fixed A/P	.086 (.020)	.086 (.015)	.073 (.026)	.081 (.013)	.094 (.022)	.075 (.012)
Fixed MC-EIG/J	.083 (.020)	.084 (.014)	.084 (.029)	.079 (.012)	.099 (.029)	.074 (.013)
Fixed MC-EIG/C	.086 (.020)	.089 (.015)	.073 (.026)	.084 (.013)	.094 (.021)	.079 (.013)
DAD/J (control)	.083 (.021)	.084 (.015)	.079 (.028)	.079 (.013)	.097 (.024)	.075 (.012)
DAD/C (control)	.087 (.021)	.088 (.015)	.074 (.027)	.083 (.012)	.096 (.021)	.077 (.013)
Step-DAD–grid/J (ICML'25)	.086 (.024)	.083 (.014)	.078 (.026)	.080 (.012)	.095 (.027)	.075 (.012)
Step-DAD–grid/C (ICML'25)	.086 (.021)	.088 (.015)	.074 (.026)	.083 (.012)	.096 (.021)	.077 (.013)
IO-SMC^2–grid/J (ICML'24)	.082 (.019)	.083 (.014)	.084 (.028)	.078 (.012)	.106 (.030)	.073 (.012)
IO-SMC^2–grid/C (ICML'24)	.086 (.020)	.087 (.014)	.073 (.026)	.082 (.013)	.095 (.022)	.076 (.012)
PASOA–grid/J (ICML'24)	.091 (.023)	.081 (.014)	.081 (.030)	.081 (.013)	.101 (.031)	.075 (.013)
PASOA–grid/C (ICML'24)	.082 (.020)	.087 (.015)	.074 (.028)	.082 (.013)	.095 (.021)	.076 (.012)
TempoGeom-R/J	.079 (.019)	.082 (.014)	.073 (.026)	.081 (.013)	.094 (.022)	.075 (.012)
TempoGeom-R/C	.079 (.019)	.087 (.016)	.073 (.026)	.084 (.013)	.094 (.021)	.079 (.013)

Table 6: Fourier, MLE: held-out primary mean (sample SD) RMSE over 30 paired graph seeds per cell. All 14 method/target/control rows are reported. Reg settings were separately selected per method before confirmation; MLE and Reg fit the same acquired data.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.125 (.037)	.125 (.026)	.085 (.026)	.112 (.020)	.136 (.047)	.100 (.021)
Fixed A/P	.126 (.038)	.166 (.042)	.086 (.027)	.148 (.033)	.135 (.046)	.134 (.034)
Fixed MC-EIG/J	.116 (.032)	.123 (.025)	.092 (.027)	.108 (.020)	.178 (.093)	.098 (.021)
Fixed MC-EIG/C	.126 (.038)	.166 (.042)	.086 (.027)	.148 (.033)	.135 (.046)	.134 (.034)
DAD/J (control)	.111 (.032)	.118 (.025)	.106 (.033)	.107 (.019)	.175 (.064)	.096 (.020)
DAD/C (control)	.126 (.035)	.145 (.033)	.086 (.026)	.128 (.025)	.142 (.043)	.115 (.026)
Step-DAD–grid/J (ICML'25)	.111 (.030)	.116 (.025)	.101 (.034)	.108 (.019)	.172 (.055)	.096 (.020)
Step-DAD–grid/C (ICML'25)	.125 (.035)	.146 (.034)	.086 (.025)	.127 (.025)	.142 (.044)	.115 (.026)
IO-SMC^2–grid/J (ICML'24)	.114 (.033)	.120 (.025)	.111 (.034)	.108 (.021)	.167 (.041)	.098 (.021)
IO-SMC^2–grid/C (ICML'24)	.125 (.037)	.124 (.027)	.086 (.027)	.112 (.019)	.139 (.047)	.099 (.020)
PASOA–grid/J (ICML'24)	.131 (.040)	.116 (.026)	.100 (.033)	.115 (.022)	.255 (.186)	.103 (.022)
PASOA–grid/C (ICML'24)	.116 (.033)	.136 (.032)	.090 (.029)	.128 (.030)	.159 (.058)	.114 (.029)
TempoGeom-R/J	.104 (.029)	.114 (.023)	.087 (.027)	.112 (.020)	.136 (.047)	.100 (.021)
TempoGeom-R/C	.104 (.029)	.136 (.032)	.086 (.027)	.133 (.030)	.135 (.046)	.121 (.028)

Table 7: Fourier, Reg: held-out primary mean (sample SD) RMSE over 30 paired graph seeds per cell. All 14 method/target/control rows are reported. Reg settings were separately selected per method before confirmation; MLE and Reg fit the same acquired data.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.087 (.033)	.099 (.018)	.070 (.017)	.086 (.014)	.086 (.027)	.083 (.014)
Fixed A/P	.087 (.034)	.106 (.018)	.071 (.017)	.093 (.016)	.086 (.027)	.091 (.016)
Fixed MC-EIG/J	.083 (.031)	.097 (.018)	.073 (.019)	.084 (.013)	.100 (.029)	.082 (.014)
Fixed MC-EIG/C	.087 (.034)	.106 (.018)	.071 (.017)	.093 (.016)	.086 (.027)	.091 (.016)
DAD/J (control)	.081 (.030)	.095 (.018)	.082 (.019)	.083 (.014)	.098 (.036)	.081 (.014)
DAD/C (control)	.087 (.033)	.104 (.018)	.070 (.017)	.090 (.014)	.087 (.026)	.087 (.016)
Step-DAD–grid/J (ICML'25)	.079 (.030)	.094 (.018)	.080 (.019)	.083 (.014)	.097 (.036)	.081 (.015)
Step-DAD–grid/C (ICML'25)	.087 (.033)	.104 (.018)	.071 (.017)	.090 (.014)	.087 (.026)	.087 (.016)
IO-SMC^2–grid/J (ICML'24)	.082 (.031)	.096 (.017)	.081 (.019)	.084 (.014)	.100 (.027)	.082 (.015)
IO-SMC^2–grid/C (ICML'24)	.087 (.033)	.098 (.019)	.070 (.017)	.086 (.013)	.088 (.027)	.083 (.014)
PASOA–grid/J (ICML'24)	.089 (.039)	.091 (.016)	.080 (.018)	.084 (.016)	.098 (.033)	.083 (.013)
PASOA–grid/C (ICML'24)	.083 (.031)	.098 (.016)	.072 (.017)	.090 (.014)	.090 (.028)	.088 (.015)
TempoGeom-R/J	.077 (.028)	.093 (.017)	.070 (.018)	.086 (.014)	.086 (.027)	.083 (.014)
TempoGeom-R/C	.077 (.028)	.098 (.019)	.070 (.018)	.093 (.015)	.085 (.026)	.088 (.017)

Table 8: Complete primary comparison counts. Each cell pools six external variants and six related conditions as a descriptive count, not as independent replications. Lower mean is distinct from Holm-supported lower error; adjustment uses the full 288-test family.

Proposed (Reg)	External fit	Metric	Lower mean	Lower, Holm	Higher, Holm
TempoGeom-R/J	MLE	Clock	36/36	34/36	0/36
TempoGeom-R/J	MLE	Affinity	36/36	36/36	0/36
TempoGeom-R/J	Reg	Clock	34/36	9/36	0/36
TempoGeom-R/J	Reg	Affinity	22/36	8/36	4/36
TempoGeom-R/C	MLE	Clock	36/36	34/36	0/36
TempoGeom-R/C	MLE	Affinity	36/36	30/36	0/36
TempoGeom-R/C	Reg	Clock	35/36	11/36	0/36
TempoGeom-R/C	Reg	Affinity	3/36	0/36	14/36

Table 9: Balanced descriptive primary summary, regularized estimators for every method. Each graph-seed block first averages its six related conditions, then 30 blocks are summarized. Joint MSE is squared clock RMSE plus squared affinity RMSE, the criterion used in development. Its SD is across graph-seed blocks. These aggregates are exploratory summaries, not an additional primary testing family or per-condition dominance claim.

Design (Reg)	Clock RMSE	Affinity RMSE	Joint MSE (SD)
Rank-bound	0.08251	0.08668	0.01523 (0.00276)
Fixed A/P	0.08292	0.08858	0.01571 (0.00284)
Fixed MC-EIG/J	0.08674	0.08345	0.01550 (0.00284)
Fixed MC-EIG/C	0.08271	0.09034	0.01597 (0.00295)
DAD/J (control)	0.08658	0.08298	0.01539 (0.00269)
DAD/C (control)	0.08342	0.08816	0.01568 (0.00282)
Step-DAD–grid/J (ICML'25)	0.08571	0.08262	0.01520 (0.00270)
Step-DAD–grid/C (ICML'25)	0.08340	0.08827	0.01569 (0.00282)
IO-SMC^2–grid/J (ICML'24)	0.08935	0.08252	0.01579 (0.00311)
IO-SMC^2–grid/C (ICML'24)	0.08295	0.08529	0.01506 (0.00274)
PASOA–grid/J (ICML'24)	0.08988	0.08237	0.01599 (0.00344)
PASOA–grid/C (ICML'24)	0.08257	0.08679	0.01524 (0.00259)
TempoGeom-R/J	0.07995	0.08324	0.01415 (0.00239)
TempoGeom-R/C	0.07978	0.08823	0.01502 (0.00274)

Table 10: Mean realized first- and last-pair probe gaps, averaged over the 30 graph seeds and 15 dyads. The first decision is prior-only; later decisions may use acquired counts. These are descriptive decision diagnostics, not generating-parameter inputs.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design/family	First	Last	First	Last	First	Last
Rank-bound / Poly	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156
Rank-bound / Fourier	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156
TempoGeom-R/J / Poly	0.0156	0.0592	0.0156	0.0161	0.0156	0.0156
TempoGeom-R/J / Fourier	0.0156	0.0588	0.0156	0.0163	0.0156	0.0156
TempoGeom-R/C / Poly	0.0156	0.0592	0.0156	0.0159	0.0156	0.0156
TempoGeom-R/C / Fourier	0.0156	0.0585	0.0156	0.0160	0.0156	0.0156

Table 11: Polynomial, MLE: additional-scale confirmation, mean (sample SD) RMSE over ten fresh paired graph seeds per cell. Scale 16 belongs to the design prior but not the recovery-development grid; scale 48 also lies outside the discrete prior support. All settings remain frozen.

	$\lambda = 16$		$\lambda = 48$	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.085 (.015)	.135 (.033)	.098 (.032)	.123 (.023)
Fixed A/P	.086 (.015)	.106 (.021)	.098 (.033)	.099 (.019)
Fixed MC-EIG/J	.092 (.020)	.102 (.022)	.118 (.054)	.097 (.017)
Fixed MC-EIG/C	.085 (.015)	.135 (.033)	.098 (.032)	.123 (.023)
DAD/J (control)	.091 (.019)	.104 (.022)	.102 (.034)	.099 (.018)
DAD/C (control)	.085 (.015)	.119 (.027)	.099 (.031)	.110 (.021)
Step-DAD–grid/J (ICML'25)	.090 (.019)	.104 (.023)	.101 (.035)	.100 (.019)
Step-DAD–grid/C (ICML'25)	.084 (.014)	.120 (.027)	.099 (.031)	.110 (.021)
IO-SMC^2–grid/J (ICML'24)	.091 (.017)	.102 (.020)	.131 (.026)	.095 (.017)
IO-SMC^2–grid/C (ICML'24)	.087 (.016)	.118 (.027)	.104 (.043)	.109 (.021)
PASOA–grid/J (ICML'24)	.091 (.030)	.107 (.023)	.120 (.042)	.107 (.021)
PASOA–grid/C (ICML'24)	.082 (.015)	.114 (.022)	.103 (.042)	.109 (.018)
TempoGeom-R/J	.081 (.015)	.102 (.021)	.097 (.033)	.098 (.019)
TempoGeom-R/C	.083 (.015)	.125 (.030)	.097 (.034)	.117 (.021)

Table 12: Polynomial, Reg: additional-scale confirmation, mean (sample SD) RMSE over ten fresh paired graph seeds per cell. Scale 16 belongs to the design prior but not the recovery-development grid; scale 48 also lies outside the discrete prior support. All settings remain frozen.

	$\lambda = 16$		$\lambda = 48$	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.071 (.012)	.090 (.016)	.081 (.026)	.083 (.015)
Fixed A/P	.071 (.012)	.084 (.015)	.081 (.026)	.076 (.016)
Fixed MC-EIG/J	.077 (.014)	.081 (.014)	.091 (.028)	.074 (.017)
Fixed MC-EIG/C	.071 (.012)	.090 (.016)	.081 (.026)	.083 (.015)
DAD/J (control)	.077 (.014)	.082 (.014)	.081 (.027)	.075 (.016)
DAD/C (control)	.071 (.012)	.086 (.015)	.081 (.025)	.080 (.015)
Step-DAD–grid/J (ICML'25)	.075 (.015)	.082 (.015)	.083 (.027)	.075 (.016)
Step-DAD–grid/C (ICML'25)	.070 (.012)	.087 (.015)	.081 (.025)	.080 (.015)
IO-SMC^2–grid/J (ICML'24)	.073 (.013)	.083 (.014)	.098 (.026)	.074 (.016)
IO-SMC^2–grid/C (ICML'24)	.071 (.012)	.087 (.017)	.085 (.029)	.079 (.016)
PASOA–grid/J (ICML'24)	.077 (.018)	.081 (.013)	.085 (.028)	.077 (.016)
PASOA–grid/C (ICML'24)	.071 (.013)	.084 (.014)	.083 (.031)	.079 (.015)
TempoGeom-R/J	.069 (.013)	.082 (.014)	.081 (.026)	.076 (.017)
TempoGeom-R/C	.070 (.013)	.085 (.015)	.080 (.027)	.082 (.015)

Table 13: Fourier, MLE: additional-scale confirmation, mean (sample SD) RMSE over ten fresh paired graph seeds per cell. Scale 16 belongs to the design prior but not the recovery-development grid; scale 48 also lies outside the discrete prior support. All settings remain frozen.

	$\lambda = 16$		$\lambda = 48$	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.087 (.031)	.115 (.015)	.116 (.036)	.110 (.021)
Fixed A/P	.085 (.028)	.156 (.026)	.114 (.037)	.147 (.037)
Fixed MC-EIG/J	.084 (.029)	.114 (.016)	.132 (.038)	.108 (.021)
Fixed MC-EIG/C	.085 (.028)	.156 (.026)	.114 (.037)	.147 (.037)
DAD/J (control)	.093 (.037)	.109 (.014)	.156 (.051)	.106 (.019)
DAD/C (control)	.086 (.032)	.138 (.018)	.116 (.034)	.127 (.029)
Step-DAD–grid/J (ICML'25)	.091 (.036)	.110 (.014)	.144 (.035)	.106 (.019)
Step-DAD–grid/C (ICML'25)	.087 (.032)	.138 (.019)	.116 (.034)	.128 (.031)
IO-SMC^2–grid/J (ICML'24)	.096 (.031)	.112 (.014)	.133 (.042)	.108 (.022)
IO-SMC^2–grid/C (ICML'24)	.087 (.031)	.115 (.016)	.118 (.038)	.109 (.020)
PASOA–grid/J (ICML'24)	.096 (.026)	.115 (.025)	.146 (.043)	.120 (.025)
PASOA–grid/C (ICML'24)	.083 (.027)	.134 (.012)	.133 (.044)	.126 (.030)
TempoGeom-R/J	.081 (.030)	.110 (.016)	.116 (.036)	.110 (.021)
TempoGeom-R/C	.082 (.029)	.132 (.034)	.113 (.034)	.131 (.040)

Table 14: Fourier, Reg: additional-scale confirmation, mean (sample SD) RMSE over ten fresh paired graph seeds per cell. Scale 16 belongs to the design prior but not the recovery-development grid; scale 48 also lies outside the discrete prior support. All settings remain frozen.

	$\lambda = 16$		$\lambda = 48$	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.064 (.018)	.092 (.010)	.076 (.020)	.087 (.014)
Fixed A/P	.063 (.016)	.101 (.012)	.075 (.020)	.098 (.012)
Fixed MC-EIG/J	.062 (.017)	.090 (.011)	.081 (.024)	.086 (.014)
Fixed MC-EIG/C	.063 (.016)	.101 (.012)	.075 (.020)	.098 (.012)
DAD/J (control)	.068 (.019)	.088 (.012)	.088 (.019)	.086 (.013)
DAD/C (control)	.063 (.016)	.098 (.012)	.075 (.021)	.093 (.014)
Step-DAD–grid/J (ICML'25)	.067 (.017)	.089 (.012)	.084 (.019)	.086 (.013)
Step-DAD–grid/C (ICML'25)	.063 (.017)	.097 (.012)	.075 (.021)	.092 (.015)
IO-SMC^2–grid/J (ICML'24)	.068 (.018)	.090 (.010)	.088 (.024)	.087 (.013)
IO-SMC^2–grid/C (ICML'24)	.064 (.018)	.092 (.011)	.076 (.021)	.087 (.014)
PASOA–grid/J (ICML'24)	.070 (.019)	.091 (.014)	.077 (.022)	.091 (.015)
PASOA–grid/C (ICML'24)	.065 (.016)	.095 (.015)	.082 (.025)	.092 (.012)
TempoGeom-R/J	.062 (.017)	.088 (.011)	.076 (.020)	.087 (.014)
TempoGeom-R/C	.062 (.017)	.090 (.015)	.074 (.020)	.087 (.016)

Table 15: Doubled node heterogeneity, MLE, scale 32: mean (sample SD) RMSE over ten fresh paired graph seeds per family. The clock contrasts and affinity coefficients are doubled; the common clock level, estimator and design prior are unchanged. This limited shift is not a general robustness guarantee.

	Polynomial		Fourier	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.097 (.020)	.146 (.029)	.101 (.034)	.118 (.020)
Fixed A/P	.100 (.022)	.114 (.023)	.101 (.033)	.158 (.031)
Fixed MC-EIG/J	.114 (.021)	.110 (.020)	.114 (.038)	.115 (.020)
Fixed MC-EIG/C	.097 (.020)	.146 (.029)	.101 (.033)	.158 (.031)
DAD/J (control)	.113 (.022)	.112 (.022)	.118 (.041)	.111 (.021)
DAD/C (control)	.099 (.019)	.127 (.024)	.100 (.029)	.139 (.025)
Step-DAD–grid/J (ICML'25)	.097 (.014)	.112 (.022)	.113 (.043)	.112 (.020)
Step-DAD–grid/C (ICML'25)	.099 (.019)	.127 (.025)	.101 (.029)	.140 (.029)
IO-SMC^2–grid/J (ICML'24)	.117 (.027)	.105 (.019)	.115 (.038)	.111 (.018)
IO-SMC^2–grid/C (ICML'24)	.100 (.020)	.122 (.023)	.102 (.033)	.118 (.019)
PASOA–grid/J (ICML'24)	.112 (.023)	.118 (.021)	.115 (.041)	.129 (.029)
PASOA–grid/C (ICML'24)	.098 (.019)	.123 (.027)	.105 (.030)	.133 (.026)
TempoGeom-R/J	.102 (.022)	.113 (.022)	.101 (.034)	.118 (.020)
TempoGeom-R/C	.099 (.020)	.138 (.027)	.101 (.032)	.146 (.033)

Table 16: Doubled node heterogeneity, Reg, scale 32: mean (sample SD) RMSE over ten fresh paired graph seeds per family. The clock contrasts and affinity coefficients are doubled; the common clock level, estimator and design prior are unchanged. This limited shift is not a general robustness guarantee.

	Polynomial		Fourier	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	.097 (.018)	.124 (.023)	.110 (.034)	.129 (.026)
Fixed A/P	.099 (.018)	.113 (.022)	.111 (.036)	.152 (.033)
Fixed MC-EIG/J	.117 (.016)	.110 (.021)	.120 (.034)	.124 (.026)
Fixed MC-EIG/C	.097 (.018)	.124 (.023)	.111 (.036)	.152 (.033)
DAD/J (control)	.109 (.017)	.111 (.021)	.129 (.049)	.122 (.024)
DAD/C (control)	.098 (.017)	.119 (.020)	.111 (.036)	.144 (.030)
Step-DAD–grid/J (ICML'25)	.102 (.020)	.111 (.020)	.125 (.046)	.123 (.024)
Step-DAD–grid/C (ICML'25)	.098 (.017)	.119 (.020)	.111 (.036)	.145 (.031)
IO-SMC^2–grid/J (ICML'24)	.114 (.022)	.108 (.020)	.126 (.036)	.125 (.023)
IO-SMC^2–grid/C (ICML'24)	.099 (.016)	.115 (.022)	.110 (.034)	.129 (.025)
PASOA–grid/J (ICML'24)	.111 (.018)	.117 (.021)	.122 (.038)	.131 (.029)
PASOA–grid/C (ICML'24)	.100 (.018)	.119 (.023)	.112 (.037)	.141 (.033)
TempoGeom-R/J	.100 (.020)	.112 (.021)	.111 (.033)	.128 (.026)
TempoGeom-R/C	.099 (.019)	.125 (.020)	.110 (.035)	.144 (.032)

A.5 Residual-head prediction: construction and evaluation

The auxiliary TempoGeom-H head maps causal history features to affinity/clock coefficients and returns a transition probability p_H. We add a zero-initialized residual network r_ω and form

$$p_{HR} = \mathrm{sigmoid}\left(\mathrm{logit}(p_H) + \frac{h}{h+.5}\, 5\tanh\{r_\omega(z)/5\}\right). \tag{24}$$

The inputs z contain the existing history embedding, three causal transition-proxy features, the current state, and $\log(1+h)$. The physical branch and residual branch are trained jointly with prediction NLL and λ_r times the mean squared logit correction. The bounded residual vanishes at zero gap. The resulting predictor is evaluated for endpoint-state accuracy; a Markov-semigroup interpretation and identifiable clock estimation require separate analysis. Predictive NLL therefore does not establish the information results in Theorems 1 or 5.

To control feature access, an additional Proxy-MLP uses the same history width and all added proxy inputs. It is trained independently and reported alongside TempoGeom-H, the independently tuned MLP, the count mixture and all five external adapters, each refitted on the confirmation graphs. These are task-adapted predictive comparisons, separate from the acquisition benchmark.

A single additive-clock development graph with seed 88000 selects the residual and Proxy-MLP configurations by chronological validation NLL. Residual candidates use learning rates $.001, .003, .01$, width $C = 16$, and $\lambda_r \in \{.001, .01\}$; Proxy-MLP uses the same three rates and width. The study summary reports nine full-budget development trials. The selected residual uses rate $.01$ and $\lambda_r = .01$; Proxy-MLP uses rate $.003$. Other models reuse the configurations selected in the predictive study of Appendix AB, without confirmation-set retuning. Hyperparameter-search sizes therefore differ across models. The parent rate head uses its selected learning rate $.003$, whereas HR uses $.01$ and a different explicit regularizer. Despite shared base initialization, this compares differently tuned predictive pipelines rather than isolating a residual-architecture effect.

Held-out confirmation uses seeds 88010–88014 in each of the homogeneous, additive and nonadditive conditions, with 32 nodes and 496 dyads. All ten models are newly trained on each graph, yielding 150 fits. The chronological splits $[0, 40]$, $[40, 50]$, $[50, 60]$, half-unit observations, 1,200-update maximum, batch size 128, validation every 100 updates and four-checkpoint early stopping are shared across models. Test labels never train or select a model. Each graph supplies one selected checkpoint; repeated endpoints do not inflate the sampling unit. The primary predictive family contains 27 paired NLL comparisons (the residual against nine other rows in three conditions), with Holm adjustment across all 27. With only five paired graph seeds per condition, uncertainty remains substantial. The finite-update cap is not a convergence certificate.

All 150 confirmation fits complete. TempoGeom-HR has mean NLL .40820, .37832, and .37207 in the homogeneous, additive, and nonadditive conditions. These are lower than every recent external adapter's mean in every condition, and 13/15 corresponding differences are supported by the planned paired-t tests after the single 27-test Holm adjustment. With only five paired graphs per condition, this is a parametric, assumption-sensitive result; the post-hoc sign-flip sensitivity analysis is reported separately below. The exceptions are homogeneous FreeDyG (adjusted $p = .0517$) and additive CRAFT ($p = .0609$); neither meets the adjusted significance threshold, and nonsignificance does not establish equivalence.

The H head gives .40991, .37943, and .37209. HR's three mean reductions are .001702, .001111, and .000019; none is supported after adjustment ($p = .525, .713, 1$, respectively). The independently tuned MLP has smaller homogeneous-condition mean NLL, .40688 versus .40820 for HR. Proxy-MLP is also marginally lower there (.40817), with adjusted $p = 1$. HR has the lowest reported mean in the other two conditions, but no adjusted comparison with either MLP supports a superiority claim. These endpoint-prediction differences do not measure clock-representation quality.

Optimization was evaluated within a fixed update budget. Seventy-two of 150 fits reach 1,200 updates, including all 15 H fits, 8 HR fits, 12 independently tuned MLP fits and 9 Proxy-MLP fits. All capped outcomes are included. A completed fit at the update cap indicates budget-limited execution rather than verified convergence. The selected residual checkpoint's tighter-kernel recheck changes test NLL by at most 9.56×10^{-8} and an individual probability by at most 2.41×10^{-5}. Kernel sensitivity, graph-sampling uncertainty and optimizer convergence are distinct diagnostics.

The complete tables report NLL, Brier score, AUC, average precision and calibration error, with sample SDs over graph seeds, followed by every primary NLL comparison. These predictive outcomes are not used to choose between the clock-acquisition methods or to retune the frozen configurations.

Table 17: Held-out predictive NLL: mean (sample SD) over five paired graph seeds per condition. All rows and conditions are retained. Lower is better.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.4312 (0.0205)	0.4333 (0.0188)	0.4311 (0.0269)
GraSSP–panel (AAAI'24)	0.4283 (0.0200)	0.4295 (0.0228)	0.4246 (0.0277)
FreeDyG–state (ICLR'24)	0.4300 (0.0243)	0.4246 (0.0260)	0.4215 (0.0321)
CRAFT–state (NeurIPS'25)	0.4227 (0.0240)	0.4012 (0.0270)	0.3986 (0.0278)
TAMI–GraphMixer (NeurIPS'25)	0.4285 (0.0225)	0.4052 (0.0220)	0.4032 (0.0236)
TempoGeom-H	0.4099 (0.0238)	0.3794 (0.0234)	0.3721 (0.0280)
Independently tuned MLP	0.4069 (0.0226)	0.3801 (0.0228)	0.3741 (0.0253)
Count mixture	0.4158 (0.0221)	0.3844 (0.0240)	0.3732 (0.0281)
Proxy-MLP control	0.4082 (0.0224)	0.3813 (0.0235)	0.3736 (0.0266)
TempoGeom-HR	0.4082 (0.0233)	0.3783 (0.0237)	0.3721 (0.0294)

Table 18: Held-out predictive BRIER: mean (sample SD) over five paired graph seeds per condition. All rows and conditions are retained. Lower is better.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.1347 (0.0075)	0.1352 (0.0077)	0.1343 (0.0105)
GraSSP–panel (AAAI'24)	0.1334 (0.0077)	0.1335 (0.0095)	0.1316 (0.0107)
FreeDyG–state (ICLR'24)	0.1336 (0.0088)	0.1313 (0.0100)	0.1302 (0.0119)
CRAFT–state (NeurIPS'25)	0.1316 (0.0088)	0.1247 (0.0106)	0.1241 (0.0102)
TAMI–GraphMixer (NeurIPS'25)	0.1335 (0.0083)	0.1259 (0.0089)	0.1248 (0.0094)
TempoGeom-H	0.1276 (0.0083)	0.1173 (0.0087)	0.1144 (0.0104)
Independently tuned MLP	0.1268 (0.0080)	0.1173 (0.0083)	0.1153 (0.0096)
Count mixture	0.1295 (0.0079)	0.1191 (0.0087)	0.1151 (0.0103)
Proxy-MLP control	0.1270 (0.0079)	0.1179 (0.0086)	0.1153 (0.0099)
TempoGeom-HR	0.1272 (0.0083)	0.1168 (0.0089)	0.1144 (0.0109)

Table 19: Held-out predictive AUC: mean (sample SD) over five paired graph seeds per condition. All rows and conditions are retained. Higher is better.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.8547 (0.0055)	0.8542 (0.0073)	0.8573 (0.0113)
GraSSP–panel (AAAI'24)	0.8553 (0.0033)	0.8579 (0.0093)	0.8611 (0.0143)
FreeDyG–state (ICLR'24)	0.8449 (0.0128)	0.8500 (0.0113)	0.8500 (0.0203)
CRAFT–state (NeurIPS'25)	0.8577 (0.0064)	0.8784 (0.0101)	0.8816 (0.0107)
TAMI–GraphMixer (NeurIPS'25)	0.8506 (0.0038)	0.8734 (0.0105)	0.8778 (0.0083)
TempoGeom-H	0.8689 (0.0038)	0.8911 (0.0084)	0.8956 (0.0101)
Independently tuned MLP	0.8714 (0.0028)	0.8905 (0.0091)	0.8961 (0.0098)
Count mixture	0.8644 (0.0037)	0.8883 (0.0079)	0.8958 (0.0105)
Proxy-MLP control	0.8697 (0.0029)	0.8902 (0.0077)	0.8957 (0.0098)
TempoGeom-HR	0.8698 (0.0039)	0.8917 (0.0090)	0.8961 (0.0104)

Table 20: Held-out predictive AP: mean (sample SD) over five paired graph seeds per condition. All rows and conditions are retained. Higher is better.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.7233 (0.0333)	0.7220 (0.0272)	0.7291 (0.0272)
GraSSP–panel (AAAI'24)	0.7219 (0.0321)	0.7325 (0.0242)	0.7366 (0.0255)
FreeDyG–state (ICLR'24)	0.7038 (0.0279)	0.7180 (0.0199)	0.7266 (0.0355)
CRAFT–state (NeurIPS'25)	0.7323 (0.0214)	0.7848 (0.0271)	0.7959 (0.0234)
TAMI–GraphMixer (NeurIPS'25)	0.7171 (0.0346)	0.7795 (0.0244)	0.7887 (0.0230)
TempoGeom-H	0.7479 (0.0266)	0.8024 (0.0221)	0.8168 (0.0257)
Independently tuned MLP	0.7499 (0.0292)	0.8008 (0.0281)	0.8131 (0.0258)
Count mixture	0.7375 (0.0277)	0.7956 (0.0205)	0.8135 (0.0276)
Proxy-MLP control	0.7489 (0.0267)	0.8003 (0.0210)	0.8152 (0.0249)
TempoGeom-HR	0.7464 (0.0266)	0.8028 (0.0251)	0.8158 (0.0246)

Table 21: Held-out predictive ECE10: mean (sample SD) over five paired graph seeds per condition. All rows and conditions are retained. Lower is better.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.0308 (0.0051)	0.0393 (0.0047)	0.0396 (0.0036)
GraSSP–panel (AAAI'24)	0.0422 (0.0047)	0.0462 (0.0035)	0.0417 (0.0044)
FreeDyG–state (ICLR'24)	0.0271 (0.0079)	0.0249 (0.0088)	0.0192 (0.0066)
CRAFT–state (NeurIPS'25)	0.0323 (0.0076)	0.0313 (0.0043)	0.0330 (0.0081)
TAMI–GraphMixer (NeurIPS'25)	0.0315 (0.0074)	0.0289 (0.0049)	0.0377 (0.0073)
TempoGeom-H	0.0157 (0.0066)	0.0147 (0.0017)	0.0145 (0.0052)
Independently tuned MLP	0.0157 (0.0054)	0.0158 (0.0031)	0.0211 (0.0089)
Count mixture	0.0179 (0.0039)	0.0141 (0.0028)	0.0148 (0.0036)
Proxy-MLP control	0.0149 (0.0038)	0.0168 (0.0072)	0.0183 (0.0030)
TempoGeom-HR	0.0100 (0.0026)	0.0154 (0.0033)	0.0165 (0.0041)

Table 22: Complete predictive NLL comparisons, TempoGeom-HR minus comparator. Negative differences favor HR. Intervals are descriptive unadjusted 95% paired intervals; the final column adjusts all 27 two-sided tests together. Five graph-seed blocks do not certify equivalence or convergence.

Comparator	Difference	95% interval	p	p_{Holm}
Homogeneous				
ℓ_1LD–panel (ICLR'26)	-0.0230	[-0.0296, -0.0164]	0.000643	0.0108
GraSSP–panel (AAAI'24)	-0.0201	[-0.0258, -0.0143]	0.000634	0.0108
FreeDyG–state (ICLR'24)	-0.0218	[-0.0319, -0.0116]	0.00398	0.0517
CRAFT–state (NeurIPS'25)	-0.0145	[-0.0184, -0.0105]	0.000524	0.00996
TAMI–GraphMixer (NeurIPS'25)	-0.0203	[-0.0222, -0.0184]	7.51e-06	0.000203
TempoGeom-H	-0.0017	[-0.0034, +0.0000]	0.0525	0.525
Independently tuned MLP	+0.0013	[-0.0012, +0.0039]	0.223	1
Count mixture	-0.0076	[-0.0102, -0.0049]	0.00136	0.019
Proxy-MLP control	+0.0000	[-0.0020, +0.0021]	0.968	1
Additive				
ℓ_1LD–panel (ICLR'26)	-0.0550	[-0.0632, -0.0468]	4.98e-05	0.0012
GraSSP–panel (AAAI'24)	-0.0512	[-0.0563, -0.0460]	1.01e-05	0.000262
FreeDyG–state (ICLR'24)	-0.0462	[-0.0561, -0.0364]	0.000198	0.00417
CRAFT–state (NeurIPS'25)	-0.0229	[-0.0346, -0.0113]	0.00547	0.0609
TAMI–GraphMixer (NeurIPS'25)	-0.0269	[-0.0345, -0.0193]	0.000587	0.0106
TempoGeom-H	-0.0011	[-0.0024, +0.0002]	0.0793	0.713
Independently tuned MLP	-0.0018	[-0.0040, +0.0004]	0.0908	0.726
Count mixture	-0.0060	[-0.0090, -0.0030]	0.00507	0.0609
Proxy-MLP control	-0.0030	[-0.0070, +0.0010]	0.104	0.726
Nonadditive				
ℓ_1LD–panel (ICLR'26)	-0.0591	[-0.0663, -0.0519]	2.19e-05	0.000548
GraSSP–panel (AAAI'24)	-0.0525	[-0.0622, -0.0428]	0.000113	0.00248
FreeDyG–state (ICLR'24)	-0.0494	[-0.0569, -0.0418]	5.43e-05	0.00125
CRAFT–state (NeurIPS'25)	-0.0265	[-0.0355, -0.0175]	0.0012	0.018
TAMI–GraphMixer (NeurIPS'25)	-0.0312	[-0.0394, -0.0229]	0.000467	0.00934
TempoGeom-H	-0.0000	[-0.0047, +0.0047]	0.991	1
Independently tuned MLP	-0.0021	[-0.0076, +0.0035]	0.359	1
Count mixture	-0.0012	[-0.0048, +0.0025]	0.429	1
Proxy-MLP control	-0.0015	[-0.0065, +0.0035]	0.448	1

Post-hoc small-sample sensitivity. The primary predictive analysis uses paired-t tests with Holm adjustment across 27 comparisons. To assess sensitivity to distributional assumptions with five independent graph seeds, this post-hoc analysis enumerates all within-pair sign exchanges of each graph-level NLL difference d_i. With $n = 5$, it computes

$$p_{\mathrm{sf}} = 2^{-5} \sum_{\epsilon \in \{-1,1\}^5} \mathbf{1}\left\{ \left| \sum_{i=1}^{5} \epsilon_i d_i \right| \geq \left| \sum_{i=1}^{5} d_i \right| \right\}.$$

Validity of this exact test requires appropriate within-pair exchangeability or joint sign symmetry under the null. The observed sign pattern and its negative imply $p_{\mathrm{sf}} \geq 2/32 = .0625$. None of the 27 unadjusted sign-flip p-values is below .05 (Table 23); Holm adjustment cannot change that conclusion. The planned parametric and post-hoc sign-flip tests rely on different null assumptions; their contrast illustrates the assumption sensitivity of the predictive findings. This analysis uses the same models, checkpoints and fits and is separate from the primary testing family. More independent graph realizations, not more endpoints from the same graphs, would be needed for stronger predictive evidence.

Table 23: Post-hoc exact two-sided graph-level sign-flip sensitivity for HR minus each comparator. All 27 comparisons use five paired graphs and all 32 sign patterns. Values are unadjusted p-values from the post-hoc analysis, reported separately from the planned paired-t/Holm results.

Comparator	Homogeneous	Additive	Nonadditive
CRAFT–state (NeurIPS'25)	0.0625	0.0625	0.0625
Count mixture	0.0625	0.0625	0.4375
FreeDyG–state (ICLR'24)	0.0625	0.0625	0.0625
GraSSP–panel (AAAI'24)	0.0625	0.0625	0.0625
ℓ_1LD–panel (ICLR'26)	0.0625	0.0625	0.0625
Proxy-MLP control	0.9375	0.2500	0.5000
Independently tuned MLP	0.2500	0.1250	0.4375
TAMI–GraphMixer (NeurIPS'25)	0.0625	0.0625	0.0625
TempoGeom-H	0.0625	0.1250	1.0000

A.6 Reported numerical checks and artifact coverage

The following execution counts and numerical diagnostics are drawn from the archived study-level audit reports. The artifact-coverage paragraph below distinguishes these reports from the underlying execution records.

The primary acquisition summary reports 5,040 completed fit labels on 2,520 acquisition labels and 2,179 distinct acquired datasets. The additional-scale and doubled-heterogeneity checks complete another 1,120 and 560 fits, on 488 and 241 distinct acquired datasets respectively. Development and estimator calibration contribute 2,880 development fitting outcomes. Thus 9,600 graph-estimator fits are reported for this study, with tuning outcomes kept separate from confirmation. None has a failed selected optimizer; the only primary parameter-boundary exceptions are the two unregularized PASOA/J cases reported above. No secondary fit reaches a parameter boundary. Maximum realized-schedule log-probability discrepancy between independent ODE generation and likelihood quadrature is 2.39×10^{-11}. Every recorded acquisition has 9,000 transitions, minimum gap at least $1/64$, and minimum anchor singular value at least $.50435$.

The study account states that each RMSE was recomputed from the saved graph estimate and generating parameters, which are supplied only to simulation and scoring, never to the planner. Physical uniform streams match across methods within each graph/condition. The acquired-data identity check includes the affinity family as well as schedules, counts, graph and parameters. The reported read-only replay of 180 risk-planner decisions exactly matches their logged actions, posterior weights and risk values using only recorded past counts and the declared prior.

The predictive summary reports nine development trials and 150 newly trained confirmation fits, with confirmation NLL and Brier scores recomputed from saved per-record predictions. Targets and record indices agree across methods on each graph. A relocated-dependency audit reacquires and refits two development graph cases with both estimators: raw counts, transition probabilities and

estimates match exactly in all four fits. A separate relocated checkpoint replay reproduces the first 256 predictions of a selected HR model exactly. These reported deterministic replay checks are not additional independent scientific replications. The reported graph-estimator audit checks exact agreement with MLE at zero precision, centered-penalty invariance under a common shift, all 24 gradient coordinates in both affinity families against finite differences, and full-inverse/Schur-inverse agreement. The predictive audit checks zero-initialized agreement with the rate-head probability, a zero-gap residual, and finite gradients in both branches. A tighter kernel tolerance is evaluated at every selected residual checkpoint without using it to select a different model. These are reported automated implementation checks, subject to the artifact coverage below.

The study account reports that an input-fingerprint check halted the predictive launch *before* confirmation data generation or fitting. During development, shared raw-data generation was switched from concurrent writes to sequential execution. The selected configurations remained unchanged and were re-hashed with the sequential-generation code before the first confirmation fit; both acquisition operating points were frozen before confirmation, without using test outcomes. These chronology claims are based on the archived audit reports; the original freeze records and stopped-launch log are not included.

Artifact coverage. The artifact contains the numerical tables, aggregate reporting CSVs and audit reports describing development selection, execution checks and freeze-file consistency. These preserve reported protocol and audit information, not the underlying development trials or original freeze files. The primary acquisition and predictive implementations, frozen dependencies, raw observations, per-fit outputs, selected checkpoints and adaptive traces are not included, so end-to-end reproduction and record-level verification are not supported by this package. The summary checker verifies calculations supported by the aggregate CSVs; it cannot rederive paired observations, the complete 288-test primary family, exact sign-flip probabilities or saved-parameter equality. `STUDY_PATHS.json` maps the available reports and missing primary inputs, separately from the supporting-study records. Audit summaries and successful checks of those separate cohorts do not establish primary execution or freeze chronology.

B NUISANCE DIRECTIONS INVISIBLE AT THE ENDPOINTS

Singular-information inference provides the broader statistical context (Rotnitzky et al., 2000). Here the singular limit comes from affinity parameters invisible at observation endpoints; the proof tracks their rescaled dynamical effects before projecting the clock score.

Corollary 8 (Regular endpoint case)**.** *If J_0 has full column rank, K_0 is empty. With Π_J projecting onto $W^{1/2}J_0$,*

$$I_{v,\mathrm{eff}} = \lambda^{-2}\mathcal{E}^\top\mathcal{E} + O(\lambda^{-3}), \qquad \mathcal{E} = (I - \Pi_J)W^{1/2}G. \tag{25}$$

In particular, $Gh \in \mathrm{col}(J_0)$ implies $O(\lambda^{-4})$ information, an upper order rather than a universal matching rate.

B.1 DERIVATION OF THE NUISANCE EXPANSION

Under the endpoint regularity assumptions in Section 3, the boundary-layer calculation of Appendix K applies uniformly with first parameter derivatives:

$$y_{r,a} = p_r - \varepsilon A_r + O(\varepsilon^2), \quad \partial_v y = \varepsilon G + O(\varepsilon^2), \quad N_\varepsilon := \partial_\theta y = J_0 + \varepsilon J_1 + O(\varepsilon^2), \qquad \varepsilon = \lambda^{-1}.$$

The nuisance derivatives of $A = p'/q_0$ include derivatives of q_0 whenever the clock factor $h(t;\theta)$ depends on affinity. Omitting that dependence would give the wrong J_1. Starting-state effects are exponentially small under fixed positive gaps. Conditional score variances add along a chain because successive scores are martingale differences. Equivalently, one may stack starting-state rows with their probability weights; their common leading limit has the record-space weight W used in the theorem.

B.2 PROOF OF THEOREM 1

Let J_0 have rank r among d nuisance coordinates. Choose a $d \times r$ complement L_0 and a $d \times (d-r)$ basis K_0 of its nullspace. For each $\varepsilon > 0$ the matrix $T_\varepsilon = [L_0, \varepsilon^{-1}K_0]$ is invertible. Multiplying

the nuisance derivative by it does not change the nuisance score space, and

$$N_\varepsilon T_\varepsilon = \big[J_0 L_0 + O(\varepsilon),\ J_1 K_0 + O(\varepsilon)\big].$$

For fixed conditioned starting states, whitening uses $W_\varepsilon^{1/2} = W^{1/2} + O(\varepsilon)$. For expected conditional information, stack both starting states. Let D_ε replicate each record into the two rows with weights equal to the square roots of their starting-state probabilities. These probabilities may depend on λ; nevertheless $D_\varepsilon^\top D_\varepsilon = I$. Uniformity of the endpoint expansions implies

$$U_\varepsilon = D_\varepsilon U + O(\varepsilon), \qquad C_\varepsilon = \varepsilon D_\varepsilon W^{1/2} G + O(\varepsilon^2),$$

where U_ε is the rescaled, whitened nuisance derivative and C_ε is the whitened clock derivative. In the fixed-start case take $D_\varepsilon = I$. Since D_ε is an isometry, the nonzero singular values of $D_\varepsilon U$ equal those of U. Full column rank of U therefore gives the uniform projector approximation

$$\Pi_{U_\varepsilon} = D_\varepsilon \Pi_U D_\varepsilon^\top + O(\varepsilon).$$

This follows from the formula $A(A^\top A)^{-1}A^\top$ in a uniformly full-rank neighborhood. It inverts a well-conditioned limiting Gram matrix, not the original singular-limit nuisance block. Efficient information is the clock residual Gram matrix:

$$\begin{aligned} I_{v,\mathrm{eff}} &= C_\varepsilon^\top (I - \Pi_{U_\varepsilon}) C_\varepsilon \\ &= \varepsilon^2 G^\top W^{1/2} (I - \Pi_U) W^{1/2} G + O(\varepsilon^3). \end{aligned}$$

If $\mathcal{E}_* h \neq 0$, the leading coefficient is positive. If $\mathcal{E}_* h = 0$, the residual derivative itself is $O(\varepsilon^2)$, so its squared norm is $O(\varepsilon^4)$. These statements prove the theorem. They do not assert a positive quartic coefficient. □

Scope of the sufficient rank condition. The space in (5) is invariant to the choice of complement and nullspace basis: changing the complement adds a J_0-null direction and hence leaves $J_0 L_0$ unchanged up to invertible transformation; changing either basis only changes column coordinates. Full rank of J_0 is recovered with an empty K_0. If a nonzero linear combination of the columns of $J_1 K_0$ belongs to $\mathrm{col}(J_0)$, the two-scale matrix is deficient and second or higher probability derivatives may determine the limiting nuisance space. The theorem does not replace that analysis by dropping parameters. Likewise, a nuisance class with freely varying preferences at every endpoint can consume the whole clock-score space. The result is a local finite-dimensional information statement, not global recovery of nonsmooth latent embeddings or neural weights.

B.3 Proof of the scalar phase distinction

This scalar comparison uses $q = \lambda e^v$, $\eta = \theta_0 + f + \theta_1 \phi$, aligned endpoints $(1, 2, 3, 4, 5, 6)$ and staggered endpoints $(0.5, 1.5, 2.5, 4, 5, 6)$. These are distinct from the equal-gap-multiset shared-graph schedules in Corollary 2. At $\theta = 0$ and $q_0 = e^v = 1$, let $j_r = p_r(1 - p_r) > 0$ and $b_r = f'(t_r)$. Then

$$J_0 = (j, 0), \quad G = j \odot b, \quad J_1 K_0 = -j \odot b \odot \cos(2\pi t).$$

The derivative of the $\sin^2$ term in (7) is zero at half-integer endpoints. Its higher derivatives need not vanish. For the aligned design, $\cos(2\pi t) = 1$, so G belongs to the limiting nuisance span. The two nuisance columns are independent because b_r is not constant.

For the staggered design, the first three cosines are -1 and the last three are 1. If G belonged to that span, dividing each row by j_r would give $b_r = a + d\, b_r \cos(2\pi t_r)$ for fixed a, d. Distinct slopes among the first three points force $d = -1$ and $a = 0$. A nonzero slope at a last-three point then contradicts this equation. Thus the projected clock coefficient is strictly positive. This algebraic argument establishes the distinction independently of numerical singular-value thresholds.

Independent information computation. For an interval $[u, t]$ with constant total clock $q = \lambda$, the endpoint probability is computed as

$$y_a(t) = e^{-q(t-u)} a + \int_0^{q(t-u)} e^{-x} p(t - x/q)\, dx.$$

Its log-clock derivative is evaluated by differentiating this expression, including the moving upper limit:

$$\partial_v y_a = \int_0^{q(t-u)} e^{-x} p'(t - x/q) \frac{x}{q} dx + q(t-u) e^{-q(t-u)} \{p(u) - a\}.$$

Nuisance derivatives integrate $p(1-p)(1,\phi)$ with the same weight. The numerical study uses balanced conditional start-state exposures for each of the six intervals, assigning each start weight $1/2$. This is a specified conditional information design, not an estimate from six noisy realized transitions. Quadrature is truncated at integrated lag 60 only when the remaining exponential tail is negligible at the stated tolerances. QR projection of the actual sensitivity columns avoids subtracting nearly equal Schur-complement terms. The reported log slopes are descriptive slopes over $\lambda \in \{256, 512, 1024, 2048\}$, not proofs of a universal exact exponent.

C Proof of the all-orders invisible-nuisance construction

Proposition 9 (All-orders invisible nuisance)**.** *For any finite collection of nonoverlapping positive gaps $[u_r, t_r]$, take total clock $q = \lambda e^v$ at $v = 0$ and a smooth equilibrium bounded away from zero and one. There is a finite-dimensional smooth affinity nuisance family, with every time derivative of its parameter sensitivities zero near all endpoints, whose efficient conditional clock information satisfies*

$$I_{v,\mathrm{eff}} = \sum_r \frac{w_{r,0} w_{r,1}}{w_{r,0} + w_{r,1}} (\lambda \Delta_r)^2 e^{-2\lambda\Delta_r}, \qquad w_{r,a} = \frac{\mu_{r,a}}{y_{r,a}(1 - y_{r,a})}. \tag{26}$$

Here $\Delta_r = t_r - u_r$, $y_{r,a}$ is the connectedness probability conditional on start a, and normalized start-state exposures satisfy $\mu_{r,0} + \mu_{r,1} = 1$, with both weights bounded away from zero. Thus information is $\Theta(\lambda^2 e^{-2\lambda\Delta_{\min}})$ even for a moving equilibrium.

Fix finitely many nonoverlapping conditioned transition intervals $[u_r, t_r]$. For each interval choose a nonnegative C^∞ bump ϕ_r supported strictly inside it and positive on a subinterval; extend it by zero near both boundaries. Let

$$\eta(t;\theta) = \eta_0(t) + \sum_r \theta_r \phi_r(t), \qquad p(t;\theta) = \mathrm{sigmoid}\{\eta(t;\theta)\}, \qquad q = \lambda e^v.$$

The base η_0 can be any smooth bounded moving affinity on this finite horizon. A sufficiently small fixed parameter neighborhood keeps p bounded away from zero and one. Every time derivative of $\partial_{\theta_r} p$ vanishes in an endpoint neighborhood. Thus these directions are invisible not only to J_0, J_1, but to every algebraic endpoint term obtained by repeated boundary-layer expansion.

Conditioning on the starting state removes parameter dependence of that initial state from the score. The nuisance derivative for interval r is zero for other interval bumps and, for either starting state, equals

$$b_r(\lambda) = \int_{u_r}^{t_r} \lambda e^{-\lambda(t_r - s)} p_0(s)(1 - p_0(s)) \phi_r(s) ds > 0.$$

Its small size does not remove its direction from a local Fisher projection. The two connectedness probabilities satisfy the exact identity $y_{r,1} - y_{r,0} = e^{-\lambda\Delta_r}$. Consequently their derivatives with respect to v obey $g_{r,1} - g_{r,0} = -\lambda\Delta_r e^{-\lambda\Delta_r}$. The nuisance column for this interval spans $(1,1)^\top$. Profiling the weighted two-row clock score gives

$$\min_z \{w_{r,0}(g_{r,0} - z)^2 + w_{r,1}(g_{r,1} - z)^2\} = \frac{w_{r,0} w_{r,1}}{w_{r,0} + w_{r,1}} (g_{r,0} - g_{r,1})^2.$$

Independent interval-specific nuisance coordinates permit this minimization separately in each interval. Summing proves (26). With finitely many gaps, endpoint probabilities bounded away from zero and one, and normalized start exposures bounded below, the harmonic prefactors remain bounded above and below by positive constants. The shortest gap determines the exponential order stated in Proposition 9. If one start state is never exposed in an interval, that interval instead has zero residual information after profiling its free bump.

This is a pointwise local Fisher statement at each finite λ. The nuisance velocities needed to imitate a clock score can grow as $1/b_r(\lambda)$. It does not assert uniform local asymptotic normality or indistinguishability of fixed-size clock alternatives within a λ-independent bounded nuisance perturbation. Such uniform or finite-sample claims require additional control of the parameter neighborhood.

The construction uses smooth but nonanalytic compactly supported functions. It does not contradict a full-rank finite-jet theorem or establish the same pathology for every analytic finite-dimensional family. It shows that smooth endpoint expansions alone do not justify deleting directions whose leading jets vanish. A numerical audit uses three unit intervals, $\eta_0(t) = 0.2t$, balanced starting-state weights, and the standard bump $\exp\{-1/[x(1-x)]\}$ on each interval's middle half, where $x \in (0,1)$ is its rescaled coordinate. Direct sensitivity projection is compared with (26) for $\lambda = 1, 2, 4, 8, 16$. Larger-clock values use the stable profile formula, since subtracting nearly collinear score columns can obscure exponentially small information in floating point.

D MATCHING INFORMATION RATES AND SHARED-GRAPH PHASE DESIGN

D.1 A SECOND RESIDUAL COEFFICIENT

Theorem 1 gives a quartic upper order when the first residual vanishes. The following refinement identifies when that order is attained. It concerns a rescaled nuisance matrix, so its columns are well-conditioned in the limit even when the original nuisance derivatives are not. We prove Lemma 10, including its matrix and expected-conditional-information forms.

Lemma 10 (Matching residual coefficient)**.** *Let the whitened clock derivative and an invertibly rescaled whitened nuisance matrix satisfy*

$$C_\varepsilon = \varepsilon C_1 + \varepsilon^2 C_2 + O(\varepsilon^3), \qquad U_\varepsilon = U_0 + \varepsilon U_1 + O(\varepsilon^2),$$

with U_0 of full column rank. For a fixed clock direction h with $C_1 h = U_0 a_h$, define $R_2 h = (I - \Pi_{U_0})(C_2 h - U_1 a_h)$. Then

$$h^\top I_{\mathrm{eff}} h = \varepsilon^4 \|R_2 h\|^2 + O(\varepsilon^5).$$

In particular, $R_2 h \neq 0$ gives a matching $\Theta(\varepsilon^4)$ rate. When $C_1 = U_0 A$, the corresponding matrix identity is $I_{\mathrm{eff}} = \varepsilon^4 R_2^\top R_2 + O(\varepsilon^5)$ with $R_2 = (I - \Pi_{U_0})(C_2 - U_1 A)$.

Proof. Subtracting an exact nuisance combination leaves the profiled residual unchanged:

$$(I - \Pi_{U_\varepsilon}) C_\varepsilon h = (I - \Pi_{U_\varepsilon})(C_\varepsilon h - \varepsilon U_\varepsilon a_h) = \varepsilon^2 R_2 h + O(\varepsilon^3).$$

Here $\Pi_{U_\varepsilon} = \Pi_{U_0} + O(\varepsilon)$ follows from full column rank. Squaring the residual proves the assertion. The same calculation columnwise proves the matrix identity. For expected conditional information, the record-to-start-state replication isometry from Appendix B applies to the residual as well. Its norm-preserving leading factor does not change the coefficient, even when starting-state probabilities depend on λ. □

This lemma does not assert a positive second coefficient for every nuisance model. It supplies the coefficient that must be checked after leading cancellation. The explicit families below are analytic in time and parameters; their endpoint expansions therefore have the required differentiated remainders.

D.2 PROOF OF THE SCALAR MATCHING COEFFICIENT

Lemma 11 (A matching quartic coefficient)**.** *For $q = \lambda e^v$, $\eta = \theta_0 + f + \theta_1 \phi$, at $v = \theta_0 = \theta_1 = 0$ and endpoints $t_r = r$, $r = 1, \ldots, 6$,*

$$I_{v,\mathrm{eff}} = C_4 \lambda^{-4} + O(\lambda^{-5}), \quad C_4 = \|(I - \Pi_{[\sqrt{j},\, \sqrt{j} \odot b]})(\sqrt{j} \odot d^2)\|^2 > 0, \tag{27}$$

where $d_r = t_r - 3$, $b_r = 0.12 d_r$, and $j_r = \mathrm{sigmoid}(f(t_r))(1 - \mathrm{sigmoid}(f(t_r)))$. Here $C_4 \approx 9.22437$.

For the original construction (7), let $\varepsilon = \lambda^{-1}$, $d = t - 3$, $b = f' = 0.12d$, $k = f'' = 0.12$, $p = \mathrm{sigmoid}(f)$, and $j = p(1-p)$. The constant-total-rate solution has the differentiated expansion

$$y = p - \varepsilon p' + \varepsilon^2 p'' + O(\varepsilon^3), \qquad \partial_v y = \varepsilon p' - 2\varepsilon^2 p'' + O(\varepsilon^3).$$

At an integer endpoint, $\phi = 0$, $\phi' = b$, and $\phi'' = 2k + d^2$. Differentiating $p = \mathrm{sigmoid}(f + \theta_0 + \theta_1 \phi)$ gives

$$\partial_{\theta_1} y = -\varepsilon j b + \varepsilon^2 \{ j\phi'' + 2j'\phi' \} + O(\varepsilon^3).$$

Since $p'' = jk + j'b$, the exact derivative combination cancels its first-order term:

$$\partial_v y + \partial_{\theta_1} y = \varepsilon^2 j d^2 + O(\varepsilon^3).$$

The limiting whitened nuisance span is generated by $\sqrt{j}$ and $\sqrt{j} \odot b$. Lemma 10 therefore gives the coefficient in Lemma 11. It is strictly positive: if $d_r^2 = A + Db_r$ at the six integer endpoints, a quadratic polynomial would agree with an affine polynomial at six distinct points, which is impossible. The coefficient is evaluated from this formula, not fitted from a log–log slope. Its numerical value is approximately 9.2243686017.

D.3 Proof of the shared-graph equal-gap-multiset corollary

Write B for the unsigned incidence matrix, $d = t - 3$, $b = 0.12d$, $k = 0.12$, $p = \mathrm{sigmoid}(f)$, $j = p(1-p)$, and $q_0 = 2\cosh(f/2)$. The model in Corollary 2 is a subfamily of (1): its formation and dissolution rates are

$$a_e = \lambda e^{(Bv)_e + \eta_e/2}, \qquad b_e = \lambda e^{(Bv)_e - \eta_e/2}.$$

The total rate depends on the unknown affinity. Thus the result does not replace the original total-rate factor by a constant. At $v = \alpha = \gamma = 0$, all dyads share the same time profile. Define

$$A = \frac{p'}{q_0} = \frac{jb}{q_0}, \qquad H_2 = \frac{A'}{q_0} = \frac{j}{q_0^2}\left\{k + \tfrac{3}{2}(1-2p)b^2\right\}.$$

The second equality uses $(\log q_0)' = \frac{1}{2} b \tanh(f/2)$ and $j' = j(1-2p)b$. Uniformly over conditioned starting states,

$$y = p - \varepsilon A + \varepsilon^2 H_2 + O(\varepsilon^3), \qquad \partial_v y = (\varepsilon A - 2\varepsilon^2 H_2)B + O(\varepsilon^3).$$

At a half-integer or integer endpoint, put $s = \cos(2\pi t) \in \{-1, 1\}$. Direct differentiation gives

$$\phi = 0, \quad \phi' = sb, \quad \phi'' = 2sk + d^2, \qquad \partial_\gamma A = sAB, \quad \partial_\gamma H_2 = \left(2sH_2 + \frac{jd^2}{q_0^2}\right) B.$$

The last identity includes the affinity derivative of the total rate. Consequently, when all endpoint signs equal s,

$$\partial_v y + s\partial_\gamma y = s\varepsilon^2 \frac{jd^2}{q_0^2} B + O(\varepsilon^3). \tag{28}$$

The offset nuisance has leading derivative jB, while the invisible nuisance has leading derivative $-s\varepsilon AB$. For a full-column-rank B, the limiting nuisance matrix has full column rank $2n$: its temporal columns are independent and its spatial columns are those of B.

For the aligned design $T_A = (0.5, 1.5, 2.5, 3.5, 4.5, 5.5)$, all signs equal -1. Set

$$u_0 = \sqrt{j}, \quad u_1 = \frac{\sqrt{j}\, b}{q_0}, \quad z_2 = \frac{\sqrt{j}\, d^2}{q_0^2}, \qquad \kappa_A = \|(I - \Pi_{[u_0, u_1]})z_2\|^2.$$

The temporal–spatial product structure and (28) give

$$I^A_{v,\mathrm{eff}} = \lambda^{-4} \kappa_A B^\top B + O(\lambda^{-5}).$$

To verify $\kappa_A > 0$, note that u_0, z_2 are even and u_1 is odd at the paired values $d = \pm 0.5, \pm 1.5, \pm 2.5$. A relation $z_2 = au_0 + cu_1$ would therefore require $c = 0$ and d^2/q_0^2 to be constant at three distinct squared distances. But $x/[4\cosh^2((-0.3 + 0.06x)/2)]$ is strictly increasing on $[0.25, 6.25]$: its logarithmic derivative is $1/x - 0.06\tanh((-0.3 + 0.06x)/2) > 1/6.25 - 0.06 > 0$. The relation is impossible.

For $T_S = (1, 2, 3, 3.5, 4.5, 5.5)$, define the temporal vectors at these endpoints and put

$$\kappa_S = \|(I - \Pi_{[u_0,\, s\odot u_1]})u_1\|^2.$$

Theorem 1 yields $I^S_{v,\mathrm{eff}} = \lambda^{-2}\kappa_S B^\top B + O(\lambda^{-3})$. If $u_1 = au_0 + c(s \odot u_1)$, the endpoint $t = 3$ forces $a = 0$. The nonzero u_1 at $t = 1$ then forces $c = 1$, whereas the nonzero u_1 at $t = 3.5$ forces $c = -1$. Thus $\kappa_S > 0$. The two nuisance temporal columns are likewise independent.

Both schedules start at zero, end at 5.5, use six transition intervals, and have gap multiset $\{0.5, 1, 1, 1, 1, 1\}$. Only the placement of the short interval differs. Numerical evaluation of the explicit coefficients gives $\kappa_A \approx 0.5779255788$ and $\kappa_S \approx 0.01123445552$. The graph factors are positive definite precisely when B has full column rank. An exact $Bh = 0$ clock gauge remains unidentifiable under either schedule; observation timing does not eliminate a spatial gauge.

Perturbations and local information. The matching identities are pointwise statements at the specified symmetric parameter and exact schedules. The recovery experiment separately varies shared clock and affinity offsets and adds known timing jitter. Fixed nonzero jitter can make the nuisance visible at the endpoints and change the limiting rank, so the exact aligned quartic law is not asserted uniformly over jitter as $\lambda \to \infty$. Likewise, inverse Fisher information is a local scale diagnostic, not a promised finite-sample confidence interval or an exact MSE formula for bounded nonlinear estimators.

D.4 Nonzero invisible affinities: an exact leading matrix and a persistent bottleneck

We prove Proposition 3. This result concerns the same shared-affinity family, not a different constant-total-rate surrogate. Write $c_e = (Bv)_e$, $a_e = (B\alpha)_e$, $g_e = (B\gamma)_e$, $d = t - 3$, $b = 0.12d$, and $k = 0.12$. At the aligned half-integer endpoints,

$$\eta_e = f + a_e, \quad \eta'_e = b(1 - g_e), \quad \eta''_e = k + g_e(-2k + d^2), \quad q_{0,e} = 2e^{c_e}\cosh((f + a_e)/2).$$

Define $p_e = \mathrm{sigmoid}(f + a_e)$ and $j_e = p_e(1 - p_e)$. The factors $p_e, j_e, q_{0,e}, \eta_e$ and η''_e are even in d, while b and η'_e are odd. Affinity nuisance remains endpoint-invisible because $\phi = 0$, even when its true coefficient is nonzero.

Exact first-order profile. The collapsed, whitened leading clock and nuisance blocks at endpoint r are

$$C_{1,r} = \mathrm{diag}\left(\frac{\sqrt{j_{er}}b_r}{q_{0,er}}\right)(I - D_g)B, \qquad D_g = \mathrm{diag}(B\gamma),$$

$$U_{0,r} = \left[\mathrm{diag}(\sqrt{j_{er}})B,\ \mathrm{diag}\left(\frac{\sqrt{j_{er}}b_r}{q_{0,er}}\right)B\right].$$

The intercept block is even over the six paired endpoints, whereas the clock and rescaled invisible-nuisance blocks are odd. Their cross inner products are therefore zero. Both nuisance blocks have full column rank when B does, and together they have rank $2n$. Profiling the remaining odd block gives, for any clock direction h,

$$h^\top L_A h = \min_z \sum_e \omega_e\left[((I - D_g)Bh)_e - (Bz)_e\right]^2$$

$$= h^\top B^\top D_g\left[\Omega - \Omega B(B^\top \Omega B)^{-1}B^\top\Omega\right]D_g Bh, \qquad \omega_e = \sum_r \frac{j_{er}b_r^2}{q_{0,er}^2}, \quad \Omega = \mathrm{diag}(\omega_e).$$

Here $\omega_e > 0$. The change from $I - D_g$ to D_g is legitimate because the Bh component is itself in the nuisance column space. Theorem 1 proves the leading-information expansion. Since Ω is positive definite,

$$\ker L_A = \{h : D_g Bh \in \mathrm{col}(B)\}.$$

The common-clock vector is always in this kernel: $D_g B\mathbf{1} = 2B\gamma$. When $v = \alpha = 0$, $\Omega = \|u_1\|^2 I$ and the formula reduces to

$$L_A = \|u_1\|^2 B^\top D_g(I - \Pi_B)D_g B.$$

For fixed v, α, its dependence on γ is exactly quadratic. In particular, it vanishes for a common γ, not only at zero.

Generic rank on a complete observation graph. For K_n, $n \geq 6$, the kernel always contains $\mathbf{1}$, so the rank is at most $n - 1$. We exhibit a rank-$n - 1$ witness, then use the nonvanishing-minor criterion. Set $\gamma_i = 0$ on three vertices and $\gamma_i = 1$ on all others. Suppose $D_g Bh = Bz$. Subtract the additive diagonal terms by writing $w_i = z_i - \gamma_i h_i$. The edge equations are

$$\gamma_i h_j + \gamma_j h_i = w_i + w_j \qquad (i \neq j).$$

The triangle on the three zero vertices forces their w_i to be zero. Edges from a zero vertex to a one vertex force all zero-vertex h_i to be a common t and all one-vertex $w_j = t$. On the clique of at least three one vertices, $h_i + h_j = 2t$ forces every remaining $h_i = t$. Thus h is constant, proving rank $n - 1$ at this witness. The entries of $(I - \Pi_B)D_g B$ are linear in γ; a nonzero $(n - 1)$-minor is a polynomial that can vanish only on a proper algebraic subset. Hence the same rank holds for generic γ. The weighted kernel is identical, so this argument does not require $v = \alpha = 0$. The six-node witness is additionally checked in exact rational arithmetic by the artifact.

Matching bottleneck as a special case. Theorem 4, proved on the whole leading kernel in Appendix E, gives the matching spectrum here with $r = n-1$. For the coefficient used in the numerical checks, put $h_0 = \mathbf{1}/\sqrt{n}$ and

$$H_{er} = -\frac{2}{\sqrt{n}} \frac{\sqrt{j_{er}}}{q_{0,er}^2}[d_r^2 + g_e(d_r^2 - 2k)], \qquad R_{20} = (I - \Pi_{U_0})H.$$

Then $h_0^\top I^A_{v,\mathrm{eff}} h_0 = \lambda^{-4}\|R_{20}\|^2 + O(\lambda^{-5})$ and, at the origin, $\|R_{20}\|^2 = 2(n-1)\kappa_A > 0$. The whole-kernel proof includes the affinity derivative of q_0, cancellation before whitening, and the even/odd argument excluding rotating-eigenvector cancellation; no separate special-case argument is needed.

Explicit symmetric coefficient and independent numerical checks. At $v = \alpha = 0$, let $\widetilde{z}_2 = (I - \Pi_{u_0})z_2$ and $\widetilde{z}_3 = (I - \Pi_{u_0})\{\sqrt{j}(d^2 - 2k)/q_0^2\}$. Then

$$\|R_{20}\|^2 = \frac{4}{n}\left[|E|\|\widetilde{z}_2\|^2 + 2(\mathbf{1}^\top g)\langle \widetilde{z}_2, \widetilde{z}_3\rangle + \|g\|^2\|\widetilde{z}_3\|^2\right].$$

For the fixed independently seeded perturbation direction in the artifact, amplitudes 0.05 and 0.2 give coefficients 5.88598 and 6.37952 on K_6. A separate endpoint-layer quadrature verifies finite-clock spectra, while exact rational arithmetic checks the rank witness. These calculations check, rather than replace, the proof.

E Observation topology and matching information spectra

This section proves Theorem 4 for an arbitrary fixed full-column-rank observation matrix B. All information matrices are expected conditional information per replicated graph path: initial states are conditioned upon, and subsequent starting-state exposures are averaged without differentiating their densities. The positive-gap expansions are uniform in the two starting states. Thus the norm-preserving start-state replication in Appendix B applies throughout.

E.1 Exact rank and the observation-budget restriction

Let $M = \Omega^{1/2}B$ and $N = \Omega^{1/2}D_gB$. Proposition 3 gives

$$L_A = N^\top(I - \Pi_M)N, \qquad r = \mathrm{rank}[(I - \Pi_M)N] = \mathrm{rank}[B, D_gB] - n.$$

The last identity follows by decomposing the column space of $[M, N]$ into the n-dimensional space of M and its orthogonal residual; multiplication by the invertible $\Omega^{1/2}$ does not change rank. The residual projector has rank $m - n$, giving $r \le m - n$.

Let $\mathbf{1}_C$ be the node indicator of any connected component C. Every row belongs to a single component, so

$$D_gB\mathbf{1}_C = 2B(\gamma \odot \mathbf{1}_C).$$

The k component indicators are independent kernel directions, proving $r \le n - k$. Full column rank implies that every component is nonbipartite. Hence a connected graph requires $m \ge 2n - 1$ to permit $n - 1$ quadratic directions. This counts observed dyad risk sets, not the number of dyads currently in state one. In particular, every absent dyad in a fully observed snapshot contributes a risk set.

E.2 Matching quartic rates on the whole kernel

The rank bound alone is only a quadratic-coefficient statement. We next prove the matching eigenvalue rates, including when the kernel has dimension larger than one. Use the notation of Appendix D.4, with $k_0 = f'' = 0.12$ (distinct from the component count). Stack endpoint rows, and put

$$F_{er} = \frac{\sqrt{j_{er}}b_r}{q_{0,er}}, \qquad H_{er} = -\frac{\sqrt{j_{er}}}{q_{0,er}^2}[d_r^2 + g_e(d_r^2 - 2k_0)].$$

Let $\mathcal{K} = \ker L_A$. For $h \in \mathcal{K}$, the unique vector z_h with $Bz_h = (I - D_g)Bh$ exists because $D_gBh \in \mathrm{col}(B)$ and B is injective. The whitened exact derivative combination $C_\varepsilon h - N_{\gamma,\varepsilon} z_h$ has zero first coefficient. Direct differentiation of

$$y = p - \varepsilon \frac{jb(1-g)}{q_0} + \varepsilon^2 \frac{j}{q_0^2}\left[k_0 + g(-2k_0 + d^2) + \frac{3}{2}(1-2p)b^2(1-g)^2\right] + O(\varepsilon^3)$$

gives its second coefficient $H_{er}(Bh)_e$. In particular, the terms involving $b^2(1-g)^2$ cancel. Parameter derivatives of q_0 are included in this formula. The first-order correction to the whitening factor also cancels, since it multiplies a zero leading derivative combination. Define the linear map

$$\mathcal{R}h = (I - \Pi_{U_0})\{H_{er}(Bh)_e\}_{er}, \qquad U_{0,r} = [\mathrm{diag}(\sqrt{j_{er}})B,\ \mathrm{diag}(F_{er})B].$$

This map is defined for every h, even though it is the second profiled coefficient only on $\mathcal{K}$. At the origin,

$$\|\mathcal{R}h\|^2 = \kappa_A \|Bh\|^2 \geq \kappa_A \sigma_{\min}(B)^2 \|h\|^2.$$

The limiting nuisance matrix stays full rank near the origin: its offset block is even in d and its invisible block is odd, and each has rank n. Continuity of the projector and H therefore gives a neighborhood and a constant $c_B > 0$ with $\|\mathcal{R}h\| \geq c_B\|h\|$ for all h. This controls the second coefficient on any kernel that arises in that neighborhood; it does not assume the kernel dimension stays fixed as the parameters change.

For each fixed parameter in the neighborhood, choose an orthonormal basis for $\mathcal{K} \oplus \mathcal{K}^\perp$. The second residual $\mathcal{R}h$ is even in time; the leading profiled clock residual on $\mathcal{K}^\perp$ is odd. Their inner product is zero. Lemma 10 and the expansion of the profiled clock matrix yield

$$I^A_{v,\mathrm{eff}} = \begin{pmatrix} \varepsilon^4 A_{\mathcal{K}} + O(\varepsilon^5) & O(\varepsilon^4) \\ O(\varepsilon^4) & \varepsilon^2 A_\perp + O(\varepsilon^3) \end{pmatrix}, \qquad A_{\mathcal{K}} \succ 0, \quad A_\perp \succ 0,$$

where the empty block is omitted when $r = 0$. Its Schur complement differs from $\varepsilon^4 A_{\mathcal{K}}$ by $O(\varepsilon^5)$; the off-diagonal correction is $O(\varepsilon^6)$. This proves $n - r$ quartic and r quadratic eigenvalues. Parity is essential: a positive second coefficient along fixed directions would not alone exclude cancellation by $O(\varepsilon)$ rotations of eigenvectors. On T_S, the origin's positive-definite leading information and full-rank nuisance matrix persist by continuity. All n eigenvalues are quadratic there. Constants for positive quadratic eigenvalues are pointwise in the fixed parameter, not uniform as a rank-changing nuisance perturbation tends to zero simultaneously with $\lambda \to \infty$.

E.3 Sparse examples and an exact minimal-budget certificate

A connected odd unicyclic observation graph has $m = n$ and full-column-rank unsigned incidence, so $r = 0$: all directions remain quartic near the origin, including nonzero invisible affinities. To illustrate how added risk sets change this conclusion, begin with the six-node graph

$$E_6 = \{12, 13, 23, 34, 45, 56\},$$

and append edges $14, 15, 16, 24, 25$ in that order. For $\gamma = (0, 1, 2, 3, 4, 5)$, exact rational elimination gives ranks $r = 0, 1, 2, 3, 4, 5$ for $m = 6, 7, 8, 9, 10, 11$. For $m = 6, 8, 11$, the first m columns of $[B, D_gB]$ have determinants -2, 8, and -192, respectively, in the displayed edge order. These explicit nonzero minors certify the three principal sparse examples directly. The executable certificate additionally stores each edge list, selected minor, exact determinant, rank, and witness. Since the relevant minors are polynomial in γ, these witness ranks hold generically, and scaling the witness to an arbitrarily small nonzero amplitude preserves rank. Thus the connected necessary budget $m = 2n - 1$ is attained for this six-node construction. It is not asserted that every graph with that many edges attains the upper bound.

For completeness, adding observed dyads at fixed times and fixed exposure cannot decrease the unnormalized efficient information for existing clock directions. For each h, efficient information is a minimum over nuisance velocities of a sum of squared whitened residuals. Adding nonnegative summands makes that minimum no smaller. This statement does not apply to information divided by the changing record count, to reallocating a fixed measurement budget, or to turning an already observed zero state into a one state.

F TIMING VISIBILITY, NONCOMMUTING LIMITS, AND A CALIBRATED SHIFT

All results in this section concern the shared phase family (8) at $v = \alpha = \gamma = 0$ with a fixed full-column-rank B. Observation times are known exactly. Perturbing a measurement schedule is not the same statistical problem as observing corrupted timestamps.

F.1 STATEMENT OF THE CALIBRATED CROSSOVER

A fixed perturbation can make the endpoint nuisance block regular, while a perturbation shrinking with the clock has a different limit (Appendix F).

Theorem 12 (Joint timing–clock limit)**.** *For full-column-rank B, at $v = \alpha = \gamma = 0$, use $T_\lambda = T_A + z/\lambda$ with fixed $z = \rho e_1$, where e_1 selects the first endpoint. Evaluate* (10) *on T_A and put $e = -\sqrt{j}\, b \odot e_1$, $a = (I - \Pi_{u_0})u_1$, $b_* = (I - \Pi_{u_0})e$. Then*

$$I^{T_\lambda}_{v,\mathrm{eff}} = \lambda^{-2}\kappa(\rho)B^\top B + O(\lambda^{-3}),$$
$$\kappa(\rho) = \frac{\rho^2(\|a\|^2\|b_*\|^2 - \langle a, b_*\rangle^2)}{\|a + \rho b_*\|^2}. \tag{29}$$

Here a, b_* are independent and $\langle a, b_*\rangle < 0$. The positive ratio $\rho_* = -\|a\|^2/\langle a, b_*\rangle \approx 1.39368$ attains $\kappa(\rho_*) = \|a\|^2 \approx 0.0156459$, the leading known-nuisance coefficient for endpoints converging to T_A.

Rescaling the invisible nuisance gives the limiting column $u_1 + \rho e$. At ρ_*, its nuisance residual is orthogonal to the clock residual. Exact alignment, fixed jitter, and clock-scaled jitter therefore have different limits; this optimum is specific to the stated symmetric family. Appendix H.1 tests its sensitivity to a larger fitted nuisance space.

F.2 FIXED NONZERO TIMING PERTURBATIONS

For fixed endpoints T, the leading whitened derivatives of the two affinity blocks are $u_0 \otimes B$ and $(\sqrt{j}\,\phi) \otimes B$; the clock derivative is $\varepsilon u_1 \otimes B$. If $\phi(T)$ is nonconstant, the endpoint nuisance matrix has full column rank $2n$. The regular case of Theorem 1 gives

$$I^T_{v,\mathrm{eff}} = \lambda^{-2}\kappa(T)B^\top B + O(\lambda^{-3}), \quad \kappa(T) = \|(I - \Pi_{[\sqrt{j}, \sqrt{j}\phi]})\sqrt{j}f'/q_0\|^2. \tag{30}$$

At $T = T_A + \delta e_1$, all but the first entries of ϕ remain zero. Since $\phi'(0.5) = -f'(0.5) \neq 0$, its first entry is nonzero for every sufficiently small nonzero δ. The nuisance span is then $\mathrm{span}\{u_0, e_1\}$, independent of the nonzero magnitude of that entry. If u_1 belonged to this span, $b_r/q_{0,r}$ would be constant on the five unchanged endpoints. Paired unchanged endpoints have opposite nonzero signs, which rules this out. Hence $\kappa(T) > 0$ for every such fixed shift. A nonzero three-column minor at this example also proves positivity for generic sufficiently small perturbations of all endpoints: the exceptional set is contained in the zero set of a nonzero real-analytic minor.

As $\delta \to 0$ through nonzero values, this fixed-design coefficient approaches the positive value obtained by projecting on $[u_0, e_1]$. At $\delta = 0$, in contrast, the appropriate two-scale nuisance span is $[u_0, u_1]$, and the quadratic coefficient is zero. There is no contradiction: convergence in λ is not uniform near the endpoint-rank change.

F.3 THE JOINT $1/\lambda$ TIMING SCALE

Put $\varepsilon = \lambda^{-1}$ and $T_\varepsilon = T_A + \rho\varepsilon e_1$. Smoothness and positive limiting gaps make the differentiated endpoint expansion uniform for bounded ρ. Expanding $\phi(T_\varepsilon)$ gives

$$\phi(T_\varepsilon) = -\rho\varepsilon\, b \odot e_1 + O(\varepsilon^2).$$

Consequently, the whitened clock, offset, and invisible-affinity columns have respective leading forms

$$\varepsilon u_1 \otimes B, \qquad u_0 \otimes B, \qquad \varepsilon(u_1 + \rho e) \otimes B, \qquad e = -\sqrt{j}\, b \odot e_1.$$

The two limiting nuisance temporal columns are independent for every finite real ρ. After removing u_0, write $a = (I - \Pi_{u_0})u_1$ and $b_* = (I - \Pi_{u_0})e$. The vector a is odd on three endpoint pairs, whereas

the nonzero odd part of b_* is supported only on the first/last pair. Thus a, b_* are independent. Their inner product is $u_{1,1}e_1 < 0$, using the even/odd orthogonality $u_0^\top u_1 = 0$. Projecting a onto $a + \rho b_*$ gives

$$\|a\|^2 - \frac{\langle a, a + \rho b_* \rangle^2}{\|a + \rho b_*\|^2} = \frac{\rho^2(\|a\|^2\|b_*\|^2 - \langle a, b_* \rangle^2)}{\|a + \rho b_*\|^2}.$$

The stable-projector argument of Theorem 1, applied to this triangular array, proves Theorem 12. For every fixed nonzero ρ this coefficient is positive. At $\rho = 0$ the formula gives a zero quadratic coefficient; Corollary 2, rather than the displayed remainder alone, supplies the matching quartic rate.

The efficient coefficient cannot exceed $\|a\|^2 = \|u_1\|^2$, the known-nuisance leading coefficient at these limiting endpoints. At

$$\rho_* = -\|a\|^2/\langle a, b_* \rangle > 0,$$

the clock residual is orthogonal to $a + \rho_* b_*$, and equality holds. This proves optimality within the one-endpoint ρ/λ family and attains the limiting known-nuisance information. It does not optimize over arbitrary finite schedules, nuisance families, unknown clock scales, or observation costs. The shift preserves endpoint count and final horizon, but not the gap multiset shared by T_A and T_S.

A nominal-scale robustness bound. Let $\chi = \langle a, b_* \rangle^2/(\|a\|^2\|b_*\|^2) \approx 0.4309084965$. If the true regime is a fixed factor $s > 0$ of the nominal regime used to choose the shift, then the actual joint-limit ratio is $s\rho_*$. Substitution gives

$$\frac{\kappa(s\rho_*)}{\|a\|^2} = \frac{s^2(1 - \chi)}{s^2 + \chi(1 - 2s)}.$$

This function increases to one for $s \le 1$ and decreases for $s \ge 1$ (differentiate, using $0 < \chi < 1$). On $s \in [1/2, 2]$ the smaller endpoint value is $1 - \chi \approx 0.5690915$. Thus multiplicative nominal-scale misspecification bounded by a factor of two preserves at least this fraction of leading known-nuisance information at the symmetric model. The statement concerns the triangular-array information coefficient, not estimated-clock risk or a data-driven pilot-selection procedure.

F.4 The smaller $1/\lambda^2$ scale

If $T_\varepsilon = T_A + s\varepsilon^2 e_1$ for fixed s, the nuisance endpoint contribution enters the second coefficient. Changes to the leading clock lag are only $O(\varepsilon^3)$. Subtracting the exact invisible-affinity column therefore gives

$$C_\varepsilon - N_{\gamma,\varepsilon} = -\varepsilon^2(z_2 + se) \otimes B + O(\varepsilon^3).$$

Whitening corrections multiplying the canceled first term vanish. Lemma 10 implies

$$I_{v,\mathrm{eff}}^{T_\varepsilon} = \varepsilon^4 \kappa_4(s) B^\top B + O(\varepsilon^5), \quad \kappa_4(s) = \|(I - \Pi_{[u_0,u_1]})(z_2 + se)\|^2. \tag{31}$$

This coefficient is positive for every real s. The residual of z_2 is nonzero and even. The residual of e has a nonzero odd component because a single-pair odd spike is not proportional to u_1, which is nonzero on all three pairs. These two residuals are independent. At $s = 1$, $\kappa_4(s) \approx 0.712751$, compared with $\kappa_4(0) \approx 0.577926$. The exact, fixed-shift, $1/\lambda$-shift, and $1/\lambda^2$-shift experiments test different limits; none may be inferred from a finite-clock risk comparison alone.

G A local exposure threshold and its limitations

Let $\beta = (v, \theta)$ and suppose the finite-λ conditional model is differentiable with nonsingular joint information after removing true gauges. A locally unbiased estimate satisfies $\mathbb{E}_\beta \widehat{v} = v$, $\partial_v \mathbb{E}_\beta \widehat{v} = I_n$, and $\partial_\theta \mathbb{E}_\beta \widehat{v} = 0$ at the parameter under consideration. The score identity and the matrix Cauchy–Schwarz inequality give

$$\mathrm{Cov}(\widehat{v}) \succeq (RI_{v,\mathrm{eff}})^{-1}.$$

Equivalently this follows from the clock block of the joint Cramér–Rao matrix and the Schur complement. Independent replicated graph paths multiply the information by R; the dependent transitions within each path must first be combined through expected conditional information. Multiplication by B and tracing proves

$$\mathbb{E}\|B\widehat{v} - Bv\|^2/m \ge \frac{\mathrm{tr}(BI_{v,\mathrm{eff}}^{-1}B^\top)}{mR}.$$

The deterministic homogeneous estimate $\widehat{v} = 0$ has squared field error $r_0^2 = \|Bv\|^2/m$. Write $V = \mathrm{tr}(BI_{v,\mathrm{eff}}^{-1}B^\top)/m$. For $r_0 > 0$, strict improvement by a locally unbiased estimate therefore requires

$$R > R_{\mathrm{loc}} := V/r_0^2. \tag{32}$$

This truth-dependent threshold is not an implementable confidence guarantee or a hyperparameter-selection rule. At $v = 0$ the homogeneous reference is already exact, so this comparison has no positive threshold interpretation.

For fixed parameters and fixed full-column-rank B, its smallest and largest singular values bound the trace above by positive multiples of $\mathrm{tr}(I_{v,\mathrm{eff}}^{-1})$. Theorem 4 gives at least one quartic eigenvalue on every aligned nonempty observation graph, and no eigenvalue decays faster than quartically in its neighborhood. Hence $V_A = \Theta(\lambda^4)$. On the staggered schedule all eigenvalues are quadratic, so $V_S = \Theta(\lambda^2)$. These are fixed-parameter information orders. They neither establish a joint large-R, large-λ likelihood-normality theorem nor impose a lower bound on arbitrary biased estimators. Bounded MLEs and ridge estimates may violate local unbiasedness; their actual risk, boundary frequency, and penalty sensitivity are consequently measured separately.

H AFFINITY UNCERTAINTY AND ANCHORED SHORT-LAG OBSERVATIONS

This section distinguishes two sources of clock information. The initial-state contrast of a two-state transition is exactly

$$P_{11}(u,t) - P_{01}(u,t) = \exp\{-\textstyle\int_u^t q(s)ds\}. \tag{33}$$

For a fixed positive gap it vanishes exponentially as the clock grows. Long-gap clock information instead comes from the equilibrium's small lag, which additional affinity parameters may explain. A short probe keeps the contrast nonvanishing. Its affinity dependence must still be accounted for; treating the total rate as affinity independent would not prove the result below.

H.1 A VISIBLE-NUISANCE CEILING AT AN UNCHANGED TRUE TRAJECTORY

Enlarge (8) by a node-shared linear trend:

$$\eta_e(t) = f(t) + (B\alpha)_e + \phi(t)(B\gamma)_e + (t-3)(B\beta)_e. \tag{34}$$

Even at $\beta = 0$, fitting this additional unknown block changes the nuisance space. Let $w = \sqrt{j}\,(t-3)$ on T_A, using the vectors from Theorem 12.

Proposition 13 (Visible-nuisance ceiling)**.** *At $v = \alpha = \gamma = \beta = 0$ and $T_\lambda = T_A + \rho e_1/\lambda$, the leading coefficient is*

$$\kappa_{\mathrm{lin}}(\rho) = \|(I - \Pi_{[u_0,w,u_1+\rho e]})u_1\|^2 \le \|(I - \Pi_{[u_0,w]})u_1\|^2 \approx 5.35571\,10^{-8}. \tag{35}$$

At the original ρ_* it is approximately $4.28687\,10^{-8}$, versus 0.0156459 before adding the unknown trend. Retuning ρ cannot exceed the displayed ceiling.

The true trajectory is identical, yet the leading information coefficient falls by over five orders of magnitude. The added visible column nearly explains $u_1 = \sqrt{j}\,.12(t-3)/q_0$, because q_0 varies little. The projection proof (Appendix H.1) gives a sensitivity result, not a finite-sample risk ratio or a failure of Theorem 12's prescribed-family statement. Useful design must account for the *fitted* nuisance family, not only its true coefficient values.

For the proof, let $d = t-3$, $w = \sqrt{j}\,d$, and evaluate the vectors in (10) at T_A. For $T_\lambda = T_A + \rho e_1/\lambda$, expansion at the moving endpoint gives the leading whitened temporal columns

$$\lambda^{-1}u_1, \qquad u_0, \quad w, \quad \lambda^{-1}(u_1 + \rho e), \qquad e = -\sqrt{j}\,b \odot e_1.$$

The spatial factor is B for each block. The columns $u_0, w, u_1 + \rho e$ are independent for every finite ρ. To see this, an equality $u_1 + \rho e = Au_0 + Dw$ on the five unchanged endpoints would require $.12d/q_0 = A + Dd$ there. Opposite unchanged endpoints force $A = 0$, but q_0 is not equal at the distinct nonzero radii. The full-column-rank rescaling argument of Theorem 1 therefore applies.

Proof of Proposition 13. Projecting the clock column onto these nuisance columns gives the coefficient stated above. Since $\mathrm{span}\{u_0, w\}$ is contained in the full nuisance span, its residual norm is an upper bound for every ρ. With

$$a_t = (I - \Pi_{[u_0,w]})u_1, \qquad b_t = (I - \Pi_{[u_0,w]})e,$$

projection onto $a_t + \rho b_t$ gives

$$\kappa_{\mathrm{lin}}(\rho) = \|a_t\|^2 - \frac{\langle a_t, a_t + \rho b_t\rangle^2}{\|a_t + \rho b_t\|^2}.$$

In this example a_t, b_t are independent and $\langle a_t, b_t\rangle \neq 0$. Hence $\rho_t = -\|a_t\|^2/\langle a_t, b_t\rangle$ attains the ceiling. The vectors are explicit elementary functions at six stated times, so the coefficients are reproducible without fitting or simulation:

$$\begin{aligned} \kappa(\rho_*) &\approx 0.0156458652986, & \kappa_{\mathrm{lin}}(\rho_*) &\approx 4.28686584018\,10^{-8}, \\ \|a_t\|^2 &\approx 5.35571323936\,10^{-8}, & \rho_t &\approx 0.0050045143. \end{aligned}$$

The artifact additionally evaluates the two small coefficients with 60-decimal arithmetic. The result is a leading local-information calculation, not a finite-sample risk ratio. Adding an unknown trend does not contradict a theorem that prescribes the original nuisance family. □

The mechanism is transparent: $u_1 = \sqrt{j}\,(.12d)/q_0$ is nearly proportional to the new visible trend w over these endpoints. The small residual arises because $q_0 = 2\cosh(f/2)$ varies little, rather than from an implementation failure. A favorable exponent can therefore coexist with a very small coefficient. Enlarging a nuisance space can only reduce efficient information; the appropriate coefficient must be recomputed for the fitted family, not only for the data-generating coefficient values.

H.2 Probe experiment and regularity assumptions

Fix a full-column-rank observation matrix B and deterministic, finite-dimensional affinity family $\eta_e(t;\theta)$ independent of v. The rates are

$$a_e(t) = \Lambda e^{(Bv)_e + \eta_e(t;\theta)/2}, \qquad b_e(t) = \Lambda e^{(Bv)_e - \eta_e(t;\theta)/2}.$$

The common regime index Λ is a nominal scale for the design, and unknown common or heterogeneous departures are part of v. Choose L distinct anchor limits $a_1 < \cdots < a_L = H$, separated from zero and each other by positive fixed gaps. Each anchor is followed by a probe separated by x_l/Λ, $x_l > 0$. The last pair ends at H: its times are $(H - x_L/\Lambda, H)$. Earlier pairs can have times $(a_l, a_l + x_l/\Lambda)$. Thus there are exactly $2L$ observed transitions, in addition to the conditioned state at zero.

On a compact parameter set assume: affinities have uniformly bounded time derivatives through order three and first parameter derivatives of those time derivatives on $[0, H]$; the finite-dimensional parameter set is compact and the normalized rates are uniformly positive; and the stacked anchor Jacobian

$$\mathcal{H} = \begin{bmatrix} H_1 \\ \vdots \\ H_L \end{bmatrix}, \qquad H_l = \partial_\theta(\eta_e(a_l;\theta))_e, \qquad \sigma_{\min}(\mathcal{H}) \geq s_0 > 0$$

is injective after removing actual gauges. All dimensions remain fixed as Λ grows. These hypotheses imply uniform interior equilibrium probabilities and upper and lower bounds on the local clock factors.

The injectivity assumption has substantive content. It excludes affinity parameters invisible at every anchor, whose smaller scores could otherwise survive nuisance rescaling. For $\eta = f + \sum_{j=1}^{k} \psi_j(t) B\theta_j$, it is equivalent to full column rank of B and of $\Phi = (\psi_j(a_l))_{lj}$. In particular $L \geq k$ is needed for this sufficient construction; it is not claimed as a universal minimal observation count. For the three bases $1, \phi, d$ and anchors $.75, 2.75, 5.5$, Φ is invertible. In the column order $(1, \phi, d)$,

$$\det \Phi = -\frac{3}{10\pi} - \frac{109}{64\pi^2} < 0.$$

The artifact verifies this expression and its numerical value -0.268055606746. Thus the added linear trend and nonzero phase coefficients satisfy the anchor-rank condition, without their values being known.

H.3 LIMIT INFORMATION AND ITS POSITIVE LOWER BOUND

For a single dyad at anchor l, write $\eta = \eta_e(a_l;\theta)$, $p = \mathrm{sigmoid}(\eta)$, $j = p(1-p)$, and

$$r = x_l\, 2e^{(Bv)_e}\cosh(\eta/2), \qquad z = e^{-r}, \qquad y_a = p + (a-p)z, \quad \mu_0 = 1-p,\ \mu_1 = p.$$

In the frozen local experiment, the anchor is Bernoulli(p) and the probe has conditional success probability y_a given anchor state a. In local coordinates (c,η), where $c = (Bv)_e$,

$$d_a = \partial_c y_a = (p-a)rz, \qquad e_a = \partial_\eta y_a = j(1-z) + \tfrac{1}{2}\tanh(\eta/2)d_a.$$

Define the conditional information and the anchor contribution by

$$F = \sum_{a=0}^{1} \frac{\mu_a}{y_a(1-y_a)} \begin{bmatrix} d_a^2 & d_a e_a \\ d_a e_a & e_a^2 \end{bmatrix}, \qquad \overline{F} = F + \begin{bmatrix} 0 & 0 \\ 0 & j \end{bmatrix}, \quad w = F_{cc} - \frac{F_{c\eta}^2}{F_{\eta\eta}+j}. \tag{36}$$

Both starting states are represented with their limiting exposures; these exposure probabilities are not differentiated again in the conditional score. The anchor density has already supplied j.

Proof of Theorem 5. For the long gap preceding each anchor, the connectedness probability is $p + O(\Lambda^{-1})$ uniformly in its initial state; its clock derivative is $O(\Lambda^{-1})$ and its nuisance derivative is $jH_l + O(\Lambda^{-1})$. Rescale time within the probe by $s = \Lambda(t - a_l)$. Its forward equation converges, together with first parameter derivatives, to the constant-coefficient equation at the anchor. Bounded mixed derivatives and variation of constants give uniform $O(\Lambda^{-1})$ errors for both values and first derivatives on $0 \le s \le x_l$. The slightly shifted final anchor obeys the same expansion.

Scores of successive conditional transitions are orthogonal martingale increments. Summing expected conditional information therefore converges to the information of the anchor/probe blocks just described, with an $O(\Lambda^{-1})$ remainder. Set

$$M = \sum_l B^\top \mathrm{diag}(F_{cc,el})B,$$
$$C = \sum_l B^\top \mathrm{diag}(F_{c\eta,el})H_l,$$
$$N = \sum_l H_l^\top \mathrm{diag}(F_{\eta\eta,el} + j_{el})H_l.$$

The anchor-rank condition and $j_{el} > 0$ imply $N \succ 0$, uniformly on the compact set. Therefore its inverse is stable, there is no hidden rescaled nuisance column, and

$$I_{v,\mathrm{eff}}(\Lambda) = I_* + O(\Lambda^{-1}), \qquad I_* = M - CN^{-1}C^\top. \tag{37}$$

To lower-bound I_*, enlarge the limiting nuisance family to one unrelated affinity value per observed dyad and anchor. Minimization over that larger space separates over the pairs and gives

$$I_* \succeq B^\top \mathrm{diag}\left(\sum_l w_{el}\right) B.$$

Each weight is strictly positive. Indeed, F is positive semidefinite and $F_{cc} > 0$, so

$$w = \frac{F_{cc}(F_{\eta\eta}+j) - F_{c\eta}^2}{F_{\eta\eta}+j} \ge \frac{jF_{cc}}{F_{\eta\eta}+j} > 0.$$

Continuity on the compact set yields a positive minimum. With B injective, the lower bound is positive definite. The upper bound $I_* \preceq M$ is finite. Hence all n clock-information eigenvalues are $\Theta(1)$, uniformly over the stated compact set, rather than quadratic or quartic vanishing orders. □

At $\eta = 0$, $F_{c\eta} = 0$ and the weight reduces to $w = r^2/(e^{2r}-1)$. This familiar two-state clock weight explains why probes should not be either too short or too long in local clock units. We do not claim that high-frequency Markov-rate estimation itself is new. The result supplies a nuisance-aware sufficient condition and graph-matrix lower bound for a fixed-count schedule when affinity parameters are jointly unknown.

H.4 Unknown nominal scale and acquisition cost

The theorem is uniform in a bounded common shift of v. Consequently, a design using scale Λ remains in its stated regime when the physical common scale is $s\Lambda$ for any fixed s in a declared compact interval. It does not require access to the true s. More generally, geometric probe widths supply a range of local exposures r: if every admissible dyad has at least one probe with $r \in [r_-, r_+] \subset (0, \infty)$ and all affinities are in a compact interval, the displayed positive-weight bound supplies a common information floor. The other anchors still ensure nuisance rank.

The supporting unknown-scale experiment uses this schedule, based on a prespecified acquisition range rather than a pilot estimate or the true clock:

$$(0.75,\ 0.8125,\ 2.75,\ 2.765625,\ 5.49609375,\ 5.5).$$

Its widths $1/16, 1/64, 1/256$ stay fixed across physical multipliers $8, 32, 64$ and heterogeneous truths; the clock reference remains 32, with a common shift estimated in v. Separate asymptotic diagnostics use widths $(2, .5, .125)/\Lambda$ and the same anchor limits. In contrast, the fixed A/P control in Section 5.3 uses three $1/64$ probes and satisfies that comparison's minimum-gap constraint.

This intervention preserves six observations and final horizon, not the original gap multiset or a hardware-resolution cost. Known, resolvable short gaps are an acquisition requirement. The fixed schedule has minimum gap $1/256$, whereas the original staggered schedule has minimum gap $.5$. Arbitrary timestamp noise, an unbounded unknown rate range, or an affinity family violating anchor rank is outside the guarantee. A broader nuisance basis calls for checking the anchor Jacobian and, where necessary, adding anchors rather than silently deleting the unobserved nuisance directions.

I Quantitative design margins and the exposure bridge

All matrix norms in this section are spectral norms. Nuisance coordinates have a fixed scaling; an arbitrary rescaling changes both the perturbation size and the smallest singular value. Consequently, an isolated unnormalized condition number is not an invariant measure of identification.

I.1 Proof of the finite-clock stability margin

Let $P_* = \Pi_{N_*}$ and $P_\Lambda = \Pi_{N_\Lambda}$. Weyl's inequality gives $\sigma_{\min}(N_\Lambda) \geq s - \epsilon \geq s/2$. If U_Λ is an orthonormal basis for $\mathrm{col}(N_\Lambda)$, then

$$\|(I - P_*)U_\Lambda\| \leq \|(I - P_*)N_\Lambda\|\,\|N_\Lambda^+\| \leq \frac{\epsilon}{s - \epsilon} \leq \frac{2\epsilon}{s}.$$

The analogous bound with $*$ and Λ interchanged is at most ϵ/s. Since $P_\Lambda - P_* = (I - P_*)P_\Lambda - P_*(I - P_\Lambda)$, $\|P_\Lambda - P_*\| \leq 3\epsilon/s \leq 4\epsilon/s$. With $E_\Lambda = (I - P_\Lambda)C_\Lambda$ and $E_* = (I - P_*)C_*$, it follows that

$$\|E_\Lambda - E_*\| \leq \epsilon + 4\epsilon M/s = a.$$

Expanding $E_\Lambda^\top E_\Lambda - E_*^\top E_*$ gives the bound $(2M + a)a$ in Proposition 6. Anchor-score rows of N_* contain $\mathrm{diag}(\sqrt{j_{el}})\mathcal{H}$, so $N_*^\top N_* \succeq j_{\min}\mathcal{H}^\top\mathcal{H}$. This proves the stated lower bound for s. The $O(\Lambda^{-1})$ score approximations are those proved in Appendix H.3; limiting probe probabilities and starting-state exposures stay in the interior on the compact parameter set.

A graph-scaled finite bound. Let $\underline{w} = \min_e \sum_l w_{el}$ and $b = \lambda_{\min}(B^\top B) > 0$. If $(2M + a)a \leq \underline{w}b/2$, then

$$I_{v,\mathrm{eff}}(\Lambda) \succeq \underline{w}B^\top B - (\underline{w}b/2)I \succeq (\underline{w}/2)B^\top B.$$

The constants depend on the stated compact class and fixed graph. This is not a dimension-uniform conclusion when the graph grows. Numerical evaluations of a frozen surrogate alone do not certify the perturbation size ϵ or the displayed sufficient inequality.

I.2 Proof of the fixed-resolution boundary

Consider a deterministic family of schedules in $[0, H]$ with a fixed number K of gaps, every one at least $\delta > 0$. Uniform derivative bounds on the whole time domain and a uniform lower rate bound

make the long-gap expansion uniform over this schedule class. Affinity does not depend on v, hence

$$\partial_v P_{a1}(u,t) = O(\Lambda^{-1})$$

uniformly in the starting state, endpoints, and parameters. For large Λ, all endpoint conditional probabilities remain uniformly bounded away from zero and one. The expected unprofiled clock information, a finite sum of squared whitened derivatives, is therefore $O(\Lambda^{-2})$. Efficient information is smaller in the positive-semidefinite order. This argument does not require invertibility of an endpoint nuisance block.

If each affinity is constant in time, p_e is independent of both time and v, and

$$P_{a1}(u,t) = p_e + (a - p_e)e^{-H_e(u,t)}, \qquad \partial_v P_{a1} = -(a - p_e)e^{-H_e}\partial_v H_e.$$

For the shared clock parameterization, $H_e \geq q_{\min}\Lambda\delta$ and $\|\partial_v H_e\| \leq C\Lambda$ because the horizon and normalized parameter class are bounded. Squaring the whitened derivatives proves $O(\Lambda^2 e^{-2q_{\min}\Lambda\delta})$. The conclusion concerns known deterministic observation times. It neither proves a result for adaptive stopping rules nor models timestamp corruption.

I.3 Regular estimation and clock-field risk

Fix the graph, a finite physical clock scale, and an identifiable interior point of the correctly specified finite-dimensional family. Treat one replicated graph path as one independent observation, and use its conditional likelihood given the initial states. Conditional transition scores are martingale increments, so their cross terms have expectation zero. Their sum has mean zero and covariance equal to the expected conditional Fisher information $\mathcal{I}$. The initial-state distribution determines expected exposure but is not differentiated as an extra likelihood term. With uniformly positive probabilities and smooth rates on a local compact neighborhood, the score moments are finite and the expected negative Hessian equals $\mathcal{I}$.

For a consistent local likelihood root $\widehat{\zeta}_R$, $\zeta = (v, \theta)$, with a nonsingular limiting local information matrix, Taylor expansion of its score equation yields

$$\sqrt{R}(\widehat{\zeta}_R - \zeta_0) = \mathcal{I}^{-1}\frac{1}{\sqrt{R}}\sum_{r=1}^{R} s_r(\zeta_0) + o_p(1) \implies N(0, \mathcal{I}^{-1}).$$

Here consistency, an interior root with high probability, and the local uniform Hessian law are assumptions on the selected estimator; a numerical termination flag does not establish them. Block inversion gives the clock covariance $I_{v,\text{eff}}^{-1}$. If $R\|B(\widehat{v} - v)\|^2$ is uniformly integrable, then

$$\lim_{R\to\infty} R\,\mathbb{E}\frac{\|B(\widehat{v} - v)\|^2}{m} = \frac{\operatorname{tr}(BI_{v,\text{eff}}^{-1}B^\top)}{m} = V.$$

If $I_{v,\text{eff}} \succeq \kappa B^\top B$, matrix inversion reverses the positive-definite order. Since $B(B^\top B)^{-1}B^\top$ is a rank-n projector,

$$V \leq \frac{1}{m\kappa}\operatorname{tr}\{B(B^\top B)^{-1}B^\top\} = \frac{n}{m\kappa}.$$

This bridge explains the reported information-based scale without treating it as a finite-sample bound on biased, boundary-constrained, or misspecified estimates. In the latter cases neither efficient-information inversion nor an empirical ratio near one certifies estimator efficiency.

Acquisition specification. Appendices L.3 and L.4 define the common feasible sets and all acquisition objectives used in the primary experiments. Restricted-prior protocols are reported separately in Appendices Q and R.

J Spatial pooling, nuisance parameters, and graph topology

J.1 Conditional dyad information

For one homogeneous dyad define $p = \operatorname{sigmoid}(\eta)$, $q = 2e^c\cosh(\eta/2)$, $x = q\Delta$, and $z = e^{-x}$. Conditional flip probabilities from states zero and one are

$$r_0 = p(1 - z), \qquad r_1 = (1 - p)(1 - z).$$

For parameters (c, η), let the two rows of D be the gradients of r_0, r_1. The derivative columns are

$$D_{\cdot c} = (pxz, (1-p)xz)^\top, \tag{38}$$

$$D_{\cdot \eta} = (p(1-p)(1-z), -p(1-p)(1-z))^\top + (p - \tfrac{1}{2})D_{\cdot c}. \tag{39}$$

The second term reflects $\partial_\eta \log q = p - 1/2$; dropping it would use a different clock convention. For m transitions, averaged over stationary starting states,

$$F = mD^\top \operatorname{diag}\left\{\frac{1-p}{r_0(1-r_0)}, \frac{p}{r_1(1-r_1)}\right\} D. \tag{40}$$

Although successive states are correlated, conditional likelihood scores are martingale differences. Their cross-covariances vanish, so the expected information is the sum over transitions. Stationarity supplies the weights in (40), not an independence approximation. Initial-state densities are conditioned away; adding a parameter-dependent initial distribution would add information not included here.

Both weights are positive for finite positive rates and $m > 0$. The c derivative has two positive entries, whereas subtracting $(p - 1/2)D_{\cdot c}$ from the affinity derivative gives entries with opposite signs. The columns are independent. Therefore F is positive definite and its Schur complement $a^I - (b^I)^2/d^I$ is strictly positive. At $p = 1/2$, the cross term vanishes and

$$a^I = m\frac{x^2}{e^{2x} - 1}.$$

For general known p, the clock information is

$$m\frac{x^2 z^2 p(1-p)}{1-z}\left\{\frac{1}{1-p+pz} + \frac{1}{p+(1-p)z}\right\}.$$

This specializes the conditional-information calculation discussed by Radaelli et al. (2023).

Theorem 14 (Nuisance-adjusted graph information)**.** *For finite positive rates, positive gaps and observation counts, $\widetilde{w}_e = a_e^I - (b_e^I)^2/d_e^I > 0$. Profiling any differentiable affinity parameterization independent of the clock parameters gives*

$$I_{v,\mathrm{eff}} \succeq B^\top \operatorname{diag}(\widetilde{w}_e) B. \tag{41}$$

Equality holds for unrestricted dyad affinities. Efficient information is nonsingular exactly when every connected component is nonbipartite; bipartite components admit an exact alternating clock shift.

J.2 Proof of Theorem 14

Let B be the unsigned incidence matrix, R the affinity Jacobian, and let a, b, d collect the entries of the dyad Fisher matrices. Independence conditional on latent parameters gives the block information

$$I_{vv} = B^\top \operatorname{diag}(a) B, \quad I_{v\theta} = B^\top \operatorname{diag}(b) R, \quad I_{\theta\theta} = R^\top \operatorname{diag}(d) R.$$

After removing redundant nuisance coordinates, or equivalently using the orthogonal score projection and a Moore–Penrose inverse, the efficient information is

$$I_{v,\mathrm{eff}} = B^\top \operatorname{diag}(a) B - B^\top \operatorname{diag}(b) R [R^\top \operatorname{diag}(d) R]^\dagger R^\top \operatorname{diag}(b) B. \tag{42}$$

For any clock direction h, the score-projection characterization gives

$$\begin{aligned} h^\top I_{v,\mathrm{eff}} h &= \min_z \sum_e \{a_e (Bh)_e^2 + 2b_e (Bh)_e (Rz)_e + d_e (Rz)_e^2\} \\ &= \sum_e (a_e - b_e^2/d_e)(Bh)_e^2 + \min_z \sum_e d_e \{(Rz)_e + (b_e/d_e)(Bh)_e\}^2 \\ &\geq h^\top B^\top \operatorname{diag}(\widetilde{w}) B h. \end{aligned}$$

When every dyad affinity is free, choose $R = I$ and make the last sum zero, proving equality. Restricting nuisance directions cannot reduce the efficient clock information.

For positive weights, the nullspace of $B^\top \operatorname{diag}(\widetilde{w})B$ equals the nullspace of B. Its elements satisfy $h_i + h_j = 0$ on every observed edge. On a connected component, all values are determined by one root value with signs alternating along paths. An odd cycle forces that value to zero. On a bipartite component, a nonzero value on one side and its negative on the other is a null vector; an isolated vertex is also unobserved. Hence the weighted signless Laplacian is nonsingular exactly when every component is nonbipartite, a classical graph fact (Cvetković et al., 2007).

On a bipartite component, adding ϵh to v leaves Bv exactly unchanged, without changing affinities. Thus every rate and every observation law is unchanged, and no affinity restriction repairs that clock ambiguity. On nonbipartite components the positive lower bound establishes nonsingular efficient information. This proves the theorem's information criterion, not global recovery of the latent coordinates. □

$$\inf_{\widehat{v}_i} \max_{v' \in \{v, v+\log 2\, e_i\}} \mathbb{P}_{v'}\{|\widehat{v}_i - v'_i| \geq \tfrac{1}{2}\log 2\} \geq \tfrac{1}{2}\left(1 - \sqrt{\tfrac{1}{2}\sum_{e \ni i} m_e e^{-2q_e\Delta_e}}\right)_+ . \tag{43}$$

J.3 Proof of the network resolution lower bound

Fix all affinities at zero and let the initial-state law be the same under both alternatives. Shift v_i by $\log 2$. Each incident dyad's total rate doubles; all other dyads are unchanged. For an incident dyad let $z = e^{-q_e\Delta_e}$. Its flip probabilities under the two models are $f = (1-z)/2$ and $g = (1-z^2)/2$. The elementary Bernoulli divergence bound gives

$$D_{\mathrm{Ber}}(f\|g) \leq \frac{(f-g)^2}{g(1-g)} = \frac{z^2(1-z)^2}{1-z^4} \leq z^2.$$

The last inequality follows from $(1-z)^2 \leq 1 - z^4$ for $0 < z < 1$. The Markov chain rule and conditional dyad independence yield

$$\mathrm{KL}(P_v \| P_{v+\log 2 e_i}) \leq \sum_{e \ni i} m_e e^{-2q_e\Delta_e} =: D_*.$$

Pinsker's inequality bounds total variation by $\sqrt{D_*/2}$. Any estimator of v_i induces a test by thresholding at $v_i + \log 2/2$. Under either alternative, a testing error implies an absolute estimation error at least $\log 2/2$. The sum of the two testing error probabilities is at least $1 - \mathrm{TV}$. Taking half the sum and truncating the bound at zero proves (43). If $D_* \leq dme^{-2x} \leq 1/2$, the lower bound is at least $1/4$.

The lower bound is deliberately proved in an easier known-affinity submodel. Unknown affinities cannot make that restricted problem easier. It concerns stationary or homogeneous-rate coarse sampling, and must not be extrapolated to all moving-equilibrium processes; the next section identifies a distinct fast-clock regime.

K Regular endpoint information: derivation and special cases

K.1 Precise regularity and the parameterization

There are finitely many dyads and finitely many observation gaps, fixed independently of λ. Every gap has length at least $\Delta_{\min} > 0$. In a neighborhood of the parameter point under study, write

$$q_{e,\lambda}(t; v, \theta) = \lambda e^{X_e(t)^\top v} h_e(t;\theta), \quad p_e(t;\theta) = \mathrm{sigmoid}(\eta_e(t;\theta)),$$

where X_e is fixed and $h_e > 0$. The geometric-mean clock in (1) has $h_e = 2\cosh(\eta_e/2)$. The alternative convention $h_e = 1$ is also covered and is used only in an explicit confounding diagnostic. The regime index λ is not jointly estimated as a redundant additional intercept.

Assume the following local conditions. The function $q_{e,0}$ is uniformly bounded above and below by positive constants on the finite horizon. The functions $p_e, q_{e,0}$ are piecewise smooth and absolutely continuous on each gap, with bounded first parameter derivatives and appropriate piecewise time

derivatives. Every observation endpoint has a common positive-length left neighborhood in which p_e has bounded derivatives through order three in time, $q_{e,0}$ through order two, and their first parameter derivatives satisfy the same bounds. Endpoint values and derivatives are continuous in the parameter neighborhood. These conditions permit uniform differentiation of the expansions below. Assume additionally $p_r \in [\epsilon, 1-\epsilon]$ and that $J = J_0$ has full column rank as a substantive assumption for Corollary 8.

These conditions allow piecewise-linear trajectories and activities when the last active facet before each endpoint is stable in a parameter neighborhood. They do not automatically hold at a persistent facet tie or for every singular latent-coordinate parameterization. All conclusions are local to a regular chart. At an observation endpoint that is a known motion knot, time derivatives are taken from the left.

K.2 A BOUNDARY-LAYER EXPANSION

Let $y_a(t)$ be the connected-state probability conditional on starting state $a \in \{0, 1\}$ at u. Put $z(t) = y_a(t) - p(t)$ and $Q(s, v) = \int_s^v q_0(t)dt$. Variation of constants gives

$$y_a(v) = p(v) + e^{-\lambda Q(u,v)}\{a - p(u)\} - \int_u^v e^{-\lambda Q(s,v)} p'(s)ds. \tag{44}$$

Pieces separated from v contribute an exponentially small amount, uniformly with first parameter derivatives, because q_0 is bounded below. On the smooth endpoint neighborhood change variable $z = Q(s, v)$; since $q_0 > 0$, its inverse $s = t(z)$ exists. The remaining integral is

$$\int_0^{Q(u,v)} e^{-\lambda z} f(z)dz, \qquad f(z) = \frac{p'(t(z))}{q_0(t(z))}.$$

The expression across earlier nonsmooth pieces can be treated separately; only the neighborhood of $z = 0$ determines its algebraic asymptotics. A first-order Taylor bound gives $f(z) = f(0) + O(z)$, uniformly with first parameter derivatives. Integrating 1 and z against $e^{-\lambda z}$ yields

$$y_a(v) = p(v) - \frac{A}{\lambda} + O(\lambda^{-2}), \qquad A = \frac{p'(v^-)}{q_0(v)}. \tag{45}$$

Parameter differentiation is justified by the uniform derivative assumptions. The derivative of the exponentially small boundary term introduces at most polynomial powers of λ, which remain exponentially small. Since p does not depend on v and $\nabla_v \log q_0 = X$, one obtains

$$\nabla_v y_a(v)^\top = \lambda^{-1} A X(v)^\top + O(\lambda^{-2}), \qquad \nabla_\theta y_a(v)^\top = \nabla_\theta p(v)^\top + O(\lambda^{-1}). \tag{46}$$

The bounds are uniform over the two starting states. This elementary binary expansion is consistent with the broader adiabatic framework of Avron et al. (2012); the result below concerns the additional observation and nuisance geometry.

K.3 REGULAR ENDPOINT CASE: PROOF OF COROLLARY 8

When $J_0 = J$ has full column rank, take $L_0 = I$ and let K_0 have no columns in Theorem 1. Then $U = W^{1/2}J$ and $\mathcal{E}_* = (I - \Pi_{W^{1/2}J})W^{1/2}G = \mathcal{E}$. The endpoint expansion above and the projection argument in Appendix B.2 give

$$I_{v,\text{eff}} = \lambda^{-2}\mathcal{E}^\top\mathcal{E} + O(\lambda^{-3}).$$

The same argument covers fixed starting states and expected conditional information through the norm-preserving start-state replication. A nonzero $\mathcal{E}h$ gives a matching $\Theta(\lambda^{-2})$ rate. If $\mathcal{E}h = 0$, the residual score is $O(\lambda^{-2})$ and its squared norm is $O(\lambda^{-4})$. This is only an upper order; a positive quartic coefficient requires the additional calculation in Appendix D. □

K.4 INTERPRETATION AND LIMITING CASES

Known affinity. With no nuisance columns, $\Pi_J = 0$ and the retained clock information is a weighted Gram matrix of the endpoint lags times the clock design. Pooling multiple endpoint times or dyads can resolve clock directions that a single observation does not excite.

A nuisance parameter can absorb a moving-equilibrium lag. Suppose q_0 is constant and $\eta(t) = \theta_0 + bt$, where b is known and θ_0 unknown. Then $A_r = (b/q_0)p_r(1-p_r)$ and $J_r = p_r(1-p_r)$. The leading lag vector lies in the offset tangent space. A varying p is therefore insufficient to retain λ^{-2} efficient clock information. In the original geometric-mean parameterization, q_0 also changes with η; the multiple-time lag vector need not lie in the affine-offset/slope tangent space. Both cases are evaluated numerically.

Very flexible affinities. If one nuisance parameter can freely perturb each endpoint preference, the endpoint Jacobian spans the entire record space. Then $\mathcal{E} = 0$. Additional temporal variation does not automatically solve the joint identification problem; restrictions or shared structure in affinity are substantive assumptions.

Stationary versus higher-order excitation. If p is constant throughout a gap, (44) contains only its exponentially damped initial-state term. With $q_\lambda = \lambda q_0$, the squared log-clock derivative is then of order $\lambda^2 e^{-2\lambda Q(u,v)}$. In contrast, $p'(v) = 0$ alone does *not* imply exponential information loss: higher nonzero endpoint derivatives can yield higher-order polynomial lags. The theorem isolates the first-order excitation coefficient, not every possible boundary-layer order.

Topology remains relevant. When $X_e = e_i + e_j$, a bipartite alternating clock direction has $X_e^\top h = 0$ on every edge at every time. Then $Gh = 0$ regardless of excitation. Temporal variation does not remove a structural invariance of the clock design. For a nonbipartite graph, nuisance projection can still remove particular lag directions. The spatial and temporal criteria therefore address different obstructions.

L Prior support, matched conditioning, and finite-clock transfer

This section separates three requirements that are not interchangeable: representing uncertainty in every fitted affinity direction, constraining temporal conditioning, and checking whether a frozen-pair surrogate transfers to the finite-clock experiment.

L.1 A missing direction cannot be restored by prior reweighting

The restricted acquisition prior uses four affinity triples

$$(0,0,0), \quad (.25,.15,-.15), \quad (-.25,-.15,.15), \quad (.25,-.15,.15),$$

crossed with four clock scales. Every triple (α,β,γ) satisfies $\beta+\gamma=0$. Its affinity covariance therefore has rank two, whereas the recovery generator varies all three coefficient blocks independently. The following elementary observation explains why this matters specifically for finite-prior information objectives; it is a support diagnostic, not an optimal-design theorem.

Proposition 15 (Support is preserved by finite-KL reweighting)**.** *Let a prior π be supported on an affine set $\{\theta : A\theta = b\}$. For any Q satisfying $\mathrm{KL}(Q\|\pi) < \infty$, and for any posterior formed from π using a nonnegative likelihood with positive evidence, $A\theta = b$ almost surely. Thus $\mathrm{Var}_Q(A\theta) = 0$, and the mutual information between $A\theta$ and any acquired outcome is zero. In contrast, the conditional Fisher information computed by differentiating a likelihood in the full ambient parameter space need not vanish in the missing directions.*

Proof. Finite KL divergence implies $Q \ll \pi$, so any set with zero prior mass has zero Q mass. A posterior with positive evidence is also absolutely continuous with respect to π. The asserted variance and mutual information vanish because $A\theta$ is deterministic under each such law. For the contrast, a Bernoulli observation with success probability $\mathrm{sigmoid}(\theta)$ has ambient Fisher information $\mathrm{sigmoid}(\theta)(1-\mathrm{sigmoid}(\theta)) > 0$, even if the working prior is the point mass at $\theta = 0$. □

Consequently, a larger finite-KL ambiguity radius cannot add the excluded direction to REIG. KW-D and KW-Ds retain all three local affinity derivatives at each scenario, and the Rank-bound conditioning matrix also retains three columns. Giving every method the same rank-two list therefore

does not give their objectives the same scope of nuisance uncertainty. Appendices Q–R report these restricted-prior results separately from the full-affine-rank-prior acquisition-criterion comparison in Appendix M.

L.2 Moment-matched, full-affine-rank tensor priors

A dyad coefficient is the sum of two independent $U[-a, a]$ node coefficients, with $a = .15$. Its mean is zero, variance $2a^2/3 = .015$, and fourth moment $16a^4/15 = .00054$. We approximate each of the three independent dyad coefficients by either of two fixed positive-weight rules:

marginal rule	nodes	weights
2-point	$\{-\sqrt{.015}, \sqrt{.015}\}$	$\{1/2, 1/2\}$
3-point	$\{-\sqrt{.036}, 0, \sqrt{.036}\}$	$\{5/24, 7/12, 5/24\}$.

Taking their tensor products and crossing with the four equally weighted scales $8, 16, 32, 64$ gives 32 and 108 scenarios. Both affinity covariances are $.015I_3$, with rank three and condition number one in the fixed coefficient coordinates. The three-point marginal additionally matches the fourth moment (and all odd moments through degree five). These rules approximate the same independent triangular-coefficient prior; increasing the scenario count does not change its mean or covariance. A finite discrete prior still does not have continuous support, and neither rule is a posterior approximation guarantee for the graph model. The restricted-prior and full-affine-rank-prior studies share graph recovery distributions, an unknown common clock shift, and the likelihood estimator. They differ in their working priors, so outcome differences cannot be attributed solely to covariance rank in a controlled causal sense.

L.3 A common feasible set isolates the conditioning safeguard

For each basis and prior rule, every method selects from the same finite candidate set $\mathcal{D}$. All candidates have six post-initial observations, initial time zero, horizon 5.5, and minimum gap $\delta = 1/64$. For $D = (0, t_1, \ldots, t_6)$ define

$$A(D) = (t_1, t_3, t_5), \quad h(D) = (t_2 - t_1, t_4 - t_3, t_6 - t_5),$$
$$\mathcal{D}_\tau = \{D \in \mathcal{D} : \sigma_{\min}(\Phi(A(D))) \geq \tau\}.$$

The same $\mathcal{D}_\tau$ is used by Rank-bound, KW-Ds, KW-D, REIG and EPIG. The values $\tau = 0, .1, .25, .5$ are fixed before recovery. Every candidate produced by any objective's local search is offered to every other objective, including candidates produced under a different threshold. Thus the final selection is the exact best score in the shared finite pool, but is not a certificate of a continuous global optimum. The full six-gap simplex is available to all objectives: a candidate is not required to be a separated-anchor asymptotic design. In particular, this temporal singular-value safeguard alone does not verify all assumptions of Theorem 5. Detailed constraint ablations are reported in Appendix M, not counted as additional external methods.

L.4 Self-contained acquisition objectives

All objectives maximize a scalar score of one complete six-transition dyad panel; the selected schedule is then transferred to the shared-parameter replicated graph. This is a task-specific criterion comparison, not optimization of the full graph Bayesian likelihood or reproduction of each source paper's training pipeline. KW-Ds and KW-D are two criteria from one paper, not two independent external papers. Conditioning variants and the fixed anchor/probe schedule are internal controls.

Write scenarios as $\xi_s = (\ell_s, \alpha_s, \beta_s, \gamma_s)$ with weights π_s from Appendix L.2, affinity $\eta_s(t) = -.2 + (\alpha_s, \beta_s, \gamma_s)\psi(t)$, and physical clock $e^{c_s} = \ell_s$. For $D = (0, t_1, \ldots, t_6)$, let $P_{s,r}$ be the exact transition matrix and let $\mu_{s,r}(a)$ be the probability of the starting state at transition r, propagated forward from $\mu_{s,1}(1) = \mathrm{sigmoid}(\eta_s(0))$. The initial state is conditioned on in every likelihood; its information is not rewarded. Expectations over its law do not differentiate that law as an extra observation.

Rank-bound. For anchors $a_l = t_{2l-1}$ and widths $h_l = t_{2l} - t_{2l-1}$, set $r_{sl} = 2\ell_s h_l \cosh(\eta_s(a_l)/2)$ in (16). With w_{sl} from (15),

$$J_{\text{bound}}(D) = \min_s \sum_{l=1}^{3} w_{sl}.$$

This is a worst-scenario frozen-pair surrogate. Scenario weights enter the other criteria but not this minimum. A scalar-scenario minimum is not a verified uniform graph-information floor over unseen parameters.

Nuisance-adjusted and joint Fisher criteria. For $y_{sra} = P_{s,r}(a,1)$ form the 12-by-4 conditional-score matrix with row

$$S_s(r,a) = \sqrt{\frac{\mu_{s,r}(a)}{y_{sra}(1-y_{sra})}}\partial_{(c,\alpha,\beta,\gamma)} y_{sra}.$$

Partition $S_s = [C_s, N_s]$ into one clock and three affinity columns. The implemented adaptations of Kyriakopoulos and Wolf (2015) are

$$J_{\text{KW-Ds}}(D) = \sum_s \pi_s \|(I - \Pi_{N_s})C_s\|^2, \qquad J_{\text{KW-D}}(D) = \log\det\left(\sum_s \pi_s S_s^\top S_s\right).$$

Thus KW-Ds averages *scenario-wise profiled* information; KW-D takes the determinant *after* averaging full information. They are not interchangeable expected-log-determinant objectives. All three ambient nuisance derivatives are retained. The implementation uses score SVDs, evaluating the joint log determinant as twice the sum of log singular values of the stacked matrices $\sqrt{\pi_s}S_s$. It introduces no determinant ridge or probability renormalization.

Conditional paths and robust information gain. For initial state a, let $p_a = \sum_s \pi_s \mu_{s,1}(a)$ and $\pi_{s|a} = \pi_s \mu_{s,1}(a)/p_a$. Enumerate all $y = (y_1, \ldots, y_6) \in \{0,1\}^6$, with $y_0 = a$, and define

$$L_{s,a}(y) = \prod_{r=1}^{6} P_{s,r}(y_{r-1}, y_r), \qquad \overline{L}_a(y) = \sum_s \pi_{s|a} L_{s,a}(y), \qquad k_{s,a} = \sum_y L_{s,a}(y) \log \frac{L_{s,a}(y)}{\overline{L}_a(y)}.$$

The REIG adapter (Go and Isaac, 2022, Equations (14) and (18)) uses radius $\rho = .1$:

$$J_{\text{REIG}}(D) = \sum_{a=0}^{1} p_a \min_{q \in \Delta:\, \text{KL}(q\|\pi_{\cdot|a}) \le \rho} \sum_s q_s k_{s,a}.$$

The nominal mixture $\overline{L}_a$ is held fixed inside this minimization: this is the source criterion's affine robustification, not minimization of a recomputed mutual information for every q. The exponential tilt $q_s \propto \pi_{s|a}\exp(-k_{s,a}/\nu)$ solves the interior case; a one-dimensional root search chooses ν to meet the KL constraint. The minimum-support boundary case is handled explicitly. Nominal EIG $\sum_a p_a \sum_s \pi_{s|a} k_{s,a}$ is saved as a diagnostic, not an extra external baseline.

Prediction-oriented information gain. For each of 21 equally spaced times u on $[0, 5.5]$, define an independent equilibrium-state target $Z_u \mid \xi_s \sim \text{Bernoulli}(\text{sigmoid}(\eta_s(u)))$. Conditional independence here is part of the *target definition*, not a forecast of the acquired chain's next state. Put

$$J_{a,u}(y,z) = \sum_s \pi_{s|a} L_{s,a}(y) \mathbb{P}(Z_u = z \mid \xi_s), \quad m_{a,u}(z) = \sum_s \pi_{s|a} \mathbb{P}(Z_u = z \mid \xi_s).$$

The EPIG adapter (Bickford Smith et al., 2023) is

$$J_{\text{EPIG}}(D) = \sum_a p_a \frac{1}{21} \sum_u \sum_{y,z} J_{a,u}(y,z) \log \frac{J_{a,u}(y,z)}{\overline{L}_a(y) m_{a,u}(z)}.$$

Path sums are exact over 64 outcomes; transition integrals use checked quadrature. No Monte Carlo EIG estimator or neural surrogate is used. The recovery predictor is the common graph maximum-likelihood fit rather than the source paper's Bayesian predictor; the comparison evaluates EPIG's acquisition objective within this specified estimation pipeline. Selected objective values are rechecked at doubled quadrature order. The indexed acquisition-objectives code is checked independently for probabilities, derivatives, path enumeration and the robust dual.

L.5 Pointwise finite-clock transfer diagnostics

For a fixed graph, parameters and schedule, form a comparison experiment by replacing each transition into t_{2l-1} by an independent Bernoulli anchor with probability $p_e(t_{2l-1})$, and freezing affinity at that anchor during the following probe. Let C_f, N_f be its score matrices and C, N the finite-clock score matrices, using identical conditional starting-state exposure conventions and row ordering. This comparison law exists for any positive gaps; it approximates the true law only when anchors mix and affinity moves little over probes.

Normalize graph-clock coordinates using $G = B^\top B$ and $W = G^{-1/2}$, and normalize each node-affinity block by the same W. Define $\widetilde{C} = CW$, $\widetilde{N} = N\,\mathrm{diag}(W, W, W)$ and similarly for f. The efficient information in these coordinates is $\widetilde{I} = G^{-1/2} I_{v,\mathrm{eff}} G^{-1/2}$. The enlarged-nuisance argument for independent anchor values gives

$$\widetilde{I}_f \succeq \kappa_f I_n, \qquad \kappa_f = \min_e \sum_l w_{el}.$$

If $s = \sigma_{\min}(\widetilde{N}_f) > 0$, $\epsilon = \max\{\|\widetilde{C} - \widetilde{C}_f\|, \|\widetilde{N} - \widetilde{N}_f\|\} \le s/2$, $M = \|\widetilde{C}_f\|$, and $a = \epsilon(1 + 4M/s)$, the proof of Proposition 6 applies without an asymptotic limit and yields

$$\lambda_{\min}(\widetilde{I}) \ge \kappa_f - (2M + a)a. \tag{47}$$

This is a deterministic, pointwise sufficient inequality. A numerical value of κ_f alone does not establish it, and a negative right-hand side or failure of $\epsilon \le s/2$ is inconclusive. Our held-out checks report the actual finite-clock information, surrogate, remainder and sufficient margin separately. They use floating-point refinement and are not interval-arithmetic certificates or bounds uniform over the unobserved parameter space.

L.6 Validity, tightness, and ranking are different questions

For a fixed parameter and candidate D, write $g(D) = \lambda_{\min}(\widetilde{I}_D)$, $f(D) = \lambda_{\min}(\widetilde{I}_{f,D})$, and $\kappa(D) = \min_e \sum_l w_{el}$. The frozen model gives $\kappa(D) \le f(D)$; this alone gives neither $\kappa(D) \le g(D)$ nor order preservation between candidates. When the perturbation condition in (47) holds with remainder Δ_D, Weyl's inequality supplies the interval

$$L_D = \max\{0, \kappa(D) - \Delta_D\} \le g(D) \le U_D = f(D) + \Delta_D.$$

Thus $L_D > U_{D'}$ certifies a pointwise ranking, whereas $\kappa(D) > \kappa(D')$ does not. If such intervals are valid for every candidate in a finite pool and $\widehat{D}$ maximizes L_D, then

$$0 \le \max_D g(D) - g(\widehat{D}) \le \max_D U_D - L_{\widehat{D}}.$$

Both statements follow directly by substituting the lower and upper inequalities. Constructing these intervals from actual score differences requires the parameters and the finite-clock computation. They are therefore *oracle diagnostics*, evaluated at specified parameters rather than uniformly over unknown parameters. Applying a validated bound uniformly over a declared parameter set would require additional analysis. Appendix N evaluates ranking directly, including cases where the larger surrogate selects worse actual information.

M Acquisition-criterion comparisons: protocols and complete outcomes

This section reports the complete acquisition-criterion comparison and sensitivity outcomes, separately from the finite-clock planner evaluation in Appendix A. The artifact collection indexed in `STUDY_PATHS.json` contains protocols, candidate pools, counts, parameters, starts, attempts, diagnostics and aggregation code. Restricted-prior results are reported separately and are not pooled with this experiment.

Table 24: Matched full-affine-rank prior, shared conditioning threshold $\tau = .5$, K_6, $\lambda = 32$, $R = 100$. Entries are mean RMSE (sample SD) across 30 paired seeds; lower is better. All methods share the candidate pool, count, horizon and gap floor. Constrained variants are not additional external baselines.

	Polynomial		Fourier	
Design	Clock	Affinity	Clock	Affinity
Rank-bound	0.095 (0.026)	0.140 (0.027)	0.091 (0.023)	0.108 (0.020)
KW-Ds (HSB'14)	0.091 (0.025)	0.123 (0.028)	0.092 (0.021)	0.111 (0.021)
KW-D (HSB'14)	0.199 (0.145)	0.088 (0.020)	0.202 (0.247)	0.093 (0.018)
REIG (UAI'22)	0.085 (0.025)	0.188 (0.040)	0.089 (0.023)	0.107 (0.020)
EPIG (AISTATS'23)	1.251 (0.280)	0.138 (0.037)	1.553 (0.096)	0.134 (0.029)

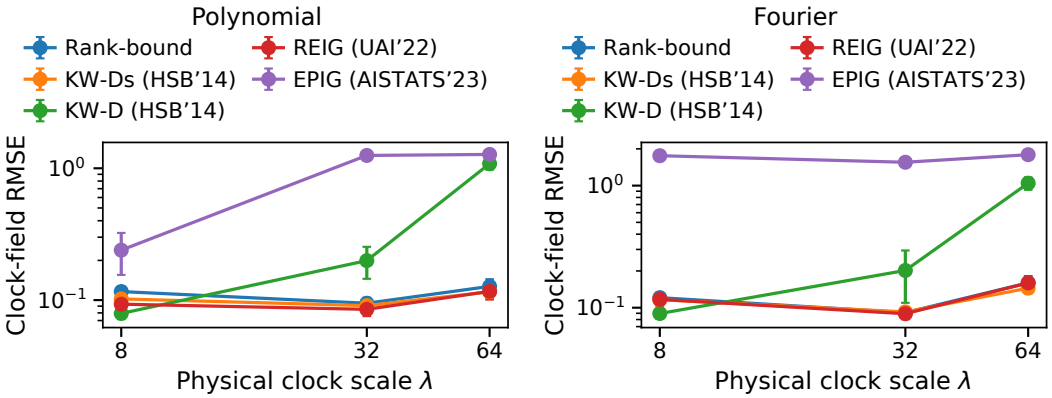

Figure 2: All three physical scales under the common 108-scenario prior and $\tau = .5$, on K_6 with $R = 100$ and gap floor $1/64$. Mean log-clock RMSE with marginal 95% intervals over 30 paired seeds (logarithmic vertical axis). These are not the 156-contrast simultaneous intervals. Every selected outcome is retained.

M.1 Design, generation and fitting protocol

Chronology and candidate access. The protocol was fixed before design selection and recovery; it was not externally preregistered. The two priors in Appendix L.2 use fixed positive weights. For each family/prior, an initial bank contains 512 scrambled Sobol six-gap allocations, 256 proposed structured short-pair schedules (only feasible ones retained), the restricted-prior schedules and a uniform schedule. Softmax coordinates allocate the excess length $5.5 - 6\delta$ above six minimum gaps. All methods share this initial bank. At each threshold $0, .5$, each of the five objectives receives two Nelder–Mead starts and a cap of 160 evaluations per start. Every evaluated feasible candidate is retained in the common pool. Final selection follows the shared finite-pool rule in Appendix L.3 at all four thresholds. The final pools contain 3,125/3,108 polynomial and 3,217/3,209 Fourier candidates for 32/108 scenarios. Search logs, local termination statuses and all candidate scores are saved.

The score-SVD implementation avoids condition-number squaring in the joint Fisher determinant. Selected scores pass the 10^{-8} objective-refinement gate; no method receives shorter gaps or additional observations. A Gram-determinant pilot fails the same gate and is excluded from these selections; its code, outputs and failed check are retained in the indexed pilot records.

Independent realizations and count data. The primary seeds are 13000–13029, disjoint from the restricted-prior 11000–11009 and 12000–12029 sets. For each seed, six node clock deviations and three independent six-node affinity blocks follow $U[-.15, .15]$, with streams $\texttt{seed} + 50000j$, $j = 0, 1, 2, 3$. On K_6, the true clock coordinates are $v_i = v_i^{\mathrm{dev}} + \frac{1}{2}\log(\lambda/32)$. Inference fixes the reference 32, not the true λ. For each family, scale $8, 32, 64$, objective and threshold $0, .5$, a separate conditional-panel dataset is generated with $R = 100$ paths, 15 dyads and six transitions, hence

Table 25: Selected 108-scenario schedules. The conditioned initial state is at zero and the last observation at 5.5. All shown intervals obey $\delta = 1/64$ before rounding; the condition is on t_1, t_3, t_5. The exact schedules and every candidate score are saved in the artifact.

Family/τ	Design	t_1	t_2	t_3	t_4	t_5	t_6	$\sigma_{\min}$
Poly/0	Rank-bound	1.8177	1.8333	1.8489	1.8646	5.4844	5.5000	0.0114
Poly/0	KW-Ds (HSB'14)	5.4219	5.4375	5.4531	5.4688	5.4844	5.5000	4.31×10^{-5}
Poly/0	KW-D (HSB'14)	0.0418	0.0876	2.7587	2.8015	5.4625	5.5000	1.0267
Poly/0	REIG (UAI'22)	0.0156	5.4375	5.4531	5.4687	5.4844	5.5000	0.0171
Poly/0	EPIG (AISTATS'23)	0.1427	0.3615	5.0195	5.2038	5.3547	5.5000	0.1873
Poly/0.5	Rank-bound	1.0268	1.0424	2.2843	2.2999	5.4844	5.5000	0.5000
Poly/0.5	KW-Ds (HSB'14)	0.0156	3.6727	3.6883	3.7039	5.4844	5.5000	0.8632
Poly/0.5	KW-D (HSB'14)	0.0418	0.0876	2.7587	2.8015	5.4625	5.5000	1.0267
Poly/0.5	REIG (UAI'22)	0.0156	0.0315	0.9469	5.4685	5.4842	5.5000	0.5000
Poly/0.5	EPIG (AISTATS'23)	0.1420	0.3058	4.3991	4.7613	5.3338	5.5000	0.5036
Fourier/0	Rank-bound	1.8277	1.8433	3.6427	3.6583	5.4844	5.5000	1.2173
Fourier/0	KW-Ds (HSB'14)	0.0156	5.4375	5.4531	5.4687	5.4844	5.5000	3.68×10^{-4}
Fourier/0	KW-D (HSB'14)	1.7978	1.8331	3.6312	3.6665	5.4647	5.5000	1.2246
Fourier/0	REIG (UAI'22)	0.0156	5.4375	5.4531	5.4687	5.4844	5.5000	3.68×10^{-4}
Fourier/0	EPIG (AISTATS'23)	0.4957	0.8845	4.7236	4.9924	5.2785	5.5000	0.1553
Fourier/0.5	Rank-bound	1.8277	1.8433	3.6427	3.6583	5.4844	5.5000	1.2173
Fourier/0.5	KW-Ds (HSB'14)	2.0882	2.1038	3.4630	3.4786	5.4844	5.5000	0.9987
Fourier/0.5	KW-D (HSB'14)	1.7978	1.8331	3.6312	3.6665	5.4647	5.5000	1.2246
Fourier/0.5	REIG (UAI'22)	1.8314	1.8470	3.6438	3.6594	5.4844	5.5000	1.2161
Fourier/0.5	EPIG (AISTATS'23)	0.4473	0.8879	3.7010	4.4666	4.9776	5.5000	0.5009

9,000 observed transitions. Initial states are drawn from the true equilibrium law; the likelihood conditions on their realized values. A sequential count simulation preserves state continuity rather than generating unrelated transitions.

Generation uses an independently coded DOP853 forward ODE with relative/absolute tolerances $2 \times 10^{-12}/2 \times 10^{-14}$. Its probabilities are checked against endpoint quadrature before generating counts. The random count stream is $\texttt{seed} + 10000\lambda + R + 10^6 n + 10^8 \texttt{family}$, shared across designs within a condition. Shared parameter and random streams create paired comparisons; they do not make records within a seed independent. Each data file retains the graph, truth, schedule, probabilities, counts and stream identifiers. Identical selected schedules can give identical arrays under different method labels; the number of files is not a count of independent samples.

Recovery estimator and numerical configuration. All 24 node parameters are estimated, with clock bounds $[-1.6, 1.6]^6$ and affinity bounds $[-.7, .7]^{18}$ in the primary experiment. Both the primary and follow-up drivers call `fit` in `robustness_extension/code/sensitivity_model.py`, using `acquisition_study/code/model.py` and the safe quadrature in `robustness_extension/code/panel_kernel.py`.

The objective is signed binomial-deviance least squares for the conditional transition counts, with each row weighted by its observed starting-state count divided by R. It has the same minimizer as the conditional negative log likelihood. Analytic probability derivatives give the residual Jacobian; a fourth-order local KL expansion avoids cancellation when an interior empirical probability differs from its fitted value by less than 10^{-4}. Bounded trust-region reflective least squares uses Jacobian-based coordinate scaling, `ftol=1e-10` and `xtol=gtol=1e-9`. Starts are zero and a Gaussian draw from stream $\texttt{seed} + 600000$, with coordinate standard deviations .22 for clocks and .08 for affinities, clipped inside the common bounds by 10^{-8}. Neither start uses truth or outcomes. The selected start minimizes training deviance. A selected unsuccessful status triggers reruns from the same starts at caps 1,200 then 4,800 after the initial 300 evaluations per start; every attempt remains.

The positive endpoint-layer quadrature uses outer orders $o \in \{40, 80, 160\}$ with inner order $\max(12, \lfloor o/3 \rfloor)$. Refinement replaces o by $2o$ and recomputes the inner order. A nonfinite or nonpositive probability, or row-mass error at least 2×10^{-10}, triggers geometric panel subdivision of the same integral; probabilities and derivatives are not renormalized. The tail cutoff uses a lower integrated-clock bound of 80. At a selected fit, maximum changes in log probabilities and probability derivatives must be below 10^{-10} and 10^{-9}, respectively, between the base and refined rules;

Table 26: Complete matched primary experiment: 108 scenarios, 30 paired seeds per cell, K_6, $R = 100$. Mean (SD) clock/affinity RMSE. All 1,800 fits are retained.

Design	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
	Clock	Affinity	Clock	Affinity	Clock	Affinity
Polynomial, $\tau = 0$						
Rank-bound	0.094 (0.027)	0.664 (0.103)	0.078 (0.021)	0.694 (0.117)	0.115 (0.032)	0.705 (0.108)
KW-Ds (HSB'14)	0.082 (0.026)	1.044 (0.112)	0.067 (0.022)	1.001 (0.132)	0.093 (0.028)	0.987 (0.090)
KW-D (HSB'14)	0.079 (0.020)	0.107 (0.021)	0.199 (0.145)	0.088 (0.020)	1.083 (0.290)	0.086 (0.014)
REIG (UAI'22)	0.093 (0.030)	0.861 (0.122)	0.077 (0.017)	0.852 (0.126)	0.096 (0.040)	0.885 (0.130)
EPIG (AISTATS'23)	0.202 (0.225)	0.286 (0.075)	1.264 (0.246)	0.285 (0.074)	1.258 (0.166)	0.274 (0.074)
Polynomial, $\tau = 0.5$						
Rank-bound	0.116 (0.029)	0.150 (0.031)	0.095 (0.026)	0.140 (0.027)	0.127 (0.046)	0.126 (0.032)
KW-Ds (HSB'14)	0.102 (0.028)	0.185 (0.036)	0.091 (0.025)	0.123 (0.028)	0.116 (0.035)	0.100 (0.019)
KW-D (HSB'14)	0.079 (0.020)	0.107 (0.021)	0.199 (0.145)	0.088 (0.020)	1.083 (0.290)	0.086 (0.014)
REIG (UAI'22)	0.093 (0.028)	0.236 (0.051)	0.085 (0.025)	0.188 (0.040)	0.116 (0.042)	0.188 (0.053)
EPIG (AISTATS'23)	0.239 (0.225)	0.134 (0.036)	1.251 (0.280)	0.138 (0.037)	1.274 (0.127)	0.134 (0.033)
Fourier, $\tau = 0$						
Rank-bound	0.121 (0.033)	0.118 (0.020)	0.091 (0.023)	0.108 (0.020)	0.160 (0.058)	0.095 (0.017)
KW-Ds (HSB'14)	0.087 (0.025)	1.192 (0.170)	0.074 (0.023)	1.307 (0.193)	0.101 (0.028)	1.286 (0.189)
KW-D (HSB'14)	0.090 (0.023)	0.115 (0.018)	0.202 (0.247)	0.093 (0.018)	1.047 (0.327)	0.088 (0.017)
REIG (UAI'22)	0.087 (0.025)	1.192 (0.170)	0.074 (0.023)	1.307 (0.193)	0.101 (0.028)	1.286 (0.189)
EPIG (AISTATS'23)	1.565 (0.468)	0.343 (0.094)	1.346 (0.214)	0.344 (0.087)	1.488 (0.200)	0.310 (0.099)
Fourier, $\tau = 0.5$						
Rank-bound	0.121 (0.033)	0.118 (0.020)	0.091 (0.023)	0.108 (0.020)	0.160 (0.058)	0.095 (0.017)
KW-Ds (HSB'14)	0.116 (0.035)	0.124 (0.020)	0.092 (0.021)	0.111 (0.021)	0.145 (0.038)	0.102 (0.018)
KW-D (HSB'14)	0.090 (0.023)	0.115 (0.018)	0.202 (0.247)	0.093 (0.018)	1.047 (0.327)	0.088 (0.017)
REIG (UAI'22)	0.117 (0.034)	0.117 (0.019)	0.089 (0.023)	0.107 (0.020)	0.160 (0.058)	0.095 (0.017)
EPIG (AISTATS'23)	1.758 (0.450)	0.143 (0.029)	1.553 (0.096)	0.134 (0.029)	1.792 (0.207)	0.139 (0.038)

otherwise both starts are rerun at the next outer tier. If all precision tiers are exhausted, the fit fails. A coordinate within 10^{-3} of its fitting bound is flagged as a boundary estimate.

Retained outcomes. All 1,800 primary fits execute. One selected EPIG fit (Fourier, scale 64, $\tau = 0$, seed 13009) reaches the final cap; its clock/affinity RMSEs are $1.344/.452$. It is not excluded or relabeled converged. Across primary fits, 378 hit a clock bound and 457 an affinity bound; categories can overlap and are not counts of independent failures. All 700 secondary fits terminate successfully. These flags characterize the bounded estimator, not global optimality or identifiability.

Metrics and inference. Clock RMSE is $\|B(\widehat{v} - v)\|/\sqrt{m}$, including the unknown common shift. Affinity RMSE averages squared logit-trajectory errors over all dyads and 21 equally spaced horizon times; it is not restricted to observed times. The prediction metrics use the corresponding sigmoid probabilities. The primary inferential family has 156 seed-paired contrasts: four external criteria minus Rank-bound, at two thresholds, two families, three scales and two errors (96), plus $\tau = .5$ minus zero for five methods, two families, three scales and two errors (60). Each uses 30 paired seeds and multiplier $t_{29,1-.05/(2\cdot 156)} = 4.081486$. These are approximate simultaneous Student-t intervals under their usual distributional assumptions, not distribution-free coverage. All contrasts are saved, including inconclusive and unfavorable results. No seeds are pooled with the restricted-prior studies.

M.2 What the common conditioning safeguard contributes

The comparison changes only the objective within a given common feasible set. At scale 32, polynomial Rank-bound, KW-Ds and REIG all exhibit poor affinity recovery without that safeguard. Their respective mean affinity errors drop from $.694, 1.001, .852$ to $.140, .123, .188$ at $\tau = .5$. KW-D retains its well-conditioned selected schedule. In Fourier, unconstrained Rank-bound already selects a well-conditioned schedule; its result is therefore exactly unchanged. KW-Ds and REIG change

Table 27: Central paired differences ($\lambda = 32$), with simultaneous Bonferroni-adjusted intervals over all 156 prespecified primary contrasts. Negative favors the first alternative: the method rather than Rank-bound, or $\tau = .5$ rather than zero. A zero-width interval occurs for an unchanged selected schedule.

Family	Comparison	Clock: difference [interval]	Affinity: difference [interval]
Poly	KW-Ds (HSB'14) − RB	-0.004 [-0.026, 0.018]	-0.017 [-0.043, 0.009]
Poly	KW-D (HSB'14) − RB	0.105 [-0.009, 0.218]	-0.051 [-0.077, -0.025]
Poly	REIG (UAI'22) − RB	-0.010 [-0.036, 0.017]	0.049 [0.022, 0.075]
Poly	EPIG (AISTATS'23) − RB	1.157 [0.943, 1.371]	-0.002 [-0.029, 0.026]
Poly	Rank-bound: .5 − 0	0.017 [-0.002, 0.035]	-0.555 [-0.634, -0.476]
Poly	KW-Ds (HSB'14): .5 − 0	0.023 [0.002, 0.045]	-0.878 [-0.977, -0.779]
Poly	KW-D (HSB'14): .5 − 0	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
Poly	REIG (UAI'22): .5 − 0	0.008 [-0.011, 0.026]	-0.663 [-0.762, -0.565]
Poly	EPIG (AISTATS'23): .5 − 0	-0.012 [-0.245, 0.220]	-0.147 [-0.190, -0.104]
Fourier	KW-Ds (HSB'14) − RB	0.001 [-0.012, 0.014]	0.003 [-0.005, 0.011]
Fourier	KW-D (HSB'14) − RB	0.111 [-0.078, 0.299]	-0.015 [-0.030, 0.000]
Fourier	REIG (UAI'22) − RB	-0.002 [-0.009, 0.005]	-0.001 [-0.004, 0.003]
Fourier	EPIG (AISTATS'23) − RB	1.462 [1.388, 1.537]	0.027 [0.001, 0.052]
Fourier	Rank-bound: .5 − 0	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
Fourier	KW-Ds (HSB'14): .5 − 0	0.018 [-0.002, 0.038]	-1.196 [-1.344, -1.048]
Fourier	KW-D (HSB'14): .5 − 0	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
Fourier	REIG (UAI'22): .5 − 0	0.015 [-0.006, 0.036]	-1.200 [-1.349, -1.051]
Fourier	EPIG (AISTATS'23): .5 − 0	0.207 [0.042, 0.372]	-0.210 [-0.264, -0.155]

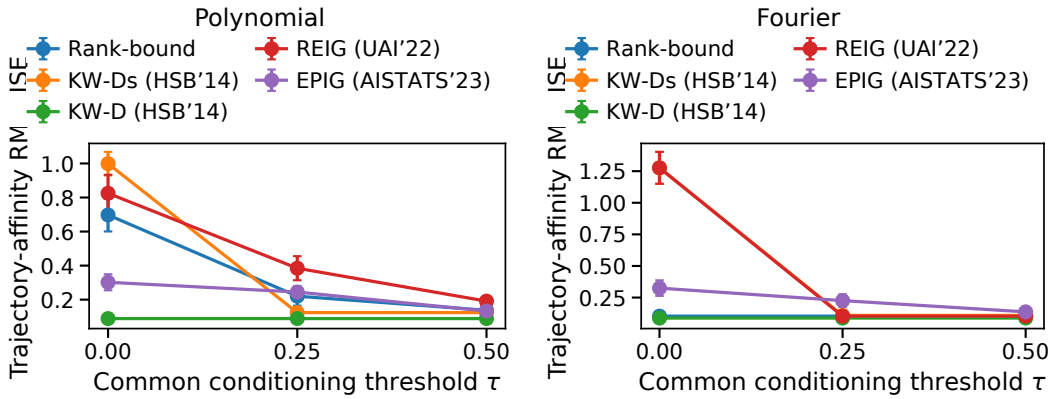

Figure 3: Common-threshold ablation, scale 32 and the same ten paired seeds at every point. Bars are marginal 95% Student-t intervals, not the 156-contrast simultaneous intervals. The safeguard benefits several objectives; unchanged selections produce unchanged outcomes.

from the same clustered schedule to well-conditioned alternatives and recover affinities comparably to Rank-bound.

The adjusted central intervals do not distinguish Rank-bound clock error from either safeguarded KW-Ds or REIG in either family. Safeguarded KW-D has a supported polynomial affinity advantage, while safeguarded REIG has higher polynomial affinity error than Rank-bound. These mixed results indicate that both conditioning and target choice affect recovery. The full scale grid and its boundary outcomes remain visible in Table 26.

For a more granular threshold sensitivity, the first ten primary seeds are also fitted at $\tau = .25$ with the 108-scenario prior. Figure 3 restricts every displayed threshold to that same ten-seed subset. Thus it does not connect 30-seed endpoints to a ten-seed midpoint as if they were one cohort. The $\tau = .1$ design and finite-information diagnostics are saved, but no recovery at .1 is claimed.

Table 28: Sensitivity analyses on the first ten primary seeds, $\lambda = 32$, $\tau = .5$. Each entry is mean clock/affinity RMSE. All columns except the 32-scenario column use the 108-scenario design. The wide-box and timestamp columns reuse the same physical primary datasets; these are not independent replications.

Family	Design	Primary subset	32 scenarios	Box $b = 2.8$	Timestamp error
Poly	Rank-bound	0.087/0.137	0.087/0.137	0.087/0.137	0.151/0.137
Poly	KW-Ds (HSB'14)	0.078/0.124	0.078/0.124	0.078/0.124	0.110/0.125
Poly	KW-D (HSB'14)	0.171/0.089	0.171/0.089	0.171/0.089	0.171/0.089
Poly	REIG (UAI'22)	0.084/0.191	0.084/0.191	0.084/0.191	0.102/0.190
Poly	EPIG (AISTATS'23)	1.135/0.133	1.187/0.136	1.114/0.133	1.129/0.133
Fourier	Rank-bound	0.090/0.103	0.090/0.103	0.090/0.103	0.147/0.103
Fourier	KW-Ds (HSB'14)	0.092/0.109	0.092/0.109	0.092/0.109	0.146/0.109
Fourier	KW-D (HSB'14)	0.160/0.088	0.160/0.088	0.160/0.088	0.169/0.088
Fourier	REIG (UAI'22)	0.090/0.103	0.090/0.103	0.090/0.103	0.147/0.103
Fourier	EPIG (AISTATS'23)	1.549/0.137	1.456/0.140	1.563/0.138	1.540/0.137

M.3 Prior count, estimator bounds and timestamp error

The 700 secondary fits comprise 200 fits using the 32-scenario prior at both primary thresholds; 100 using the 108-scenario prior at $\tau = .25$; 200 wider-box refits of fixed primary data; and 200 corrupted-timestamp refits of fixed primary data. All use the first ten primary seeds and scale 32. There are 2,100 distinct data keys across the 2,500 selected fits, not 2,500 distinct datasets. Complete means, SDs, paired differences and descriptive unadjusted 95% intervals are in the CSV summaries. These exploratory intervals are separate from the primary 156-contrast family and do not establish additional independent replications.

The two priors share first and second moments by construction; only the 108-scenario rule also matches fourth moments. Similar results here support stability to this particular discretization change, not adequacy of every discrete prior or convergence to a continuous-prior optimum. Full affine rank is necessary to avoid the demonstrated missing linear direction, but does not ensure representative tails, nonlinear coverage or prior calibration.

Widening the affinity box from .7 to 2.8 keeps the physical dataset, clock bounds, initial starts and objective fixed. With $\tau = 0$, polynomial Rank-bound affinity RMSE rises from .697 to 2.313, and KW-Ds from .999 to 3.946. Fourier KW-Ds and REIG rise from 1.276 to 5.062. With $\tau = .5$, all these methods' central estimates remain close to their standard-box values. Hence the error magnitude of the unconditioned fits is substantially estimator-dependent; these comparisons do not establish a bound-independent superiority ratio. The restricted-prior eight-start checks apply only to that study's datasets, not to the full-affine-rank-prior fits reported here.

For timestamp measurement error, data are still generated at the exact physical schedule. Only reported interior times $t_1, \ldots, t_5$ are perturbed by independent $U[-1/256, 1/256]$ errors from stream $\texttt{seed} + 801000$; times zero and 5.5 stay fixed. Because $1/256 = \delta/4$, order is preserved. The physical hardware floor holds before measurement corruption, not necessarily for the recorded intervals. Fits treat those corrupted timestamps as exact. This is a deliberately misspecified estimator, not an error-aware procedure or an additional clock-identification theorem. On safeguarded Rank-bound the subset clock errors rise from $.087/.090$ to $.151/.147$ in polynomial/Fourier models, while affinity errors remain near $.137/.103$. Short probes are therefore sensitive to timestamp accuracy even when temporal affinity conditioning is good.

M.4 Prediction target and finite-clock diagnostics

For the independent equilibrium Bernoulli target, we score each likelihood-fitted predictor on the same 21-time grid. Excess log loss is the average Bernoulli KL divergence from the true target to that predictor; excess Brier risk is the mean squared probability difference. This removes irreducible target noise without creating extra outcome samples. It evaluates the target motivating EPIG, but with the common graph likelihood estimator rather than a scenario-posterior Bayesian predictor. The lowest central mean excess log loss is attained by KW-D. EPIG generally improves on the

Table 29: Secondary predictive target: mean excess log loss against the true equilibrium Bernoulli target, $\lambda = 32$, 30 seeds, 108 scenarios. All designs use the same fitted likelihood predictor; these are not source-paper Bayesian predictors.

	Polynomial		Fourier	
Design	$\tau = 0$	$\tau = .5$	$\tau = 0$	$\tau = .5$
Rank-bound	0.0550	0.0025	0.0015	0.0015
KW-Ds (HSB'14)	0.1138	0.0019	0.1822	0.0016
KW-D (HSB'14)	0.0010	0.0010	0.0011	0.0011
REIG (UAI'22)	0.0845	0.0045	0.1822	0.0014
EPIG (AISTATS'23)	0.0105	0.0025	0.0149	0.0023

Table 30: Finite-clock transfer on 20 held-out graph parameters per family (108-scenario designs, $\tau = .5$). κ_f is the frozen-pair surrogate; ℓ is the actual minimum eigenvalue in graph-normalized coordinates. Values are means. Positive counts pointwise sufficient margins satisfying both perturbation conditions. These are floating-point diagnostics, not uniform certificates.

	Polynomial			Fourier		
Design	κ_f	ℓ	Positive	κ_f	ℓ	Positive
Rank-bound	0.379	0.395	13/20	0.379	0.395	19/20
KW-Ds (HSB'14)	0.252	0.526	0/20	0.379	0.395	17/20
KW-D (HSB'14)	0.223	0.323	0/20	0.270	0.289	19/20
REIG (UAI'22)	0.250	0.523	0/20	0.379	0.395	19/20
EPIG (AISTATS'23)	0.002	0.008	0/20	$< 10^{-4}$	$< 10^{-4}$	0/20

unconstrained clock-focused criteria on its own target despite having poor clock estimates within the common-estimator pipeline.

The held-out finite-clock audit uses seeds 14000–14019, independent of the recovery seeds. For each, a nominal scale is drawn log-uniform on $[8, 64]$ using stream $\texttt{seed} + 810000$, and graph coefficients are drawn as before. Both families, both prior orders, all five objectives and four thresholds give 1,600 labeled evaluations; identical schedules reuse the same computed information and do not add independent cases. No held-out information value changes a selected schedule.

Each case applies the normalized comparison in Appendix L.5, reporting the actual and frozen minimum efficient eigenvalues, dyad-minimum surrogate, both score errors, nuisance singular value, perturbation condition and sufficient margin. The frozen lower bound and every applicable perturbation inequality are checked with doubled quadrature order. Table 30 retains successful and unsuccessful margins; their pointwise interpretation is specified with (47).

M.5 Execution and reproducibility checks

The executable audit requires complete specification coverage and checks every raw count array, conditional state continuity, random-stream regeneration, generating coordinates, selected-design hash, truth-independent starts, bounds, objective selection, final likelihood, all reported errors and doubled-order probabilities/derivatives. Independently coded forward ODE and finite-difference checks examine selected prior-design cases for both families and both scenario counts. Separately enumerated paths and a robust dual verify REIG and EPIG values. Mutation tests reject a corrupted count, nonfinite counts and missing coverage. These are automated numerical and record-consistency checks.

Table 31 summarizes the primary-study checks. Appendix N.4 specifies the 12 stratified regeneration cases and the distinction between regeneration, fixed-data refitting and cached replay. These are verification reruns of existing seeds, not additional independent test realizations.

All code dependencies and raw evidence are supplied under the artifact root. The primary-summary entry point in `STUDY_PATHS.json` generates the primary and sensitivity tables and complete CSV summaries only when both cohorts are complete. Local-search caps, the retained fit cap and

Table 31: Primary-study automated verification. Numerical entries are maximum observed discrepancies, not interval enclosures. Independent ODE/derivative/criterion checks cover 12 selected prior-design cases; count, fit, metric and likelihood replay covers the complete primary and sensitivity cohorts.

Check	Observed value
Count arrays / selected fits replayed	2,100 / 2,500
Recomputed metric discrepancy	0
Conditional log-likelihood discrepancy	1.16×10^{-9}
Fitted log-probability refinement	6.70×10^{-13}
Fitted derivative refinement	5.04×10^{-13}
Independent prior log-probability discrepancy	2.05×10^{-12}
Independent finite-difference derivative discrepancy	6.54×10^{-11}
Independent information-criterion discrepancy	1.12×10^{-13}
Count-regeneration checks / changed estimates	12 / 0
Retained final capped primary fits	1

the failed numerical pilot remain documented; forecasting exceptions are documented in Appendices Q.8 and T.4.

N CANDIDATE RANKING AND FIXED-BUDGET GENERALIZATION

This section reports exploratory tests designed to distinguish the information-preserving timing principle from the performance of its Rank-bound instantiation. Code and raw evidence are in the indexed design-validation collection. Candidate indices, graph budgets, schedules, seeds and stress-test definitions were recorded before evaluating these outcomes. The study follows analysis of the primary results and is not externally preregistered. None of its outcomes changes the primary candidate pools, scores, schedules or recovery results.

Main-text ranking summary and its scope. A frozen information floor is not automatically an acquisition ranking. On the 20 held-out parameter seeds of the finite-clock audit, safeguarded polynomial Rank-bound has mean surrogate .379 and actual normalized minimum information .395, while KW-Ds has the smaller surrogate .252 but larger actual information .526. The sufficient perturbation margin is positive at 13/20 Rank-bound points and no KW-Ds points: failing that conservative test does not imply poor recovery. Full successful and unsuccessful margins remain in Appendix M.4.

To test ranking directly, we sample 128 feasible candidates per family uniformly from the common pool, add its five preselected objective winners, and freeze those indices before evaluating 20 disjoint parameter seeds with scales log-uniform on $[8, 64]$. Actual graph information and local inverse-Fisher variance use the generating parameters only for this held-out audit; they never alter any selected design. Table 32 reports actual minimum-information retention relative to the best candidate in the audited subset, and the corresponding local variance inflation. These are oracle comparisons within a finite subset, not regret against a continuous optimum or observed finite-sample risk ratios. The distinction is substantive: polynomial Rank-bound retains .701 of the best audited information, with 1.448 local variance inflation; KW-Ds retains .933 with 1.087 inflation. In Fourier, Rank-bound, KW-Ds and REIG all retain about .914, with 1.102 inflation. Mean within-seed Spearman correlation between the prior bound and actual information is positive over all candidates ($.588/.583$ for polynomial/Fourier), but drops to $-.686/-.170$ in the top score quartile. A coarse separation of poor designs therefore need not provide a useful ordering near the top. Warm pointwise frozen-pair evaluation is cheaper than actual graph-information evaluation in this implementation (Appendix N.1); this excludes uncertainty integration, common-pool construction and margin verification, and is not an end-to-end speedup or cheap certification result.

Table 32: Finite-clock ranking audit: 133 candidates per family, 20 disjoint held-out parameter seeds. Mean information retention relative to the best audited candidate (higher is better); mean local variance inflation relative to the smallest audited variance (lower is better). These oracle quantities are not empirical RMSE ratios. Appendix N.1 reports dispersion and scope.

	Information retention		Local variance inflation	
Design	Polynomial	Fourier	Polynomial	Fourier
Rank-bound	0.701	0.914	1.448	1.102
KW-Ds (HSB'14)	0.933	0.915	1.087	1.101
KW-D (HSB'14)	0.483	0.566	6.502	3.181
REIG (UAI'22)	0.924	0.914	1.092	1.102
EPIG (AISTATS'23)	0.013	3.72×10^{-6}	4.75×10^{7}	2.30×10^{8}

N.1 Candidate sample, oracle quantities and independent units

For each family, use the 108-scenario common pool and $\tau = .5$. Uniformly sample 128 feasible candidate indices without replacement using streams 92001 and 92002; add the indices of the five already selected objective winners. A duplicated selected index is counted only once. The resulting pool hashes, all indices, acquisition scores, conditioning values and times are saved before computing any held-out information. The two complete feasible pools contain 1,667 polynomial and 2,211 Fourier candidates. The sampled candidates need not be independent, distinct in qualitative timing pattern, or representative of all continuous schedules; local-search descendants remain part of the stated finite-pool population.

Seeds 15000–15019 determine node parameters by the primary generator and nominal scales by $\log\lambda \sim U[\log 8, \log 64]$ from stream $\texttt{seed} + 810000$. Seeds are paired across families. All candidate statistics for a given seed share those parameters, and candidates are not treated as independent inferential replicates. There are 20 independent parameter seeds, disjoint from both the finite-clock information audit and primary recovery cohorts. The matrices are normalized by $G = B^\top B$ exactly as in Appendix L.5.

For every candidate and seed, the artifact records $g(D) = \lambda_{\min}(G^{-1/2} I_{v,\mathrm{eff}} G^{-1/2})$, its frozen counterpart, the pointwise dyad-minimum surrogate, both score perturbations, nuisance singular values, the remainder condition and sufficient margin. It also computes

$$V(D) = \frac{1}{m}\,\mathrm{tr}\{(G^{-1/2} I_{v,\mathrm{eff}} G^{-1/2})^{-1}\},$$

which equals the field-variance scale in (18) before dividing by R. Per-seed information retention is $g(D)/\max_{D'\in\mathcal{A}} g(D')$, and local variance inflation is $V(D)/\min_{D'\in\mathcal{A}} V(D')$, where $\mathcal{A}$ is the sampled set plus winners. The two denominators may correspond to different designs. Both comparisons know the generating parameters and are *oracle diagnostics*; they are not new acquisition algorithms, finite-sample risk guarantees, or uniform parameter-set certificates.

Spearman correlations compare the prior Rank-bound score with actual information over $\mathcal{A}$, over its top score quartile, and with $-V(D)$. A separate correlation compares the pointwise surrogate with actual information. All correlations are computed within each parameter seed before summarizing across seeds. Restricting to the top quartile asks about ordering near more competitive surrogate values, not merely separating very poor from reasonable designs. No candidate-level standard error is used. All descriptive intervals and dispersion measures here are separate from the primary 156-contrast inferential family.

Numerical validity and timing scope. Each point uses transition probabilities and derivatives at quadrature orders $(80, 32)$ and $(160, 64)$. Probability mass, positivity, probability/derivative refinement, the frozen lower bound and every applicable perturbation inequality are enforced. An absent positive margin remains an inconclusive sufficient test. Warm single-thread timings compare the frozen-pair minimum with actual normalized graph-information evaluation at the same parameters and selected schedules. Each timing is a median of 15 blocks of 20 calls, alternating order after warm-up. They exclude JIT compilation, design-pool construction, averaging over parameter uncertainty, independent ODE generation, refinement and margin verification. Shared-host timings do not

Table 33: Within-seed Spearman correlation, summarized by mean (sample SD) over 20 parameter seeds. The same 133 candidate indices are evaluated at every seed. A favorable global correlation does not imply accurate ranking near the top of the surrogate scale.

Comparison	Polynomial	Fourier
Prior score vs. actual information	0.588 (0.388)	0.583 (0.415)
Same, restricted to top score quartile	-0.686 (0.110)	-0.170 (0.650)
Prior score vs. negative local variance	0.582 (0.384)	0.554 (0.430)
Pointwise surrogate vs. actual information	0.732 (0.221)	0.977 (0.024)

Table 34: Mean (sample SD) of the per-seed oracle ratios underlying Table 32. EPIG's near-zero clock information produces extreme local variances; these are mathematical local scales, not predictions of bounded-estimator error.

Family	Design	Information retention	Variance inflation
Polynomial	Rank-bound	0.701 (0.096)	1.448 (0.234)
Polynomial	KW-Ds (HSB'14)	0.933 (0.127)	1.087 (0.175)
Polynomial	KW-D (HSB'14)	0.483 (0.376)	6.502 (11.167)
Polynomial	REIG (UAI'22)	0.924 (0.120)	1.092 (0.170)
Polynomial	EPIG (AISTATS'23)	0.013 (0.033)	4.75×10^{7} (1.80×10^{8})
Fourier	Rank-bound	0.914 (0.132)	1.102 (0.175)
Fourier	KW-Ds (HSB'14)	0.915 (0.132)	1.101 (0.176)
Fourier	KW-D (HSB'14)	0.566 (0.345)	3.181 (3.919)
Fourier	REIG (UAI'22)	0.914 (0.132)	1.102 (0.175)
Fourier	EPIG (AISTATS'23)	3.72×10^{-6} (1.51×10^{-5})	2.30×10^{8} (2.77×10^{8})

establish end-to-end search speed or a cheap certified screening algorithm; in particular, the primary common-pool search evaluates all objectives for every candidate.

N.2 Fixed-total-exposure sparse graphs and a nonoptimized control

The graphs use the exact edge order

$$E_6 = \{12, 13, 23, 34, 45, 56\}, \qquad E_{11} = E_6 \cup \{14, 15, 16, 24, 25\},$$

with K_6 as the complete graph. All have six nodes and full-column-rank unsigned incidence. Their distinct rank budgets motivate the graph selection, but the polynomial/Fourier recovery experiment does not instantiate the phase family of Theorem 4. In particular, it does not claim to validate quartic rates for polynomial affinities or shrinking probes at a fixed hardware floor.

At scale 32, each graph receives exactly 1,650 dyad-paths: 275 paths per dyad on G_6, 150 on G_{11} and 110 on K_6. Each path contributes six conditional transitions, hence 9,900 transitions for every objective and graph, plus its conditioned initial state. This controls total acquisition exposure, not just observations per path. We do not invoke the fixed-per-dyad information monotonicity result under this reallocation.

The five optimized schedules are read unchanged from the 108-scenario, $\tau = .5$ files. The *fixed anchor/probe* control uses

$$D_{\text{fixed}} = (0, .75, .75 + \delta, 2.75, 2.75 + \delta, 5.5 - \delta, 5.5), \qquad \delta = 1/64.$$

Its anchor singular values are .747747 (polynomial) and .568215 (Fourier). These times are specified directly without using an objective score or any recovery outcome; the control is not relabeled as an external method. It is separate from the primary common search pool, so no claim is made that the primary pool choices optimize over this additional control.

Seeds 16000–16019 are disjoint from the primary recovery and information cohorts. Node parameters use the four independent uniform streams from Appendix M.1; inference uses reference clock 32 and the complete fitting configuration in Appendix M.1, including all 24 parameters, common bounds, two starts and caps $300, 1200, 4800$. Conditional path counts use stream $\texttt{seed} + 320000 + R + 6000000 + 10^8\texttt{family} + 2 \cdot 10^8\texttt{misspec}$. Methods share streams within

Table 35: Median across the five selected schedules of each schedule's warm median kernel time (milliseconds); the last column is the median paired kernel-time ratio. Frozen-pair evaluation is cheaper here, but full uncertainty integration, pool construction and certification are excluded. These are not end-to-end acquisition speedups.

Family	Frozen pair (ms)	Actual information (ms)	Kernel ratio
Polynomial	0.143	19.393	136.0
Fourier	0.139	27.625	197.9

Table 36: Fixed total exposure: 9,900 transitions, scale 32, shared $\tau = .5$. Mean clock RMSE on 20 disjoint paired seeds; lower is better. Fixed A/P is an unoptimized internal control, not an external baseline. Sample SDs, all 720 fit outcomes and the 120-fit omitted-basis test are in Appendix N.

	Polynomial			Fourier		
Design	G_6	G_{11}	K_6	G_6	G_{11}	K_6
Rank-bound	0.091	0.093	0.089	0.082	0.098	0.083
KW-Ds (HSB'14)	0.074	0.077	0.083	0.080	0.083	0.086
KW-D (HSB'14)	0.228	0.275	0.174	0.156	0.161	0.185
REIG (UAI'22)	0.075	0.080	0.070	0.083	0.099	0.083
EPIG (AISTATS'23)	1.136	1.150	1.384	1.733	1.520	1.536
Fixed A/P	0.084	0.086	0.089	0.087	0.080	0.080

a condition; different graphs reuse parameter seeds but have different R and corresponding count streams. The 720 correctly specified fits cross these 20 seeds with two families, three graphs and six methods. Independent DOP853 generation is checked against endpoint quadrature; state continuity, all starts and every termination status are saved. No selected cap or boundary estimate is excluded.

Table 37: Correctly specified fixed-budget recovery. Entries are mean RMSE (sample SD), with 20 paired seeds in every row. The final column counts clock-boundary / nuisance-boundary / selected-capped fits. Every exception is retained.

Family	Graph	Design	Clock	Affinity	Boundaries / cap
Poly.	G_6	Rank-bound	0.091 (0.025)	0.129 (0.020)	0 / 0 / 0
Poly.	G_6	KW-Ds (HSB'14)	0.074 (0.029)	0.121 (0.016)	0 / 0 / 0
Poly.	G_6	KW-D (HSB'14)	0.228 (0.155)	0.084 (0.013)	2 / 0 / 0
Poly.	G_6	REIG (UAI'22)	0.075 (0.022)	0.179 (0.048)	0 / 4 / 0
Poly.	G_6	EPIG (AISTATS'23)	1.136 (0.460)	0.130 (0.027)	10 / 0 / 0
Poly.	G_6	Fixed A/P	0.084 (0.023)	0.096 (0.015)	0 / 0 / 0
Poly.	G_{11}	Rank-bound	0.093 (0.026)	0.141 (0.031)	0 / 0 / 0
Poly.	G_{11}	KW-Ds (HSB'14)	0.077 (0.022)	0.113 (0.026)	0 / 0 / 0
Poly.	G_{11}	KW-D (HSB'14)	0.275 (0.216)	0.074 (0.014)	1 / 0 / 0
Poly.	G_{11}	REIG (UAI'22)	0.080 (0.026)	0.158 (0.047)	0 / 0 / 0
Poly.	G_{11}	EPIG (AISTATS'23)	1.150 (0.355)	0.129 (0.036)	17 / 0 / 0
Poly.	G_{11}	Fixed A/P	0.086 (0.029)	0.101 (0.019)	0 / 0 / 0
Poly.	K_6	Rank-bound	0.089 (0.023)	0.133 (0.020)	0 / 0 / 0
Poly.	K_6	KW-Ds (HSB'14)	0.083 (0.020)	0.104 (0.025)	0 / 0 / 0
Poly.	K_6	KW-D (HSB'14)	0.174 (0.086)	0.079 (0.014)	0 / 0 / 0
Poly.	K_6	REIG (UAI'22)	0.070 (0.023)	0.178 (0.045)	0 / 0 / 0
Poly.	K_6	EPIG (AISTATS'23)	1.384 (0.280)	0.128 (0.028)	19 / 0 / 0
Poly.	K_6	Fixed A/P	0.089 (0.027)	0.101 (0.021)	0 / 0 / 0
Fourier	G_6	Rank-bound	0.082 (0.033)	0.097 (0.014)	0 / 0 / 0
Fourier	G_6	KW-Ds (HSB'14)	0.080 (0.031)	0.104 (0.016)	0 / 0 / 0
Fourier	G_6	KW-D (HSB'14)	0.156 (0.066)	0.087 (0.014)	0 / 0 / 0
Fourier	G_6	REIG (UAI'22)	0.083 (0.030)	0.096 (0.015)	0 / 0 / 0
Fourier	G_6	EPIG (AISTATS'23)	1.733 (0.429)	0.136 (0.021)	16 / 0 / 0
Fourier	G_6	Fixed A/P	0.087 (0.025)	0.133 (0.030)	0 / 3 / 0
Fourier	G_{11}	Rank-bound	0.098 (0.023)	0.094 (0.020)	0 / 0 / 0
Fourier	G_{11}	KW-Ds (HSB'14)	0.083 (0.021)	0.099 (0.019)	0 / 0 / 0
Fourier	G_{11}	KW-D (HSB'14)	0.161 (0.066)	0.093 (0.014)	0 / 0 / 0

Family	Graph	Design	Clock	Affinity	Boundaries / cap
Fourier	G_{11}	REIG (UAI'22)	0.099 (0.023)	0.095 (0.020)	0 / 0 / 0
Fourier	G_{11}	EPIG (AISTATS'23)	1.520 (0.185)	0.130 (0.025)	18 / 0 / 0
Fourier	G_{11}	Fixed A/P	0.080 (0.025)	0.153 (0.041)	0 / 0 / 0
Fourier	K_6	Rank-bound	0.083 (0.030)	0.098 (0.018)	0 / 0 / 0
Fourier	K_6	KW-Ds (HSB'14)	0.086 (0.020)	0.098 (0.014)	0 / 0 / 0
Fourier	K_6	KW-D (HSB'14)	0.185 (0.062)	0.089 (0.011)	0 / 0 / 0
Fourier	K_6	REIG (UAI'22)	0.083 (0.029)	0.098 (0.018)	0 / 0 / 0
Fourier	K_6	EPIG (AISTATS'23)	1.536 (0.165)	0.131 (0.032)	19 / 0 / 1
Fourier	K_6	Fixed A/P	0.080 (0.021)	0.146 (0.033)	0 / 0 / 0

Statistical interpretation. Tables report mean and sample SD of the per-realization clock-field and time-grid affinity RMSEs. Complete seed-paired method differences, conditional on graph/family, are saved with descriptive unadjusted 95% Student-t intervals. They are exploratory and are not folded into the primary simultaneous-contrast family. The same 20 seeds are reused across all graph and misspecification conditions; 840 fit labels are not 840 independent realizations. We report every method rather than selecting a favorable subset or interpreting non-significant differences as equivalence.

The local variance prediction is computed at the generating parameters only for correctly specified models. Its comparison with mean squared clock error is descriptive: regular asymptotic covariance is not a finite-sample theorem for bounded nonconvex estimation, and boundary cases can invalidate that interpretation.

Table 38: Mean squared clock-field error divided by the mean local inverse-Fisher variance prediction, using the generating parameters and actual exposure. Ratios near one are a descriptive check, not a finite-sample guarantee. For ill-conditioned or boundary-limited fits, notably EPIG clock recovery, an enormous local variance can make this ratio tiny without implying efficient or accurate estimation.

	Polynomial			Fourier		
Design	G_6	G_{11}	K_6	G_6	G_{11}	K_6
Rank-bound	1.119	1.173	1.058	0.969	1.278	0.969
KW-Ds (HSB'14)	1.043	1.064	1.211	0.918	0.907	0.968
KW-D (HSB'14)	2.463	4.214	1.310	1.200	1.312	1.643
REIG (UAI'22)	1.013	1.180	0.910	0.977	1.283	0.970
EPIG (AISTATS'23)	5.63×10^{-5}	0.001	0.001	1.29×10^{-10}	4.77×10^{-7}	7.68×10^{-7}
Fixed A/P	0.943	1.033	1.088	1.033	0.874	0.864

N.3 Omitted-cubic stress test

On polynomial K_6 at the same total budget, generate an additional node-shared coefficient for the cubic Legendre basis $(5s^3-3s)/2$, independently $U[-.25,.25]$ using stream $\texttt{seed}+290000$. Keep every acquisition schedule, the three-basis working prior, and the fitted three-basis likelihood unchanged. The resulting 120 fits use the same 20 seeds and all six methods. This deliberately violates correct specification; the temporal singular-value check concerns the *fitted* basis and cannot certify absence of bias from omitted variation. The full true four-basis trajectory, including its omitted component, is used to score affinity error. No inverse-Fisher risk claim is made for these misspecified fits.

N.4 Execution audit and cache-safe reproducibility

The design-validation driver rejects nonempty output directories and cached fits. Its audit enforces complete coverage and repeats the consistency, regeneration and estimation checks of Appendix M.5, adding fixed-budget exposure and predictive-error checks. Regression tests distinguish optimization from replay and check empty-output acceptance, cache rejection in both optimization modes, explicit replay and uniqueness of the 12 reproduction cases.

Table 39: Omitted-cubic stress on polynomial K_6, 9,900 transitions per fit. Mean RMSE (sample SD) on the same 20 seeds; clean and misspecified count streams differ. The true four-basis trajectory is used for affinity error. All 120 misspecified fits reach the selected-start stopping criterion; 16 EPIG fits hit a clock boundary, with no other boundary flags in this cohort.

	Clock		Affinity	
Design	Correct	Omitted cubic	Correct	Omitted cubic
Rank-bound	0.089 (0.023)	0.087 (0.030)	0.133 (0.020)	0.175 (0.038)
KW-Ds (HSB'14)	0.083 (0.020)	0.077 (0.027)	0.104 (0.025)	0.202 (0.045)
KW-D (HSB'14)	0.174 (0.086)	0.171 (0.045)	0.079 (0.014)	0.151 (0.023)
REIG (UAI'22)	0.070 (0.023)	0.077 (0.027)	0.178 (0.045)	0.311 (0.072)
EPIG (AISTATS'23)	1.384 (0.280)	1.166 (0.252)	0.128 (0.028)	0.226 (0.071)
Fixed A/P	0.089 (0.027)	0.079 (0.034)	0.101 (0.021)	0.145 (0.023)

The 12 stratified checks rerun existing primary seeds, including the capped Fourier/EPIG case; they are not independent replications. The indexed regeneration-check entry point provides `regenerate` (default: regenerate counts and optimize), `refit` (optimize copied fixed counts), and `replay` (read saved fits without optimization). Both optimization modes reject nonempty outputs; neither initializes from saved estimates nor selects by recovery error. Regeneration reproduces parameters and reported values exactly in the recorded environment, without establishing global optimality.

Table 40: Design-validation checks; audit modes are defined in Appendix N.4. Boundary and convergence indicators are separate. The retained cap is Fourier/K_6/EPIG, seed 16007.

Check	Result
Optimization fits / independent recovery seeds	840 / 20
Correctly specified / omitted-cubic fits	720 / 120
Transitions per fit / planned coverage	9,900 / 840 of 840
Selected capped fits (retained)	1
Clock / nuisance boundary flags	118 / 7
Maximum metric / log-likelihood replay error	0.000 / 1.85×10^{-9}
Maximum generator log-probability refinement	3.43×10^{-13}
Candidate information evaluations / parameter seeds	5,320 / 20
Maximum ranking probability / derivative refinement	2.57×10^{-13} / 1.93×10^{-13}
Count-regeneration checks / cached fit reads	12 / 0
Regeneration parameter / metric discrepancy	0 / 0
Cache-mode regression tests	5 passed

O RECENT POLICY-BASED ACQUISITION METHODS: MATCHED GRID ADAPTERS

This study compares observation-design criteria with Step-DAD (Hedman et al., 2025) and Inside-Out SMC^2 (Iqbal et al., 2024); KW-Ds is an internal reference. Relative to the candidate-pool study in Appendix M, it uses a common restricted timing grid, a different initial-state convention, disjoint recovery seeds and adaptive decisions. The implementations are source-informed task adapters rather than executions of unmodified upstream software or reproductions of the published benchmarks. Joint and clock-targeted variants are two objectives for each of two source algorithms. Appendix P reports the post-hoc PASOA comparison on the same seeds.

O.1 COMMON ACTIONS, DATA ACCESS, AND PHYSICAL BUDGET

Every method observes $R = 100$ paths on each of the 15 dyads of K_6, at six post-initial times on $[0, 5.5]$, with minimum gap $\delta = 1/64$. Each dyad starts with exactly 50 paths in each state, independently of the parameters; this initial condition is conditioned out. There are 9,000 transitions per graph realization for every method. A policy makes three decisions per dyad. Each decision commits to both an anchor and its follow-up probe *before either outcome is observed*. First anchors are $.75, 1, 1.25$, second anchors are $2.25, 2.75, 3.25$, and probe widths are $\delta\{1, 2, 4, 8, 16\}$. The final pair is $(5.5 - d, 5.5)$. Feasible complete schedules satisfy $\sigma_{\min}(\Phi(t_1, t_3, t_5)) \geq .5$; masks exclude prefixes with no feasible completion. The grid has 875 polynomial and 1,050 Fourier schedules.

The grid supplies anchor/probe structure to *every* comparator. This experiment tests optimization and feedback within that class, not whether an unconstrained learned policy discovers the structure. Rank-bound and KW-Ds are reoptimized exactly on this grid; their schedules and recovery values are computed for this grid. Fixed A/P uses $(.75, .75 + \delta, 2.75, 2.75 + \delta, 5.5 - \delta, 5.5)$. The Monte Carlo EIG controls screen every candidate using 256 prior rollouts, then reevaluate the union of the top 32 candidates for each target using 8,192 rollouts. They are not exact EIG maximizers.

Policies receive only that dyad's past observation times and transition counts, in chronological order; future slots are zero. All adaptive methods have the same observation access. They do not pool graph-wide posterior information or receive generating parameters. Fixed schedules cannot use feedback. The estimator, in contrast, jointly fits shared node parameters across all dyads. Thus adaptive and fixed comparisons have equal physical budgets, not identical feedback or computational resources.

O.2 FINITE-PRIOR INFERENCE AND ALGORITHMIC ADAPTATIONS

Design uses a 108-point marginal cubature: four equally weighted scales $8, 16, 32, 64$ crossed with 27 affinity-coefficient scenarios. Each coefficient has support $\{-\sqrt{.036}, 0, \sqrt{.036}\}$ and weights $(5/24, 7/12, 5/24)$. The scalar-panel posterior is evaluated by enumeration. This removes inner Monte Carlo error but does not approximate an arbitrarily fine continuous posterior: full affine rank is not continuous support, and posterior concentration or calibration for off-support truths is not guaranteed.

For history h and a committed pair, two sequential transition-count likelihoods update all 108 weights. The joint target concerns clock and affinity together. The clock target first marginalizes the 27 nuisance scenarios within each clock value, using their conditional weights given h. Neither target assumes that affinities are known. The four-valued clock target is explicitly a task-specific finite-prior objective; recovery is subsequently scored against *continuous* node-shared generating parameters.

Step-DAD–grid (ICML'25). The adapter retains offline sequential prior-contrastive estimation (sPCE), followed by posterior-conditioned online policy refinement. A masked categorical policy uses a two-hidden-layer width-64 tanh network on the padded chronological history; discrete-action score-function gradients avoid differentiating through sampled binary outcomes. Offline batches contain 128 trajectories and 1,023 contrastive draws. Negatives are shared within a batch but independent of its generating parameters and simulated outcomes. After the first pair, exact finite-posterior inference replaces approximate posterior sampling, and independent full-policy copies are

fine-tuned for the remaining two pairs, separately for each dyad. The online batch is 32 and learning rate .001. Prior-only development selects 64 or 256 refinement updates. The matched DAD control uses the identical initial checkpoint without this refinement. Both joint (J) and nuisance-marginalized clock (C) targets are retained.

Inside-Out SMC^2–grid (ICML'24). The adapter retains the risk-sensitive Feynman–Kac objective, conditional outer SMC with a retained reference trajectory, and Markovian score-climbing policy updates. Four parallel chains use 64 outer particles each. Exact enumeration replaces the inner parameter filter. At each pair the potential is $\exp(r_k)$, where r_k is the KL divergence from the updated to the preceding joint or clock-marginal posterior. Both likelihood and predictive-normalization terms are included; the additive-constant-noise simplification is inapplicable to these binomial observations. The policy gradient averages terminal particle scores with their normalized weights. The risk parameter is one. This is a finite-support, exact-inner specialization, not a performance test of approximate inner filtering on the source dynamical benchmarks.

Both methods use the same history-network architecture and a fixed 2% uniform mixture over feasible actions in training and deployment. Each family/target/method crosses two learning rates $(.001, .0003)$ with two training seeds, for 32 runs of 8,000 updates. Checkpoints at $500, 1000, 2000, 4000, 8000$ updates are compared on 4,096 prior rollouts, then the best checkpoint across the four runs is frozen. This selection uses target-specific EIG estimates, not recovery-test RMSE. A separate 16,384-rollout prior audit is retained. Online refinement is selected on 128 independent observed prefixes with 128 continuation rollouts per prefix. The no-refinement result is retained but is not an eligible replacement for a Step-DAD refinement. These finite training budgets and validation choices are not convergence certificates.

Table 41: Prior-only policy selection and independent 16,384-rollout audit (nats, Monte Carlo SE). Each row selects among four training runs. Fine-tuning updates are selected on separate prior prefixes. The audit for Step-DAD here concerns its initial policy, before online refinement. Cost is the sum of elapsed seconds for its four candidate runs on the shared host.

Family	Policy/target	Rate	Step	Audit EIG	MC SE	Refine	Cost (s)
Poly.	Step-DAD/J	0.0003	8000	2.306	0.0021	256	333.5
Poly.	Step-DAD/C	0.0003	8000	1.322	0.0011	256	338.2
Poly.	IO-SMC^2/J	0.0010	2000	2.277	0.0022	0	690.1
Poly.	IO-SMC^2/C	0.0003	2000	1.319	0.0011	0	851.9
Fourier	Step-DAD/J	0.0010	2000	2.343	0.0023	256	390.8
Fourier	Step-DAD/C	0.0010	1000	1.320	0.0011	64	341.2
Fourier	IO-SMC^2/J	0.0010	500	2.338	0.0023	0	743.7
Fourier	IO-SMC^2/C	0.0010	2000	1.318	0.0012	0	765.8

O.3 Continuous-truth confirmation and common inference

Seeds 28000–28029 generate continuous node parameters with the same uniform-coefficient function as the primary study, at scales $8, 32, 64$ and in both basis families. No confirmation outcomes enter offline training, posterior-refinement development, schedule selection or checkpoint selection. Protocol, code and selected inputs are hashed before generating this cohort; the study is not externally preregistered. Each family/scale cell has 30 graph realizations. The same parameter seeds are reused across conditions and methods; the resulting 1,980 method/condition fits are not 1,980 independent realizations.

An independent dense DOP853 integration generates transition probabilities at each selected interval. Common uniform draws, transformed by inverse binomial CDFs, pair methods' transition-count noise. Counts preserve state continuity, including observed absences. Policies see only the realized history; generating probabilities enter the simulator and validation, never policy inputs. Final counts, selected times, uniforms, posterior diagnostics, truth, estimates, optimization attempts, and refinement traces are saved. Initial-state likelihoods are omitted consistently.

The common 24-parameter graph MLE uses the candidate-pool estimator's parameter bounds, two truth-independent starts, objective-based start selection, status-triggered caps $300, 1200, 4800$, ana-

lytical derivatives and quadrature-doubling tests. Only the timetable indexing is extended to allow different times on different dyads. For a fixed policy and its internal randomization, the sequential action probabilities depend on the observed history but not on the unknown generating parameters; they therefore do not add a parameter-dependent factor to the likelihood used for fitting. This is a comparison of acquisition quality under a common estimator, not of the methods' native end-to-end inference pipelines.

Tables report means and sample SDs across the 30 paired graph seeds within each family/scale cell. Saved pairwise differences use descriptive, unadjusted 95% Student-t intervals. They are separate from the primary 156-contrast family and conditional on the selected policies; no training-seed-averaged or simultaneous superiority claim is made. All method rows, boundary estimates and selected capped fits are retained.

Table 42: Polynomial matched-grid confirmation. Mean (sample SD) RMSE over 30 paired graph seeds per cell; all outcomes retained. J/C specify the design target.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.100 (.028)	.144 (.028)	.095 (.028)	.131 (.024)	.144 (.041)	.114 (.020)
KW-Ds (HSB'14)	.100 (.028)	.138 (.025)	.094 (.028)	.124 (.020)	.145 (.042)	.109 (.019)
Fixed A/P	.101 (.027)	.117 (.020)	.095 (.027)	.104 (.016)	.144 (.042)	.091 (.015)
Fixed MC-EIG/J	.090 (.027)	.113 (.019)	.114 (.031)	.100 (.015)	.182 (.067)	.089 (.015)
Fixed MC-EIG/C	.100 (.028)	.144 (.028)	.095 (.028)	.131 (.024)	.144 (.041)	.114 (.020)
DAD–grid/J (control)	.093 (.025)	.114 (.020)	.104 (.028)	.103 (.016)	.150 (.048)	.091 (.015)
DAD–grid/C (control)	.100 (.030)	.127 (.023)	.097 (.029)	.115 (.018)	.143 (.044)	.101 (.017)
Step-DAD–grid/J (ICML'25)	.099 (.028)	.110 (.019)	.100 (.026)	.104 (.016)	.160 (.052)	.092 (.015)
Step-DAD–grid/C (ICML'25)	.100 (.030)	.127 (.023)	.096 (.029)	.116 (.019)	.143 (.044)	.102 (.017)
IO-SMC^2–grid/J (ICML'24)	.092 (.026)	.111 (.020)	.119 (.034)	.099 (.017)	.185 (.071)	.089 (.015)
IO-SMC^2–grid/C (ICML'24)	.101 (.028)	.126 (.024)	.095 (.028)	.113 (.020)	.148 (.042)	.098 (.017)

Table 43: Fourier matched-grid confirmation. Mean (sample SD) RMSE over 30 paired graph seeds per cell; all outcomes retained. J/C specify the design target.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.119 (.027)	.134 (.026)	.086 (.031)	.111 (.022)	.141 (.062)	.104 (.014)
KW-Ds (HSB'14)	.119 (.026)	.172 (.043)	.086 (.031)	.142 (.035)	.142 (.060)	.135 (.024)
Fixed A/P	.118 (.026)	.178 (.043)	.086 (.032)	.146 (.035)	.143 (.062)	.141 (.025)
Fixed MC-EIG/J	.108 (.025)	.132 (.025)	.094 (.035)	.108 (.022)	.171 (.084)	.102 (.014)
Fixed MC-EIG/C	.118 (.026)	.178 (.043)	.086 (.032)	.146 (.035)	.143 (.062)	.141 (.025)
DAD–grid/J (control)	.111 (.025)	.129 (.024)	.108 (.038)	.104 (.021)	.178 (.064)	.100 (.014)
DAD–grid/C (control)	.118 (.027)	.157 (.036)	.088 (.032)	.128 (.028)	.142 (.055)	.122 (.019)
Step-DAD–grid/J (ICML'25)	.113 (.026)	.128 (.024)	.103 (.036)	.105 (.021)	.181 (.076)	.101 (.014)
Step-DAD–grid/C (ICML'25)	.118 (.027)	.157 (.034)	.088 (.033)	.128 (.026)	.142 (.056)	.121 (.021)
IO-SMC^2–grid/J (ICML'24)	.110 (.023)	.126 (.025)	.107 (.037)	.106 (.021)	.181 (.115)	.102 (.014)
IO-SMC^2–grid/C (ICML'24)	.119 (.026)	.135 (.027)	.086 (.032)	.111 (.022)	.143 (.063)	.104 (.014)

Recovery and target tradeoffs. The clock-targeted variants remain close to Rank-bound: the largest absolute mean clock-RMSE difference across the Step-DAD and Inside-Out adapters and all six conditions is .00401. Eleven of the twelve corresponding descriptive paired intervals contain zero; the exception favors Rank-bound over clock-targeted Inside-Out on polynomial scale 64 (difference $-.00401$, interval $[-.00712, -.00089]$). These unadjusted intervals do not establish simultaneous superiority or statistical equivalence. Fixed A/P also remains competitive for clock recovery, with means $.101, .095, .144$ (polynomial) and $.118, .086, .143$ (Fourier).

The joint-information variants exhibit a clock–affinity trade-off that depends on the family and scale. At scale 8, joint Inside-Out has lower mean clock RMSE than Rank-bound in both families ($.092/.110$ versus $.100/.119$). At scale 64, its means are $.185/.181$ versus $.144/.141$. In polynomial

models, both joint modern adapters reduce mean affinity error relative to Rank-bound at every tested scale: joint Inside-Out gives $.111, .099, .089$ versus $.144, .131, .114$. Clock-targeted adapters also improve polynomial affinity recovery, whereas the analogous Fourier ordering differs. All affinity results and controls remain in the full tables; a favorable clock row is not treated as joint-parameter dominance.

Refinement and timing diagnostics. Step-DAD does not uniformly improve on its matched DAD initialization. For the joint target, refinement changes polynomial clock RMSE from $.093$ to $.099$ at scale 8, $.104$ to $.100$ at scale 32, and $.150$ to $.160$ at scale 64. For the clock target, absolute mean changes are below $.0002$ in all six conditions. These outcomes apply to the selected budgets, finite prior and restricted grid. Performance in the source algorithm's published settings requires separate evaluation. Complete prior-prefix development curves, including refinements without improvement, are reported.

Recorded probe widths help explain the target distinction without changing any selected design. At polynomial scale 64, joint Inside-Out uses a mean first-probe gap $.0641$, whereas its clock-targeted variant uses $.0176$; joint and clock-targeted Step-DAD use mean final-probe gaps $.0331$ and $.0166$. These are descriptive acquired-action summaries, not oracle choices. The complete per-graph timing and posterior-effective-support diagnostics are supplied as CSV files. Finite-support posterior concentration is not a calibration claim for the continuous generating parameters.

Coverage and retained exceptions. All 1,980 fits completed, with 1,680 distinct acquired-data keys; identical selected schedules and paired uniforms can produce duplicate datasets across method labels. No selected fit remains capped and no clock estimate hits its bound. Two affinity-boundary fit labels correspond to one shared dataset: Fourier, scale 8, seed 28028, under Fixed A/P and Fixed MC-EIG/C. Both are retained. The 360 DAD/Step-DAD pre-refinement history comparisons agree exactly. No confirmation result was used to retune, replace or select a policy.

O.4 Numerical checks, computational cost, and reproducibility

Before confirmation, common-schedule probabilities and derivatives from the extended kernel match the original kernel exactly. Independent interval ODEs, dense ODE composition and quadrature agree within the recorded tolerances; finite differences check the extended analytical derivatives. Enumerated NumPy posterior updates check the float32 policy environment. A separately enumerated toy Feynman–Kac target checks the stationary first-action probability of the conditional SMC implementation using 1,500 retained sweeps. An independent discrete sPCE score-gradient test compares 131,072 simulated trajectories with exact enumeration. These checks test specific implemented components, not complete equivalence to upstream software or global optimization.

Table 44: Reported checks for the matched-grid adapters and all 1,980 confirmation fits. Numerical discrepancies and Monte Carlo checks have different meanings; neither certifies global policy or likelihood optimization.

Check	Recorded result
Complete confirmation coverage	1980/1980
Distinct acquired-data keys	1680
Physical transitions in every fit	9000
Matched DAD/Step-DAD initial histories	360
Selected fits still capped	0
Clock-boundary / affinity-boundary fits	0 / 2
Maximum true-probability log discrepancy	$2.25e-11$
Maximum fit log-probability refinement error	$2.70e-13$
Maximum fit derivative refinement error	$1.83e-13$
Finite-prior log-weight cross-check	$2.50e-05$
CSMC toy absolute error / Monte Carlo SE	0.000503 / 0.002018
sPCE gradient absolute error / Monte Carlo SE	0.000684 / 0.000754

Offline training, prior-design selection, online refinement, acquisition and downstream fitting times are recorded separately. Jobs run with one numerical thread per process on a shared CPU host, with

concurrent jobs and a four-CPU quota; there is no GPU timing or isolated-hardware speedup claim. Exact-inner finite enumeration and cached grid probabilities materially simplify this benchmark. Total offline costs include unsuccessful-to-improve checkpoints and all candidate training configurations, rather than only the selected model. Online refinement uses additional computation but no additional physical measurements.

Table 45: Per-graph elapsed seconds, mean (sample SD) over all 180 confirmation conditions. Acquisition includes policy decisions and simulation; refinement is its separately timed component. Costs are descriptive shared-host measurements, not isolated-hardware speed ratios. Offline training/selection is additional.

Design	Acquisition	Refinement	Common fit
Rank-bound	0.006 (0.001)	0.000 (0.000)	0.073 (0.024)
KW-Ds (HSB'14)	0.006 (0.001)	0.000 (0.000)	0.072 (0.023)
Fixed A/P	0.006 (0.001)	0.000 (0.000)	0.072 (0.023)
Fixed MC-EIG/J	0.006 (0.001)	0.000 (0.000)	0.075 (0.026)
Fixed MC-EIG/C	0.006 (0.002)	0.000 (0.000)	0.072 (0.022)
DAD–grid/J (control)	0.008 (0.001)	0.000 (0.000)	0.075 (0.023)
DAD–grid/C (control)	0.008 (0.002)	0.000 (0.000)	0.073 (0.021)
Step-DAD–grid/J (ICML'25)	5.467 (0.334)	5.447 (0.329)	0.073 (0.021)
Step-DAD–grid/C (ICML'25)	3.355 (2.024)	3.346 (2.024)	0.073 (0.020)
IO-SMC^2–grid/J (ICML'24)	0.008 (0.002)	0.000 (0.000)	0.076 (0.023)
IO-SMC^2–grid/C (ICML'24)	0.008 (0.001)	0.000 (0.000)	0.074 (0.022)

The `modern_baseline_study` implementation, training/checkpoint records, freeze manifests, raw datasets and per-graph estimates are not included in the artifact. The protocol, tables and audit summaries specify the reported study, but its training, recovery and record-level checks cannot be rerun from the packaged files. `STUDY_PATHS.json` records this coverage separately from the included supporting studies.

P PASOA AND MATCHED-GRID ACQUISITION COMPARISONS

This appendix evaluates PASOA (Iollo et al., 2024) alongside the two source algorithms in Appendix O. The combined acquisition comparison comprises three source algorithms, with J/C denoting joint and clock-targeted variants. The complete KW results are reported as internal references in Appendix O, separately from this comparison. The five temporal-graph methods in Section 5.4 address prediction and are evaluated separately from acquisition.

P.1 SOURCE MECHANISM AND TASK ADAPTATION

PASOA combines sequential contrastive design with tempered sequential Monte Carlo (SMC). Our independent implementation, PASOA–grid, retains posterior-informed contrastive acquisition, adaptive likelihood tempering, systematic resampling and invariant Markov moves. It uses the unchanged common action grid and finite design prior from Appendix O. The upstream source files inspected were `exptax/inference/SMC.py` and `exptax/estimators.py` in the authors' `jcopo/pasoa` repository. The inspected tree and file hashes are recorded in the experiment package. This is an independent task adaptation rather than an execution of unmodified upstream code.

There are three substantive adaptations. First, instead of a continuous design optimizer, every feasible action is evaluated using a sampled contrastive score and the largest score is selected. This exactly optimizes that finite *sample* objective, not the true expected information gain. Second, a prior-independence Metropolis kernel on 108 support points replaces a continuous random-walk move. Third, the C variant marginalizes affinity nuisance under the empirical particle measure. The continuous-posterior and continuous-action guarantees of the source paper are not transferred to these changes.

For the empirical posterior q after the observed history, candidate pair a, and L negative samples, the implemented joint score is

$$\widehat{U}_J(a) = \frac{1}{M}\sum_{b=1}^{M}\left[\log p(Y_b \mid \xi_{b0}, a) - \log\left\{\frac{1}{L+1}\sum_{\ell=0}^{L} p(Y_b \mid \xi_{b\ell}, a)\right\}\right]. \tag{48}$$

The draws satisfy $\xi_{b\ell} \sim q$ and $Y_b \sim p(\cdot \mid \xi_{b0}, a)$. Here $\xi = (v, \theta)$ is a dyad's clock/affinity parameter and Y_b contains two sequential transition-count matrices for 100 persistent paths. The starting counts are the observed counts at the preceding endpoint; the same generating parameter is used throughout the pair. Parameter-independent binomial coefficients cancel from the likelihood ratio. Distinct Monte Carlo draws are used for different candidate actions; sample score standard errors are saved, but do not correct for maximizing noisy scores.

For C, replace each likelihood by $p_q(Y \mid v, a) = \sum_\theta q(\theta \mid v)p(Y \mid v, \theta, a)$. This is a task-specific clock-information objective under q, not joint information with the generating affinity supplied as known. Empty empirical clock groups are never sampled. The posterior is maintained separately for each dyad; no outcome from another dyad or future pair is supplied to the planner.

P.2 PARTICLE UPDATE, BUDGETS, AND FREEZING

The shared 108-point prior comprises four clock support values and 27 affinity scenarios. At initialization, particles are sampled from that prior. After a selected pair has actually been observed, its likelihood is incorporated using increasing inverse temperature $\beta \in [0, 1]$. A bisection step chooses each increment to maintain effective sample size at least $0.9N$, followed by systematic resampling and Metropolis moves invariant to the *new* tempered posterior. With the original prior as independence proposal, the prior/proposal factors cancel from the acceptance ratio. The full observed log likelihood, including earlier pairs, remains in the target. No particle is moved using a generating parameter.

The primary configuration has $N = 2{,}048$ particles, $M = 256$ outer draws per action, $L = 127$ negative samples and 12 Metropolis proposals per tempering stage. The declared higher-budget configuration has $N = 8{,}192$, $M = 1{,}024$, $L = 255$ and 24 proposals. Both keep the same effective-sample-size threshold. There is no PASOA offline training or confirmation-error hyperparameter selection. The exact finite posterior is computed only after updates to audit particle total-variation error; it is not passed to acquisition.

This is a *post-hoc extension* informed by the results in Appendix O, rather than an externally pre-registered independent replication. The configuration, code and unchanged grid inputs were hashed before the new primary acquisitions. All 30 matched-grid graph seeds, 28000–28029, are reused at every family/scale cell; common physical uniform streams are verified against the archived Rank-bound records. Planner and particle random streams are separate from simulation randomness. The higher-budget check uses the first ten of those paired seeds, scale 32, both families and both targets. Results for both budgets are reported.

The six-observation, 9,000-transition physical budget, initial 50/50 state split, horizon, gap floor, affinity-conditioning requirement and downstream graph estimator are unchanged. The action grid is already structured as three anchor/probe pairs. Comparing within this restricted class does not test whether PASOA or another policy discovers that structure in an unrestricted action space. Adaptive methods commit to both times of a pair before either observation is collected.

P.3 Complete outcomes and sampling units

The extension completes 360 primary PASOA fits: two targets, two families, three scales and 30 seeds per cell. All 360 correspond to distinct acquired datasets. With the 1,800 non-KW outcomes reaggregated from the matched-grid study, the combined comparison contains 2,160 fits across 12 method/target/control rows and 1,860 distinct acquired datasets. Shared schedules and physical random streams explain repeated datasets; repeated fit labels are not independent replications. The 40 higher-budget fits and four development smoke fits are separate from that count. The 180 KW outcomes are reported separately in Appendix O.

All 360 primary and 40 higher-budget fits complete without a selected optimizer cap, a clock-boundary estimate or an affinity-boundary estimate. The matched-grid internal controls contain two affinity-boundary fit labels for one shared dataset. Every fit attempt and termination status is saved. The same bounded, two-start graph maximum-likelihood estimator is used for every acquisition method; these results isolate acquired-data effects rather than reproduce each source method's full native inferential pipeline.

Table 46: Polynomial: complete matched-grid comparison. Mean (sample SD) RMSE over 30 paired graph seeds per cell. PASOA is a post-hoc extension; the other methods reuse results from Appendix O, reaggregated without rerunning fits. KW results are reported separately there as internal references.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.100 (.028)	.144 (.028)	.095 (.028)	.131 (.024)	.144 (.041)	.114 (.020)
Fixed A/P	.101 (.027)	.117 (.020)	.095 (.027)	.104 (.016)	.144 (.042)	.091 (.015)
Fixed MC-EIG/J	.090 (.027)	.113 (.019)	.114 (.031)	.100 (.015)	.182 (.067)	.089 (.015)
Fixed MC-EIG/C	.100 (.028)	.144 (.028)	.095 (.028)	.131 (.024)	.144 (.041)	.114 (.020)
DAD–grid/J (control)	.093 (.025)	.114 (.020)	.104 (.028)	.103 (.016)	.150 (.048)	.091 (.015)
DAD–grid/C (control)	.100 (.030)	.127 (.023)	.097 (.029)	.115 (.018)	.143 (.044)	.101 (.017)
Step-DAD–grid/J (ICML'25)	.099 (.028)	.110 (.019)	.100 (.026)	.104 (.016)	.160 (.052)	.092 (.015)
Step-DAD–grid/C (ICML'25)	.100 (.030)	.127 (.023)	.096 (.029)	.116 (.019)	.143 (.044)	.102 (.017)
IO-SMC^2–grid/J (ICML'24)	.092 (.026)	.111 (.020)	.119 (.034)	.099 (.017)	.185 (.071)	.089 (.015)
IO-SMC^2–grid/C (ICML'24)	.101 (.028)	.126 (.024)	.095 (.028)	.113 (.020)	.148 (.042)	.098 (.017)
PASOA–grid/J (ICML'24)	.101 (.031)	.112 (.018)	.105 (.030)	.107 (.017)	.208 (.137)	.094 (.013)
PASOA–grid/C (ICML'24)	.091 (.027)	.120 (.022)	.097 (.029)	.113 (.017)	.150 (.049)	.100 (.017)

Tables 46–47 report all comparison rows, sample SDs and both targets. PASOA/C has smaller mean clock error than Rank-bound at scale 8 in both families, but larger mean errors at scales 32 and 64. PASOA/J illustrates why joint information is not the same objective as clock recovery: at polynomial scale 64 its affinity error is .094 versus .114 for Rank-bound, but its clock error is .208 versus .144. Clock-targeted PASOA improves polynomial affinity recovery at all three scales, while the analogous Fourier comparison favors Rank-bound. The clock–affinity trade-off depends on the family and scale.

All pairwise graph-seed differences for clock and affinity RMSE are supplied as mean differences and descriptive 95% Student-t intervals. The graph seed, not a dyad, transition, method row, or

Table 47: Fourier: complete matched-grid comparison. Mean (sample SD) RMSE over 30 paired graph seeds per cell. PASOA is a post-hoc extension; the other methods reuse results from Appendix O, reaggregated without rerunning fits. KW results are reported separately there as internal references.

	$\lambda = 8$		$\lambda = 32$		$\lambda = 64$	
Design	Clock	Affinity	Clock	Affinity	Clock	Affinity
Rank-bound	.119 (.027)	.134 (.026)	.086 (.031)	.111 (.022)	.141 (.062)	.104 (.014)
Fixed A/P	.118 (.026)	.178 (.043)	.086 (.032)	.146 (.035)	.143 (.062)	.141 (.025)
Fixed MC-EIG/J	.108 (.025)	.132 (.025)	.094 (.035)	.108 (.022)	.171 (.084)	.102 (.014)
Fixed MC-EIG/C	.118 (.026)	.178 (.043)	.086 (.032)	.146 (.035)	.143 (.062)	.141 (.025)
DAD–grid/J (control)	.111 (.025)	.129 (.024)	.108 (.038)	.104 (.021)	.178 (.064)	.100 (.014)
DAD–grid/C (control)	.118 (.027)	.157 (.036)	.088 (.032)	.128 (.028)	.142 (.055)	.122 (.019)
Step-DAD–grid/J (ICML'25)	.113 (.026)	.128 (.024)	.103 (.036)	.105 (.021)	.181 (.076)	.101 (.014)
Step-DAD–grid/C (ICML'25)	.118 (.027)	.157 (.034)	.088 (.033)	.128 (.026)	.142 (.056)	.121 (.021)
IO-SMC^2–grid/J (ICML'24)	.110 (.023)	.126 (.025)	.107 (.037)	.106 (.021)	.181 (.115)	.102 (.014)
IO-SMC^2–grid/C (ICML'24)	.119 (.026)	.135 (.027)	.086 (.032)	.111 (.022)	.143 (.063)	.104 (.014)
PASOA–grid/J (ICML'24)	.128 (.025)	.126 (.024)	.100 (.028)	.120 (.024)	.204 (.090)	.107 (.013)
PASOA–grid/C (ICML'24)	.109 (.026)	.148 (.033)	.092 (.032)	.131 (.028)	.150 (.059)	.118 (.023)

reused budget label, is the sampling unit. Family/scale conditions share underlying random streams, so the six cells are not pooled as independent repetitions. Intervals are conditional on the frozen matched-grid policies and are not adjusted for the full set of pairwise contrasts. Neither nonsignificance nor close rounded means establishes equivalence.

P.4 Computational-budget and numerical audits

Table 48: PASOA budget check at $\lambda = 32$, ten paired graph seeds per family and target. Each entry is a mean RMSE; the difference interval is higher minus primary budget. Intervals are descriptive, unadjusted 95% paired t intervals. No row is selected as a replacement for the primary results.

Family/target	Metric	Primary	Higher	Difference [95% interval]
Polynomial/J	Clock	0.0964	0.1040	$+0.0076$ $[-0.0002, +0.0153]$
Polynomial/J	Affinity	0.1154	0.1135	-0.0019 $[-0.0058, +0.0019]$
Polynomial/C	Clock	0.0922	0.0919	-0.0003 $[-0.0070, +0.0065]$
Polynomial/C	Affinity	0.1208	0.1213	$+0.0005$ $[-0.0046, +0.0056]$
Fourier/J	Clock	0.0864	0.0903	$+0.0039$ $[-0.0053, +0.0132]$
Fourier/J	Affinity	0.1211	0.1105	-0.0106 $[-0.0243, +0.0031]$
Fourier/C	Clock	0.0773	0.0756	-0.0017 $[-0.0147, +0.0114]$
Fourier/C	Affinity	0.1340	0.1272	-0.0068 $[-0.0284, +0.0147]$

The budget comparison uses precisely the same ten graph seeds on both sides, not the 30-seed primary mean versus a ten-seed subset. Larger particle and Monte Carlo budgets do not uniformly improve recovery. Every descriptive paired interval for the clock and affinity changes includes zero; this does not establish budget invariance. The comparison is only at scale 32 and does not certify convergence at other scales or remove design-grid limitations.

On these matched cases, mean full-posterior total-variation error decreases from .03892 to .01922; exact finite posteriors are audit references only. Across all primary cases, the mean is .03702 and the maximum individual dyad/update error is .07453. The maximum corresponding higher-budget error is .03585. No transition-count continuity violation or infeasible schedule is observed; the smallest anchor singular value is .50435. Independent ODE probabilities and the estimator's 80-node quadrature agree at the realized schedules to maximum absolute log-probability error 2.15×10^{-11}.

Automated checks cover vectorized versus scalar count likelihoods, zero information under identical parameter likelihoods, Metropolis detailed balance for a nonuniform prior, and SMC recovery of a known finite posterior. Final reaggregation recomputes clock/affinity errors and fitted log likelihoods from saved estimates and raw data, verifies the paired uniform streams, and checks completeness

against the full specified seed grid. Their scope is numerical implementation and record-consistency checking.

Primary acquisition/simulation loops average 2.38 seconds (J) and 2.74 seconds (C) per graph in this run, with sample SDs .53 and .78 seconds. Downstream fitting averages .087 and .083 seconds. These wall times include the planner, particle updates and data simulation, exclude construction of the precomputed transition grid, and were measured under concurrent CPU workers. Per-case and higher-budget times are archived. Concurrent workloads and the excluded grid construction limit these timings to implementation diagnostics rather than controlled end-to-end speed comparisons.

P.5 Reproduction and evidence organization

The PASOA protocol and its associated freeze manifest are included in `paper/validation/contemporary_benchmarks/`. The protocol's SHA-256 digest matches its manifest entry. This establishes consistency with the supplied manifest, not the timing of freezing relative to execution.

The PASOA implementation, referenced design arrays, raw datasets, estimates, optimizer records and particle/action diagnostics from `modern_baseline_study/pasoa_extension/` are not included. The methods, outcomes and execution checks in this appendix are therefore reported from the study account; their summary and execution commands cannot be rerun from the packaged files. Cached or resumed execution refers to continuation of a run rather than an independent replication. `STUDY_PATHS.json` distinguishes the included protocol and manifest from the missing implementation and execution records.

ADDITIONAL SUPPORTING STUDIES AND LIKELIHOOD METHODS

These appendices report restricted-prior acquisition results, phase-family diagnostics, affine-cell likelihood methods, forecasting adapters and geometric-recovery evidence, including adverse outcomes and implementation limits. Their protocols and cohorts are separate from both the full-affine-rank-prior acquisition comparison and the finite-clock evaluation in Appendix A (page 14); they are not additional baselines or replications of either study. Appendices B–N cover the information theory and mechanism studies. Checkpoint replay and fixed-data refitting are identified separately from independent experiments.

Reading guide. Appendices Q–T give supporting acquisition and recovery studies; Appendices U–W give likelihood statements, proofs and implementation details. Appendices X–Z report phase-family, topology and timing experiments. Appendices AA–AC describe forecasting protocols, execution checks and geometric diagnostics. Cross-references link the protocols, proofs and complete outcomes within this document.

Contents of the additional supporting studies

Q	Restricted-prior acquisition, hardware, and scaling	72
R	Polynomial/Fourier sensitivity and restricted-prior replication	80
S	Target-aligned predictive evaluation	87
T	Enlarged-nuisance and unknown-scale recovery	89
U	Positive likelihood control: statements	94
V	Proof and implementation of the second-order likelihood bound	95
W	Exact integration and positive observation likelihoods	96
X	Shared-graph recovery and computational-cost protocols	99
Y	Nonzero nuisance, homogeneous controls, and estimator sensitivity	102
Z	Topology, timing, and exposure experiments	107
AA	Reproducible learning protocols	113
AB	Learning results, uncertainty, and execution checks	116
AC	Geometric representation and field-recovery diagnostics	121

Q RESTRICTED-PRIOR ACQUISITION, HARDWARE, AND SCALING

This study provides restricted-prior evidence for hardware, exposure and scaling. Its acquisition prior has rank-two affinity support (Appendix L); its results are separate from the full-affine-rank-prior acquisition-criterion comparison in Appendix M.

Q.1 MODELS, STATISTICAL UNITS, AND CHRONOLOGY

This ten-seed restricted-prior study is distinct from the phase-family studies and the disjoint replication in Appendix R. Write $s(t) = 2t/5.5 - 1$. Polynomial affinity is

$$\eta_e(t) = -.2 + (B\alpha)_e + s(t)(B\beta)_e + \frac{1}{2}(3s(t)^2 - 1)(B\gamma)_e.$$

The Fourier condition instead uses $[1, \sin(2\pi t/5.5), \cos(2\pi t/5.5)]$. Every fitted coefficient is unknown. Both families are smooth, finite-dimensional models of (1), not restrictions of the specially constructed phase family. Basis normalization is fixed before any singular-value or optimization comparison. An omitted-basis experiment generates an additional $(5s^3 - 3s)(B\xi)/2$ term but still fits the three polynomial bases. Its results test one particular misspecification, not arbitrary-function robustness.

Ten paired seeds, 11000–11009, generate node clock deviations and each of three node affinity coefficient blocks independently from $U[-.15, .15]$, with streams `seed+50000`j, $j = 0, 1, 2, 3$. In the omitted-basis condition, $\xi_i \sim U[-.25, .25]$ uses stream `seed+290000`. True physical scales $\lambda \in \{8, 32, 64\}$ enter only through $v_i = v_i^{\text{dev}} + \frac{1}{2}\log(\lambda/32)$. Generation and fitting both use $c = \log 32 + Bv$; inference never receives the generating scale as an argument. No true-clock alignment, removal of the mean clock, or oracle initialization is used for errors or selection.

The initial crossed grid was fixed and its schedules searched before those recovery outcomes. It contains 770 count datasets. A follow-up joint-parameter Fisher comparator adds 110 datasets after the initial results revealed a clock/affinity tradeoff. Its motivation is therefore retrospective, but its search uses only the same prior grid, never recovery errors, generating coefficients, or test outcomes. The separate 20-dataset conditioning experiment specifies its two anchor configurations and all seeds before fitting. None of these experiments was externally preregistered. This chronology is recorded in three protocol files instead of describing all follow-ups as prospective.

The resulting 900 count datasets support 1,780 selected estimates: heterogeneous and same-data homogeneous fits for the 880 crossed datasets, and one heterogeneous fit for each of the 20 conditioning datasets. These counts are repeated conditions on ten independent realizations, not 900 independent seeds or 1,780 independent comparisons. The count representation compresses storage without reducing observation cost.

Q.2 OBSERVATION GRAPHS AND ACQUISITION BUDGETS

Each observed dyad has six transitions at known deterministic times in $[0, 5.5]$, conditioned on an initially observed state at zero. The initial state is drawn with probability $\text{sigmoid}(\eta_e(0))$, but its density is not an additional fitted likelihood term. Complete graphs use all $\binom{n}{2}$ dyads. Sparse graphs use the distinct unordered pairs $(i, i+1)$ and $(i, i+2)$ modulo n, for $n = 6, 16, 32, 64$. They have $m = 2n$ observed risk sets, are connected and contain triangles, and therefore have injective unsigned incidence. An unconnected observation is still in the risk set.

The experiment crosses replications $R = 10, 100, 1000$, graph size, physical scale, and minimum allowed gap $\delta \in \{1/256, 1/64, 1/16\}$ as specified in `protocol.json`. It is a selected factorial grid, not every possible Cartesian product: the complete six-node primary comparison varies scale and R at $\delta = 1/64$; the hardware comparison varies δ at $R = 100$ and scales 32 and 64; the sparse comparison varies n and R at scale 32; the 64-node and Fourier panels use $R = 100$. The full cell list is machine readable and all cell results appear below.

Every comparison at a stated δ enforces the same minimum-gap feasibility, count and horizon. This is a hardware-resolution constraint, not a common discrete timestamp lattice, acquisition-energy model, or timestamp-noise model. Shorter gaps than δ are not allowed to any candidate. Exposure is exactly $6mR$ observed transitions: on K_6, 9,000 at $R = 100$ and 90,000 at $R = 1000$; on the

32-node sparse graph, 3,840, 38,400 and 384,000 at the three replication levels. Increasing n at fixed R increases total acquisition, and is not advertised as free graph scaling.

Q.3 Design definitions and external provenance

The fixed template uses anchors $.75, 2.75, 5.5 - h_3$ and widths $h = \max\{(1/16, 1/64, 1/256), \delta\}$ componentwise. The matched-width randomized control has exactly these three widths and the same last anchor, with the first two anchors drawn from $[.2, 1.5]$ and $[1.8, 4.5]$ using the independent stream `seed+851000`. The intervals guarantee separation and feasibility. These internal controls are not additional external papers.

The rank-bound design solves the surrogate program in Appendix I. The first two anchors range over $[.25, 1.75]$ and $[2, 4]$; each width lies in $[\delta, .25]$; the last pair ends at 5.5. The smallest singular value of the normalized temporal basis matrix must be at least $.5$. The 16 equally weighted working scenarios combine scales $8, 16, 32, 64$ with affinity coefficient triples $(0, 0, 0)$, $(.25, .15, -.15)$, $(-.25, -.15, .15)$ and $(.25, -.15, .15)$. These working coefficients are not the recovery truths. The objective is the minimum across this finite grid of the summed frozen-pair weights. Between-grid worst-case behavior and finite-clock perturbation constants are not certified by that search.

Clock-targeted KW-Ds (HSB'14). We adapt the nuisance-adjusted prior design principle of Kyriakopoulos and Wolf (2015) to the conditional panel likelihood. For each working scenario, both starting states contribute according to their recursively propagated exposure, with no extra derivative of that exposure inside a conditional score. We profile the three local affinity directions, then maximize the mean scalar efficient clock information. This is the scalar version of a determinant of prior-mean efficient information. It differs from both the mean-log comparator and averaging a full information matrix before profiling. It also differs from the source paper's independent-cell observation model and chemical master-equation implementation.

Joint-parameter KW-D (HSB'14). The follow-up uses the same paper's full-parameter criterion: maximize the log determinant of the prior-mean 4×4 conditional information for a dyad clock and its three affinity coefficients. This targets joint reconstruction rather than clock information alone. KW-D and KW-Ds are two objectives from the same external design framework, not two independent source papers. Neither reproduces the adaptive Bayesian algorithm of Barendregt et al. (2023).

Both numerical comparators search unrestricted feasible six-gap schedules, allowing short initial observations. Gaps are parameterized as $g_i = \delta + (5.5 - 6\delta)\,\mathrm{softmax}(z)_i$, with the last logit fixed to zero. Forty initial candidates include structured schedules; KW-Ds also includes ten matched-width random schedules. Differential evolution uses 24 iterations followed by at most 400 Nelder–Mead evaluations, and the best training-free design objective determines selection. The rank-bound program uses population multiplier eight and 40 differential-evolution iterations followed by local polishing. All candidate improvements, exact schedules, scenario information, evaluation counts, elapsed times and stopping flags are retained. Finite budget exhaustion is not global optimality. Designs are validated at doubled quadrature precision before use.

Q.4 Independent data generation and estimation

Adaptive DOP853 integration of the forward probability ODE generates conditional transition probabilities for every dataset. A separately implemented positive nested quadrature verifies generating log probabilities. Replicated binary paths are then sampled sequentially via sufficient binomial counts, propagating the realized number in state one to the next interval. The noise stream is $\texttt{seed+}10000\lambda + R + 10^6 n + 10^8 f + 2\cdot 10^8 M$, where f indexes the family and M the omitted-basis condition. Streams pair conditions; this does not assert that different observation schedules contain identical continuous sample paths.

Every crossed dataset fits $4n$ heterogeneous coordinates and, separately, one homogeneous dyad clock plus all $3n$ affinity coordinates. The homogeneous offset has half-width 3.2, matching the field range of node clocks of half-width 1.6. All node affinity half-widths are .7. No ridge penalty is used. The two truth-independent starts are zero and one normal draw using `seed+600000`, with clock standard deviations .22 per node or .44 for the homogeneous dyad offset and affinity

Table 49: All fixed schedules, rounded only for display. F=0/1 denotes polynomial/Fourier; $d = 1/\delta$. Each row starts at conditioned time zero. Exact times and search diagnostics are in the artifact.

F	d	Design	t_1	t_2	t_3	t_4	t_5	t_6
0	16	template	0.75000	0.81250	2.75000	2.81250	5.43750	5.50000
0	16	bound	0.44873	0.51123	2.06112	2.12362	5.43750	5.50000
0	16	KW-Ds (HSB'14)	0.06250	0.12500	0.18750	0.25000	5.43750	5.50000
0	256	template	0.75000	0.81250	2.75000	2.76562	5.49609	5.50000
0	256	bound	0.27684	0.28624	2.00533	2.01447	5.45062	5.50000
0	256	KW-Ds (HSB'14)	0.01202	0.02404	0.03608	0.04802	5.48802	5.50000
0	64	template	0.75000	0.81250	2.75000	2.76562	5.48438	5.50000
0	64	bound	0.25000	0.26562	2.00000	2.01562	5.48438	5.50000
0	64	KW-Ds (HSB'14)	0.01562	0.03125	0.04688	0.06250	5.48438	5.50000
0	64	KW-D (HSB'14)	0.04183	0.08760	2.74890	2.79193	5.46262	5.50000
1	64	template	0.75000	0.81250	2.75000	2.76562	5.48438	5.50000
1	64	bound	0.77718	0.79281	3.81864	3.83429	5.48437	5.50000
1	64	KW-Ds (HSB'14)	0.01562	0.03125	0.04688	5.46875	5.48438	5.50000
1	64	KW-D (HSB'14)	1.79975	1.84210	3.65550	3.67112	5.48437	5.50000

standard deviation .08. Starts and bounds are identical across observation designs; no truth-based or recovery-based tuning is performed.

Signed binomial-deviance least squares minimizes the conditional negative log likelihood up to a data-only constant. A local fourth-order expansion avoids cancellation at nearly zero deviance. Trust-region reflective fitting uses analytic derivatives, Jacobian scaling, $\texttt{ftol} = 10^{-10}$ and $\texttt{xtol} = \texttt{gtol} = 10^{-9}$. Each start initially receives 300 evaluations; if the selected run reaches its cap, both initial starts are repeated with a 1200-evaluation cap. The initial capped record remains available. Selection minimizes the training objective, not recovery error. Every boundary estimate is retained; clock and nuisance boundary indicators use distance 10^{-3} from their respective boxes. A successful termination flag is not a proof of a global optimum.

The compiled kernel independently evaluates positive probabilities and local analytic derivatives. A lower normalized rate gives a tail cutoff of 80 integrated-clock units, so discarded mass is at most e^{-80}. Moving-cutoff derivative terms occur only at that exponentially small tail. Outer quadrature tiers are 40, 80 and 160, with inner order at least 12 and approximately one third of the outer order. Each selected fit must satisfy doubled-order log-probability error below 10^{-10} and derivative error below 10^{-9}. This is numerical refinement evidence, not the affine-cell exact-arithmetic certificate. Full-gap vectorized quadrature and independently integrated fitted probabilities provide additional cross-checks.

Q.5 Metrics, uncertainty, and complete results

Clock error is $\|B(\widehat{v} - v)\|/\sqrt{m}$, including the unknown common shift. Affinity error averages squared trajectory error over all observed dyads and 21 equally spaced times in $[0, 5.5]$. This measures reconstruction across the whole interval, not only at observed endpoints. In omitted-basis tests, the reference includes the omitted cubic term. We additionally retain mean clock error, local variance, selected function evaluations, status, individual start times and all fitted parameters.

Tables report means and sample standard deviations over ten paired realizations. Figures of mean RMSE use marginal Student intervals for the mean, not simultaneous bands. Pairwise differences are computed seed by seed; no family-wide significance or population-uniform guarantee is claimed. The replication plot instead displays root mean squared error across realizations to match $\sqrt{\overline{V}/R}$. The ratio $\overline{\|B(\widehat{v} - v)\|^2/m}/(\overline{V}/R)$ is descriptive: boundary, misspecified or inconsistent estimates need not obey a Cramér–Rao comparison. In particular KW-Ds often has nuisance boundaries even when its clocks are accurate.

For each matched cell and comparator D, let $d_s^{(u)} = E_{s,\mathrm{bound}}^{(u)} - E_{s,D}^{(u)}$ for target u (clock or whole-trajectory affinity). Table 52 reports $\bar{d}^{(u)} \pm t_{.975,9}\,\mathrm{SD}(d^{(u)})/\sqrt{10}$. Its eleven cells are exactly those containing all three rank-bound, KW-Ds and KW-D designs at gap floor $1/64$ under correct spec-

ification. No design is imputed in unexecuted cells. The accompanying CSV retains 270 metric-specific paired contrasts across the executed grid, including controls; it does not add realizations. These retrospective intervals are marginal, without multiplicity adjustment. For example, at polynomial K_6, scale 32 and $R = 100$, the rank-bound minus KW-D clock difference is $-.191$ with interval $[-.393, .011]$, and its affinity difference is .017 with interval $[-.006, .039]$. The mean ordering therefore does not establish a resolved difference in that cell.

Table 50: Complete polynomial results: means over ten realizations. C/S denote complete/sparse risk sets; $d = 1/\delta$; M indicates an omitted cubic term. Error is log-clock RMSE; A is trajectory-affinity RMSE; Cb/Nb are clock/nuisance boundary fractions; seconds include every executed start and status extension.

n	G	λ	R	d	Design	M	Error	SD	A	Cb	Nb	Sec.
6	C	8	100	64	bound	0	0.130	0.043	0.115	0.0	0.0	0.07
6	C	8	100	64	KW-Ds (HSB'14)	0	0.092	0.031	0.891	0.0	1.0	0.12
6	C	8	100	64	KW-D (HSB'14)	0	0.089	0.019	0.105	0.0	0.0	0.18
6	C	8	100	64	random	0	0.135	0.034	0.125	0.0	0.0	0.06
6	C	8	100	64	template	0	0.135	0.034	0.107	0.0	0.0	0.07
6	C	8	1000	64	bound	0	0.040	0.016	0.040	0.0	0.0	0.06
6	C	8	1000	64	KW-Ds (HSB'14)	0	0.026	0.005	0.719	0.0	1.0	0.11
6	C	8	1000	64	KW-D (HSB'14)	0	0.022	0.008	0.037	0.0	0.0	0.14
6	C	8	1000	64	random	0	0.028	0.006	0.040	0.0	0.0	0.06
6	C	8	1000	64	template	0	0.030	0.007	0.035	0.0	0.0	0.06
6	C	32	100	16	bound	0	0.833	0.472	0.087	0.4	0.0	0.26
6	C	32	100	16	KW-Ds (HSB'14)	0	1.002	0.635	0.582	0.5	0.7	0.79
6	C	32	100	16	template	0	1.247	0.423	0.090	0.8	0.0	0.65
6	C	32	100	64	bound	0	0.109	0.028	0.099	0.0	0.0	0.06
6	C	32	100	64	bound	1	0.099	0.037	0.194	0.0	0.0	0.07
6	C	32	100	64	KW-Ds (HSB'14)	0	0.074	0.013	0.927	0.0	1.0	0.10
6	C	32	100	64	KW-Ds (HSB'14)	1	0.057	0.019	0.988	0.0	1.0	0.12
6	C	32	100	64	KW-D (HSB'14)	0	0.300	0.264	0.083	0.1	0.0	0.29
6	C	32	100	64	random	0	0.118	0.033	0.122	0.0	0.0	0.06
6	C	32	100	64	template	0	0.119	0.039	0.107	0.0	0.0	0.06
6	C	32	100	64	template	1	0.103	0.027	0.159	0.0	0.0	0.06
6	C	32	100	64	uniform	0	1.798	0.104	0.107	1.0	0.0	0.72
6	C	32	100	256	bound	0	0.122	0.038	0.104	0.0	0.0	0.07
6	C	32	100	256	KW-Ds (HSB'14)	0	0.069	0.014	0.993	0.0	1.0	0.13
6	C	32	100	256	template	0	0.125	0.042	0.108	0.0	0.0	0.08
6	C	32	1000	64	bound	0	0.026	0.007	0.034	0.0	0.0	0.05
6	C	32	1000	64	bound	1	0.024	0.010	0.162	0.0	0.0	0.05
6	C	32	1000	64	KW-Ds (HSB'14)	0	0.018	0.006	0.793	0.0	1.0	0.10
6	C	32	1000	64	KW-Ds (HSB'14)	1	0.023	0.009	0.765	0.0	1.0	0.10
6	C	32	1000	64	KW-D (HSB'14)	0	0.051	0.024	0.028	0.0	0.0	0.13
6	C	32	1000	64	random	0	0.037	0.014	0.036	0.0	0.0	0.05
6	C	32	1000	64	template	0	0.032	0.012	0.032	0.0	0.0	0.05
6	C	32	1000	64	template	1	0.035	0.011	0.119	0.0	0.0	0.05
6	C	64	100	16	bound	0	1.087	0.120	0.099	0.7	0.0	0.49
6	C	64	100	16	KW-Ds (HSB'14)	0	0.988	0.175	0.671	0.7	1.0	1.38
6	C	64	100	16	template	0	1.148	0.187	0.093	0.9	0.0	0.49
6	C	64	100	64	bound	0	0.160	0.064	0.099	0.0	0.0	0.08
6	C	64	100	64	KW-Ds (HSB'14)	0	0.105	0.030	0.991	0.0	1.0	0.14
6	C	64	100	64	KW-D (HSB'14)	0	1.220	0.131	0.082	0.9	0.0	1.07
6	C	64	100	64	random	0	0.233	0.108	0.102	0.0	0.0	0.09
6	C	64	100	64	template	0	0.171	0.066	0.094	0.0	0.0	0.08
6	C	64	100	256	bound	0	0.130	0.063	0.100	0.0	0.0	0.08
6	C	64	100	256	KW-Ds (HSB'14)	0	0.090	0.023	0.929	0.0	1.0	0.15
6	C	64	100	256	template	0	0.125	0.038	0.095	0.0	0.0	0.10
6	C	64	1000	64	bound	0	0.042	0.020	0.027	0.0	0.0	0.06
6	C	64	1000	64	KW-Ds (HSB'14)	0	0.029	0.009	0.748	0.0	1.0	0.15
6	C	64	1000	64	KW-D (HSB'14)	0	0.780	0.388	0.024	0.8	0.0	0.63
6	C	64	1000	64	random	0	0.052	0.015	0.032	0.0	0.0	0.09
6	C	64	1000	64	template	0	0.051	0.020	0.029	0.0	0.0	0.08
6	S	32	10	64	bound	0	0.385	0.156	0.369	0.0	0.4	0.07

n	G	λ	R	d	Design	M	Error	SD	A	Cb	Nb	Sec.
6	S	32	10	64	KW-Ds (HSB'14)	0	0.284	0.102	0.951	0.0	1.0	0.09
6	S	32	10	64	template	0	0.463	0.209	0.379	0.0	0.5	0.09
6	S	32	100	64	bound	0	0.097	0.044	0.122	0.0	0.0	0.04
6	S	32	100	64	KW-Ds (HSB'14)	0	0.070	0.013	0.921	0.0	1.0	0.09
6	S	32	100	64	template	0	0.128	0.037	0.114	0.0	0.0	0.05
6	S	32	1000	64	bound	0	0.030	0.008	0.040	0.0	0.0	0.04
6	S	32	1000	64	KW-Ds (HSB'14)	0	0.024	0.010	0.730	0.0	1.0	0.07
6	S	32	1000	64	template	0	0.038	0.011	0.035	0.0	0.0	0.04
16	S	32	10	64	bound	0	0.373	0.101	0.390	0.1	0.6	0.33
16	S	32	10	64	KW-Ds (HSB'14)	0	0.285	0.050	1.034	0.0	1.0	0.37
16	S	32	10	64	template	0	0.472	0.130	0.370	0.1	0.7	0.36
16	S	32	100	64	bound	0	0.094	0.015	0.119	0.0	0.0	0.12
16	S	32	100	64	KW-Ds (HSB'14)	0	0.082	0.015	0.869	0.0	1.0	0.31
16	S	32	100	64	template	0	0.116	0.022	0.110	0.0	0.0	0.13
16	S	32	1000	64	bound	0	0.035	0.006	0.040	0.0	0.0	0.12
16	S	32	1000	64	KW-Ds (HSB'14)	0	0.026	0.005	0.794	0.0	1.0	0.27
16	S	32	1000	64	template	0	0.042	0.012	0.036	0.0	0.0	0.12
32	S	32	10	64	bound	0	0.358	0.076	0.373	0.1	0.8	0.60
32	S	32	10	64	KW-Ds (HSB'14)	0	0.271	0.038	0.957	0.0	1.0	1.21
32	S	32	10	64	KW-D (HSB'14)	0	1.264	0.139	0.317	1.0	0.0	14.04
32	S	32	10	64	template	0	0.522	0.168	0.358	0.3	0.8	1.11
32	S	32	100	64	bound	0	0.104	0.010	0.118	0.0	0.0	0.33
32	S	32	100	64	KW-Ds (HSB'14)	0	0.084	0.011	0.982	0.0	1.0	1.04
32	S	32	100	64	KW-D (HSB'14)	0	0.395	0.208	0.095	0.4	0.0	5.71
32	S	32	100	64	template	0	0.133	0.014	0.119	0.0	0.0	0.36
32	S	32	1000	64	bound	0	0.035	0.005	0.036	0.0	0.0	0.29
32	S	32	1000	64	KW-Ds (HSB'14)	0	0.028	0.003	0.777	0.0	1.0	0.83
32	S	32	1000	64	KW-D (HSB'14)	0	0.060	0.011	0.029	0.0	0.0	1.67
32	S	32	1000	64	template	0	0.042	0.006	0.037	0.0	0.0	0.30
64	S	32	100	64	bound	0	0.102	0.011	0.116	0.0	0.0	1.04
64	S	32	100	64	KW-Ds (HSB'14)	0	0.084	0.008	0.984	0.0	1.0	3.98
64	S	32	100	64	KW-D (HSB'14)	0	0.433	0.190	0.096	0.8	0.0	12.95
64	S	32	100	64	template	0	0.128	0.012	0.114	0.0	0.0	1.19

Table 51: Complete Fourier results: means over ten realizations. C/S denote complete/sparse risk sets; $d = 1/\delta$; M indicates an omitted cubic term. Error is log-clock RMSE; A is trajectory-affinity RMSE; Cb/Nb are clock/nuisance boundary fractions; seconds include every executed start and status extension.

n	G	λ	R	d	Design	M	Error	SD	A	Cb	Nb	Sec.
6	C	32	100	64	bound	0	0.089	0.036	0.137	0.0	0.0	0.09
6	C	32	100	64	KW-Ds (HSB'14)	0	0.061	0.014	1.261	0.0	1.0	0.19
6	C	32	100	64	KW-D (HSB'14)	0	0.103	0.031	0.100	0.0	0.0	0.17
6	C	32	100	64	random	0	0.109	0.025	0.143	0.0	0.0	0.09
6	C	32	100	64	template	0	0.106	0.024	0.124	0.0	0.0	0.07

Table 52: Paired differences for both recovery targets: rank-bound minus comparator, with marginal 95% Student intervals over the same ten paired realizations in each cell. Every row uses gap floor $1/64$. P/F denotes polynomial/Fourier affinity and C/S complete/sparse observation graphs. Positive differences favor the comparator for that metric. KW-Ds and KW-D are criteria from one external paper. Intervals are neither simultaneous nor joint, and rows sharing seeds are not independent replications.

Basis	Graph	λ	R	Comparator	Δ clock [95% interval]	Δ affinity [95% interval]
F	6C	32	100	KW-Ds (HSB'14)	+0.029 [+0.006, +0.051]	-1.124 [-1.208, -1.040]
F	6C	32	100	KW-D (HSB'14)	-0.014 [-0.031, +0.003]	+0.037 [+0.013, +0.062]
P	6C	8	100	KW-Ds (HSB'14)	+0.038 [-0.002, +0.079]	-0.776 [-0.822, -0.730]

Table 53: Same-data fitted homogeneous-clock controls, polynomial K_6, gap $1/64$. Every fit retains all 18 unknown affinity coordinates and receives the same initialization rule and likelihood. No true common scale is supplied.

λ	R	Design	Clock RMSE	SD	Affinity RMSE
8	100	bound	0.132	0.027	0.115
8	100	KW-Ds (HSB'14)	0.118	0.018	0.878
8	100	KW-D (HSB'14)	0.117	0.018	0.106
8	1000	bound	0.109	0.017	0.040
8	1000	KW-Ds (HSB'14)	0.109	0.018	0.683
8	1000	KW-D (HSB'14)	0.108	0.018	0.037
32	100	bound	0.118	0.018	0.099
32	100	KW-Ds (HSB'14)	0.114	0.019	0.941
32	100	KW-D (HSB'14)	0.136	0.034	0.083
32	1000	bound	0.108	0.018	0.034
32	1000	KW-Ds (HSB'14)	0.108	0.018	0.800
32	1000	KW-D (HSB'14)	0.110	0.017	0.028
64	100	bound	0.124	0.023	0.099
64	100	KW-Ds (HSB'14)	0.121	0.025	0.995
64	100	KW-D (HSB'14)	0.901	1.072	0.081
64	1000	bound	0.110	0.018	0.027
64	1000	KW-Ds (HSB'14)	0.110	0.018	0.751
64	1000	KW-D (HSB'14)	0.185	0.074	0.024

Basis	Graph	λ	R	Comparator	Δ clock [95% interval]	Δ affinity [95% interval]
P	6C	8	100	KW-D (HSB'14)	+0.041 [+0.012, +0.071]	+0.010 [-0.006, +0.026]
P	6C	8	1000	KW-Ds (HSB'14)	+0.015 [+0.003, +0.026]	-0.679 [-0.712, -0.646]
P	6C	8	1000	KW-D (HSB'14)	+0.019 [+0.007, +0.030]	+0.003 [-0.004, +0.010]
P	6C	32	100	KW-Ds (HSB'14)	+0.035 [+0.011, +0.058]	-0.827 [-0.968, -0.687]
P	6C	32	100	KW-D (HSB'14)	-0.191 [-0.393, +0.011]	+0.017 [-0.006, +0.039]
P	6C	32	1000	KW-Ds (HSB'14)	+0.008 [+0.003, +0.013]	-0.759 [-0.813, -0.705]
P	6C	32	1000	KW-D (HSB'14)	-0.026 [-0.040, -0.011]	+0.006 [+0.002, +0.010]
P	6C	64	100	KW-Ds (HSB'14)	+0.054 [+0.0004, +0.108]	-0.892 [-1.029, -0.756]
P	6C	64	100	KW-D (HSB'14)	-1.061 [-1.171, -0.950]	+0.017 [-0.0003, +0.034]
P	6C	64	1000	KW-Ds (HSB'14)	+0.014 [+0.004, +0.023]	-0.720 [-0.798, -0.643]
P	6C	64	1000	KW-D (HSB'14)	-0.737 [-1.013, -0.462]	+0.003 [-0.0002, +0.007]
P	32S	32	10	KW-Ds (HSB'14)	+0.087 [+0.022, +0.152]	-0.585 [-0.620, -0.549]
P	32S	32	10	KW-D (HSB'14)	-0.906 [-0.995, -0.817]	+0.055 [+0.033, +0.077]
P	32S	32	100	KW-Ds (HSB'14)	+0.021 [+0.011, +0.031]	-0.863 [-0.914, -0.812]
P	32S	32	100	KW-D (HSB'14)	-0.291 [-0.437, -0.144]	+0.023 [+0.011, +0.035]
P	32S	32	1000	KW-Ds (HSB'14)	+0.007 [+0.004, +0.010]	-0.740 [-0.774, -0.707]
P	32S	32	1000	KW-D (HSB'14)	-0.025 [-0.033, -0.017]	+0.007 [+0.005, +0.010]
P	64S	32	100	KW-Ds (HSB'14)	+0.018 [+0.012, +0.023]	-0.868 [-0.891, -0.844]
P	64S	32	100	KW-D (HSB'14)	-0.332 [-0.465, -0.198]	+0.020 [+0.013, +0.026]

Q.6 Controlled anchor conditioning

All three probe widths equal $1/64$. Let the last anchor be $a_3 = 5.5 - 1/64$ and the other anchors be $a_3 - 2d$ and $a_3 - d$. We vary $d \in \{2, 1, .5, .2, .1, .04, .03125\}$ at scales $8, 32, 128$. Every schedule retains six observations, horizon 5.5, identical probe widths and minimum-gap feasibility. The matrix is the normalized polynomial basis evaluated at the three anchors. Its smallest singular value decreases from $.47769$ to 4.3052×10^{-5}, while the clock-information eigenvalues do not collapse (Figure 4). Thus a poor sufficient anchor-conditioning margin is not evidence of lost clock identification.

A separately declared recovery comparison uses $d = 2$ and $.04$, scale 32, $R = 1000$ and the same ten paired seeds. Mean clock RMSE is $.02495$ and $.02360$, respectively. In contrast, affinity trajectory RMSE increases from $.04059$ to $.93225$, with nuisance boundaries in all ten clustered fits and none in the separated fits. The change concerns affinity-trajectory reconstruction from narrowly localized

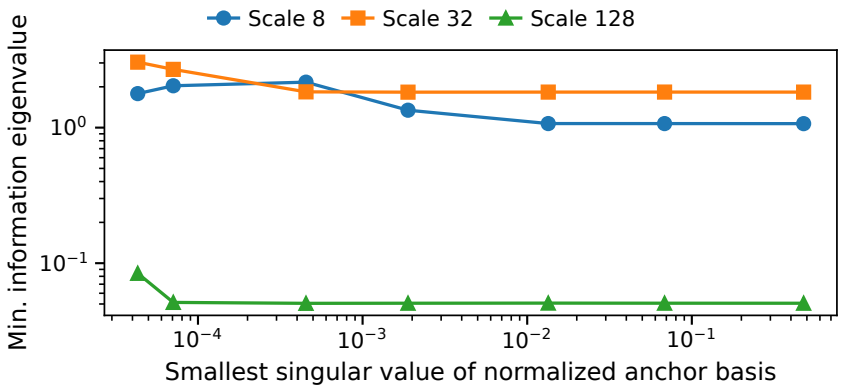

Figure 4: Width-matched anchor-conditioning diagnostic: identical short probes and progressively clustered polynomial anchors. Small anchor singular values do not force clock-information collapse. This tests the sufficient margin, not its converse.

Table 54: Polynomial affinities, K_6, common gap floor $1/64$, six post-initial transitions and horizon 5.5. Clock-field RMSE: mean $\pm$ SD over the ten paired realizations (lower is better). KW-Ds and KW-D adapt two objectives from one external source; KW-D is a retrospective follow-up. Paired contrasts are in Table 52.

λ	R	Fixed template	Rank-bound	KW-Ds (HSB'14)	KW-D (HSB'14)
8	100	0.135 ± 0.034	0.130 ± 0.043	0.092 ± 0.031	0.089 ± 0.019
8	1,000	0.030 ± 0.007	0.040 ± 0.016	0.026 ± 0.005	0.022 ± 0.008
32	100	0.119 ± 0.039	0.109 ± 0.028	0.074 ± 0.013	0.300 ± 0.264
32	1,000	0.032 ± 0.012	0.026 ± 0.007	0.018 ± 0.006	0.051 ± 0.024
64	100	0.171 ± 0.066	0.160 ± 0.064	0.105 ± 0.030	1.220 ± 0.131
64	1,000	0.051 ± 0.020	0.042 ± 0.020	0.029 ± 0.009	0.780 ± 0.388

observations, not a necessary clock-identification condition. This experiment also explains why a clock-only Fisher design should not be evaluated as though it optimized full affinity recovery.

Q.7 Hardware and scaling results

These displays concern matched physical resource limits and scaling in the restricted-prior study, not the full-affine-rank-prior common-pool comparison.

Q.8 Executed verification

Numerical replay and enforced checks. The restricted-prior acquisition audit regenerates all 900 count arrays from their probabilities, streams and initial exposures, and re-evaluates all 1,780 selected estimates. It verifies selection objectives, clock and whole-trajectory affinity errors, likelihoods, fitted coordinates, bounds, boundary flags, timing and exposure metadata, row mass, and doubled-order probabilities and derivatives. The independent quadrature is also recomputed against the generating probabilities for all datasets. Nine stratified fitted cases receive independent forward-ODE and full-gap vectorized-quadrature comparisons; both fitted families and omitted-basis data are represented. Finite differences check 144 parameter derivatives. The independent-solver checks are stratified, rather than claims of 900 ODE refits.

Each required discrepancy has a finite acceptance threshold; missing or nonfinite values, count discrepancies and incorrect metadata cause a nonzero exit. Table 56 gives the observed maxima and scope. Synthetic mutation tests deliberately inject discrepancies, missing coverage and nonfinite values: all 33 corruptions are rejected. These tests exercise validation logic without changing experimental data. The machine-readable audit and test results are in `submission_validation/strict_audit.json` and `submission_validation/gate_tests.json`.

Table 55: Sparse-graph scaling: $\lambda = 32$, $R = 100$, gap $1/64$, $m = 2n$. Rank-bound and KW-Ds are mean clock-field RMSE over ten paired realizations. Local is $\sqrt{\overline{V}/R}$ for KW-Ds, not a risk bound. Seconds sum both KW-Ds starts on a shared host.

n	m	Transitions	Rank-bound	KW-Ds (HSB'14)	Local	Seconds
6	12	7,200	0.097	0.070	0.081	0.09
16	32	19,200	0.094	0.082	0.081	0.31
32	64	38,400	0.104	0.084	0.081	1.04
64	128	76,800	0.102	0.084	0.081	3.98

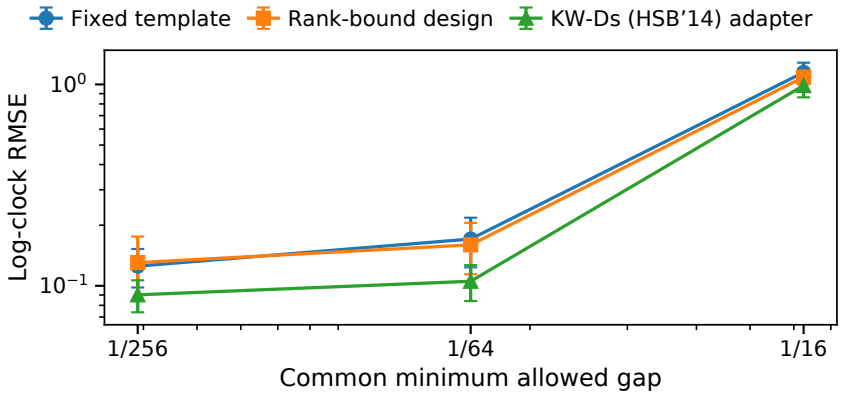

Figure 5: Restricted-prior hardware comparison, polynomial K_6, scale 64 and $R = 100$. All three designs obey the same gap floor, count and horizon. Marginal 95% intervals use ten paired seeds. Coarse timing degrades clock recovery.

Optimization reruns on fixed data. A separate execution repeats every fit in this study: two estimators on each of 880 crossed datasets and one on each of 20 conditioning datasets. It uses the specified zero and seeded random starts, common bounds, training-objective selection, and 300-evaluation budget with the same status-triggered 1200-evaluation extension. Saved estimates are used only after fitting for comparison, never for initialization or selection. All 1,780 selected parameter vectors, objectives, likelihoods and clock RMSEs reproduce exactly in the recorded environment; the largest affinity-RMSE discrepancy is 1.11×10^{-16}. Selected starts, termination and boundary flags also agree. One extension is repeated; 628 boundary-selected estimates remain (618 crossed and ten conditioning). Boundary flags can overlap and do not describe 628 independent failures.

Full starts, parameters, status and comparisons are retained in `submission_validation/refit/`. This validates optimizer reproducibility on fixed data; it neither adds independent seeds nor certifies global optimality. Design searches and the ODE data-generation workflow are not rerun in this fixed-data validation. The source datasets, fits and checkpoints are retained unchanged, rather than replaced by these verification outputs.

Theory diagnostics and separate validation scopes. A separate polynomial shrinking-probe test evaluates $\Lambda = 32, 128, 512, 2048, 8192$. The full-matrix operator discrepancy from the independently assembled frozen limit decreases from 1.06×10^{-4} to 4.04×10^{-7}, while the smallest eigenvalue approaches 1.023534636. One hundred random matrix perturbations satisfy the stability inequality (largest observed left/right ratio .03731). These numerical diagnostics complement the analytic proofs.

Acquisition checks have a separate scope from forecasting and phase-family recovery checks. The neural prediction-array replay exceeds its 2×10^{-6} threshold with a discrepancy of approximately 2.027×10^{-6}, and remains a failed strict check despite small aggregate-loss differences. One long-gap phase-family fit remains capped. Neither exception is reclassified, and the neural models were not retrained in this validation. Their records remain available with Appendix T.4.

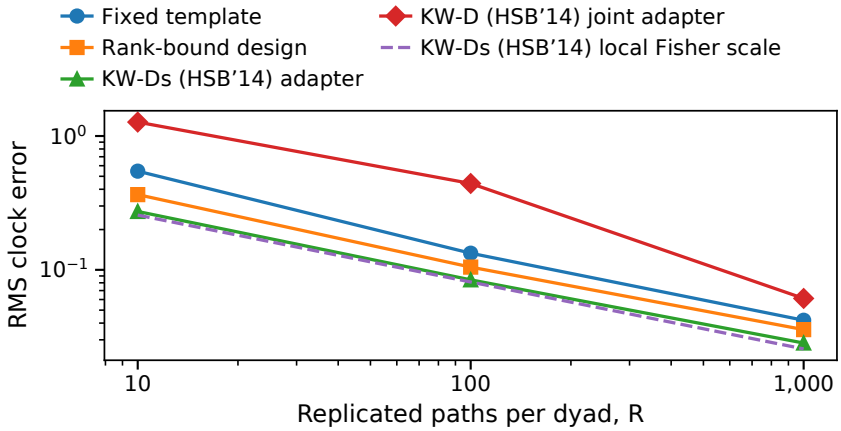

Figure 6: 32-node replication study, scale 32 and gap floor $1/64$. Curves show root mean squared error across ten paired realizations, not mean RMSE. The local Fisher scale is not a finite-sample risk bound.

Table 56: Restricted-prior study numerical replay. Maxima are absolute discrepancies except where log error is specified. Solver cross-checks use nine stratified cases each; derivatives use 144 coordinate checks. Thresholds are numerical validation tolerances, not statistical confidence levels. Optimization reruns are reported separately below.

Check	Scope	Maximum	Threshold
Saved fitted probabilities	1,780 fits	0	1×10^{-12}
Log likelihood	1,780 fits	0	1×10^{-7}
Clock RMSE	1,780 fits	0	1×10^{-10}
Affinity RMSE	1,780 fits	1.11×10^{-16}	1×10^{-10}
Row mass	1,780 fits	5.55×10^{-13}	2×10^{-10}
Doubled-order log probability	1,780 fits	6.94×10^{-13}	1×10^{-10}
Doubled-order derivative	1,780 fits	5.52×10^{-13}	1×10^{-9}
Generating-probability log error	900 datasets	3.54×10^{-13}	2×10^{-9}
Selected objective	1,780 fits	7.11×10^{-15}	1×10^{-9}
Independent ODE log error	9 checks	1.26×10^{-11}	2×10^{-8}
Vectorized quadrature log error	9 checks	5.04×10^{-13}	2×10^{-8}
Finite-difference derivative	144 checks	2.76×10^{-10}	2×10^{-7}

R Polynomial/Fourier sensitivity and restricted-prior replication

Restricted-prior status. The working prior in this restricted-prior study excludes $\beta + \gamma \neq 0$. Its results and execution records are reported separately, and method comparisons do not isolate adequately represented prior uncertainty or matched conditioning. The full-affine-rank-prior acquisition-criterion comparison is in Appendix M.

R.1 Scope and chronology

This extension addresses estimator-box dependence and acquisition-objective breadth in the polynomial and Fourier models, separately from the phase-family sensitivity study. It has three components: 360 fixed-data fits on the 60 central datasets of the ten-seed restricted-prior study; prior-only searches for two additional external criterion adapters; and 1,800 fits on 900 replication datasets generated from 30 disjoint seeds. The ten-seed study's 900 datasets and 1,780 fits are kept separate. Overall, the two restricted-prior finite-basis studies contain 1,800 unique count datasets and 3,940 selected fitted estimates, with intermediate optimization attempts recorded separately.

Table 57: Polynomial fixed-data sensitivity: all ten paired central realizations, two specified starts. b bounds every affinity-node coordinate; clock-node bounds stay ± 1.6. $\Delta\ell$ is mean conditional log-likelihood minus the baseline $b = .7$ result for the same design/data. No outcome is removed.

b	Design	Clock RMSE	Affinity RMSE	N_b	$\Delta\ell$
0.35	Rank-bound	0.1090	0.0995	0/10	0.0000
0.35	KW-Ds (HSB'14)	0.0745	0.4805	10/10	-0.4369
0.35	KW-D (HSB'14)	0.3003	0.0828	0/10	0.0000
0.7	Rank-bound	0.1090	0.0995	0/10	-0.0000
0.7	KW-Ds (HSB'14)	0.0744	0.9268	10/10	0.0000
0.7	KW-D (HSB'14)	0.3003	0.0828	0/10	0.0000
1.4	Rank-bound	0.1090	0.0995	0/10	-0.0000
1.4	KW-Ds (HSB'14)	0.0741	1.7725	10/10	0.6695
1.4	KW-D (HSB'14)	0.3003	0.0828	0/10	-0.0000
2.8	Rank-bound	0.1090	0.0995	0/10	-0.0000
2.8	KW-Ds (HSB'14)	0.0740	2.8705	9/10	1.4396
2.8	KW-D (HSB'14)	0.3003	0.0828	0/10	-0.0000

The extension is retrospective, not externally preregistered. The fixed-data bound grid and start rules were written before its fitting runs. The external objectives, prior, hardware limits and search budgets were fixed before the replication recovery outcomes; their selected schedules and source hashes were recorded before generating seeds 12000–12029. A subsequent, uniformly applied status-only optimization extension is described below. Recovery truth, test errors and boundary frequencies never select schedules, starts, hyperparameters or fitted estimates. The ten-seed cohort and the thirty-seed replication are reported separately, not pooled after observing their results.

R.2 Matched estimator bounds and additional starts

For every ten-seed-study central dataset (K_6, $\lambda = 32$, $R = 100$, $\delta = 1/64$), refit rank-bound, KW-Ds (HSB'14) and KW-D (HSB'14) at four affinity-node boxes $[-b, b]^{18}$, $b \in \{.35, .7, 1.4, 2.8\}$. The six clock-node bounds remain $[-1.6, 1.6]^6$, including the unknown common clock shift. All four boxes contain the generating affinity coefficients, which lie in $[-.15, .15]$. The design, data, likelihood, numerical gates and first two starts are identical across boxes: zero and a Gaussian start from stream `seed+600000`, with clock-node standard deviation .22 and affinity-node standard deviation .08. Starts are clipped just inside the box. Each start receives 300 evaluations, with a common status-triggered extension to 1200; none of these 360 final fits needs that extension.

The primary $2 \times 3 \times 10 \times 4 = 240$ fits use two starts. Another 120 fits use eight starts at $b = .7$ and 2.8, retaining the first two and adding six Gaussian starts with clock standard deviation .44 and affinity standard deviation $b/2$ from the same continued stream. The selected estimate always minimizes training deviance, never recovery error. Every fit and every boundary outcome is retained; all 360 selected fits terminate successfully. A coordinate is flagged at its bound when its distance to that bound is below 10^{-3}.

Tables 57–58 report clock-field error, 21-time trajectory-affinity error, affinity-boundary counts and conditional log-likelihood changes on identical data. For KW-Ds, increasing the box permits affinity motion that better fits observed transitions but worsens whole-trajectory recovery. At $b = 2.8$, polynomial KW-Ds has nine boundary fits out of ten and Fourier KW-Ds has ten; rank-bound and KW-D have none at any tested bound. In contrast, the KW-Ds clock means change by less than .0004 from $b = .7$ to 2.8. These results establish the tradeoff for the stated bounded estimators, not bound-independent affinity-error magnitudes.

The largest eight-versus-two-start change is 1.70×10^{-5} in clock RMSE and 2.32×10^{-5} in affinity RMSE (Table 59). This makes additional-start dependence an implausible explanation for the displayed error gaps on these datasets, but does not certify global optimization. The ten-seed central comparison, including all paired intervals in Table 52, is reported alongside the thirty-seed replication.

Table 58: Fourier fixed-data sensitivity: all ten paired central realizations, two specified starts. b bounds every affinity-node coordinate; clock-node bounds stay ± 1.6. $\Delta\ell$ is mean conditional log-likelihood minus the baseline $b = .7$ result for the same design/data. No outcome is removed.

b	Design	Clock RMSE	Affinity RMSE	N_b	$\Delta\ell$
0.35	Rank-bound	0.0894	0.1371	0/10	-0.0000
0.35	KW-Ds (HSB'14)	0.0605	0.6048	10/10	-0.5859
0.35	KW-D (HSB'14)	0.1035	0.0997	0/10	-0.0000
0.7	Rank-bound	0.0894	0.1371	0/10	-0.0000
0.7	KW-Ds (HSB'14)	0.0605	1.2613	10/10	-0.0000
0.7	KW-D (HSB'14)	0.1035	0.0997	0/10	0.0000
1.4	Rank-bound	0.0894	0.1371	0/10	-0.0000
1.4	KW-Ds (HSB'14)	0.0606	2.5020	10/10	0.6515
1.4	KW-D (HSB'14)	0.1035	0.0997	0/10	-0.0000
2.8	Rank-bound	0.0894	0.1371	0/10	-0.0000
2.8	KW-Ds (HSB'14)	0.0606	4.8278	10/10	1.0214
2.8	KW-D (HSB'14)	0.1035	0.0997	0/10	-0.0000

Table 59: Eight-start versus two-start fits on identical restricted-prior central data. First two starts and all other settings are unchanged. Error columns give the largest absolute paired change over ten realizations; $\overline{\Delta\ell}$ is mean eight-minus-two log likelihood.

Family	Design	b	Max clock change	Max affinity change	$\overline{\Delta\ell}$
Fourier	Rank-bound	0.7	9.9e-09	1.4e-08	-1.5e-12
Fourier	Rank-bound	2.8	5e-08	3.6e-08	6.4e-13
Fourier	KW-Ds (HSB'14)	0.7	9.8e-09	1.3e-08	1.7e-11
Fourier	KW-Ds (HSB'14)	2.8	1.8e-08	2.3e-05	4.4e-11
Fourier	KW-D (HSB'14)	0.7	7.1e-08	1.5e-08	-2e-12
Fourier	KW-D (HSB'14)	2.8	7.1e-08	1.2e-08	2.3e-12
Poly.	Rank-bound	0.7	1.4e-08	5.5e-09	-7.3e-13
Poly.	Rank-bound	2.8	3.1e-08	2.3e-08	8.2e-13
Poly.	KW-Ds (HSB'14)	0.7	2.6e-09	1.5e-07	7.3e-13
Poly.	KW-Ds (HSB'14)	2.8	9.7e-09	1.4e-06	7.3e-13
Poly.	KW-D (HSB'14)	0.7	8.3e-06	2.4e-08	2.4e-10
Poly.	KW-D (HSB'14)	2.8	1.7e-05	2.2e-08	3.9e-10

R.3 Two additional external acquisition criteria

Shared conditional experiment. Both adapters use the same 16 scalar-dyad working scenarios as the restricted-prior KW designs: scales $8, 16, 32, 64$ crossed with coefficient triples $(0, 0, 0)$, $(.25, .15, -.15)$, $(-.25, -.15, .15)$ and $(.25, -.15, .15)$, initially weighted uniformly. A query is a complete six-transition binary path $Y = (X_1, \ldots, X_6)$ with known initial state $A = X_0$. Every design has horizon 5.5, six post-initial observations, and minimum gap $1/64$. No method obtains an extra reset, observation or shorter gap.

For scenario s, let $\mu_s(a) = \Pr_s(A = a)$, $p_a = \frac{1}{16}\sum_s \mu_s(a)$, and $w_s^a = \mu_s(a)/(16p_a)$. Write

$$L_s^a(y;\xi) = \prod_{j=1}^{6} P_s(t_{j-1}, t_j)_{y_{j-1},y_j}, \qquad y_0 = a, \qquad \overline{L}^a(y;\xi) = \sum_s w_s^a L_s^a(y;\xi).$$

All $2^6 = 64$ outcomes are enumerated for each initial state. Conditioning reweights the scenario prior by the initial observation; its information is not rewarded as an extra acquired outcome. This is a scalar one-dyad, one-path design surrogate transferred to the replicated joint-graph estimator, not exact optimization of information for $R = 100$ paths and 24 graph parameters. This distinction applies to interpreting these adapters, rather than attributing the graph results to the source authors' complete pipelines.

Table 60: Clock versus trajectory-affinity recovery on K_6, scale 32, $R = 100$, common gap $1/64$. Means over ten paired realizations; lower is better in each column. KW-D targets all parameters and KW-Ds targets clocks. Affinity error covers 21 horizon-wide times; paired intervals are in Table 52.

Design	Poly. clock	Poly. affinity	Fourier clock	Fourier affinity
Rank-bound design	0.109	0.099	0.089	0.137
KW-Ds (HSB'14) adapter	0.074	0.927	0.061	1.261
KW-D (HSB'14) joint adapter	0.300	0.083	0.103	0.100

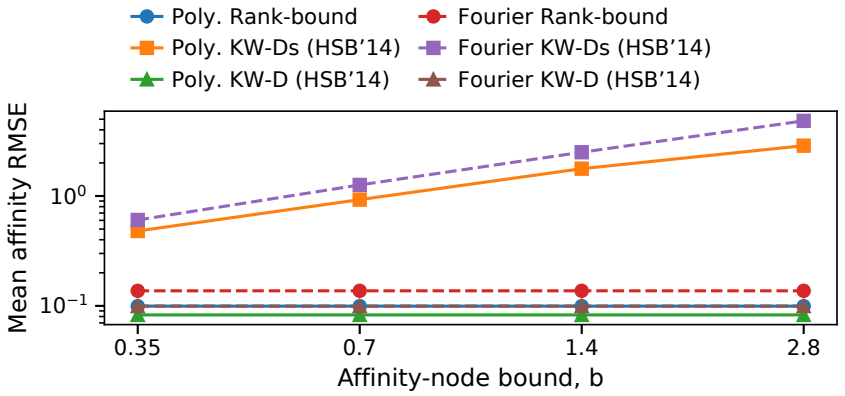

Figure 7: Restricted-prior bound sensitivity on ten paired central realizations. Schedules are fixed across bounds; the full-affine-rank-prior common-conditioning ablation is reported separately. Boxes do not add independent realizations.

REIG (UAI'22): ambiguity-set robustness. We adapt the affine robust expected-information criterion of Go and Isaac (2022), their equations (14) and (18). Define

$$d_s^a(\xi) = \sum_y L_s^a(y;\xi) \log \frac{L_s^a(y;\xi)}{\overline{L}^a(y;\xi)}.$$

The optimized criterion is

$$\sum_{a=0}^{1} p_a \min_{q \in \Delta_{16}:\, \mathrm{KL}(q\|w^a) \le .1} \sum_s q_s d_s^a(\xi). \tag{49}$$

The ambiguity radius .1 is fixed, not selected from recovery results. The inner minimum is evaluated by the exponential tilt $q_s \propto w_s^a \exp(-d_s^a/\tau)$, choosing τ to meet the KL constraint, with the appropriate endpoint cases. Its independently checked dual is

$$-\min_{\tau \ge 0} \left\{ .1\tau + \tau \log \sum_s w_s^a \exp(-d_s^a/\tau) \right\}.$$

Equation (49) is the source's affine robust criterion, not exact worst-case EIG under a simultaneously recomputed mixture. Exact finite-scenario enumeration replaces its sampling-based information estimator; no neural estimator is trained.

EPIG (AISTATS'23): prediction-oriented acquisition. Following the predictive-information objective of Bickford Smith et al. (2023), introduce a target $Z_t \sim \mathrm{Bernoulli}(\mathrm{sigmoid}(\eta_s(t)))$ independent of the query path conditional on scenario s. Its target time is uniform over 21 equally spaced points in $[0, 5.5]$. We maximize

$$\frac{1}{21} \sum_t \sum_a p_a I(Y; Z_t \mid A = a, \xi), \tag{50}$$

using the joint law

$$J_t^a(y,z) = \sum_s w_s^a L_s^a(y;\xi)[\mathrm{sigmoid}(\eta_s(t))]^z[1-\mathrm{sigmoid}(\eta_s(t))]^{1-z}.$$

The target is an independent equilibrium state, not the future state on the same measured path, a clock parameter, or an oracle trajectory observation. We report equilibrium-probability RMSE and proper predictive scores on this same target (Appendix S). Prediction-oriented information is not asserted to minimize clock or affinity-logit error.

Search and numerical evaluation. Six positive gaps are parameterized by a softmax allocation above the common floor, summing to 5.5. Each search uses a 40-candidate differential-evolution population for 24 generations, followed by up to 400 Nelder–Mead evaluations; structured rank-bound, KW-Ds, KW-D, fixed-template and randomized width-matched candidates are included. The same finite prior, feasibility constraints and outcome-independent search budget are used for both criteria. Source code records every incumbent improvement, selected times, evaluation counts and termination statuses. Search-budget termination is not a certificate of a global design optimum.

Transition integrals are evaluated at quadrature order 40 and rechecked at 80. Scenario-path probabilities sum to one, robust information does not exceed nominal EIG, and EPIG obeys the corresponding data-processing inequality. Independent ODE-based enumeration and a separately optimized REIG dual agree to below 2×10^{-14}. The sources comprise three external papers: KW-Ds and KW-D share one source, while REIG and EPIG add one each. The internal controls and endpoint-forecasting adapters are not additional external clock-acquisition baselines.

R.4 Disjoint 30-seed replication and complete results

Seeds 12000–12029 are disjoint from 11000–11009. They use the same node-coefficient distribution and shifted-clock generation rule, with scales $8, 32, 64$, complete six-node graphs, $R = 100$, and the five frozen schedules. Each of the $30\times2\times3\times5 = 900$ replication datasets is fitted at $b = .7$ and $b = 2.8$, giving 1,800 selected estimates. Initial states and transition counts are generated sequentially from independent adaptive-ODE probabilities; matched random streams pair designs within a seed and condition. A realization is reused across scales and boxes, so these are 30 independent units, not 900.

All fits initially receive two initial starts with the shared 300/1200 evaluation rule. Thirty-six selected fits require the 1200 extension; seven remain capped, all under EPIG. A separately recorded status-only completion gives *every* remaining capped fit the same 4800 budget from its initial starts, regardless of method, boundary status or recovery error. All seven then terminate, leaving zero final selected convergence failures. The initial capped attempts remain immutable, and a source-hash-checked record links each completion to its capped result. The completion changes central EPIG clock means by less than 2×10^{-5} and does not explain the target tradeoff. All 1,094 final boundary estimates are included in means, uncertainties and timing summaries.

For each primary metric—clock-field and horizon-wide affinity-logit RMSE—we report rank-bound minus each of four comparators within each family, scale and bound. The 96 contrasts are all retained. Marginal paired Student-t intervals use 29 degrees of freedom; a second set uses critical value $t_{29,1-.05/(2\cdot96)}$ for Bonferroni-adjusted simultaneous 95% coverage across this fixed family, under the same approximate t calibration. This is not a distribution-free guarantee or a correction for unreported exploratory analyses. Equilibrium-probability RMSE is secondary and is not included in the 96 primary contrasts. Standard deviations measure variation between the thirty realizations, not uncertainty between individual trajectories sharing fitted coefficients.

At the central setting, rank-bound's adjusted affinity-error differences relative to KW-Ds and REIG exclude zero in both families. KW-D has a smaller affinity error than rank-bound, also with adjusted intervals excluding zero. The polynomial clock differences between rank-bound and the three information-based comparators are inconclusive after multiplicity adjustment. These results support an inference-target trade-off. The expanded-box results retain the same interpretation: central KW-Ds affinity RMSE becomes 2.994 in the polynomial model and 4.985 in the Fourier model, while rank-bound remains $.104$ and $.166$. Full results include the high-scale KW-D clock failures and the EPIG adapter's weaker clock recovery; successful acquisition-objective optimization is not equated with successful recovery of every parameter.

Table 61: Central replication differences (rank-bound minus comparator), with Bonferroni-adjusted 95% intervals covering all 96 planned primary contrasts across two families, three scales, two bounds and two error metrics. Negative favors rank-bound. Intervals use 30 paired realizations and a Student-t approximation; they are not distribution-free guarantees. Machine-readable output also retains all marginal intervals.

Family	Comparator	Clock difference [interval]	Affinity difference [interval]
Poly.	KW-Ds (HSB'14)	0.014 [-0.004, 0.032]	-0.828 [-0.965, -0.690]
Poly.	KW-D (HSB'14)	-0.167 [-0.348, 0.014]	0.019 [0.007, 0.030]
Poly.	REIG (UAI'22)	0.014 [-0.005, 0.033]	-0.837 [-0.964, -0.710]
Poly.	EPIG (AISTATS'23)	-1.295 [-1.478, -1.113]	-0.474 [-0.551, -0.397]
Fourier	KW-Ds (HSB'14)	0.022 [0.001, 0.044]	-1.132 [-1.274, -0.990]
Fourier	KW-D (HSB'14)	-0.004 [-0.029, 0.020]	0.064 [0.035, 0.093]
Fourier	REIG (UAI'22)	0.019 [-0.003, 0.042]	-1.107 [-1.256, -0.958]
Fourier	EPIG (AISTATS'23)	-1.315 [-1.532, -1.097]	-0.028 [-0.068, 0.012]

Table 62: Polynomial independent replication, affinity-node bound 0.7: mean field RMSE (standard deviation) over 30 paired realizations. Equil. is equilibrium-probability RMSE, a secondary EPIG-aligned target; C/N are clock/affinity boundary counts out of 30. Time sums the two fitting starts, including status extensions, on a shared host. All cells use K_6, $R = 100$, six post-initial observations and minimum gap $1/64$.

λ	Design	Clock RMSE (SD)	Affinity RMSE (SD)	Equil.	C/N	Sec.
8	Rank-bound	0.116 (0.035)	0.125 (0.023)	0.031	0/0	0.07
8	KW-Ds (HSB'14)	0.088 (0.033)	0.892 (0.093)	0.187	0/30	0.12
8	KW-D (HSB'14)	0.076 (0.020)	0.108 (0.020)	0.026	0/0	0.06
8	REIG (UAI'22)	0.087 (0.032)	0.902 (0.106)	0.190	0/30	0.12
8	EPIG (AISTATS'23)	1.664 (0.549)	0.563 (0.119)	0.118	26/27	2.81
32	Rank-bound	0.085 (0.029)	0.104 (0.019)	0.025	0/0	0.05
32	KW-Ds (HSB'14)	0.070 (0.023)	0.931 (0.191)	0.194	0/30	0.11
32	KW-D (HSB'14)	0.252 (0.247)	0.085 (0.017)	0.021	2/0	0.11
32	REIG (UAI'22)	0.071 (0.023)	0.941 (0.177)	0.195	0/30	0.11
32	EPIG (AISTATS'23)	1.380 (0.255)	0.578 (0.114)	0.120	26/28	2.31
64	Rank-bound	0.170 (0.109)	0.094 (0.022)	0.023	1/0	0.09
64	KW-Ds (HSB'14)	0.115 (0.040)	0.911 (0.143)	0.194	0/30	0.12
64	KW-D (HSB'14)	1.052 (0.297)	0.081 (0.017)	0.020	26/0	0.43
64	REIG (UAI'22)	0.115 (0.039)	0.908 (0.154)	0.192	0/30	0.12
64	EPIG (AISTATS'23)	1.468 (0.159)	0.577 (0.092)	0.119	26/26	2.11

R.5 Panel-subdivision safeguard and independently executed checks

A retained pilot using single-panel endpoint quadrature encountered 48 probability-mass guard failures among 360 attempted fits, all in the wide-box, eight-start stress test. These were numerical execution failures, not failed statistical estimates. Extreme starts concentrate the transition integral more sharply than a single outer panel resolves. The pilot and its code are preserved in `robustness_extension/results/single_panel_start_stress/` and `code/sensitivity_model_single_panel.py`; they are not relabeled as successful runs.

The panel-subdivision implementation uses the same positive transition integral, analytic derivative and e^{-80} tail truncation, with a geometric-panel fallback only when the specified probability-mass gate fails. With q_{end} the endpoint total rate, the first panel in distance from the endpoint has width $\min\{\text{integration lag}, 8/q_{\mathrm{end}}\}$; subsequent widths extend geometrically to cover the full integration lag. Internal panel-boundary contributions cancel when differentiating the integral. Neither probabilities nor derivatives are renormalized. Mass, precision-doubling and optimizer tolerances are identical across the single-panel and panel-subdivision implementations. All 360 sensitivity cases were rerun under this uniform numerical rule, with zero execution failures; the 1,780-estimate ten-seed study and its numerical code are kept separate.

Table 63: Polynomial independent replication, affinity-node bound 2.8: mean field RMSE (standard deviation) over 30 paired realizations. Equil. is equilibrium-probability RMSE, a secondary EPIG-aligned target; C/N are clock/affinity boundary counts out of 30. Time sums the two fitting starts, including status extensions, on a shared host. All cells use K_6, $R = 100$, six post-initial observations and minimum gap $1/64$.

λ	Design	Clock RMSE (SD)	Affinity RMSE (SD)	Equil.	C/N	Sec.
8	Rank-bound	0.116 (0.035)	0.125 (0.023)	0.031	0/0	0.07
8	KW-Ds (HSB'14)	0.089 (0.034)	3.020 (0.420)	0.330	0/30	0.24
8	KW-D (HSB'14)	0.076 (0.020)	0.108 (0.020)	0.026	0/0	0.06
8	REIG (UAI'22)	0.088 (0.032)	3.008 (0.432)	0.331	0/30	0.25
8	EPIG (AISTATS'23)	1.712 (0.601)	0.731 (0.243)	0.138	27/1	0.53
32	Rank-bound	0.085 (0.029)	0.104 (0.019)	0.025	0/0	0.05
32	KW-Ds (HSB'14)	0.070 (0.023)	2.994 (0.622)	0.330	0/30	0.11
32	KW-D (HSB'14)	0.252 (0.247)	0.085 (0.017)	0.021	2/0	0.11
32	REIG (UAI'22)	0.071 (0.023)	3.014 (0.625)	0.332	0/30	0.11
32	EPIG (AISTATS'23)	1.360 (0.206)	0.800 (0.221)	0.148	28/0	1.24
64	Rank-bound	0.170 (0.109)	0.094 (0.022)	0.023	1/0	0.09
64	KW-Ds (HSB'14)	0.115 (0.040)	2.875 (0.566)	0.337	0/28	0.12
64	KW-D (HSB'14)	1.052 (0.297)	0.081 (0.017)	0.020	26/0	0.42
64	REIG (UAI'22)	0.115 (0.040)	2.891 (0.646)	0.331	0/28	0.11
64	EPIG (AISTATS'23)	1.447 (0.123)	0.772 (0.223)	0.144	29/0	0.52

Table 64: Fourier independent replication, affinity-node bound 0.7: mean field RMSE (standard deviation) over 30 paired realizations. Equil. is equilibrium-probability RMSE, a secondary EPIG-aligned target; C/N are clock/affinity boundary counts out of 30. Time sums the two fitting starts, including status extensions, on a shared host. All cells use K_6, $R = 100$, six post-initial observations and minimum gap $1/64$.

λ	Design	Clock RMSE (SD)	Affinity RMSE (SD)	Equil.	C/N	Sec.
8	Rank-bound	0.099 (0.031)	0.188 (0.048)	0.046	0/0	0.08
8	KW-Ds (HSB'14)	0.091 (0.030)	1.250 (0.177)	0.227	0/30	0.19
8	KW-D (HSB'14)	0.099 (0.032)	0.118 (0.021)	0.029	0/0	0.09
8	REIG (UAI'22)	0.082 (0.023)	1.181 (0.118)	0.219	0/30	0.19
8	EPIG (AISTATS'23)	1.148 (0.772)	0.211 (0.049)	0.051	16/0	0.42
32	Rank-bound	0.090 (0.028)	0.166 (0.040)	0.040	0/0	0.07
32	KW-Ds (HSB'14)	0.068 (0.018)	1.298 (0.190)	0.234	0/30	0.16
32	KW-D (HSB'14)	0.095 (0.028)	0.102 (0.021)	0.025	0/0	0.07
32	REIG (UAI'22)	0.071 (0.017)	1.273 (0.186)	0.230	0/30	0.18
32	EPIG (AISTATS'23)	1.405 (0.316)	0.193 (0.041)	0.047	27/0	2.50
64	Rank-bound	0.123 (0.053)	0.162 (0.034)	0.040	0/0	0.10
64	KW-Ds (HSB'14)	0.096 (0.035)	1.305 (0.168)	0.239	0/30	0.17
64	KW-D (HSB'14)	0.172 (0.130)	0.095 (0.014)	0.023	1/0	0.12
64	REIG (UAI'22)	0.102 (0.033)	1.273 (0.175)	0.230	0/30	0.19
64	EPIG (AISTATS'23)	1.402 (0.134)	0.192 (0.037)	0.047	25/0	2.04

The extension audit regenerates all 900 replication count datasets and recomputes both primary metrics, saved likelihoods, bound flags, initial starts and precision-doubling errors for all 2,160 final extension estimates. It also verifies the frozen design hashes and rejects above-threshold or nonfinite values in 20 scalar gate tests. Separate numerical checks cover ten fitted or extreme-initial-parameter cases, including the largest single-panel mass discrepancies, with 80 finite-difference derivative comparisons. Independent adaptive Radau integration gives a maximum transition log-probability discrepancy 8.06×10^{-10}; the largest derivative discrepancy is 3.14×10^{-8}, and the largest panel-order-doubling log discrepancy is 3.73×10^{-13}. Four independent ODE/path-enumeration and dual-criterion checks validate the two acquisition objectives. These are floating-point numerical cross-checks, not rigorous enclosures or global-optimization certificates.

The ten-seed study's count/estimate replay, 28 numerical gates and 33 corrupted-output rejection tests also pass in a separate rerun. Forecasting exceptions remain separate: the recorded prediction-

Table 65: Fourier independent replication, affinity-node bound 2.8: mean field RMSE (standard deviation) over 30 paired realizations. Equil. is equilibrium-probability RMSE, a secondary EPIG-aligned target; C/N are clock/affinity boundary counts out of 30. Time sums the two fitting starts, including status extensions, on a shared host. All cells use K_6, $R = 100$, six post-initial observations and minimum gap $1/64$.

λ	Design	Clock RMSE (SD)	Affinity RMSE (SD)	Equil.	C/N	Sec.
8	Rank-bound	0.099 (0.031)	0.188 (0.048)	0.046	0/0	0.08
8	KW-Ds (HSB'14)	0.092 (0.030)	5.002 (0.799)	0.347	0/30	0.49
8	KW-D (HSB'14)	0.099 (0.032)	0.118 (0.021)	0.029	0/0	0.09
8	REIG (UAI'22)	0.082 (0.023)	5.020 (0.771)	0.344	0/30	0.65
8	EPIG (AISTATS'23)	1.268 (0.794)	0.211 (0.049)	0.051	19/0	0.50
32	Rank-bound	0.090 (0.028)	0.166 (0.040)	0.040	0/0	0.07
32	KW-Ds (HSB'14)	0.068 (0.019)	4.985 (0.730)	0.340	0/30	0.20
32	KW-D (HSB'14)	0.095 (0.028)	0.102 (0.021)	0.025	0/0	0.07
32	REIG (UAI'22)	0.071 (0.017)	4.993 (0.765)	0.346	0/30	0.20
32	EPIG (AISTATS'23)	1.416 (0.301)	0.195 (0.041)	0.047	29/0	1.77
64	Rank-bound	0.123 (0.053)	0.162 (0.034)	0.040	0/0	0.10
64	KW-Ds (HSB'14)	0.096 (0.035)	4.955 (0.631)	0.348	0/30	0.19
64	KW-D (HSB'14)	0.172 (0.130)	0.095 (0.014)	0.023	1/0	0.12
64	REIG (UAI'22)	0.103 (0.033)	4.992 (0.756)	0.342	0/30	0.21
64	EPIG (AISTATS'23)	1.406 (0.125)	0.192 (0.037)	0.047	28/0	2.73

array discrepancy 2.027×10^{-6} remains above its 2×10^{-6} threshold. Neural training is not part of these validation checks. The artifact reports ten-seed results, pilot numerical failures, sensitivity and replication results, and status-only completions separately.

S TARGET-ALIGNED PREDICTIVE EVALUATION

These secondary scores use the restricted-prior designs. Predictive scores for the full-affine-rank-prior comparison are in Appendix M.4.

Target and estimand. The EPIG adapter in Appendix R targets an independent equilibrium edge state, not the next state of the measured trajectory. We evaluate this exact target with the same frozen joint-graph fits used in the recovery comparison. For each observed dyad e and each of the 21 equally spaced horizon-wide times t, put

$$p_{et} = \mathrm{sigmoid}(\eta_e(t)), \qquad \widehat{p}_{et} = \mathrm{sigmoid}(\widehat{\eta}_e(t)).$$

The target is $Z_{et} \sim \mathrm{Bernoulli}(p_{et})$. Its expected excess logarithmic loss and excess Brier risk, averaged over the $21m$ dyad–time pairs, are

$$L_{\mathrm{pred}} = \frac{1}{21m} \sum_{e,t} \mathrm{KL}(\mathrm{Bernoulli}(p_{et}) \,\|\, \mathrm{Bernoulli}(\widehat{p}_{et})), \tag{51}$$

$$B_{\mathrm{pred}} = \frac{1}{21m} \sum_{e,t} (p_{et} - \widehat{p}_{et})^2. \tag{52}$$

Both subtract the oracle risk for the same target, so lower is better and zero denotes the correct predictive probability. Expectations are evaluated analytically, without sampling extra labels. Truth is used only for scoring, not for fitting, selection, or acquisition. The predictor is the common constrained graph MLE, not a posterior-predictive implementation of the source paper. Thus this comparison evaluates the transferred acquisition criteria within the stated pipeline.

Relation to affinity and clock errors. For $A(x) = \log(1 + e^x)$, the Bernoulli KL equals $A(\widehat{\eta}) - A(\eta) - A'(\eta)(\widehat{\eta} - \eta)$. Since $0 \le A''(x) \le 1/4$ and $|\,\mathrm{sigmoid}'(x)| \le 1/4$,

$$0 \le L_{\mathrm{pred}} \le \frac{1}{8} E_\eta^2, \qquad 0 \le B_{\mathrm{pred}} \le \frac{1}{16} E_\eta^2, \qquad E_\eta^2 = \frac{1}{21m} \sum_{e,t} (\widehat{\eta}_e(t) - \eta_e(t))^2.$$

Table 66: Independent 30-seed replication: K_6, scale 32, $R = 100$, gap $1/64$, affinity bound .7. Clock/affinity are mean RMSE; Pred. is mean excess log loss (10^{-3} nats) on the independent equilibrium-state target; lower is better. N_b counts affinity-boundary fits. Secondary predictive scores use the same frozen graph fits (Appendix S).

	Polynomial				Fourier			
Design	Clock	Affinity	Pred.	N_b	Clock	Affinity	Pred.	N_b
Rank-bound	0.085	0.104	1.36	0/30	0.090	0.166	3.55	0/30
KW-Ds (HSB'14)	0.070	0.931	101.35	30/30	0.068	1.298	176.56	30/30
KW-D (HSB'14)	0.252	0.085	0.91	0/30	0.095	0.102	1.32	0/30
REIG (UAI'22)	0.071	0.941	102.89	30/30	0.071	1.273	171.51	30/30
EPIG (AISTATS'23)	1.380	0.578	38.51	28/30	1.405	0.193	4.76	0/30

Table 67: Target-aligned scores at the central setting: means (sample SDs) over 30 paired realizations. KL is excess logarithmic loss; Brier is excess Brier risk, both multiplied by 10^3. Probability RMSE is averaged after taking the root within each realization. These secondary scores do not add independent data or replace the primary contrasts.

Family	Design	10^3 KL (SD)	10^3 Brier (SD)	Prob. RMSE
Poly.	Rank-bound	1.358 (0.475)	0.663 (0.230)	0.0253
Poly.	KW-Ds (HSB'14)	101.350 (38.480)	38.808 (14.652)	0.1935
Poly.	KW-D (HSB'14)	0.914 (0.347)	0.448 (0.171)	0.0208
Poly.	REIG (UAI'22)	102.889 (40.917)	39.387 (16.021)	0.1952
Poly.	EPIG (AISTATS'23)	38.511 (13.563)	14.783 (4.258)	0.1202
Fourier	Rank-bound	3.553 (1.622)	1.724 (0.787)	0.0404
Fourier	KW-Ds (HSB'14)	176.562 (50.017)	55.283 (12.693)	0.2337
Fourier	KW-D (HSB'14)	1.318 (0.512)	0.645 (0.249)	0.0249
Fourier	REIG (UAI'22)	171.511 (58.496)	54.452 (20.704)	0.2298
Fourier	EPIG (AISTATS'23)	4.756 (1.934)	2.291 (0.917)	0.0469

These elementary bounds explain the connection to affinity-logit error. In contrast, the equilibrium target contains no clock parameter: both predictive risks can vanish while the clock estimate is wrong. Predictive scores therefore cannot certify clock identification, and clock error alone cannot assess the EPIG target.

Complete coverage and uncertainty. All 1,800 final estimates from the 900 disjoint-seed datasets are scored: two families, three scales, two affinity boxes, five designs, and thirty paired realizations. Boundary estimates and source-checked status completions remain included. This retrospective secondary analysis reuses saved fits, without new datasets, independent realizations, optimization or neural training. Its predictive comparisons are separate from the 96 primary recovery contrasts. Machine-readable results retain every cell, sample standard deviations, marginal Student intervals, and 96 paired contrasts for the two predictive scores. Bonferroni intervals over this secondary family use the same approximate Student calibration as Appendix R.4; they do not provide study-wide confirmatory coverage or correct exploratory choices.

Findings. At the central setting, KW-D has the smallest mean excess log loss in both families. Rank-bound is next, followed by EPIG; KW-Ds and REIG have much larger predictive losses despite smaller clock-error means (Table 66). EPIG's predictive loss is therefore not inferred from its poor clock recovery. The common-estimator experiment supports a target-dependent trade-off between acquisition and prediction objectives. Table 67 also reports excess Brier risk and the equilibrium-probability RMSE.

Recomputation. The standalone `release_tools/predictive_target.py` reads the frozen estimates and generating coefficients, checks the source hashes of status completions, and writes all metrics and tables to a separate directory. It computes KL using the log-partition formula and independently via Bernoulli probabilities; complementary probabilities are evaluated as $\mathrm{sigmoid}(-\eta)$ to avoid subtractive cancellation. The largest cross-formula discrepancy

is 3.56×10^{-15}, and recomputed probability RMSE agrees with the existing records within 5.56×10^{-17}. The expected-Brier identity is checked separately. These are deterministic numerical checks of the secondary analysis, not new statistical replications.

T ENLARGED-NUISANCE AND UNKNOWN-SCALE RECOVERY

T.1 EXPERIMENTAL QUESTION AND EXECUTABLE SPECIFICATION

The `uncertainty_study` experiment tests two distinct statements. At an unchanged symmetric truth, additional unknown affinity functions can sharply reduce long-gap clock information. Under jointly nonzero nuisance coefficients and unknown common scale, short-gap schedules can still support useful fitted clock recovery. The information diagnostic and recovery study use their own stated parameters; no symmetric coefficient is presented as measured estimator risk.

All implementations, fixed design specifications, raw count arrays, true probabilities, all fitted parameters, all initializations, objective values, termination flags, numerical refinements, and individual errors are retained. The complete study uses seeds 9900–9909, distinct from the phase-recovery seeds. The experiment was not preregistered. Search designs and bounds were fixed before these recovery fits; no outcome from this grid selects a schedule or estimator.

Truth and model. For each seed, a NumPy generator draws six $v_i^{\mathrm{dev}} \sim U[-.15, .15]$ followed by six $\alpha_i \sim U[-.12, .12]$. A second stream with seed plus 500000 draws $\gamma_i \sim U[-.1, .1]$ and $\beta_i \sim U[-.04, .04]$. The zero-nuisance control sets both latter vectors to zero, but still fits all their coordinates. All models use f and ϕ in (7) plus the linear basis $t - 3$. Thus this is uncertainty about coefficients in an enlarged finite family, not a misspecified-model or arbitrary-function experiment.

The data generator encodes the true physical scale $\lambda \in \{8, 32, 64\}$ as $v = v^{\mathrm{dev}} + \frac{1}{2}\log(\lambda/32)\mathbf{1}$. The forward model in both generation and fitting is $c = \log 32 + Bv$. The physical scale is never supplied as an estimator argument, and no truth-based centering or Procrustes-type alignment is used when computing log-clock-field error. The common scale is identifiable through v whenever the observation likelihood supplies information; it is not a second redundant fitted intercept.

Designs and acquisition resources. All five schedules are fixed across the three physical scales. Each has six observed transitions per dyad, an initial conditioned state at zero, and final time 5.5. The staggered, calibrated-nominal, uniform, and anchor–probe times are specified in this manuscript and `designs.json`. The exact numerical search times are

$$(.0123299998729,\ .0246369317467,\ 1.0057956323288,$$
$$1.0184497266576,\ 5.4873229301966,\ 5.5).$$

Their minimum gap is approximately .01230693, versus $1/256 = .00390625$ for the geometric probe template, .5 for staggering, $.5 + \rho_*/32$ for the nominal shift's first gap (its minimum gap), and $5.5/6$ for uniform observations. The numerical search allows a minimum gap of .002 for every candidate. Equal horizon and count do not mean equal hardware requirements: short-gap methods require finer time resolution and do not preserve the original gap multiset. No timestamp noise or acquisition-energy model is simulated.

The geometric template uses three separated anchors with full-rank evaluation matrix for $[1, \phi, t-3]$, followed by probe widths $1/16, 1/64, 1/256$. This is a transparent fixed schedule for the declared prior bracket $[8, 64]$, not a fitted estimate of the optimal spacing or a promise outside that bracket. In Theorem 5's separate asymptotic experiment, widths shrink with Λ; this finite study keeps them unchanged.

Numerical Fisher-design comparator. The search implements a D_s-type criterion with nuisance-adjusted *conditional panel* information. A working dyad has an unknown clock and three local affinity coefficients. For each candidate six-time schedule, exact conditional starting-state exposures are propagated along the path, clock and nuisance score rows are whitened, and a QR residual gives the efficient clock information. The objective is

$$\frac{1}{16}\sum_{s \in \mathcal{S}} \log I_{c,\mathrm{eff}}(T; s).$$

The working scenarios combine scales $\{8, 16, 32, 64\}$ with local (α, γ, β) patterns $(0, 0, 0)$, $(.15, .2, .04)$, $(-.15, -.2, -.04)$, and $(.15, -.2, -.04)$, equally weighted. These are design scenarios, not the ten generated node-vector truths. The same numerical origin 32 is used throughout.

This is a criterion-level adaptation of the nuisance-adjusted Fisher-design principle, not a reproduction of Kyriakopoulos and Wolf (2015). Their chemical examples use independent measurements and a prior-averaged information criterion; here observations are successive conditional transitions and the objective averages log efficient scalar information. Nor is this the online posterior-variance policy of Barendregt et al. (2023). That work addresses an unknown-rate sampling problem through Bayesian updates; our comparison is an offline prior-grid design with a declared uncertainty range.

Two differential-evolution runs use seeds 27001 and 27002, 30 candidate vectors and 22 generations, with 15 randomized clustered initial schedules and the four fixed controls among the initial population. Five interior endpoints are sorted, and the final endpoint is fixed. A 600-evaluation Nelder–Mead refinement starts from the best searched schedule. The minimum-gap constraint is enforced by a penalty, and every retained schedule is checked for validity. The selected design uses 1,923 objective evaluations in 129.19 seconds in the release runtime; stopping messages and the improvement trace are stored. This is a budgeted heuristic search, not a globally certified optimum. Selection uses design information, never recovery errors or realized counts. Selected scenario information is re-evaluated at doubled quadrature order.

T.2 Count generation and estimators

Independent adaptive DOP853 integration of the forward equation generates transition probabilities. A separate positive nested quadrature checks them; its maximum log discrepancy over all 280 datasets is 5.45×10^{-13}. Initial states are drawn from the specified equilibrium at zero and then conditioned upon in estimation; their density is not included in the fitted conditional likelihood. Each path's subsequent state counts are simulated by sequential binomial draws, preserving Markov starting-state exposures. Noise seeds are $s + 10000\lambda + R$, with an additional 9000000 for the zero-nuisance control. Shared streams pair schedules; the 280 condition-specific datasets are not 280 independent statistical units.

The grid contains 150 complete-graph datasets at $R = 10^4$ (five designs, three scales, ten seeds), 60 at $R = 10^3$ (two short-gap designs, three scales, ten seeds), 50 sparse-graph datasets (five designs, scale 32, ten seeds), and 20 zero-nuisance controls (nominal shift and probes, scale 32, ten seeds). The sparse graph is G_6 with risk sets $\{12, 13, 23, 34, 45, 56\}$, not a graph with six realized positive states. Exposure is $90R$ dyad transitions on K_6 and $36R$ on G_6. Sufficient counts reduce storage and computation, not the number of measured transitions.

Each dataset fits a 24-coordinate shared-clock model and a 19-coordinate homogeneous model with one common clock and all 18 affinity coordinates. Both use the same likelihood and nuisance family. The coordinate half-widths for the full fit are $(1.6, .8, 2, .4)$ for $(v, \alpha, \gamma, \beta)$; the homogeneous log-clock offset has half-width 3.2, preserving the same physical clock range for a constant field. No ridge penalty is used in the enlarged-nuisance study. Both fits have one zero and two random truth-independent starts; random scales and streams are explicit in `model.py`. No start is initialized at the truth.

Signed binomial deviance least squares with analytic derivatives minimizes the conditional negative log likelihood up to a parameter-independent constant. The implementation uses a local series for near-zero binomial deviance to avoid cancellation. Trust-region reflective optimization uses `ftol` $= 10^{-11}$, `xtol` $=$ `gtol` $= 10^{-10}$, and an initial budget of 1000 function evaluations per start. The selected start minimizes the training objective, regardless of its recovery error. All starts and convergence diagnostics remain available.

Four selected fits reached the initial cap. Each was rerun from the same initial three starts with a uniform 6000-evaluation cap, triggered only by termination status. Initial records are preserved separately. Three then terminate; one nominal-shift G_6 joint fit at seed 9908 still reaches the cap. That estimate remains in all tables. Successful termination is not global optimality, and high boundary fractions for long-gap fits remain part of the results. The search comparator sometimes reaches an affinity boundary despite accurate clock recovery; its clock and nuisance boundary flags are separately recorded. No method is granted a recovery-dependent bound or budget.

Table 68: Complete recovery grid: G6, nonzero nuisance generation, part 1. Every nuisance coordinate is fitted in both amplitude settings. Bound columns give fractions with any active fitting bound. Success is the selected start's optimizer termination status, not global optimality.

Scale	R	Design	Fit	RMSE	Bound	Success
32	10,000	Anchor–probe	common	0.0997 ± 0.0237	0.0	10/10
32	10,000	Anchor–probe	joint	0.0198 ± 0.0055	0.0	10/10
32	10,000	Nominal shift	common	1.9443 ± 0.8902	0.8	10/10
32	10,000	Nominal shift	joint	1.5062 ± 0.3996	1.0	9/10
32	10,000	D_s search	common	0.0993 ± 0.0240	1.0	10/10
32	10,000	D_s search	joint	0.0115 ± 0.0042	1.0	10/10
32	10,000	Staggered	common	1.1267 ± 0.5765	0.6	10/10
32	10,000	Staggered	joint	1.3243 ± 0.2749	1.0	10/10
32	10,000	Uniform	common	1.6610 ± 0.8780	0.4	10/10
32	10,000	Uniform	joint	1.8365 ± 0.5647	1.0	10/10

The positive endpoint-layer quadrature truncates at a dyad-specific integrated-clock lower bound of 80, bounding omitted mass by e^{-80}. Its analytic local derivatives omit only the exponentially small moving-tail contribution. Outer orders 48, 96, and 192 are available, each checked against doubling, with inner order at least 24 and half the outer order. Final log-probability changes must be below 10^{-10} and derivative changes below 10^{-9}. Maximum observed fitted changes are 7.68×10^{-12} and 1.63×10^{-11}. This nonlinear-phase computation is audited separately from the exact-arithmetic affine-cell likelihood theorem.

T.3 Complete estimation results

Clock-field error is $\|B(\widehat{v} - v)\|/\sqrt{m}$, including unknown common offset. Tables show means and sample standard deviations over ten paired realizations. All displayed intervals are marginal paired 95% Student intervals, not simultaneous bands or guarantees over a parameter class. Detailed CSVs also contain common-offset error, centered-field error, same-data fitted homogeneous comparisons, and reference-only errors for the prior origin and true-scale zero-offset estimate. The latter is not used as the principal empirical comparator.

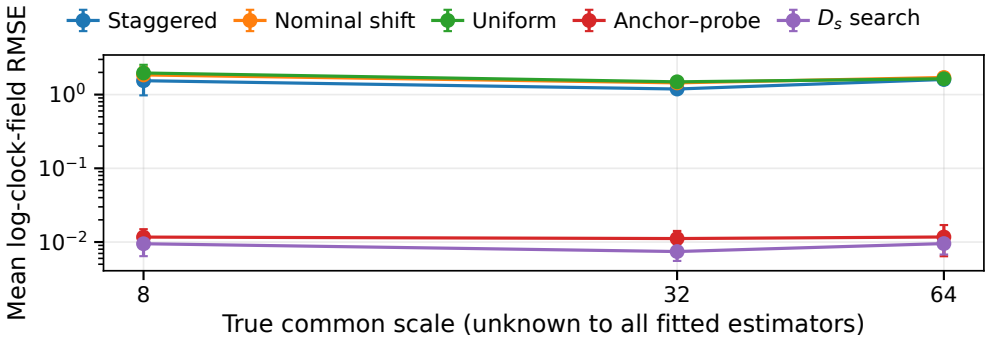

Figure 8: All five designs at the three unknown scales, enlarged nonzero nuisances, $R = 10^4$ and K_6. Error bars are across-realization SDs, not confidence intervals. Long-gap boundary solutions are retained. The numerical search is competitive with and often better than the fixed probe template.

T.4 Structural calculations and independent audits

The artifact evaluates 60 finite-information cells: five schedules, six physical scales $8, 16, 32, 64, 128, 512$, and two nuisance families, at otherwise unchanged symmetric truths. The main figure shows only the declared design bracket. Information is computed by projecting whitened score matrices, not subtracting nearly equal Fisher blocks. In the enlarged family at scale 32, the smallest eigenvalues are approximately 2.38×10^{-10} (staggered), 1.90×10^{-10} (nominal shift),

Table 69: Complete recovery grid: K6, nonzero nuisance generation, part 1. Every nuisance coordinate is fitted in both amplitude settings. Bound columns give fractions with any active fitting bound. Success is the selected start's optimizer termination status, not global optimality.

Scale	R	Design	Fit	RMSE	Bound	Success
8	1,000	Anchor–probe	common	0.1058 ± 0.0148	0.0	10/10
8	1,000	Anchor–probe	joint	0.0355 ± 0.0111	0.0	10/10
8	1,000	D_s search	common	0.1054 ± 0.0161	1.0	10/10
8	1,000	D_s search	joint	0.0324 ± 0.0088	0.9	10/10
32	1,000	Anchor–probe	common	0.1055 ± 0.0157	0.0	10/10
32	1,000	Anchor–probe	joint	0.0411 ± 0.0090	0.0	10/10
32	1,000	D_s search	common	0.1049 ± 0.0158	1.0	10/10
32	1,000	D_s search	joint	0.0235 ± 0.0046	1.0	10/10
64	1,000	Anchor–probe	common	0.1064 ± 0.0143	0.0	10/10
64	1,000	Anchor–probe	joint	0.0379 ± 0.0089	0.0	10/10
64	1,000	D_s search	common	0.1053 ± 0.0158	1.0	10/10
64	1,000	D_s search	joint	0.0266 ± 0.0069	1.0	10/10
8	10,000	Anchor–probe	common	0.1046 ± 0.0157	0.0	10/10
8	10,000	Anchor–probe	joint	0.0117 ± 0.0033	0.0	10/10
8	10,000	Nominal shift	common	3.1605 ± 1.8046	0.5	10/10
8	10,000	Nominal shift	joint	1.8484 ± 0.5132	0.8	10/10
8	10,000	D_s search	common	0.1046 ± 0.0158	0.0	10/10
8	10,000	D_s search	joint	0.0095 ± 0.0030	0.0	10/10
8	10,000	Staggered	common	0.6253 ± 0.4325	0.6	10/10
8	10,000	Staggered	joint	1.5432 ± 0.5672	1.0	10/10
8	10,000	Uniform	common	3.1683 ± 1.8513	0.6	10/10
8	10,000	Uniform	joint	1.9641 ± 0.5720	1.0	10/10
32	10,000	Anchor–probe	common	0.1047 ± 0.0157	0.0	10/10
32	10,000	Anchor–probe	joint	0.0111 ± 0.0030	0.0	10/10

Table 70: Complete recovery grid: K6, nonzero nuisance generation, part 2. Every nuisance coordinate is fitted in both amplitude settings. Bound columns give fractions with any active fitting bound. Success is the selected start's optimizer termination status, not global optimality.

Scale	R	Design	Fit	RMSE	Bound	Success
32	10,000	Nominal shift	common	2.2071 ± 1.0463	0.4	10/10
32	10,000	Nominal shift	joint	1.4427 ± 0.2402	0.8	10/10
32	10,000	D_s search	common	0.1046 ± 0.0157	0.6	10/10
32	10,000	D_s search	joint	0.0074 ± 0.0019	0.7	10/10
32	10,000	Staggered	common	0.6744 ± 0.5479	0.4	10/10
32	10,000	Staggered	joint	1.1905 ± 0.1314	0.7	10/10
32	10,000	Uniform	common	2.4814 ± 0.7204	0.5	10/10
32	10,000	Uniform	joint	1.4902 ± 0.1605	0.9	10/10
64	10,000	Anchor–probe	common	0.1046 ± 0.0158	0.0	10/10
64	10,000	Anchor–probe	joint	0.0117 ± 0.0053	0.0	10/10
64	10,000	Nominal shift	common	2.1531 ± 0.7327	0.3	10/10
64	10,000	Nominal shift	joint	1.7060 ± 0.2805	0.7	10/10
64	10,000	D_s search	common	0.1047 ± 0.0157	0.5	10/10
64	10,000	D_s search	joint	0.0095 ± 0.0028	0.5	10/10
64	10,000	Staggered	common	1.5893 ± 0.7140	0.5	10/10
64	10,000	Staggered	joint	1.6032 ± 0.2242	0.8	10/10
64	10,000	Uniform	common	2.1661 ± 0.5374	0.5	10/10
64	10,000	Uniform	joint	1.6396 ± 0.1928	0.8	10/10

2.06×10^{-10} (uniform), 1.020 (probes), and 2.573 (numerical search). These are numerical information diagnostics, not lower bounds on biased estimator risk.

A separate five-scale experiment tests Theorem 5. At $\Lambda = 32, 128, 512, 2048, 8192$, probe multipliers are $(2, .5, .125)$ and the three anchors tend to $.75, 2.75, 5.5$. The fixed node phase vector is $(.02, -.03, .01, -.02, .04, -.01)$ and the fixed linear-trend vector is $(.01, -.015, .005, .007, -.003, .011)$. The finite smallest eigenvalue changes from 1.019380 to

Table 71: Complete recovery grid: K6, zero nuisance generation, part 1. Every nuisance coordinate is fitted in both amplitude settings. Bound columns give fractions with any active fitting bound. Success is the selected start's optimizer termination status, not global optimality.

Scale	R	Design	Fit	RMSE	Bound	Success
32	10,000	Anchor–probe	common	0.1046 ± 0.0157	0.0	10/10
32	10,000	Anchor–probe	joint	0.0098 ± 0.0033	0.0	10/10
32	10,000	Nominal shift	common	1.9048 ± 1.1963	0.4	10/10
32	10,000	Nominal shift	joint	1.3883 ± 0.1361	0.9	10/10

1.019308, approaching the exact limiting eigenvalue 1.019307844; the independently evaluated frozen-pair lower bound is 1.019307842. This experiment changes the widths with Λ and must not be confused with the fixed-schedule unknown-scale grid.

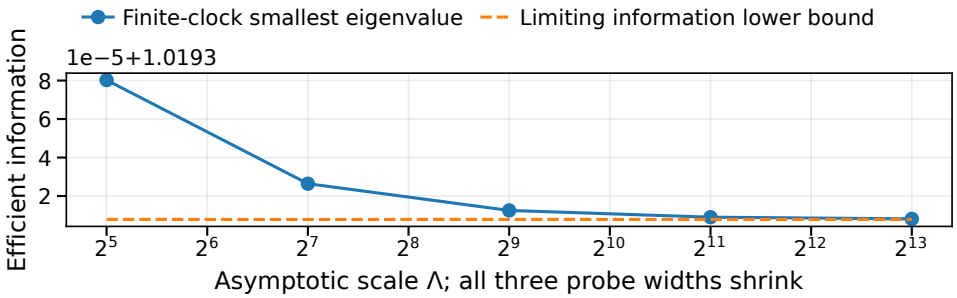

Figure 9: Shrinking probes with fixed nonzero nuisance coefficients. The nonvanishing limit and the limiting lower bound are distinct from a fixed physical schedule tested at changing true scales.

All 560 selected estimates are replayed against their saved probabilities, likelihoods and errors, and all 280 sequential count arrays are regenerated. Every design retains initial zero, six positive gaps and horizon 5.5. All selected-start decisions are checked against their candidate objectives, including cap-reaching fits. A separate 60-decimal computation verifies the small trend coefficients and the closed-form anchor determinant. Twenty-four parameter derivative finite differences and selected independent Radau truth/fitted integrations provide additional checks; precise discrepancies, software versions and command logs are recorded in `uncertainty_study/results/audit.json`. These numerical checks complement the analytic information results.

Forecasting checkpoint replay. This audit re-evaluates saved neural checkpoints without retraining. The 30-checkpoint primary subset and all 105 additional-study checkpoints were re-evaluated. One additional-study checkpoint exceeds the specified 2×10^{-6} predictive-array tolerance, reaching 2.027×10^{-6}; the strict audit therefore fails, and its log is retained. A continuation diagnostic records the tolerance violation and evaluates all 345 prediction blocks: maximum NLL and validation-NLL changes are 2.53×10^{-9} and 1.66×10^{-9}, and all selected steps agree. Source checkpoints and scientific records remain unchanged; replay outputs are stored separately. The replay shows near-agreement while documenting a tolerance violation.

T.5 Likelihood profiles and the role of optimization

The common log-clock profile imposes $2\,\mathrm{mean}(v) = g$ on a fixed 13-point grid $g \in [-1.5, 1.5]$. All five clock-contrast coordinates and all 18 nuisance coordinates are refitted under the specified coordinate bounds. The same seed-9900 K_6, scale-32, $R = 10^4$ nonzero-nuisance case is used for the nominal and probe schedules. Each constrained point uses the unconstrained fitted estimate, the previous constrained point and a zero-nuisance constant-clock initialization. Infeasible contrasts are reset without consulting the truth. The minimum training NLL is retained, with all starts and constraint errors. SLSQP uses a 2000-iteration budget. Profiles are numerical constrained optima, not globally certified likelihood curves, and are not assigned Wilks-calibrated confidence intervals.

Table 72: Unknown moving geometry: the same three state-path datasets, 12 nodes, $\lambda = 16$, 512 training path repetitions, and two validation candidates per configuration. Clock RMSE and test NLL are mean $\pm$ SD. The full five-dataset study is separate (Section AC).

Geometry bins	Update cap	Clock-field RMSE	Test NLL
3	800	0.794 ± 0.041	0.5130 ± 0.0003
3	2,400	0.511 ± 0.100	0.5130 ± 0.0003
6	800	0.811 ± 0.036	0.5131 ± 0.0004
6	2,400	0.585 ± 0.099	0.5131 ± 0.0004

Table 73: Prediction controls rather than native-input rankings. Positive paired NLL differences favor the affine head. Intervals include numerical-bound enlargement. Matched-input experiments use five additive-clock graphs and 2,400 updates. The strict gap-4 probe uses a separate ten-graph cohort, excludes gap 4 from training/validation, and matches input pathways.

Experiment	Comparator	NLL difference [95% interval]
Matched summaries	GraphMixer (ICLR'23)	$0.00114\ [-0.00028, 0.00257]$
Matched summaries	DyGFormer (NeurIPS'23)	$0.00371\ [0.00017, 0.00725]$
Matched summaries	Constant head	$-0.00027\ [-0.00124, 0.00071]$
Strict gap-4 probe	Constant head	$-0.01026\ [-0.01829, -0.00223]$
Strict gap-4 probe	MLP	$-0.00216\ [-0.01113, 0.00680]$

Figure 10 subtracts each schedule's minimum over this fixed grid. Across the full grid, the nominal-shift curve spans 11.45 total twice-NLL units. At $g = -.25$ and .25, its increases are .642 and zero, compared with 2156.14 and 2605.75 for probes. The complete profile run selects 26 constrained fits from 78 candidate starts; all selected starts terminate. Replayed objectives match exactly, maximum constraint error is 3.11×10^{-15}, and the largest 96/192-order log-probability change is 6.20×10^{-13}. These computational checks do not establish global optimality or a confidence-region interpretation.

The unknown-moving-geometry budget experiment is retained separately in Table 72. It uses the same three datasets at two optimization budgets, with nearly unchanged displayed test NLL and different clock errors. It therefore demonstrates inadequacy of predictive NLL as a recovery diagnostic, while leaving optimization and weak identification as distinct possible causes. It is not a proof of empirical nonidentifiability in that geometric model. The full autonomous study is documented in the artifact; Section AC separates its scope from the present results.

U Positive likelihood control: statements

Positive transition calculations and structure-preserving integration provide related numerical foundations (Arns et al., 2010; Coudière et al., 2020; Erreygers and De Bock, 2017). The bounds below concern positive assembly for this binary-rate model and are separate from its identifiability analysis. Accurate information analysis must be distinguished from accurate likelihood computation. Split piecewise-affine log-rates at facet and trajectory boundaries. Their hazard primitive is

$$\int_0^h e^{z_0+(z_1-z_0)s/h} ds = h e^{z_0}\,\mathrm{exprel}(z_1 - z_0), \qquad \mathrm{exprel}(w) = (e^w - 1)/w, \quad \mathrm{exprel}(0) = 1. \tag{53}$$

With exact integrated total clocks H_k and midpoint affinities $\widehat{\eta}_k$, set

$$E = e^{-\sum H_k}, \quad w_k = (1 - e^{-H_k}) e^{-\sum_{j>k} H_j},$$
$$\widehat{P}_{01} = \sum w_k\,\mathrm{sigmoid}(\widehat{\eta}_k), \quad \widehat{P}_{10} = \sum w_k\,\mathrm{sigmoid}(-\widehat{\eta}_k), \quad \widehat{P}_{00} = E + \widehat{P}_{10},\ \widehat{P}_{11} = E + \widehat{P}_{01}. \tag{54}$$

Theorem 16 (Rate-independent log bound). *If $|\eta - \widehat{\eta}_k| \le \delta$ on every cell, then $|\log P_{ab} - \log \widehat{P}_{ab}| \le \delta$ for every entry. For J affine pieces with total absolute affinity variation V_η, at most $J + V_\eta/(2\tau)$ cells suffice for log error τ, without a minimum probability or an upper clock bound.*

Theorem 17 (Second-order refinement). *On an affine cell of length h, let $D = |\eta'|$, $B_c = |c'|$, $q \le q_{\max}$, $\delta = Dh/2$, and $\rho = h^2\{D(B_c + D/2 + q_{\max}) + D^2 e^\delta/2\}/4$. If $\rho < 1$, the cell log*

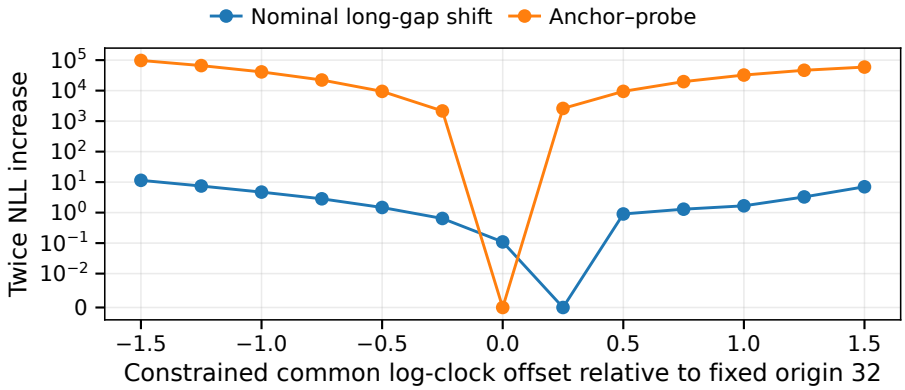

Figure 10: Common-clock profiles on the same fixed grid, with all contrasts and affinities refitted. At offsets $\pm.25$, the probe curve rises by over 2,000 total twice-NLL units, while the nominal-shift curve changes by less than one. One K_6 realization, $R = 10^4$, enlarged nonzero nuisances. These are numerical profiles, not confidence regions.

bound improves to $\zeta = \min\{\delta, -\log(1-\rho)\}$*; otherwise use* δ*. The complete positive assembly has bound* $\max_k \zeta_k$*. Bounded slopes and clocks give* $O(\tau^{-1/2})$ *subdivisions.*

The first proof uses positive weights and 1-Lipschitz $\log\operatorname{sigmoid}(\pm\eta)$; the second uses weighted midpoint cancellation. Appendices W and V give proofs and mesh rules. These are exact-arithmetic value bounds, not floating-point enclosures, gradient guarantees, or optimizer certificates. Bounds add over records and average for mean losses. Holding a causal head's coefficients fixed at a forecast origin transfers the numerical result to that conditional forecast, not an identification theorem to its neural weights.

V Proof and Implementation of the Second-Order Likelihood Bound

Consider one cell $[m - h/2, m + h/2]$ and the positive normalized weight proportional to $w(t) = q(t)\exp\{-\int_t^v q(s)ds\}$. The endpoint v may be the end of the entire observation gap. Its later-cell factor is constant within the current cell. For affine c and η,

$$(\log w)' = (\log q)' + q, \qquad (\log q)' = c' + \tfrac{1}{2}\eta' \tanh(\eta/2).$$

Therefore $|(\log w)'| \le K = |c'| + |\eta'|/2 + q_{\max}$.

Let $S = t - m$. Reflection gives, for $s \ge 0$,

$$\frac{|w(m+s) - w(m-s)|}{w(m+s) + w(m-s)} = \left|\tanh\left(\frac{\log w(m+s) - \log w(m-s)}{2}\right)\right| \le \tanh(Ks) \le Ks.$$

Integrating the numerator of $\mathbb{E}_w S$ in symmetric pairs and dividing by the common normalization shows

$$|\mathbb{E}_w S| \le K\mathbb{E}_w S^2 \le Kh^2/4. \tag{55}$$

For either $F(t) = \operatorname{sigmoid}(\eta(t))$ or $F(t) = \operatorname{sigmoid}(-\eta(t))$, direct differentiation yields

$$|F'(m)| \le DF(m), \qquad |F''(t)| \le D^2F(t) \le D^2e^{D|t-m|}F(m).$$

The last inequality uses the log-sigmoid Lipschitz property. Taylor's theorem with an integral remainder consequently gives

$$\begin{aligned}\left|\frac{\mathbb{E}_w F(t)}{F(m)} - 1\right| &\le D|\mathbb{E}_w S| + \tfrac{1}{2}D^2e^{Dh/2}\mathbb{E}_w S^2 \\ &\le \frac{h^2}{4}\left(DK + \tfrac{1}{2}D^2e^{Dh/2}\right) = \rho.\end{aligned}$$

If $\rho < 1$, the ratio lies in $[1-\rho, 1+\rho]$. Its absolute logarithm is at most $-\log(1-\rho)$. Theorem 16 also gives $Dh/2$; either valid bound may be used. Each exact cell integral is bounded multiplicatively by its positive midpoint approximation. Summing cells preserves a common multiplicative bound. Adding the unchanged nonnegative survival term preserves it for the diagonal entries. This proves Theorem 17.

V.1 An executable mesh rule

For an affine interval of length H, define a global upper bound

$$\begin{aligned} q_{\max} = \; & \exp(\max\{c(0) + \eta(0)/2, c(H) + \eta(H)/2\}) \\ & + \exp(\max\{c(0) - \eta(0)/2, c(H) - \eta(H)/2\}) \, . \end{aligned}$$

The maximum of each affine log-rate is attained at an endpoint, even when their sum has a smaller maximum. Put $K = |c'| + D/2 + q_{\max}$ and

$$M_1 = \max\{1, \lceil DH/(2\tau) \rceil\}, \quad M_2 = \max\left\{1, \lceil DH/2 \rceil, \left\lceil H\sqrt{\frac{DK + eD^2/2}{4(1-e^{-\tau})}} \right\rceil \right\}.$$

Then $M = \min(M_1, M_2)$ equal cells suffice. In the second branch $Dh/2 \le 1$ and $\rho \le 1 - e^{-\tau}$; in the first branch the original bound suffices. The implementation evaluates the tighter cellwise bounds after construction and verifies their maximum. A batch may use the largest required M among its records. Adaptive mesh decisions are detached during differentiation; the gradient audit tests the resulting fixed-mesh surrogate, not differentiability of the integer selection map. The code rejects nonfinite coefficients or excessive requested subdivisions instead of silently accepting an unmet tolerance.

V.2 Loss comparison and precision

For records $r = 1, \ldots, N$ and exact penalties, let $|\widehat{\ell}_r - \ell_r| \le \zeta_r$. Then $|\widehat{L} - L| \le \bar{\zeta} = N^{-1} \sum_r \zeta_r$. For two selected checkpoints, their exact difference is within $\bar{\zeta}_1 + \bar{\zeta}_2$ of the computed difference. No statistical averaging reduces these deterministic bounds. For a finite set of candidates with common error τ, selecting the smallest approximate validation loss gives exact validation loss at most 2τ above the smallest exact candidate loss. This is a comparison within the given candidate set, not a generalization bound or a statement about the nonconvex global optimum.

Numerical enclosure of paired-interval endpoints. Suppose a paired graph-level difference d_i is approximated by $\widehat{d}_i$ with $|d_i - \widehat{d}_i| \le b_i$, where b_i is the sum of the two mean-loss numerical bounds. For $n \ge 2$ graphs, put $t_* = t_{n-1,0.975}$. Both endpoints of the exact-data paired Student interval are within

$$B_n = \frac{1}{n} \sum_i b_i + \frac{t_* \|b\|_2}{\sqrt{n(n-1)}}$$

of the corresponding endpoints computed from $\widehat{d}$. Indeed, the mean changes by at most the first term. The sample standard deviation is a centered Euclidean norm divided by $\sqrt{n-1}$; the reverse triangle inequality and the norm-one centering projection bound its change by $\|b\|_2/\sqrt{n-1}$. Multiplication by $t_*/\sqrt{n}$ gives the second term. The paired figures and comparison CSVs use this full enlargement, not just the error in the mean. This is a deterministic discretization enclosure of an interval statistic, not a finite-sample coverage theorem for nonnormal graph-level differences or a directed-rounding guarantee.

W Exact integration and positive observation likelihoods

W.1 Piecewise-exponential clock integration

Fix a trajectory bin and write the relative position as $x + vt$. For the signed facet set $F = \{a_k, -a_k\}_{k=1}^K$,

$$\nu_A(x + vt) = \max_{f \in F} (f^\top x + t f^\top v).$$

A line is strictly maximal on an interval: its dominance region is an intersection of half-lines. Consequently, at most $2K$ lines are active. Sorting slopes, retaining the largest intercept among parallel lines, and applying a stack-based upper-hull construction takes $O(KD + K \log K)$ arithmetic operations. Clock knots, motion knots, and observation boundaries are merged before this calculation.

On each resulting piece, both log-rates $z(t) = c(t) \pm \eta(t)/2$ are affine. The endpoint exponential primitive is stated in Eq. (53) in Appendix U and follows by direct integration, with its continuous zero-slope limit. The implementation evaluates small w by its Taylor expansion and uses `expm1` otherwise. A zero-length piece contributes zero. Summing formation and dissolution integrals gives the exact integrated clock, including when either individual log-rate has zero slope.

Differentiation. Suppose no distinct signed facets have identical affine traces on a bin. The maximizer is then unique except at finitely many times. On a bounded parameter neighborhood, distances and exponentiated rates are locally Lipschitz with uniformly bounded difference quotients. Dominated convergence permits differentiation of the integrated rate through the almost-everywhere active facet. The resulting derivative is continuous on this nondegenerate domain. Equivalently, differentiating the explicit piece sum produces crossing-time boundary terms that cancel, since adjacent hazard formulas agree at the crossing. Persistent ties are excluded from an ordinary-gradient claim. None of this proves smoothness of an event log-rate at an event occurring exactly on a facet tie, or global convergence of nonconvex training.

The scalar reference implementation uses the upper hull. For simple tensor batching, `engine.py` instead splits at all pairwise signed-facet intersections, including inactive ones. Its polygonal partition may therefore have $O(K^2)$ candidate boundaries rather than the hull's $O(K)$. This implementation tradeoff is not hidden in the theoretical cost claim. For ℓ_1, the batched implementation uses coordinate zeros directly and requires at most $D + 1$ pieces within an affine bin.

W.2 Proof of Theorem 16

Let $Q(t, v) = \int_t^v q(s)ds$, $E = e^{-Q(u,v)}$, and let the positive measure

$$d\mu(t) = e^{-Q(t,v)} q(t)dt$$

have total mass $1 - E$. The scalar forward equation gives, for either initial state,

$$P_{01} = \int p(t)d\mu(t), \qquad P_{10} = \int (1 - p(t))d\mu(t), \tag{56}$$

$$P_{00} = E + \int (1 - p(t))d\mu(t), \qquad P_{11} = E + \int p(t)d\mu(t). \tag{57}$$

For a cell $C_k = [t_k, t_{k+1}]$, its measure is

$$\mu(C_k) = (1 - e^{-H_k})e^{-\sum_{j>k} H_j} = w_k.$$

This identity uses the complete integrated clock, not the clock at a representative time.

Since

$$\frac{d}{d\eta} \log \operatorname{sigmoid}(\eta) = 1 - \operatorname{sigmoid}(\eta), \qquad \frac{d}{d\eta} \log \operatorname{sigmoid}(-\eta) = -\operatorname{sigmoid}(\eta),$$

both functions are 1-Lipschitz. Thus $|\eta(t) - \widehat{\eta}_k| \le \delta$ implies

$$e^{-\delta} \operatorname{sigmoid}(\pm\widehat{\eta}_k) \le \operatorname{sigmoid}(\pm\eta(t)) \le e^{\delta} \operatorname{sigmoid}(\pm\widehat{\eta}_k).$$

Integrating against the *same* positive measure proves the bounds for both off-diagonals. Adding the nonnegative, unchanged term E preserves the bounds for both diagonals, because $e^{-\delta}E \le E \le e^{\delta}E$. Strict positivity permits taking logarithms.

For an affine-affinity piece with absolute change V_k, equal subdivision into M_k pieces makes the maximum deviation from the midpoint affinity $V_k/(2M_k)$. The stated choice of M_k therefore suffices. Since $\max\{1, \lceil x \rceil\} \le 1 + x$ for $x \ge 0$, the cell count is at most $J + V/(2\tau)$. This is an existence/algorithmic count using piece-specific refinement. The batched prototype pads pieces to the largest subdivision count in a batch, and may use more cells.

Algorithm 3 Snapshot evaluation and optional verified update

Require: Parameters ϑ, observed gaps, target per-record tolerance $\tau > 0$.
1: Split each gap at motion, activity, and norm-facet boundaries.
2: Refine piece k into $\max(1, \lceil V_k/(2\tau) \rceil)$ cells.
3: Integrate both rates exactly on each cell and form H_k.
4: Compute $\log w_k = \log(1 - e^{-H_k}) - \sum_{j>k} H_j$.
5: Assemble both transition directions positively using (54).
6: Propagate endpoint-extremum bounds; return the loss interval and surrogate loss.
7: **if** a candidate update ϑ' is supplied **then**
8: Evaluate both objective intervals, including their penalties.
9: Accept only if the candidate upper endpoint is smaller than the current lower endpoint.
10: Otherwise refine or return an unresolved/rejected proposal.
11: **end if**

For multiple observed gaps, applying the bound to each observed entry gives

$$|\ell - \widehat{\ell}| \le \sum_r \delta_r. \tag{58}$$

Consequently, a normalized loss is controlled by the corresponding normalized tolerance sum. No assumption that the states at different times are independent is made. □

Relation to absolute-error enclosures. The absolute bound $L_\eta h/8$ also follows from the positive measure, because $\mathrm{sigmoid}' \le 1/4$. A tightly propagated, probability-specific range enclosure can already be useful for rare events. Our multiplicative theorem should not be read as proving that every absolute-enclosure implementation necessarily has poor rare-event complexity. It supplies an explicit *a priori* log-likelihood bound without a probability denominator and motivates evaluating both directions positively. The subtraction failures in the numerical audit concern a particular ordinary-floating-point implementation, not an impossibility for other integrators.

W.3 Positive interval bounds and verified objective comparisons

On each affine piece, compute $\eta_{k,-}$ and $\eta_{k,+}$ from endpoint extrema. Substitute $\mathrm{sigmoid}(\eta_{k,-})$ and $\mathrm{sigmoid}(\eta_{k,+})$ into the positive expression for P_{01} to obtain lower and upper bounds. Reverse the affinity extrema for P_{10}; add the exact survivor E for the diagonal bounds. These bounds can be tighter than $e^{\pm\delta}\widehat{P}$.

For each observed transition with probability interval $[l_r, u_r]$, its negative log-likelihood lies in $[-\log u_r, -\log l_r]$. Add these intervals, normalize, and include the exactly evaluated penalty to obtain $[L_-, L_+]$. If $L_+(\vartheta') < L_-(\vartheta)$, the candidate has a smaller true objective in exact arithmetic. Otherwise refine before accepting; failure of the intervals to separate is not evidence of improvement. If a uniformly ϵ-accurate objective is minimized to within ξ of its infimum, its true objective is within $\xi + 2\epsilon$ of the true infimum. This conditional statement neither supplies such an optimizer nor certifies a derivative from a function-value error.

The research code caps the padded subdivision count at 1,024 and raises an explicit exception when that limit is exceeded. It uses float64 positive sums and stable exponential primitives; it does not implement outward-rounded interval arithmetic or a universal overflow proof. The ordinary training metrics use a 10^{-15} probability floor, whereas the full-objective polishing and interval-flag likelihood use direct observed-entry probabilities with a much smaller safety floor. The main synthetic forecast outputs remain far from the metric floor, as recorded in the execution audit. Thus, numerical interval guarantees and finite-precision audits remain distinct.

W.4 Supporting polygonal approximation result

For $D = 2$, $K \ge 2$, and uniformly spaced unsigned directions $a_k = (\cos(k\pi/K), \sin(k\pi/K))$,

$$\cos(\pi/(2K))\|z\|_2 \le \nu_K(z) \le \|z\|_2.$$

If all relative positions have Euclidean norm at most R, the distance error is bounded by $\delta_K = R(1-\cos(\pi/(2K))) \le R\pi^2/(8K^2)$. Fix identical trajectories, intercepts, clocks, and initial laws in the Euclidean and polygonal models. Each log-rate changes by at most δ_K. Writing the perturbed rate as $\widetilde{\lambda} = \lambda e^w$, the jump-process likelihood ratio and compensator identity give

$$\mathrm{KL}(P_2 \| P_K) = \sum_{e,s} \mathbb{E}_{P_2} \int Y_e^s \lambda_e^s (e^{w_e^s} - 1 - w_e^s) dt \le \tfrac{1}{2} e^{\delta_K} \delta_K^2 \mathbb{E}_{P_2} N.$$

Here N is the total transition count. Applying the same observation kernel cannot increase relative entropy. This supporting $O(K^{-4})$ bound holds for fixed parameters, not for the statistical error of fitted models, and is not the central information result.

X Shared-graph recovery and computational-cost protocols

X.1 Zero-nuisance shared-graph recovery

Model and independent experimental units. For the complete graph on six nodes, the unsigned incidence matrix $B \in \mathbb{R}^{15\times 6}$ has full column rank. We use the family (8), with the known functions f, ϕ from (7), and jointly estimate the 18 coordinates (v, α, γ). Independent realization seeds are 8800–8809. For each seed, v_i is uniform on $[-0.15, 0.15]$, α_i is uniform on $[-0.12, 0.12]$, and $\gamma = 0$. Although its true value is zero, γ is never fixed or revealed during fitting. The same heterogeneous truth is used across paired schedules, rates and path counts. The regime multiplier $\lambda \in \{8, 32\}$ is known and not an additional estimated intercept.

Each experimental unit contains $R \in \{10^5, 10^6\}$ independent replicated full graph paths observed at seven times including time zero. Each path has independent dyads conditional on the deterministic shared parameters. The evaluation grid consists of ten realizations, two rates, two sample sizes and four schedules, totaling 160 selected fits. The starting state on each dyad is drawn with probability $\mathrm{sigmoid}\{f(0) + (B\alpha)_e\}$. Its density is not included in the fitted conditional likelihood.

Exact schedules and known jitter. Aligned endpoints are $(0.5, 1.5, 2.5, 3.5, 4.5, 5.5)$; staggered endpoints are $(1, 2, 3, 3.5, 4.5, 5.5)$. Both start at zero and have gap multiset $\{0.5, 1, 1, 1, 1, 1\}$. The jittered schedules add the same seed-specific five independent $U[-0.02, 0.02]$ perturbations to the first five endpoints of each design, keeping the last endpoint fixed. These realized times are supplied to the fitter. The jittered schedules retain the count and total span but are not asserted to have identical gap multisets. Timing is fixed before observing states; neither a true-information criterion nor clock error selects it. Corollary 2 concerns exact schedules at the symmetric parameter. The heterogeneous, jittered tests evaluate finite-sample consequences beyond that pointwise construction, not a uniform extension of its asymptotic exponent.

Data generation and sufficient path counts. Transition probabilities are computed independently of the fitting routine by SciPy Radau integration of the formation and dissolution contributions, with relative tolerance 3×10^{-12} and absolute tolerance 3×10^{-14}. Survival uses a separate DOP853 integration of the total hazard. Given these high-accuracy reference probabilities, sequential binomial sampling generates the transition counts exactly: from m_a paths currently in state a, draw the number ending in state one as $\mathrm{Binomial}(m_a, P_{a1})$, then carry the resulting endpoint counts to the next interval. These sufficient counts follow the independent-Markov-path distribution rather than a Gaussian approximation; individual trajectories are not stored. The probabilities themselves are numerically integrated, not symbolically exact. Count arrays, reference probabilities, truths, estimates and times are all retained.

The sampling seed is the realization seed plus $10^4\lambda + R$. Reusing it couples the random streams across schedules, yielding paired comparisons; it does not make distinct seeds dependent. Initial and intermediate state counts are preserved rather than reset or balanced at each interval. The conditional log likelihood is $\sum_{e,r,a,b} N_{erab} \log P_{erab}$, using all available transitions without negative sampling.

Fitting and optimizer resolution. A separate positive nested Gauss–Legendre rule evaluates the general smooth transition integral and its analytic first derivatives in the dyad coordinates (c, α, γ).

This family is smooth rather than affine within each interval; consequently this fitting quadrature is validated independently and is not assigned the affine-cell error certificate of Theorem 17. Outer/inner orders start at 48/24 and may rise to 96/48 or 192/96. Each selected fit is evaluated at twice its fitting orders; a maximum log-probability change below 10^{-10} and derivative change below 10^{-9} are required, otherwise all three starts are rerun at the next tier. Probability-mass checks also trigger finer integration during fitting.

We minimize signed binomial-deviance residuals with an analytic Jacobian using bounded trust-region reflective least squares. This removes the parameter-independent saturated-entropy constant from the conditional likelihood and preserves the MLE minimizer. It avoids termination driven by a large nearly constant loss in weakly identified clock directions. For very close empirical and fitted probabilities, a fourth-order local KL expansion avoids cancellation; final selected losses are independently reevaluated from positive transition entries. The bounds are $v_i \in [-1.25, 1.25]$, $\alpha_i \in [-0.8, 0.8]$, $\gamma_i \in [-2, 2]$. All selected boundary solutions are retained. These bounds define the estimator rather than prove global identifiability.

Each fit uses the all-zero start and two Gaussian starts, with coordinate standard deviations 0.22 for clocks, 0.06 for offsets, and 0.2 for invisible affinities. Their seed is the realization seed plus 400000; starts do not use the true parameters. The optimizer uses Jacobian-based coordinate scaling, $\texttt{ftol} = 10^{-11}$, $\texttt{xtol} = 10^{-10}$, $\texttt{gtol} = 10^{-10}$ and at most 1600 function evaluations. Selection is by lowest training deviance, never by recovery error. All starts, precision-tier retries, termination messages and selected objectives are saved. Across the 160 final selected fits, all solvers report successful termination, 23 require a finer quadrature tier, and the maximum selected evaluation count is 208. This is reproducible finite-budget optimization evidence, not proof that the global likelihood maximum was found.

Metrics, uncertainty and local diagnostics. The principal error is the dyad log-clock field RMSE,

$$\left\{ |E|^{-1} \sum_{e} [(B\widehat{v})_e - (Bv^*)_e]^2 \right\}^{1/2}.$$

We also retain node-clock error, affinity-offset error, a balanced-start transition KL diagnostic, and the nuisance-profiled conditional information. The latter propagates actual expected starting-state exposures through each path, while conditioning on initial states in the likelihood. The reported local field scale is $[\mathrm{tr}\{BI_{\mathrm{eff}}^{-1}B^\top\}/(|E|R)]^{1/2}$. It is not a confidence interval or a finite-sample MSE identity for the bounded estimator. Information is evaluated at the truth only as an explanatory diagnostic; it does not choose schedules, starts, estimators or checkpoints.

Means and sample standard deviations use ten independent parameter/noise realizations. Paired differences match seeds, and their 95% intervals use $t_{9,0.975}$ times the standard error of the paired difference. They are descriptive, without multiplicity adjustment. Table 77 includes the complete grid, including the low-count regimes where paired intervals cross zero. For $R = 10^6, \lambda = 32$, mean local field scales are 0.914 (aligned) and 0.195 (staggered), versus observed mean errors 0.815 and 0.209. The aligned estimate is affected by seven boundary hits; a normal-approximation interpretation would be unwarranted. With jitter the corresponding local scales are 0.478 and 0.202, and mean errors are 0.394 and 0.202. At $R = 10^5, \lambda = 32$, both designs show large errors and frequent boundary solutions. These outcomes delimit the operational benefit rather than converting a local information theorem into an unconditional recovery claim. The zero-nuisance MLE grid is displayed with its homogeneous and ridge controls in Table 77.

Numerical cross-checks. The maximum log-probability difference between the truth quadrature and independent Radau generator across all 160 fits is 5.27×10^{-13}. The maximum selected-fit quadrature-doubling change is 3.76×10^{-11}; the largest derivative change is 5.50×10^{-10}. The graph derivative unit test compares analytic likelihood derivatives to central differences in all 18 coordinates, with maximum absolute discrepancy 7.23×10^{-10}. A separate replay script checks all saved path counts, selected starts, likelihood values and fitted probabilities, and compares eight predeclared fitted parameter cases against independent ODE integration. Numerical checks do not replace the algebraic rank and matching-rate proofs.

Table 74: CPU accuracy–cost comparison on 18 fixed problems at requested log-error 10^{-5}. Every measured error meets the target. Times are medians across problems of seven warmed repetitions, in milliseconds. Storage is live saved-tensor MiB at forward completion, not process peak memory. Radau's derivative column times an augmented forward-sensitivity solve, including the state solution.

Method	Max log error	Forward	Derivative	Saved MiB
Positive first-order	1.3×10^{-7}	7.12	1.56	6.073
Positive combined	3.3×10^{-6}	0.79	0.60	0.130
Exponential midpoint	7.2×10^{-7}	1.11	0.29	0.064
Radau (sensitivities)	6.1×10^{-8}	7.47	19.06	—

X.2 Computational cost and independent derivative accuracy

Fixed problem set and implementations. The cost grid takes every fourth case from the 48-case random affinity/clock grid (12 cases), all four rare-affinity cases at $\eta = \pm30, \pm60$, and both stiff-clock cases. Selection is deterministic and does not use observed solver timings. We test requested log-error targets 10^{-4} and 10^{-5}. Positive first-order and combined rules use the same PyTorch float64 implementation, exact affine hazard integration, direct positive entries and their respective mesh rules. The exponential midpoint comparator freezes both total rate and equilibrium within each cell, doubling the mesh until successive maximum entrywise log changes are below one quarter of the requested target. This is an empirical stopping rule, not a rigorous error certificate. Radau uses normalized off-diagonal ODE variables, relative tolerance equal to the requested target divided by 20, and absolute tolerance one hundredth of that relative tolerance. Its achieved log error is checked after solving, not assumed from its local ODE tolerance.

Independent reference and gradients. The reference solves the normalized state equations and all four affine-endpoint parameter sensitivities jointly, giving ten augmented ODE coordinates, with a tighter relative tolerance 2×10^{-11}. The independently integrated hazard supplies survival contributions. Normalization avoids letting an absolute ODE tolerance dominate the rare probabilities. The scalar derivative objective is $-\frac{1}{4}\sum_{a,b}\log P_{ab}$. For a gradient coordinate the scaled discrepancy is $|\widehat{g}_i - g_i^{\mathrm{ref}}|/(1 + |g_i^{\mathrm{ref}}|)$; the reported error is its maximum over coordinates and problems. This compares with an independently solved sensitivity equation, unlike finite differences of the same discretized objective.

At target 10^{-5}, the maximum scaled gradient discrepancies are 1.06×10^{-7} (positive first-order), 5.82×10^{-6} (combined), 7.16×10^{-7} (midpoint), and 4.00×10^{-9} (Radau sensitivities). At target 10^{-4} they are 1.05×10^{-5}, 5.79×10^{-5}, 1.11×10^{-5} and 1.16×10^{-7}, respectively. All 144 method/problem/tolerance combinations meet their measured log-error targets. These observations do not imply a gradient error theorem or uniform behavior away from the fixed grid.

Timing and memory. Measurements use one float64 CPU thread, pinned processor affinity, two warm-up evaluations and seven measured repetitions per problem. The table reports a median within each problem and then the median across the 18 problems; raw measurements and the processor/library environment are supplied. Forward timing includes automatic mesh selection. For the differentiable positive and midpoint methods, derivative timing measures reverse propagation through the selected fixed mesh. Radau's derivative timing measures an augmented forward-sensitivity solve, including the states; it is reported explicitly as a different derivative strategy. Thus its derivative column is not a matched reverse-pass speed measurement.

Saved reverse-mode storage counts unique live tensor storage retained at forward completion through autograd hooks. It is not process peak resident memory, allocation traffic or GPU memory. At 10^{-5} the combined rule reduces this metric from 6.073 to 0.130 MiB relative to the first-order rule, while midpoint uses 0.064 MiB. The combined rule's observed forward time is 0.79 ms versus 7.12 ms for first-order, 1.11 ms for midpoint and 7.47 ms for Radau. These timings characterize the tested scalar CPU workload; accelerator scaling requires separate measurement. At 10^{-4} first-order and combined forward medians are much closer (0.87 and 0.76 ms), demonstrating tolerance-dependent rather than uniform savings.

X.3 Feature-matched forecasting comparisons

Table 75: Matched-budget additive-clock comparison: five paired graphs, 2,400-update cap, batch 256, identical learning-rate grid and validation rule. Summary-augmented external encoders receive all 37 causal summary coordinates. The median selected update and cap count diagnose finite-budget behavior, not convergence.

Method / input	NLL	Selected step	Reached cap
GraphMixer (ICLR'23), positive only	0.4203 ± 0.0171	1600	3/5
GraphMixer (ICLR'23) + summaries	0.3988 ± 0.0147	1800	5/5
DyGFormer (NeurIPS'23), positive only	0.4028 ± 0.0131	2000	5/5
DyGFormer (NeurIPS'23) + summaries	0.4013 ± 0.0164	2200	5/5
TempoGeom-H	0.3976 ± 0.0150	1800	5/5

Table 75 complements the native-input external comparison in Section AB. Both external encoders receive the same 37 causal occupancy/absence/transition summary coordinates as TempoGeom-H. Their positive-only counterparts are trained at the same longer budget, so feature changes are not conflated with giving only the augmented models extra training. Learning-rate grids, batch size 256, 2,400-update caps, checkpoint selection and paired graph realizations are matched. Table 96 gives all paired intervals; validation traces and cap counts are retained. The augmented-minus-positive NLL difference is -0.02151 $[-0.02794, -0.01509]$ for GraphMixer and -0.00142 $[-0.00597, 0.00312]$ for DyGFormer. Several runs reach the cap, so these comparisons characterize finite-budget performance on the additive-clock condition rather than converged optimization.

X.4 Additional scalar and replay diagnostics

The scalar constant-total-rate phase comparison uses integer endpoints $(1, 2, 3, 4, 5, 6)$ and staggered endpoints $(0.5, 1.5, 2.5, 4, 5, 6)$. It is distinct from the equal-gap-multiset graph construction used in the main text. Endpoint-only projected coefficients are 0.06250 and 0.07997; full nuisance adjustment changes them to zero and 0.07448. The aligned scalar's quartic coefficient is $C_4 \approx 9.2243686017$. The independent finite-clock quadrature approaches this value; the coefficient is computed from the residual formula, not estimated by fitting a power law. For the affinity-dependent graph family, large-clock slopes over the retained grid are -3.99951 and -1.99995. Positive finite-clock values and endpoint-mass residuals are saved.

The all-orders-bump audit has five finite-clock checks agreeing with the analytic profile to relative error 3.2×10^{-13}. The saved primary replay covers all 180 selected checkpoints with maximum probability discrepancy 1.8×10^{-7}; the additional replay covers 105 selected checkpoints and 345 prediction blocks. These are checkpoint re-evaluations, not new training runs or official external-pipeline reproductions. The saved-record replay recomputed the primary means and control intervals from saved per-record log probabilities, separately from its 160 joint graph fits. The observation-design study is documented in Appendix Z. Discretization bounds, numerical discrepancies, paired sampling uncertainty and optimizer termination are recorded as different quantities.

Y Nonzero nuisance, homogeneous controls, and estimator sensitivity

This study distinguishes relative schedule improvement from recovery beyond a supplied nominal clock. It complements the 160 MLE fits in Appendix X.1 with 480 nonzero-nuisance and estimator-control fits, retaining every boundary solution and grid cell. The full grid is recorded in `robustness_study/protocol.json`; the grid and selection rules were fixed before these runs, but were not independently preregistered as a confirmatory study.

Y.1 Controls and nonzero-nuisance datasets

The homogeneous estimate is $\widehat{v} = 0$ with the same known λ as every fitted estimator. It neither uses the realized true vector nor estimates nuisance parameters. Its dyad log-clock-field RMSE is

Table 76: Clock-field RMSE at $R = 10^6$, exact schedules, ten paired realizations per row. Entries are mean $\pm$ SD. The homogeneous estimate has RMSE 0.133 ± 0.029 in every row. Ridge fixes $\tau = 0.3$ before fitting; all affinities remain unpenalized. A/S denote aligned/staggered. Paired intervals and every lower-count/jitter case are in Appendix Y.

λ	a_γ	Homog.	MLE A	MLE S	Ridge A	Ridge S
8	0	0.133 ± 0.029	0.073 ± 0.022	0.051 ± 0.011	0.071 ± 0.020	0.050 ± 0.011
8	0.05	0.133 ± 0.029	0.080 ± 0.024	0.050 ± 0.011	0.078 ± 0.022	0.049 ± 0.011
8	0.2	0.133 ± 0.029	0.074 ± 0.025	0.055 ± 0.009	0.071 ± 0.023	0.054 ± 0.009
32	0	0.133 ± 0.029	0.815 ± 0.270	0.209 ± 0.095	0.251 ± 0.085	0.157 ± 0.056
32	0.05	0.133 ± 0.029	0.833 ± 0.362	0.201 ± 0.094	0.265 ± 0.066	0.152 ± 0.053
32	0.2	0.133 ± 0.029	0.662 ± 0.315	0.201 ± 0.091	0.256 ± 0.077	0.151 ± 0.053

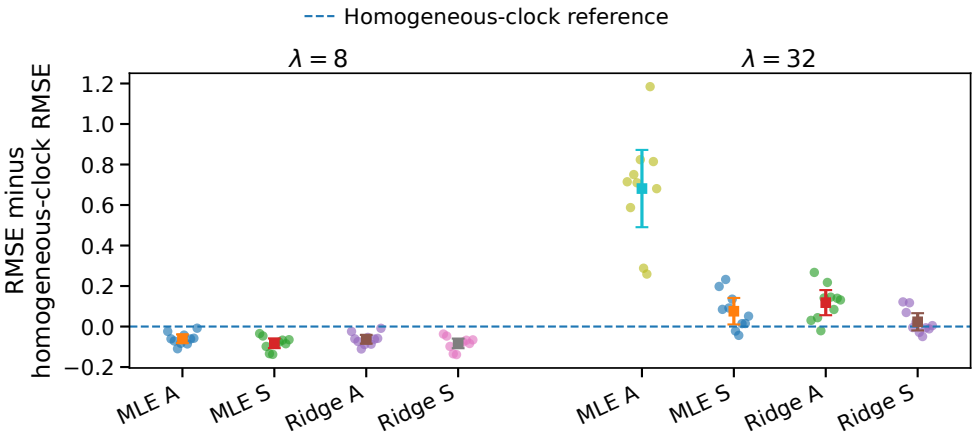

Figure 11: Every paired high-count recovery outcome relative to the homogeneous reference, including fixed-ridge and bounded-MLE controls.

computed against the saved truth. Across the ten seeds it is 0.133476 ± 0.028801, unchanged across rates, path counts, and nuisance amplitudes because the paired clock truths are unchanged. This is an intentionally strong simple reference for a narrow, centered clock-generating distribution; it is not asserted to be competitive for arbitrary clock populations.

Ridge minimizes the *total*, not average, conditional negative log likelihood plus $\|v\|^2/(2\tau^2)$. The default $\tau = 0.3$ is fixed before fitting and leaves both affinity vectors unpenalized. In the per-replicated-path signed-deviance residual, this appends $v/(\tau\sqrt{R})$, so penalty strength does not inadvertently grow with R. Selection among the same three starts minimizes the corresponding penalized objective, never clock-field error. The affinity bounds and trust-region reflective settings are otherwise identical to the MLE. No information computed at the truth selects a penalty or changes the fit.

Each of the 160 zero-nuisance datasets also receives a default-ridge fit. Nonzero-nuisance datasets retain the same v, α draws at seeds 8800–8809, but use

$$\gamma_i = a_\gamma z_i, \qquad z_i \overset{\text{iid}}{\sim} U[-1, 1], \qquad a_\gamma \in \{0.05, 0.2\},$$

with a separate generator seed equal to the realization seed plus 500000. These coordinates are not centered or disclosed to the fitter. For each amplitude, both $\lambda \in \{8, 32\}$, both exact schedules, and $R = 10^6$ give 80 nonzero-nuisance datasets. Each receives an MLE and a default-ridge fit, adding 160 fits. The amplitudes share their nuisance direction within each seed; they are sensitivity conditions, not independent replications.

Transition probabilities use independent Radau integration and sequential binomial sufficient counts, with the specified sampling seed $s + 10^4\lambda + R$. Thus paired designs and amplitudes share a random stream, whereas distinct realization seeds remain independent. Starting-state counts are propagated

through the path rather than reset at each interval. Exactly $90R$ dyad-transition records are represented by each complete dataset: 9 million at $R = 10^5$ and 90 million at $R = 10^6$. Compressed sufficient-count generation does not turn this exposure into a six-record single-graph experiment. All estimates condition on initial states and exclude their density from the likelihood.

Table 77: Complete zero-nuisance recovery grid with a homogeneous-clock reference and fixed ridge scale $\tau = 0.3$. Each group has ten paired realizations. All 160 MLE fits and their 160 same-data ridge fits are included.

R	λ	Timing	MLE A	MLE S	Ridge A	Ridge S
10^5	8	exact	0.408 ± 0.273	0.175 ± 0.050	0.157 ± 0.042	0.148 ± 0.041
10^5	8	jitter	0.291 ± 0.205	0.173 ± 0.049	0.152 ± 0.043	0.152 ± 0.046
10^5	32	exact	0.981 ± 0.234	0.782 ± 0.351	0.173 ± 0.061	0.256 ± 0.060
10^5	32	jitter	0.924 ± 0.433	0.845 ± 0.290	0.179 ± 0.078	0.276 ± 0.094
10^6	8	exact	0.073 ± 0.022	0.051 ± 0.011	0.071 ± 0.020	0.050 ± 0.011
10^6	8	jitter	0.078 ± 0.022	0.053 ± 0.012	0.076 ± 0.021	0.053 ± 0.012
10^6	32	exact	0.815 ± 0.270	0.209 ± 0.095	0.251 ± 0.085	0.157 ± 0.056
10^6	32	jitter	0.394 ± 0.192	0.202 ± 0.085	0.227 ± 0.074	0.148 ± 0.047

Table 78: Paired excess RMSE over the homogeneous-clock estimate at $R = 10^6$, exact schedules. Negative favors estimation. Intervals are descriptive 95% paired t intervals over ten seeds.

λ	a_γ	Estimator	Aligned excess [interval]	Staggered excess [interval]
8	0	MLE	-0.0604 $[-0.0818, -0.0389]$	-0.0826 $[-0.1065, -0.0587]$
8	0	ridge	-0.0628 $[-0.0845, -0.0412]$	-0.0832 $[-0.1070, -0.0594]$
8	0.05	MLE	-0.0532 $[-0.0743, -0.0321]$	-0.0835 $[-0.1070, -0.0599]$
8	0.05	ridge	-0.0555 $[-0.0769, -0.0341]$	-0.0840 $[-0.1075, -0.0606]$
8	0.2	MLE	-0.0599 $[-0.0828, -0.0370]$	-0.0786 $[-0.0996, -0.0577]$
8	0.2	ridge	-0.0626 $[-0.0851, -0.0400]$	-0.0793 $[-0.1002, -0.0584]$
32	0	MLE	$+0.6813$ $[+0.4905, +0.8720]$	$+0.0757$ $[+0.0103, +0.1411]$
32	0	ridge	$+0.1180$ $[+0.0559, +0.1800]$	$+0.0234$ $[-0.0191, +0.0659]$
32	0.05	MLE	$+0.6998$ $[+0.4515, +0.9480]$	$+0.0679$ $[+0.0085, +0.1273]$
32	0.05	ridge	$+0.1315$ $[+0.0818, +0.1812]$	$+0.0186$ $[-0.0163, +0.0535]$
32	0.2	MLE	$+0.5281$ $[+0.3151, +0.7412]$	$+0.0674$ $[+0.0045, +0.1303]$
32	0.2	ridge	$+0.1221$ $[+0.0615, +0.1828]$	$+0.0180$ $[-0.0237, +0.0597]$

Interpretation across the complete grid. At $R = 10^6, \lambda = 8$, all six exact-schedule/amplitude groups of both estimators have negative paired excess over the homogeneous reference; these are regimes where heterogeneity estimation is useful. At $\lambda = 32$, the staggered MLE is better than aligned MLE across the tested amplitudes but worse in mean than the homogeneous estimate. Default ridge reduces high-rate risk but does not establish improvement over the reference. In the high-rate jittered condition, its staggered mean is 0.148, versus the homogeneous 0.133. The lower-count outcomes remain in Table 77.

At $R = 10^5, \lambda = 32$, default ridge changes the exact-schedule MLE means of 0.981/0.782 to 0.173/0.256. The corresponding staggered-minus-aligned ridge interval is $[0.0179, 0.1471]$. With known jitter the ridge means are 0.179/0.276 and the interval is $[0.0159, 0.1788]$. In both cases the schedule ranking reverses. The bounded estimator's finite-sample risk depends on shrinkage bias, its fit to nuisance variation and nonlinear optimization, whereas the information theorem describes local likelihood sensitivity.

Y.2 Bound and penalty sensitivity

At $\lambda = 32, R = 10^6, \gamma = 0$, both exact schedules and all ten seeds are refitted after multiplying every baseline half-width $(1.25, 0.8, 2)$ for (v, α, γ) by 0.75 or 1.5. Both MLE and default ridge are tested, giving 80 additional fits. These ranges are not selected against recovery error. Bound hits mean any coordinate lies within 10^{-3} times the bound multiplier of its bound. All hits are retained.

The aligned MLE's mean increases from 0.718 through 0.815 to 0.961 as the box widens; bound hits are 8/10, 7/10, and 6/10. Staggered MLE is stable to displayed precision with no hits. Default ridge has no bound hits in these three boxes and retains means 0.251/0.157. This sensitivity establishes that a substantial component of the aligned high-rate MLE error is estimator-dependent; it should not be interpreted as a direct measurement of an inverse-Fisher risk law.

Two additional fixed scales, $\tau = 0.15$ and 0.6, are applied at $\lambda = 32$, $R = 10^6$, both exact schedules, and $a_\gamma \in \{0, 0.2\}$, giving 80 additional fits. Together with the default this is a reported sensitivity grid, not an error-tuned choice of the best penalty. At $a_\gamma = 0$, stronger shrinkage reduces staggered mean RMSE to 0.109, but its paired excess over the homogeneous estimate is -0.0248 with interval $[-0.0552, 0.0056]$. The interval includes zero; this exploratory row is not used to select the penalty or establish a recovery advantage.

Table 79: Estimator sensitivity at $R = 10^6$, $\lambda = 32$. Bound multipliers scale all three baseline parameter boxes. The homogeneous RMSE is 0.133 ± 0.029. No row is selected using recovery error.

a_γ	Bound factor	Estimator	Aligned RMSE	Staggered RMSE	Bound hits A/S
0	0.75	MLE	0.718 ± 0.201	0.209 ± 0.095	8/0
0	0.75	ridge $\tau = 0.3$	0.251 ± 0.085	0.157 ± 0.056	0/0
0	1	MLE	0.815 ± 0.270	0.209 ± 0.095	7/0
0	1	ridge $\tau = 0.3$	0.251 ± 0.085	0.157 ± 0.056	0/0
0	1.5	MLE	0.961 ± 0.367	0.209 ± 0.095	6/0
0	1.5	ridge $\tau = 0.3$	0.251 ± 0.085	0.157 ± 0.056	0/0
0	1	ridge $\tau = 0.15$	0.124 ± 0.045	0.109 ± 0.033	0/0
0	1	ridge $\tau = 0.3$	0.251 ± 0.085	0.157 ± 0.056	0/0
0	1	ridge $\tau = 0.6$	0.433 ± 0.114	0.191 ± 0.080	0/0
0.2	1	ridge $\tau = 0.15$	0.128 ± 0.044	0.106 ± 0.032	0/0
0.2	1	ridge $\tau = 0.3$	0.256 ± 0.077	0.151 ± 0.053	0/0
0.2	1	ridge $\tau = 0.6$	0.415 ± 0.105	0.183 ± 0.076	0/0

Y.3 Nonzero-nuisance spectra and the observation mechanism

Figure 12 shows the actual affinity basis and sampling endpoints. Every shown endpoint annihilates ϕ, but its derivative has phase-dependent sign. This is why endpoint visibility alone does not determine the nuisance-adjusted clock information.

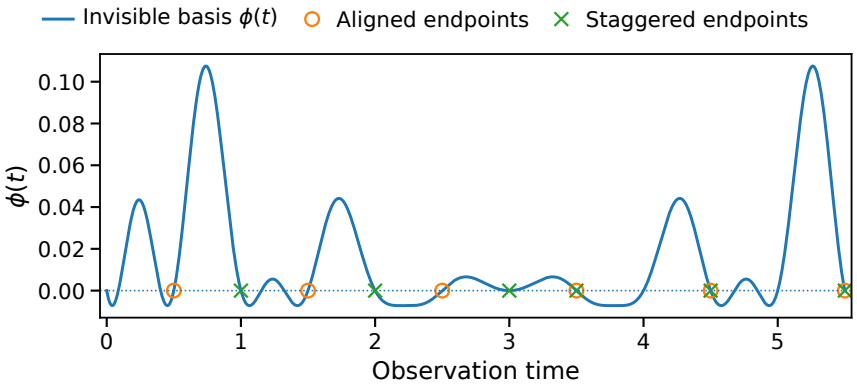

Figure 12: The exact invisible-affinity basis $\phi(t)$ and both equal-gap-multiset schedules. All marked values are zero; the lag derivative depends on $\phi'(t) = f'(t)\cos(2\pi t)$, whose signs differ across staggered endpoints. Shared markers denote shared observation times.

The spectral study fixes an independent direction z from seed 550013, tests amplitudes $0, 0.001, 0.01, 0.05, 0.2$, and evaluates $\lambda \in \{32, 128, 512, 2048, 8192\}$ under both schedules. It

checks the explicit first-order matrix separately from finite-clock information. All nonzero tested amplitudes have leading aligned rank five on K_6; the zero amplitude has rank zero. A rational-arithmetic witness and 30 additional random nonzero directions check the rank calculation. A further 30 independent cases vary all three vectors (v, α, γ) and compare the weighted formula with direct endpoint-score projection; their maximum matrix discrepancy is 4.34×10^{-18}, with leading rank five throughout. Three of these cases also receive finite-clock spectral checks.

The finite-clock calculation uses independent positive endpoint-layer quadrature, not the fitting grid or a fitted power law. Its nodes are fixed while differentiating the original integrand. A layer extending $100/(\lambda q_{0,\min})$ before the endpoint, truncated to the actual interval, resolves the endpoint contribution; survival uses the full integrated hazard. This is a numerical calculation, not an outward-rounded tail certificate. The maximum transition-mass residual across both spectral grids is below 9.55×10^{-12}. At $\lambda = 32$, four nonzero-nuisance cases agree with the separate full-interval fitting quadrature to maximum log error 1.45×10^{-13} and derivative error 9.35×10^{-14}.

For amplitudes 0.05 and 0.2, the fitted minimum-eigenvalue slopes over $\lambda \geq 512$ are approximately -4.000 for aligned sampling and -2.000 for staggered sampling. The proof, not these slopes, establishes the rates. The five other aligned eigenvalues are quadratic at fixed generic nonzero nuisance. At very small nuisance amplitude the crossover is late: the first-order coefficient is quadratic in amplitude, so finite-clock spectra need not already display their fixed-amplitude asymptotic regime. This explicitly separates limits in λ and in nuisance amplitude.

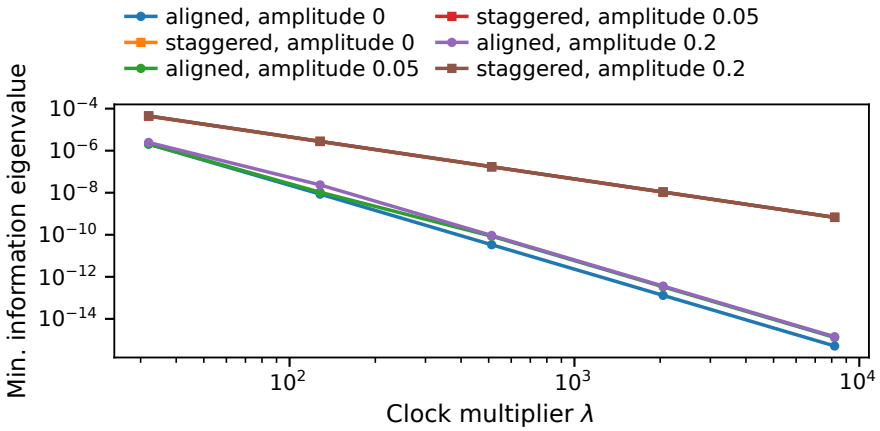

Figure 13: Minimum conditional information eigenvalue on K_6 at symmetric $v = \alpha = 0$ and nonzero invisible-affinity amplitudes. The aligned bottleneck remains quartic locally even when the leading information matrix has rank five. Staggered information remains quadratic.

Y.4 Executed validation and reproducibility

All 480 selected robustness fits terminate successfully under the supplied tolerances. Each uses three truth-independent starts and records every final start and precision-tier retry. The outer quadrature orders are 48, 96, 192, or 384, with inner order half the outer order; probability-mass checks may temporarily use a finer grid. Selected fits must pass doubled-order log-probability and derivative thresholds of 10^{-10} and 10^{-9}. Across these fits the largest changes are 7.17×10^{-11} and 9.87×10^{-10}, respectively. All data-generating probabilities, count arrays, true and fitted parameters, bounds, penalty values, objective values, and convergence messages are retained.

An independent replay reproduces all 480 saved fitted-probability arrays, log likelihoods, field errors, and selected starts in this environment. It also reproduces every one of the 240 datasets' sequential count arrays from the saved independent reference probabilities. Sixteen fitted cases are selected for independent ODE checks by fixed grid identifiers: seed 8800, both exact schedules, $\lambda = 32$, both nonzero amplitudes and both default estimators; both altered bounds for zero-nuisance MLE; and both additional ridge scales at zero nuisance. Their maximum log discrepancy is 2.19×10^{-12}. The

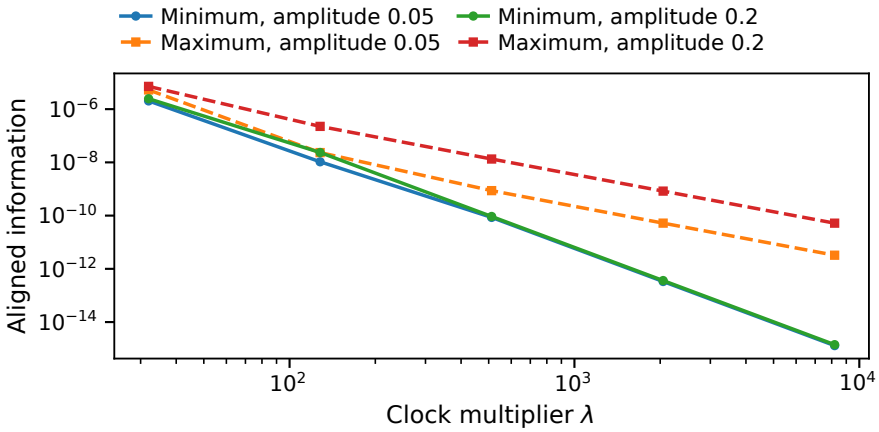

Figure 14: Minimum and maximum aligned information eigenvalues at two nonzero nuisance amplitudes. The different asymptotic slopes illustrate the mixed spectrum rather than an all-direction quartic law.

penalized objective's gradient is additionally checked against central differences at an independent nonzero-nuisance point, with maximum scaled error 7.17×10^{-10}.

The machine-readable summary retains all paired schedule differences and all excess-homogeneous differences, not only those printed in the main text. Every interval uses the ten independent seeds and is a descriptive, unadjusted paired t interval; no dyad, path count, reused amplitude, or fitted estimator is counted as an additional independent experimental unit. Forecasting checkpoints and solver measurements are reused without neural retraining or solver retiming in this study. These numerical checks assess computational consistency, not global optimizer convergence or formally verified arithmetic.

Z TOPOLOGY, TIMING, AND EXPOSURE EXPERIMENTS

The `design_study` directory contains the observation-design experiments. Scripts write raw rows, individual fits, all optimizer starts, sequential count arrays, protocol files, derived tables, and numerical audit logs. The 320 exposure-grid fits and 40 calibrated-schedule fits are distinct from the 640 recovery/robustness fits. Reusing a result or replaying a saved estimate is not counted as an optimization run. All ten parameter realizations are paired across relevant schedules and exposures; repeated conditions do not create additional independent experimental units.

Z.1 GRAPH TOPOLOGY CERTIFICATES AND FINITE SPECTRA

The edge lists in Appendix E are generated by adding the lexicographically first missing edge that maximizes the exact rank increase of $[B, D_g B]$. This construction uses an algebraic witness, not outcomes from recovery. The integer witness $\gamma = (0, 1, 2, 3, 4, 5)$ is checked by exact rational elimination with SymPy. For $m = 6, \ldots, 15$ the residual ranks are $0, 1, 2, 3, 4, 5, 5, 5, 5, 5$. Full edge lists are stored in `topology_rank_certificates.json`.

Finite spectra use four selected observation graphs with $m = 6, 8, 11, 15$, nominal clocks $\lambda \in \{128, 512, 2048, 8192\}$, fixed independently seeded $v_i, \alpha_i \sim U[-0.04, 0.04]$, and $\gamma = 0.08(-1, -0.6, -0.2, 0.2, 0.6, 1)$. Six transition records are observed on each dyad. A separate endpoint-layer quadrature computes transition probabilities and parameter derivatives, including expected starting-state exposures. The leading-kernel derivative combination is canceled before numerical projection, avoiding unnecessary subtraction of nearly identical information blocks. Singular values of the residual derivative matrix, rather than a subtracted Fisher Schur complement, give the eigenvalues.

Figures 15–18 retain every computed eigenvalue. The minimally sufficient G_{11} graph has a small quadratic coefficient: at the largest tested clocks its two smallest sorted eigenvalues still exchange asymptotic roles. Consequently, their finite-grid slopes need not be -4 and -2 yet. Exact rank and the matching-rate proof establish the limiting classification, not numerical rounding or selecting a slope window. The complete graph and odd-unicyclic examples reach their limiting patterns earlier on this grid. Constants and crossover rates can depend substantially on the graph and nuisance amplitude.

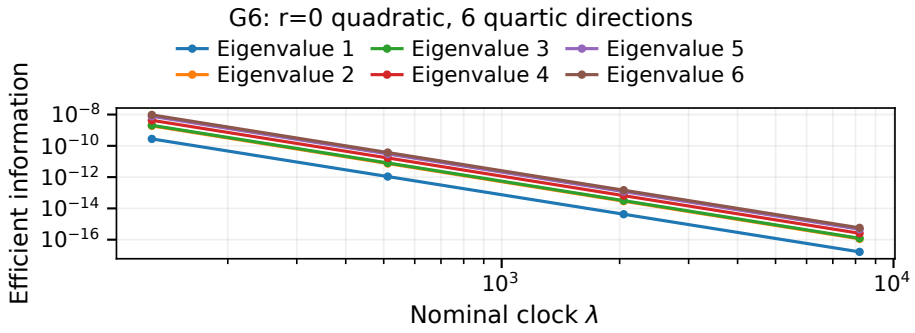

Figure 15: Odd-unicyclic observation graph: all six clock-information eigenvalues are quartic near the origin.

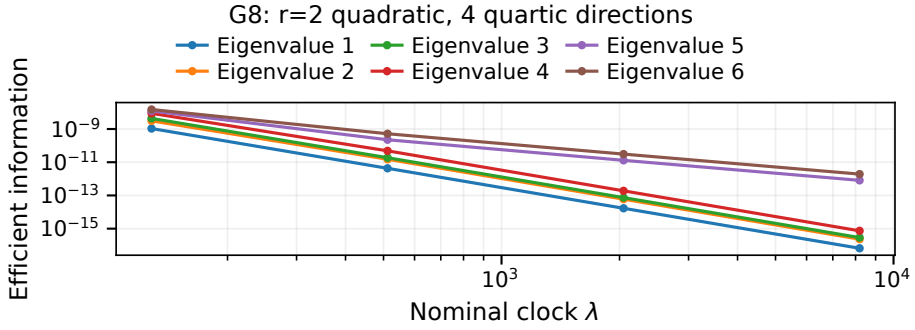

Figure 16: Eight observed dyads: two leading quadratic directions and four quartic directions.

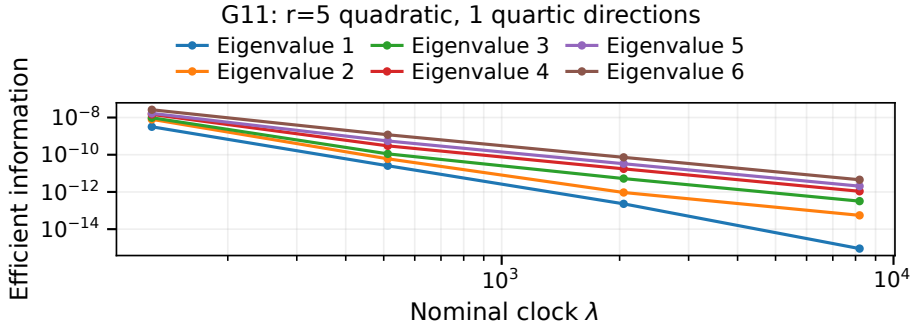

Figure 17: Minimal six-node budget attaining generic rank five. Finite eigenvalue crossover remains visible on this grid; sorted finite eigenvalues are not individually labeled by an asymptotic order.

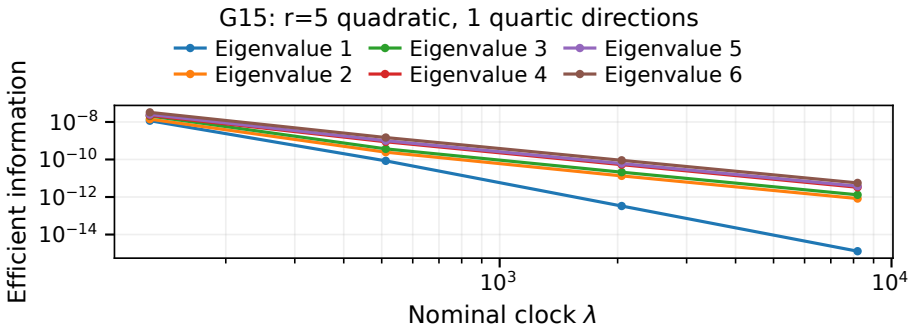

Figure 18: Complete six-node observation graph: one quartic bottleneck and five quadratic directions at the fixed nonzero nuisance.

Table 80: Regular endpoint coefficients at the symmetric parameter for the ten known-jitter seeds. The parameter draws used for recovery need not equal this symmetric parameter; their finite-clock risk is reported separately.

Seed	$\kappa(T)$
8800	0.01530388
8801	0.01397902
8802	0.01344373
8803	0.00732267
8804	0.01131929
8805	0.01086827
8806	0.01071067
8807	0.00677653
8808	0.01110912
8809	0.00475925

Z.2 Timing intervention and coefficient checks

The fixed-jitter calculation uses the ten schedule seeds 8800–8809, perturbing the first five observation endpoints independently by $U[-0.02, 0.02]$ with random seed equal to the realization seed plus 300000. The initial time remains zero and the last endpoint remains 5.5. All timestamps are known to the likelihood. The symmetric-parameter coefficients in Table 80 explain why these schedules have regular endpoint-visible information. Recovery in Appendix Y uses heterogeneous true v, α and therefore is not a direct measurement of these symmetric coefficients. For example, at $\lambda = 32, R = 10^6$, known jitter lowers aligned MLE's mean error from 0.815 to 0.394, without making it beat the homogeneous reference. A changed asymptotic order does not force useful risk at that finite exposure.

The independent finite-clock timing grid is $\lambda \in \{8, 32, 128, 512, 2048, 8192\}$. It includes exact alignment, a fixed first-endpoint shift 0.02, fixed random jitter (seed 8800), shifts ρ/λ with $\rho \in \{0.1, 1, \rho_*, 10\}$, and a shift $1/\lambda^2$. The $\rho = 10, \lambda = 8$ cell would reorder endpoints and is excluded explicitly, leaving 47 valid cases. Times include an unchanged initial zero; a first-endpoint shift changes the second entry of that seven-time array. Every shifted design keeps six transitions and horizon 5.5, but generally changes the gap multiset. All these timing-information calculations use the symmetric parameter and K_6.

The closed-form and projected coefficient calculations agree to floating-point accuracy across 141 ratios. The analytically chosen ratio is $\rho_* \approx 1.3936795893$ and attains $\kappa \approx 0.0156458653$ (Figure 19). The independent probability/sensitivity comparison uses three designs at $\lambda = 8, 32$, with maximum log error 1.45×10^{-13} and derivative discrepancy 9.27×10^{-14}. Across the high-clock information grid, maximum transition-mass discrepancy is 7.30×10^{-12}. These checks do not supply interval-arithmetic enclosures for the smallest eigenvalues.

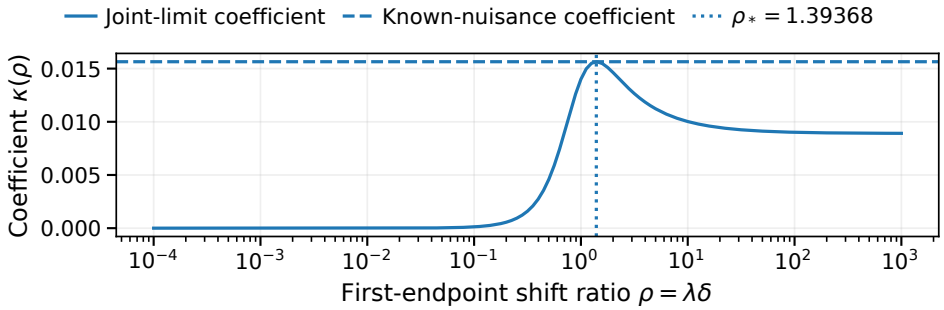

Figure 19: Analytic joint-limit coefficient for a one-endpoint shift. The calibrated ratio attains the known-nuisance leading coefficient at the symmetric model; it is not a fitted optimum or a guarantee at heterogeneous truths.

Table 81: Complete six-exposure grid. Clock-field RMSE, mean ± sample SD across ten realizations. True invisible affinity is zero but is fitted. Both affinity blocks are unpenalized; ridge fixes $\tau = 0.3$. The homogeneous mean is 0.133 in every row.

λ	$R/10^6$	MLE A	MLE S	Ridge A	Ridge S
8	0.1	0.408 ± 0.273	0.175 ± 0.050	0.157 ± 0.042	0.148 ± 0.041
8	0.3	0.192 ± 0.183	0.080 ± 0.033	0.119 ± 0.035	0.077 ± 0.030
8	1	0.073 ± 0.022	0.051 ± 0.011	0.071 ± 0.020	0.050 ± 0.011
8	3	0.042 ± 0.015	0.029 ± 0.010	0.041 ± 0.015	0.029 ± 0.010
8	10	0.027 ± 0.013	0.016 ± 0.007	0.027 ± 0.013	0.016 ± 0.006
8	30	0.013 ± 0.004	0.008 ± 0.002	0.013 ± 0.004	0.008 ± 0.002
32	0.1	0.981 ± 0.234	0.782 ± 0.351	0.173 ± 0.061	0.256 ± 0.060
32	0.3	0.838 ± 0.197	0.375 ± 0.191	0.264 ± 0.109	0.209 ± 0.077
32	1	0.815 ± 0.270	0.209 ± 0.095	0.251 ± 0.085	0.157 ± 0.056
32	3	0.887 ± 0.198	0.114 ± 0.040	0.244 ± 0.049	0.105 ± 0.037
32	10	0.455 ± 0.235	0.063 ± 0.026	0.208 ± 0.067	0.061 ± 0.024
32	30	0.244 ± 0.121	0.036 ± 0.011	0.171 ± 0.066	0.036 ± 0.011

Z.3 Six-exposure recovery grid

The protocol file fixes the six exposure values, both clock rates, both exact schedules, all ten shared parameter seeds, bounds, three starts, and the MLE/fixed-ridge estimators. At $R = 10^5, 10^6$, the estimates and counts are reused. At the four additional exposures, independent sequential counts are generated from the saved Radau probabilities, using noise seed $s + 10000\lambda + R$. The same seed formula pairs schedules within a realization. No count array is treated as extra independent graph-level replication. There are 160 additional count datasets and two selected estimates per dataset, yielding 320 selected fits. Each selection retains all three candidate parameters, objective values, termination flags, and precision refinements.

The bounded MLE and fixed ridge use the unchanged `fit_general` implementation, parameter-box half-widths $(1.25, 0.8, 2)$ for (v, α, γ), and ridge scale $\tau = 0.3$. Both affinity blocks remain unpenalized. Signed-deviance least squares avoids catastrophic objective subtraction, and quadrature doubling controls fitted probability accuracy. Every grid cell and boundary solution is retained. The six-exposure report has 480 method/realization rows, not 480 independent samples. Different exposures use different count draws; a nonmonotone empirical mean, especially for constrained weakly identified MLE, is not removed or smoothed away.

Table 82 compares empirical root-mean-square field error with the local-unbiased information scale. These are different quantities: local SD is evaluated at the true parameter and is not an estimated confidence interval. Median R_{loc} at $\lambda = 8$ is 3.07×10^5 (aligned) versus 1.04×10^5 (staggered); at $\lambda = 32$ it is 4.48×10^7 versus 1.95×10^6. The high-clock staggered range across truths is 1.52×10^6–4.36×10^6, compared with 2.72×10^7–9.81×10^7 for alignment. This is an explanatory

Table 82: MLE diagnostics for every exposure. RMS is the square root of the across-realization mean squared field error; local SD averages $\sqrt{V/R}$ at the truth and is not an empirical uncertainty interval. Boundary fraction is the fraction of selected estimates hitting any parameter bound. The last column is the paired excess-RMSE interval over the homogeneous estimate.

λ	$R/10^6$	Design	RMS	Local SD	Boundary	Excess interval
8	0.1	A	0.483	0.240	0.1	$[+0.087, +0.462]$
8	0.1	S	0.181	0.147	0.0	$[-0.005, +0.088]$
8	0.3	A	0.259	0.139	0.1	$[-0.080, +0.197]$
8	0.3	S	0.086	0.085	0.0	$[-0.086, -0.021]$
8	1	A	0.076	0.076	0.0	$[-0.082, -0.039]$
8	1	S	0.052	0.046	0.0	$[-0.107, -0.059]$
8	3	A	0.044	0.044	0.0	$[-0.113, -0.070]$
8	3	S	0.030	0.027	0.0	$[-0.125, -0.085]$
8	10	A	0.029	0.024	0.0	$[-0.123, -0.091]$
8	10	S	0.017	0.015	0.0	$[-0.135, -0.101]$
8	30	A	0.014	0.014	0.0	$[-0.140, -0.100]$
8	30	S	0.009	0.008	0.0	$[-0.146, -0.105]$
32	0.1	A	1.006	2.890	0.8	$[+0.691, +1.005]$
32	0.1	S	0.850	0.615	0.6	$[+0.401, +0.897]$
32	0.3	A	0.858	1.669	0.7	$[+0.571, +0.838]$
32	0.3	S	0.417	0.355	0.1	$[+0.112, +0.372]$
32	1	A	0.854	0.914	0.7	$[+0.491, +0.872]$
32	1	S	0.228	0.195	0.0	$[+0.010, +0.141]$
32	3	A	0.906	0.528	0.6	$[+0.605, +0.901]$
32	3	S	0.120	0.112	0.0	$[-0.055, +0.016]$
32	10	A	0.507	0.289	0.5	$[+0.143, +0.500]$
32	10	S	0.068	0.062	0.0	$[-0.099, -0.042]$
32	30	A	0.270	0.167	0.0	$[+0.035, +0.186]$
32	30	S	0.038	0.036	0.0	$[-0.117, -0.077]$

local-unbiased comparison; it does not set a prospectively valid stopping rule and is not used for estimator tuning.

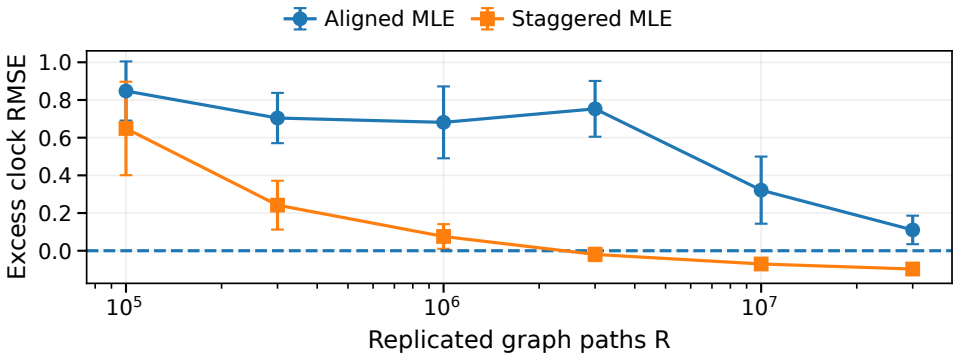

Figure 20: Descriptive paired excess-error intervals at all six high-clock exposures. The zero line denotes the homogeneous estimator's error, not zero absolute error. Intervals are marginal, not a simultaneous confidence band or a guarantee of a universal exposure threshold.

Z.4 Recovery under the analytically calibrated schedule

The timing intervention is also fitted at the ten heterogeneous truths, rather than evaluated only by its symmetric information coefficient. The nominal clock is $\lambda = 32$; the first observation is moved from 0.5 to $0.5 + \rho_*/32$, keeping initial zero and all five subsequent observations unchanged. The ratio is obtained from Theorem 12, without choosing among ratios by validation or recovery error. Both rates remain affinity-dependent, and all 18 parameters are estimated. Independent Radau

Table 83: Calibrated first-endpoint shift $\rho_*/32$ at heterogeneous clocks/offset affinities and true $\gamma = 0$, fitted jointly. Ten paired realizations; 40 joint fits. Intervals compare each fitted RMSE with the homogeneous reference and with the same-method staggered estimate, respectively. The theoretical optimum is at the symmetric parameter, not at these unknown heterogeneous truths.

$R/10^6$	Method	RMSE	Excess over homogeneous	Calibrated minus staggered
1	MLE	0.177	$[+0.001, +0.087]$	$[-0.078, +0.014]$
1	ridge	0.130	$[-0.034, +0.027]$	$[-0.059, +0.005]$
10	MLE	0.050	$[-0.103, -0.065]$	$[-0.029, +0.002]$
10	ridge	0.048	$[-0.105, -0.066]$	$[-0.028, +0.002]$

Table 84: Observed-risk-set budget at six nodes. Ranks have exact rational witnesses; the spectral orders are Theorem 4, not classifications from fitted slopes. G_6 is odd unicyclic; successive graphs add specified dyads. Staggering gives six quadratic directions in every row.

Graph	Dyads m	Rank r	Quadratic	Quartic
G_6	6	0	0	6
G_8	8	2	2	4
G_{11}	11	5	5	1
$K_6 = G_{15}$	15	5	5	1

probabilities and sequential counts are generated at $R = 10^6, 10^7$. This gives 20 count datasets and 40 selected fits, with the same MLE, fixed ridge, bounds and three-start training-objective rule.

The optimum is local to the symmetric limiting model; at these finite heterogeneous truths it is not presumed to dominate staggering. Table 83 reports both reference and schedule comparisons, including uncertainty, rather than selecting a favorable estimator or exposure. Replicated-path exposure is unchanged relative to the paired staggered comparisons; only observation times change.

Z.5 Replay and independent checks

All 320 exposure-grid estimates are replayed; the saved probabilities, likelihoods and clock-field errors reproduce exactly, and all 160 count arrays regenerate. All selected fits report successful termination, which is not a certificate of global optimality. Independent ODE evaluation of 32 selected fitted cases has maximum log-probability discrepancy 4.99×10^{-13}. The calibrated-schedule script separately replays all 40 fits, regenerates its 20 datasets, and checks four fitted cases with independent ODE integration (maximum log discrepancy 1.87×10^{-13}). Full numeric tolerances and observed discrepancies are stored in the two audit JSON files. The manuscript compiles independently of the experiments; tables are generated from recorded results.

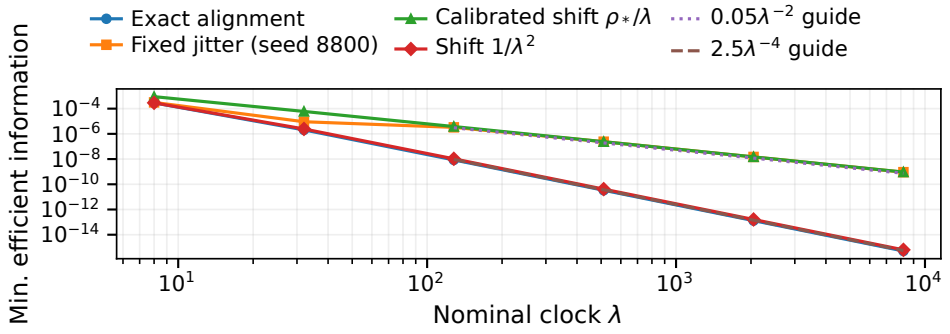

Figure 21: Fixed and shrinking timing experiments. All exact schedules and local coefficients are specified above.

AA Reproducible learning protocols

AA.1 Conditional state-prediction task and independent data generation

The primary benchmark uses 32 nodes, all 496 undirected dyads, horizon $[0, 60]$, and observations every 0.5 units. Initial centered positions are uniform on $[-0.8, 0.8]^2$. Six successive knot increments are Gaussian with standard deviation 0.24, clipped to $[-2, 2]$ before recentering. Knot times are $0, 10, \ldots, 60$. The affinity is $\alpha - 2\|r_i - r_j\|_1$, where α is twice the median initial distance. The homogeneous clock is $c = -0.4$. In the additive condition, balanced random node groups initialize activity at ± 0.7; Gaussian knot increments of standard deviation 0.12 are clipped to $[-1.1, 1.1]$ and centered. The nonadditive condition adds $(0.8 + 0.25 \sin(t_k/12))b_i b_j$ at each knot for independently shuffled balanced signs b_i.

The generator simulates event times by exact integrated-hazard inversion, refining at coordinate crossings. It samples snapshots from these event histories, independently of the approximate panel likelihood used for fitting. All hidden changes, censoring intervals, dyads and endpoint states are retained. Training records end at or before 40; validation records lie within $[40, 50]$; test records lie within $[50, 60]$. No record straddles a split boundary. There are 39,680 training records and 9,920 validation and test records at the primary gap. No negative subsampling, fictitious event label, future transition count or generating rate enters a learner.

The global primary hyperparameter development graph has seed 7300, under additive clocks. Confirmation graphs have seeds 7400–7404 in each of the three clock conditions. Configurations are fixed before confirmation; each confirmation graph uses only its own validation interval to choose an optimization checkpoint. The same graph realization is shared across models to enable paired comparisons. The three conditions use paired underlying random streams, so their 15 entries are not treated as 15 independent draws when testing a pooled condition-independent hypothesis. Graph-level results and all candidate traces are saved.

AA.2 Strictly causal feature construction

At each origin, features are constructed *before* consuming its observed graph. The current dyad state is supplied separately to the final decoder or transition-row selection. Exponential history factors are $0.5, 0.85, 0.97$. The 16 dyad features contain three occupancy summaries, six directional smoothed transition frequencies, recency, endpoint degree sum and difference, common-neighbor summary, two smoothed degree combinations, and normalized history length. Node summaries have nine coordinates: three smoothed degrees, last observed degree, a triangle statistic, recent flip fraction, history length, normalized time, and a constant. The transition proxy has three deterministic coordinates: empirical equilibrium log-odds, a geometric-mean clock proxy, and its homogeneous nuisance-adjusted information proxy. Frequencies receive pseudocounts and are clipped only within the proxy construction, as specified in `common.py`. These proxies are computed from past observations, not from generating parameters.

Positive-interaction histories store up to L past neighbors and their ages. Ties at a common snapshot time use a deterministic hash ordering fixed independently of the labels. Node IDs are transductive inputs; no inductive unseen-node claim is made. Missing observations in the separate autonomous study increase gaps rather than creating absences. In the primary complete panel, zero is an observed disconnected state.

A future-perturbation audit randomizes every state after each of five origins and confirms identical model inputs, excluding the changed target, and identical predictions for the five tested head/encoder families. This checks the implemented causal data path; it does not justify treating overlapping forecast targets as independent observations.

AA.3 Published-method adapters and primary optimization

The eight published-method comparisons are ℓ_1LD–panel (ICLR'26) (Lai et al., 2026), GraSSP–panel (AAAI'24) (Çelikkanat et al., 2024), GraphMixer–state (ICLR'23) (Cong et al., 2023), DyGFormer–state (NeurIPS'23) (Yu et al., 2023), FreeDyG–state (ICLR'24) (Tian et al., 2024), CRAFT–state (NeurIPS'25) (Yi et al., 2025), TAMI–GraphMixer (NeurIPS'25) (Yu et al., 2025),

and EdgeBank–state (NeurIPS'22) (Poursafaei et al., 2022). Full references are in the bibliography. The source-informed implementations retain their named mechanism and adapt their output to an endpoint-state probability given the current state and gap. They do not reproduce each paper's published datasets, ranking losses, negative samplers, default widths, or complete official training pipeline.

The two survival adapters retain the published hazard families under the panel observation model; they do not infer fictitious event times. They use training-only occupancy initialization, four fitted motion bins, and constant clocks. The autonomous TempoGeom control adds node clocks. Their fields are continued constantly after time 40. Their test kernels therefore have constant rates; numerical claims for general varying squared-distance clocks are kept separate.

GraphMixer retains its time/link mixing blocks and node encoder. DyGFormer uses neighbor co-occurrence encodings, four channels, patching, and two self-attention blocks. FreeDyG uses frequency mixing and its neighbor-frequency mechanism. CRAFT uses destination-conditioned cross-attention without a persistent learned memory. TAMI augments the GraphMixer backbone with logarithmic time encoding and detached per-dyad history aggregation. EdgeBank–state fits a calibrated combination of current state and past seen-link indicators, selecting a history window from $\{0.5, 2, 10, 10^{10}\}$ on validation data. The artifact records module locations and inspected source provenance. All reported scores are for these task-specific adapters.

Primary neural optimization uses Adam, at most 1,200 updates of 128 records, gradient clipping at 10, and validation every 100 updates. Four nonimproving validations, with minimum improvement 10^{-5}, trigger early stopping. This cap is at most 3.87 training-set equivalents; it is not a convergence certificate. TAMI traverses chronological chunks, resets pair memory at a training sweep, and replays history causally at evaluation. Other models sample training records uniformly. Standard weight decay is 10^{-5} unless selected otherwise; autonomous models use their selected smoothness penalty. The primary affine head also has the stated 10^{-4} mean-square parameter penalty. These model-specific regularizers differ across objectives.

Development selection first screens the saved model-specific learning-rate, width, dropout or smoothness grid for 80 updates, then refits two candidates for 600 updates. History models additionally compare $L = 16$ and $L = 64$ with full-budget refits after a three-learning-rate long-history screen. Every selected configuration and executed candidate trace is included. Main conclusions about input access or budget use the separately controlled experiment below rather than assuming this broad primary screen equates optimization difficulty.

The primary MLP is independently tuned and has a potentially different width from TempoGeom-H; it is labeled as an independently tuned MLP control, not as an exact-width decoder ablation. The exact-width comparisons below fix the encoder architecture and candidate grid. Internal controls are reported separately from the eight external methods.

AA.4 Matched history heads and multiple horizons

The full-head comparison fixes the common history width to 32, corresponding to $C = 16$. TempoGeom-H maps node and dyad histories to an origin-specific metric term, free affinity correction, and node/pair clock terms. Its affinity intercept and initial clock are bounded through $8\tanh(\cdot/8)$ and $5\tanh(\cdot/5)$; affinity and clock slopes are bounded by 0.5 and 0.3. A proxy shrinkage factor $f/(f + 0.2)$ scales pair-clock corrections. This is an empirical regularizer, not the nuisance projection of Theorem 1. TempoGeom-C uses the same history encoder and origin coefficient construction with both slopes fixed at zero. The MLP uses the same-width encoder followed by a horizon/state classifier. The rate heads have direct proxy and node-clock pathways, so the stricter decoder-only probe below also controls those pathways.

A development graph with seed 7600 selects learning rate from $\{0.001, 0.003, 0.01\}$. Ten independent additive-clock confirmation graphs use seeds 7610–7619. Training and validation jointly use gaps $0.5, 1, 2$; gap 4 is absent from both fitting and selection. Records crossing split boundaries are removed separately at each gap. The joint training set has 117,056 records. A fixed selected checkpoint is then evaluated at all four gaps, with 9,920, 9,424, 8,432 and 6,448 test records respectively. The repeated targets constitute a multi-horizon predictive objective, not a joint likelihood of independent paths.

All matched studies use at most 2,400 Adam updates of batch size 256, weight decay 10^{-5}, gradient clipping at 10, validation every 200 updates, and patience six with improvement threshold 10^{-5}. Learning rates are selected on the development graph, and checkpoints on each graph's validation interval. The maximum sample exposure is 5.25 training-set equivalents for the multi-horizon objective, and 15.48 equivalents at the primary gap. The same pseudorandom batch stream is paired across decoder families until their stopping times diverge. Unused parameters are excluded from the optimizer and active parameter count.

AA.5 A DECODER-ONLY PROBE WITH FULLY MATCHED PATHWAYS

To remove the full-head proxy/metric-pathway difference, a separate probe gives all three decoders exactly the same 32-dimensional history encoding concatenated with the same three proxy coordinates. Its ID, node, and pair encoder construction is identical through the final history representation. The affine decoder is a linear four-coefficient rate head with the bounds stated above, and the constant decoder is its two-coefficient version. The MLP concatenates current state and log horizon to the identical history/proxy vector and uses a width-32 hidden layer. There is no separate metric pathway, node-clock bypass, or information gate in this probe.

Development seed 7700 and ten confirmation seeds 7710–7719 are disjoint from the full-head study. The learning-rate grid, update budget, validation rule, gaps, splits and precision are identical. This is an internal mechanism experiment, not an additional published baseline. It was designed after inspecting the full-head study and uses a separate graph cohort.

AA.6 FEATURE ACCESS AND EXTENDED OPTIMIZATION BUDGETS

GraphMixer and DyGFormer are each evaluated with their positive-history state decoder and with an augmented decoder that receives all 37 causal summary coordinates used by the history-head family: 16 dyad features, both ordered nine-dimensional node vectors, and three proxies. The underlying published encoder is unchanged. Canonically ordered node IDs remain available through the encoder. In particular, the two node summary vectors are not compressed into only their sums and absolute differences; that compression could lose node-to-feature association. The augmented and positive-history-only state heads have the same width 64. The extra input coordinates increase the augmented head's parameter count; the resource table reports active counts. This is an input-access control, not an exactly parameter-matched architecture comparison. All preprocessing uses training or strictly past records.

This matched-control experiment uses the additive-clock condition and all five primary graph seeds 7400–7404. It is not advertised as an all-condition feature equivalence test. Development seed 7500 and the same three-learning-rate grid select configurations independently for each positive-history-only/augmented encoder. Both variants use the 2,400-update, batch-256 protocol, identical evaluation records, and the same validation checkpoint rule. Thus their comparison holds the maximum training exposure fixed, unlike comparing a long-budget feature-augmented model only with the short-budget primary run. Additional full history heads and the same-width MLP use this same development graph and long-budget primary-gap protocol. Learning curves and selected/stopped iteration counts are included; finite budget checks do not establish global optimization.

AA.7 NUMERICAL EVALUATION AND UNCERTAINTY

Primary training, validation and initial test tolerances are $0.02, 0.005, 0.002$. All selected primary checkpoints are replayed from saved parameters. Final affine-head test losses are recomputed using the combined bound at 10^{-4}, then checked at 10^{-5} without reselection. Observed endpoint entries are logged directly. Neural logits use stable Bernoulli cross-entropy; predicted probabilities are also retained for calibration and ranking metrics. Constant post-training autonomous fields have closed-form transition probabilities. A varying squared-distance model is not given an exact-affine-clock certificate.

The matched studies train at tolerance 0.02, validate at 0.001, and test at 10^{-4}. Per-record log probabilities and numerical error bounds are saved, so a mean-loss difference is bounded by both models' mean numerical bounds. Paired-interval endpoints are further enlarged for the possible change in sample standard deviation, using the bound in Appendix V. Mean NLL, Brier score, ROC

Table 85: Primary endpoint-state prediction: NLL mean ± SD over five paired graph realizations per condition. The first eight rows are published-method task adapters with their native summaries. The final two rows are internal controls, not additional external baselines. Feature-matched long-budget comparisons are discussed in the text and detailed in Appendix X.3.

Method	Homogeneous	Additive	Nonadditive
ℓ_1LD–panel (ICLR'26)	0.4430 ± 0.0223	0.4538 ± 0.0100	0.4496 ± 0.0107
GraSSP–panel (AAAI'24)	0.4359 ± 0.0181	0.4488 ± 0.0152	0.4427 ± 0.0122
GraphMixer–state (ICLR'23)	0.4431 ± 0.0194	0.4243 ± 0.0137	0.4209 ± 0.0109
DyGFormer–state (NeurIPS'23)	0.4244 ± 0.0244	0.4114 ± 0.0218	0.4117 ± 0.0168
FreeDyG–state (ICLR'24)	0.4340 ± 0.0146	0.4477 ± 0.0171	0.4431 ± 0.0145
CRAFT–state (NeurIPS'25)	0.4309 ± 0.0180	0.4224 ± 0.0134	0.4164 ± 0.0081
TAMI–GraphMixer (NeurIPS'25)	0.4393 ± 0.0177	0.4252 ± 0.0116	0.4244 ± 0.0119
EdgeBank–state (NeurIPS'22)	0.4414 ± 0.0158	0.4404 ± 0.0135	0.4375 ± 0.0096
TempoGeom-H	0.4172 ± 0.0186	0.3978 ± 0.0148	0.3929 ± 0.0107
Independently tuned MLP (control)	0.4146 ± 0.0180	0.3974 ± 0.0135	0.3923 ± 0.0095
Count mixture (control)	0.4231 ± 0.0193	0.4018 ± 0.0146	0.3942 ± 0.0095

AUC, average precision, and equal-width ten-bin calibration error are reported. Standard deviations use independent graph realizations, not endpoint records. Paired intervals use a two-sided Student-t multiplier on graph-level differences; these small-sample intervals are descriptive and not multiplicity-adjusted. Both numerical and statistical uncertainty are reported for close comparisons.

The kernel audit uses 48 random affine problems, four rare-transition problems and two stiff problems, each at three tolerances and under two mesh rules. The reference independently integrates both off-diagonal forward equations with Radau at relative tolerance 2×10^{-11} and absolute tolerance 2×10^{-16}, then adds the analytic survival term. It never obtains a rare reference entry by subtracting from one. Forty-eight derivative checks compare automatic differentiation with centered differences on fixed 32-cell meshes. Cell-count comparisons are independent of the concurrent CPU timing conditions; no GPU or isolated-hardware speed ranking is claimed.

AB Learning results, uncertainty, and execution checks

The native-input comparison below separates eight external adapters, TempoGeom-H and the two most informative internal controls. The fuller internal comparisons below are not counted as external methods. All aggregate metrics are regenerated from per-graph JSON records; per-record predictions, log probabilities and numerical error bounds are included in the executable artifact. A graph, not an endpoint record, is the sampling unit for the reported standard deviations and paired intervals. Small-sample intervals are descriptive, not multiplicity-adjusted significance claims.

Table 86: Internal primary controls, using the same five graphs as Table 85. The independently tuned MLP has its own selected width; it is not an exact-width decoder ablation.

Internal control	Homogeneous	Additive	Nonadditive
Autonomous TempoGeom	0.4368 ± 0.0192	0.4198 ± 0.0077	0.4211 ± 0.0104
TempoGeom-H	0.4172 ± 0.0186	0.3978 ± 0.0148	0.3929 ± 0.0107
Independently tuned MLP	0.4146 ± 0.0180	0.3974 ± 0.0135	0.3923 ± 0.0095
Count mixture	0.4231 ± 0.0193	0.4018 ± 0.0146	0.3942 ± 0.0095

AB.1 Decoder-only controls and held-out horizons

The full-head study shares the encoder width but retains rate-specific metric/proxy pathways. The strict decoder-only probe gives all three heads the identical history/proxy representation and uses a disjoint cohort. The three gap-specific objectives overlap in their endpoint targets; their losses are

Table 87: Secondary primary metrics under homogeneous clocks, mean ± graph-level SD; five graphs. These measure different predictive properties, not parameter recovery.

Method	Brier ↓	AUC ↑	AP ↑	ECE ↓
ℓ_1LD–panel (ICLR'26)	0.139 ± 0.008	0.847 ± 0.006	0.723 ± 0.018	0.039 ± 0.009
GraSSP–panel (AAAI'24)	0.137 ± 0.006	0.852 ± 0.006	0.726 ± 0.018	0.043 ± 0.005
GraphMixer–state (ICLR'23)	0.138 ± 0.007	0.834 ± 0.012	0.697 ± 0.021	0.036 ± 0.012
DyGFormer–state (NeurIPS'23)	0.132 ± 0.008	0.859 ± 0.009	0.735 ± 0.012	0.019 ± 0.006
FreeDyG–state (ICLR'24)	0.135 ± 0.006	0.842 ± 0.007	0.711 ± 0.029	0.015 ± 0.005
CRAFT–state (NeurIPS'25)	0.135 ± 0.007	0.852 ± 0.009	0.725 ± 0.027	0.030 ± 0.009
TAMI–GraphMixer (NeurIPS'25)	0.137 ± 0.006	0.844 ± 0.007	0.712 ± 0.026	0.035 ± 0.006
EdgeBank–state (NeurIPS'22)	0.137 ± 0.006	0.832 ± 0.005	0.670 ± 0.021	0.035 ± 0.009
TempoGeom-H	0.130 ± 0.007	0.865 ± 0.006	0.750 ± 0.020	0.012 ± 0.003
Independently tuned MLP	0.130 ± 0.007	0.867 ± 0.005	0.751 ± 0.020	0.015 ± 0.003
Count mixture	0.132 ± 0.007	0.860 ± 0.007	0.742 ± 0.020	0.020 ± 0.006
Autonomous TempoGeom	0.137 ± 0.007	0.852 ± 0.006	0.723 ± 0.015	0.038 ± 0.006

Table 88: Secondary primary metrics under additive clocks, mean ± graph-level SD; five graphs. These measure different predictive properties, not parameter recovery.

Method	Brier ↓	AUC ↑	AP ↑	ECE ↓
ℓ_1LD–panel (ICLR'26)	0.144 ± 0.003	0.843 ± 0.007	0.717 ± 0.035	0.029 ± 0.004
GraSSP–panel (AAAI'24)	0.141 ± 0.005	0.846 ± 0.008	0.725 ± 0.024	0.037 ± 0.011
GraphMixer–state (ICLR'23)	0.134 ± 0.005	0.867 ± 0.006	0.772 ± 0.023	0.033 ± 0.006
DyGFormer–state (NeurIPS'23)	0.129 ± 0.008	0.874 ± 0.011	0.776 ± 0.017	0.017 ± 0.004
FreeDyG–state (ICLR'24)	0.140 ± 0.006	0.834 ± 0.011	0.699 ± 0.028	0.021 ± 0.005
CRAFT–state (NeurIPS'25)	0.133 ± 0.005	0.868 ± 0.006	0.773 ± 0.020	0.030 ± 0.010
TAMI–GraphMixer (NeurIPS'25)	0.134 ± 0.004	0.864 ± 0.006	0.769 ± 0.020	0.027 ± 0.009
EdgeBank–state (NeurIPS'22)	0.138 ± 0.005	0.840 ± 0.005	0.680 ± 0.022	0.027 ± 0.010
TempoGeom-H	0.124 ± 0.005	0.883 ± 0.005	0.796 ± 0.018	0.014 ± 0.003
Independently tuned MLP	0.124 ± 0.005	0.885 ± 0.003	0.795 ± 0.016	0.020 ± 0.009
Count mixture	0.126 ± 0.005	0.881 ± 0.005	0.789 ± 0.021	0.014 ± 0.001
Autonomous TempoGeom	0.133 ± 0.002	0.870 ± 0.005	0.776 ± 0.026	0.024 ± 0.007

Table 89: Secondary primary metrics under nonadditive clocks, mean ± graph-level SD; five graphs. These measure different predictive properties, not parameter recovery.

Method	Brier ↓	AUC ↑	AP ↑	ECE ↓
ℓ_1LD–panel (ICLR'26)	0.142 ± 0.004	0.843 ± 0.007	0.705 ± 0.038	0.033 ± 0.005
GraSSP–panel (AAAI'24)	0.139 ± 0.004	0.849 ± 0.006	0.715 ± 0.032	0.040 ± 0.010
GraphMixer–state (ICLR'23)	0.132 ± 0.004	0.866 ± 0.008	0.767 ± 0.040	0.030 ± 0.010
DyGFormer–state (NeurIPS'23)	0.129 ± 0.006	0.872 ± 0.006	0.768 ± 0.025	0.021 ± 0.005
FreeDyG–state (ICLR'24)	0.139 ± 0.005	0.837 ± 0.011	0.695 ± 0.033	0.020 ± 0.002
CRAFT–state (NeurIPS'25)	0.131 ± 0.002	0.870 ± 0.005	0.767 ± 0.029	0.031 ± 0.012
TAMI–GraphMixer (NeurIPS'25)	0.134 ± 0.004	0.864 ± 0.007	0.764 ± 0.033	0.034 ± 0.008
EdgeBank–state (NeurIPS'22)	0.137 ± 0.004	0.840 ± 0.007	0.673 ± 0.032	0.033 ± 0.007
TempoGeom-H	0.123 ± 0.004	0.884 ± 0.006	0.792 ± 0.033	0.014 ± 0.003
Independently tuned MLP	0.123 ± 0.003	0.885 ± 0.006	0.796 ± 0.031	0.019 ± 0.007
Count mixture	0.123 ± 0.004	0.883 ± 0.006	0.792 ± 0.032	0.012 ± 0.002
Autonomous TempoGeom	0.133 ± 0.004	0.866 ± 0.006	0.765 ± 0.033	0.029 ± 0.004

Table 90: Saved primary selected configurations. Model-specific smoothness, dropout, shrinkage and weight decay, and all executed development candidates, are preserved in the JSON records. A channel width is not a common representation dimension across architectures.

Method	Learning rate	Channel width	History
ℓ_1LD–panel (ICLR'26)	0.025	–	16
GraSSP–panel (AAAI'24)	0.025	–	16
GraphMixer–state (ICLR'23)	0.012	16	64
DyGFormer–state (NeurIPS'23)	0.006	8	64
FreeDyG–state (ICLR'24)	0.006	8	64
CRAFT–state (NeurIPS'25)	0.012	8	64
TAMI–GraphMixer (NeurIPS'25)	0.006	16	64
EdgeBank–state (NeurIPS'22)	window on validation	–	–
TempoGeom-H	0.003	16	16
Independently tuned MLP	0.01	8	16
Count mixture	0.03	8	16
Autonomous TempoGeom	0.025	–	16

not treated as independent path likelihoods. Training and validation exclude gap 4. Every graph retains its validation-selected checkpoint unchanged for all four test gaps.

Table 91: Full-head multi-horizon NLL, mean $\pm$ sample SD over ten independent graphs. Gap 4 is unseen during training and validation. The two studies use disjoint graph cohorts.

Gap	Affine	Constant	MLP
0.5	0.3857 ± 0.0216	0.3846 ± 0.0220	0.3926 ± 0.0244
1	0.4531 ± 0.0277	0.4519 ± 0.0273	0.4573 ± 0.0301
2	0.4989 ± 0.0320	0.4972 ± 0.0301	0.5034 ± 0.0337
4 (unseen)	0.5256 ± 0.0344	0.5230 ± 0.0325	0.5311 ± 0.0356

Table 92: Comparator minus affine NLL in the full-head study. Positive favors the affine decoder. Intervals use ten graph-level paired differences and are unadjusted for multiplicity. Subtracting the last column from the lower endpoint and adding it to the upper endpoint encloses numerical changes to both the mean and sample standard deviation.

Gap	Comparator	Mean difference	Paired 95% interval	Num. half-width
0.5	constant	-0.00115	$[-0.00457, 0.00227]$	2.9×10^{-5}
0.5	mlp	0.00689	$[0.00323, 0.01055]$	2.9×10^{-5}
1	constant	-0.00124	$[-0.00643, 0.00395]$	2.9×10^{-5}
1	mlp	0.00416	$[-0.00028, 0.00859]$	2.9×10^{-5}
2	constant	-0.00177	$[-0.00813, 0.00459]$	3.0×10^{-5}
2	mlp	0.00449	$[-0.00049, 0.00947]$	3.0×10^{-5}
4	constant	-0.00268	$[-0.01031, 0.00496]$	3.0×10^{-5}
4	mlp	0.00541	$[-0.00002, 0.01083]$	3.0×10^{-5}

AB.2 Matched causal summaries and optimization budgets

The positive-only and augmented GraphMixer/DyGFormer pairs both use the same maximum budget, candidate grid and validation rule. Thus input availability is not confounded with comparing only a long-budget augmented run to a short-budget positive-only run. The comparison is confined to the additive condition. Comparing these longer positive-only runs with the short-budget primary study changes both the optimization cap and development selection, and is reported as a stress test rather than an isolated update-count treatment.

Table 93: Strict decoder-only probe multi-horizon NLL, mean ± sample SD over ten independent graphs. Gap 4 is unseen during training and validation. The two studies use disjoint graph cohorts.

Gap	Affine	Constant	MLP
0.5	0.3953 ± 0.0223	0.3936 ± 0.0217	0.4037 ± 0.0235
1	0.4607 ± 0.0271	0.4596 ± 0.0259	0.4634 ± 0.0280
2	0.5059 ± 0.0306	0.5037 ± 0.0284	0.5089 ± 0.0315
4 (unseen)	0.5401 ± 0.0337	0.5299 ± 0.0299	0.5380 ± 0.0330

Table 94: Comparator minus affine NLL in the strict decoder-only study. Positive favors the affine decoder. Intervals use ten graph-level paired differences and are unadjusted for multiplicity. Subtracting the last column from the lower endpoint and adding it to the upper endpoint encloses numerical changes to both the mean and sample standard deviation.

Gap	Comparator	Mean difference	Paired 95% interval	Num. half-width
0.5	constant	-0.00170	$[-0.00354, 0.00013]$	3.2×10^{-5}
0.5	mlp	0.00836	$[0.00407, 0.01266]$	3.2×10^{-5}
1	constant	-0.00111	$[-0.00352, 0.00130]$	3.3×10^{-5}
1	mlp	0.00271	$[-0.00090, 0.00632]$	3.3×10^{-5}
2	constant	-0.00215	$[-0.00597, 0.00167]$	3.3×10^{-5}
2	mlp	0.00307	$[-0.00157, 0.00772]$	3.3×10^{-5}
4	constant	-0.01026	$[-0.01826, -0.00226]$	3.3×10^{-5}
4	mlp	-0.00216	$[-0.01109, 0.00677]$	3.3×10^{-5}

Table 95: Single-gap internal head controls on the five additive graphs under the same longer budget as Table 75. These are internal controls, not additional external baselines.

Internal matched head	NLL	Selected step	Reached cap
TempoGeom-H	0.3976 ± 0.0150	1800	5/5
TempoGeom-C	0.3974 ± 0.0145	2000	5/5
Width-matched MLP	0.4004 ± 0.0132	1800	4/5

Table 96: Paired additive-clock comparisons under the matched long budget. Lower differences favor the first named model. Numerical enlargement is recorded in the accompanying CSV.

Comparison	Mean difference	Paired 95% interval
GraphMixer (ICLR'23) augmented minus positive	-0.02151	$[-0.02794, -0.01509]$
DyGFormer (NeurIPS'23) augmented minus positive	-0.00142	$[-0.00597, 0.00312]$
TempoGeom-C minus TempoGeom-H	-0.00027	$[-0.00119, 0.00065]$
Width-matched MLP minus TempoGeom-H	0.00275	$[-0.00204, 0.00754]$
GraphMixer (ICLR'23) + summaries minus TempoGeom-H	0.00114	$[-0.00023, 0.00251]$
DyGFormer (NeurIPS'23) + summaries minus TempoGeom-H	0.00371	$[0.00022, 0.00720]$

Table 97: Standard-budget and extended-budget positive-history fits on the same additive graphs. The extended run also retunes the learning rate on a separate development graph; this is a budget-adequacy stress test, not an isolated causal effect of update count. Difference is extended minus primary NLL.

Positive-only encoder	Primary cap	Long cap	Paired difference
GraphMixer (ICLR'23)	0.4243	0.4203	-0.00398
DyGFormer (NeurIPS'23)	0.4114	0.4028	-0.00860

Table 98: Median resource records for the five matched additive fits. Parameters count trainable tensors in the declared architecture, not an empirically estimated effective dimension. Adding all 37 summaries increases each external prediction head by 2,368 parameters; the positive and augmented models are input-controlled, not exactly capacity-matched. Fits ran concurrently on a four-core CPU quota, so elapsed times are execution records rather than an isolated speed ranking.

Model / input	Parameters	Fit seconds	Stopped step
GraphMixer (ICLR'23), positive only	19,409	264.7	2400
GraphMixer (ICLR'23) + summaries	21,777	268.9	2400
DyGFormer (NeurIPS'23), positive only	32,025	555.4	2400
DyGFormer (NeurIPS'23) + summaries	34,393	559.6	2400
TempoGeom-H	5,803	15.2	2400
TempoGeom-C	5,737	9.9	2400
Width-matched MLP	6,689	6.6	2400

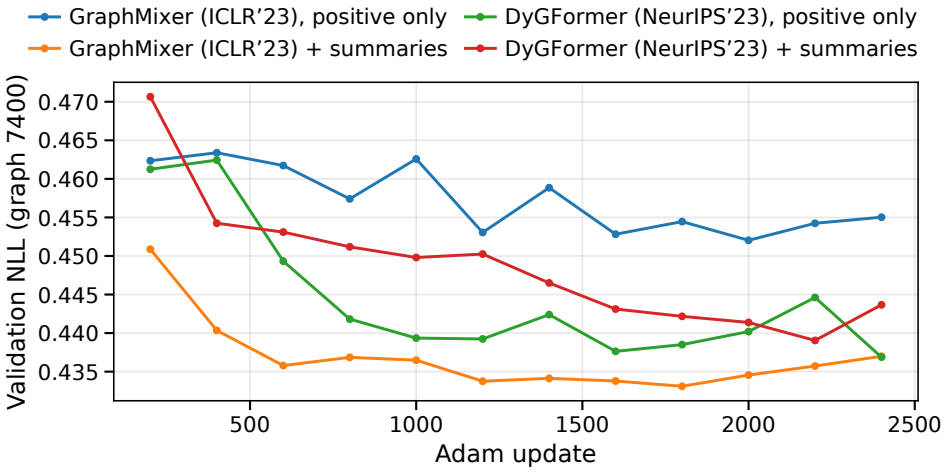

Figure 22: Complete validation traces for the fixed illustrative graph seed 7400. All other graph traces and stopped/selected updates are in the artifact. These are validation losses, not test-driven checkpoint choices.

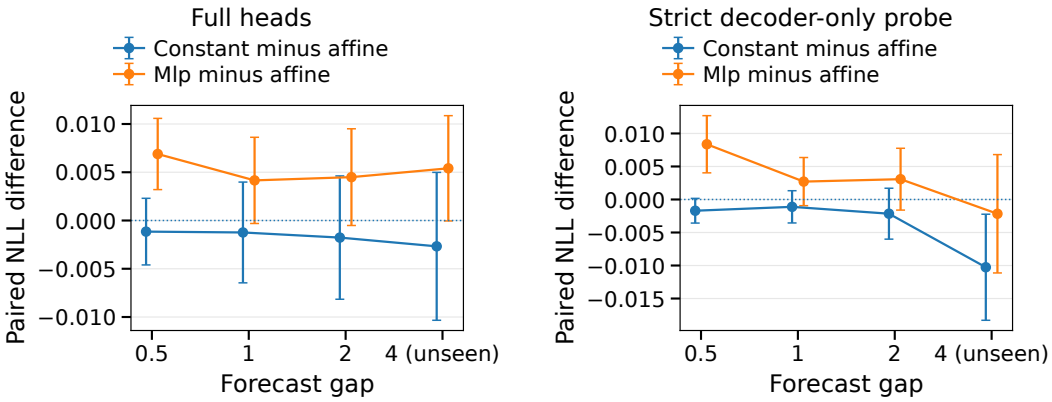

Figure 23: Paired comparator-minus-affine NLL in two disjoint ten-graph cohorts. Error bars enclose the graph-level 95% paired intervals after numerical enlargement of their centers and standard deviations; positive favors affine. Full-head and strict decoder-only probes are different tests and are not pooled.

AB.3 Likelihood precision and executable audits

Table 99: Median subdivision counts over 54 affine-rate cases. The ratio is the median of per-case ratios, not the ratio of medians. All 324 kernel evaluations satisfy their computed log-error bound against the numerical reference.

Tolerance	First rule	Combined rule	Median ratio
0.001	250	46.5	7.77
0.0001	2500	146	24.71
1×10^{-5}	25000	460.5	78.13

The first-order and combined rules are applied to exactly the same cases. The reported ratios compare subdivision counts, not concurrent CPU runtimes. Rare off-diagonal entries are computed directly in the reference and in the positive construction. The gradient test keeps the mesh fixed. Final prediction audits do not retrain or reselect checkpoints, and are not reproductions of external authors' published pipelines. Exact-arithmetic bounds are not presented as certified floating-point intervals. Machine-readable summaries give every input case, per-record bound, forward-check discrepancy, seed-level metric, and all validation traces.

The four rare and two stiff cases are additionally checked by independent positive quadrature at 50 and 80 decimal working digits. Their maximum log-probability disagreement is 1.5×10^{-33}; the Radau reference differs from the 80-digit calculation by at most 1.8×10^{-11}. All 36 kernel comparisons on these six cases satisfy their recorded bounds. Agreement at two precisions is a numerical stability audit, not an outward-rounded interval proof.

AC Geometric representation and field-recovery diagnostics

Latent-distance representations (Hoff et al., 2002; Çelikkanat et al., 2022) and formation–duration separation (Krivitsky and Handcock, 2014; Butts, 2024) supply modeling context. This section specifies the geometric representation used in the field-recovery diagnostics; the main information theorems use their own stated affinity families. One geometric representation uses centered piecewise-affine positions and activities:

$$\eta_{ij} = \alpha - 2\nu_A(r_i - r_j), \quad c_{ij} = g + u_i + u_j, \qquad \nu_A(z) = \max_k |a_k^\top z|. \tag{59}$$

Here positions and activities are centered, and trajectories are piecewise affine. Exact facet integration supports likelihood evaluation but does not establish unique geometric or clock recovery.

The phase family and polynomial/Fourier families are specified directly through (1); they are not claimed to be restrictions of this geometric representation.

The geometric-recovery study changes clock RMSE from .794 to .511 while displayed test NLL stays .5130 (Table 72). This tests optimizer budget, not geometric nonidentifiability. Its five-dataset field-recovery results, complete-path generation, interval-occupancy controls, candidate histories and implementation are available in the research artifact, separate from the conditional-panel acquisition study.

The eight endpoint-state adapters and matched-input/unseen-horizon controls in Appendices AA–AB remain forecasting comparisons, not additional clock-design baselines.