

What a Line Forgets

Hindsight Oracles under Hidden Dependency Geometry

J. R. Landers 

August 2026

Abstract

A sequential trace records what happened, but not what had to happen. It can therefore hide the dependency graph that determined which decisions were still reachable. We ask how much this hidden geometry can cost a hindsight rule that knows the realized rewards but must move locally.

For a dependency poset, every serial trace is a path through the lattice of feasible execution prefixes. We prove a sharp oracle inequality whose reachability term, under Lipschitz scores, reduces to a single distance κ_t from the current prefix to the next feasible frontier. We also quantify the cost of choosing both a graph and a comparator after seeing outcomes, and prove that the graph-selection term is unavoidable. Repeated independent traces recover the true order when every incomparable pair is eventually seen in both orientations, turning κ_t into an observable certificate. Exact experiments on 300 random posets show that this certificate can become exact before the full graph is recovered.

1 THE QUESTION INSIDE THE TRACE

A line is a faithful record of occurrence and an unreliable portrait of necessity. If operation u appears before operation v , the trace does not say whether u enabled v or merely happened to win a scheduling choice. The distinction is familiar in distributed systems, trace semantics, and process discovery, where causal order is partial even when logs are written sequentially [11, 12, 13]. It is equally familiar in precedence-constrained scheduling: the constraint graph is normally given first, and the schedule is analyzed afterward [15, 16, 17]. Here the direction is reversed. We see schedules and must reason about a decision rule whose feasible graph is itself uncertain.

This inverse problem meets a second asymmetry. A hindsight oracle may know every realized reward and still choose greedily among states reachable from its present prefix. The globally best state at a layer may have become inaccessible because of an earlier choice. Classical metrical task systems and smoothed online optimization compare causal rules with offline optima on known state spaces and charge movement explicitly [18, 19, 20, 21]. They do not usually ask which state space an order-only trace supports.

The paper in two sentences. A sequential trace tells us what happened but not what had to happen, so it hides the decision graph faced by a hindsight rule. We prove that one geometric distance controls this loss of reachability, and repeated traces can recover that distance and certify the rule, sometimes before they recover the entire graph.

The two problems fit together through order ideals: feasible sets of operations that could already have been completed. A dependency poset Q organizes these sets into a distributive lattice. Its maximal chains are exactly the serial executions, so a trace is literally a path through the hidden decision graph, not merely an analogy for one. The relevant obstruction is then geometric: after the rule has committed to a prefix G , how far is the best state at the next layer from every state the rule can still reach?

Contribution. The answer is an exact quantity,

$$\kappa_t(Q, G) = \max_{I \in \mathcal{I}_t(Q)} |G \setminus I|.$$

It is simultaneously (i) the worst projection distance from rank t to the successor frontier of G , (ii) a sharp upper control on lost log score for Lipschitz objectives, and (iii) a statistic that repeated traces can eventually reveal. Around this identity we prove four results.

1. A sharp oracle inequality decomposes the log gap into movement, surrogate distortion, approximate local maximization, and geometric inaccessibility.
2. A uniform extension charges post-outcome selection of a compatible geometry and policy. The charge scales with their joint description length, and a Gaussian family shows that the geometry contribution cannot generally be removed.
3. The ideal-projection theorem gives the closed form for κ_t , proves dependency monotonicity, identifies a trace-only minimax barrier, and attains equality in the Lipschitz accessibility bound.
4. Independent serial traces recover the hidden order at a rate controlled by rare orientations of incompatible pairs. On recovery, the trace-computable $\hat{\kappa}_t$ becomes the true radius and yields a certified plug-in oracle.

The argument therefore has three moves. First, the traces define all dependency graphs that remain possible. Second, each possible graph turns the same serial record into a path through a different lattice of feasible prefixes. Third, the projection theorem converts the oracle's inaccessible alternatives into the single radius κ_t , after which repeated traces determine when that radius may be trusted. This sequence, from trace to graph to distance to certificate, is the narrative spine of the paper.

The mathematical ingredients have distinguished histories. Finite distributive lattices are represented by order ideals [3, 4]; linear extensions, dimension, and their enumeration are classical [6, 7, 8]; lower sets of a poset form a poset antimatroid [10, 9]. Bayesian work has inferred partial orders from queue-like rankings, often restricting the order class because counting linear extensions is intractable in general [14]. PAC-Bayesian and minimum-description-length arguments attach a price to data-dependent selection [22, 23]. Our claim is not that these ingredients are new in isolation. It is that their synthesis exposes a missing variable in hindsight analysis: before charging how a rule traversed the graph, one must charge what the trace failed to say about the graph.

The paper is standalone. It develops the inverse theory of serialization introduced in [1] and the hindsight traversal analysis of [2], but neither is needed to follow the definitions or proofs below.

2 FROM SERIAL TRACES TO CANDIDATE GEOMETRIES

Let V be a labeled set of n operations. A *strict dependency poset* is $P = (V, <_P)$, where $<_P$ is irreflexive and transitive. The relation $u <_P v$ means that u is intrinsically required before v .

Definition 1 (Linear extension). A serial trace $\sigma = (\sigma(1), \dots, \sigma(n))$ is a linear extension of P if

$$u <_P v \implies \sigma^{-1}(u) < \sigma^{-1}(v).$$

The set of all linear extensions is denoted $\mathcal{L}(P)$.

Suppose the same latent system produces k observed traces

$$\Sigma_k = (\sigma_1, \dots, \sigma_k).$$

Write $<_{\sigma_j}$ for the strict total order induced by trace j and define the *agreement order*

$$A_k = \bigcap_{j=1}^k <_{\sigma_j}. \tag{1}$$

Thus $u <_{A_k} v$ precisely when every observed trace puts u before v .

Operationally, a reversal is decisive but an agreement is not. If one trace places u before v and another places v before u , the pair cannot be causally ordered. If every observed trace agrees, the order may reflect a true dependency or merely a scheduling preference that has not yet been reversed.

Definition 2 (Compatibility fiber). The deterministic fiber of geometries compatible with Σ_k is

$$\mathcal{F}(\Sigma_k) = \{Q : Q \text{ is a partial order on } V, \sigma_j \in \mathcal{L}(Q) \text{ for all } j\}. \quad (2)$$

Write $N_k = |\mathcal{F}(\Sigma_k)|$.

Proposition 3 (The inverse fiber). *For every collection of traces,*

$$\mathcal{F}(\Sigma_k) = \{Q : Q \text{ is a partial order and } Q \subseteq A_k\}. \quad (3)$$

If $A_k = P$ and P has height at most two with m strict comparabilities, then

$$|\mathcal{F}(\Sigma_k)| = 2^m. \quad (4)$$

Proof. If every σ_j extends Q , each comparison in Q occurs with the same orientation in every trace, hence $Q \subseteq A_k$. Conversely, if $Q \subseteq A_k$, every comparison in Q appears in every σ_j , so every trace extends Q . This proves (3).

Now suppose $A_k = P$ has height at most two. No strict chain contains three elements, so no subset of the m comparisons can contain distinct x, y, z with both $x < y$ and $y < z$. Transitivity is therefore automatic for every subset. Every subset is irreflexive and asymmetric, hence a partial order, and different subsets give different orders. There are 2^m subsets. $\square$

The proposition separates two notions that are easy to conflate. Agreement-order recovery means $A_k = P$; it identifies every true comparison. It does not collapse the deterministic fiber to $\{P\}$. Every suborder of P still makes the observed traces legal. Likelihood, provenance, or a prior can prefer P , but legality alone cannot.

3 THE CANONICAL DECISION GRAPH

The candidate order Q now becomes a decision space. A subset $I \subseteq V$ is an *order ideal* of Q when

$$v \in I, u <_Q v \implies u \in I.$$

Let

$$\mathcal{I}_t(Q) = \{I \subseteq V : I \text{ is an ideal of } Q, |I| = t\}. \quad (5)$$

Ordered by inclusion, all ideals form the finite distributive lattice associated with Q [3, 4]. They also form the feasible sets of a poset antimatroid [10].

An ideal is thus a feasible execution prefix: whenever it contains an operation, it also contains everything that operation requires. Moving upward by one element means executing one operation whose prerequisites are already present.

Definition 4 (Order-ideal decision graph). The graph $\mathcal{G}(Q)$ has layer $\mathcal{I}_t(Q)$, root $\emptyset$, and an edge

$$I \longrightarrow I \cup \{v\} \quad (6)$$

exactly when all predecessors of v have been completed:

$$\text{pred}_Q(v) := \{u : u <_Q v\} \subseteq I. \quad (7)$$

The vertices are completed-operation prefixes; an edge executes one currently available operation.

Lemma 5 (Paths are serializations). *Root-to- V paths in $\mathcal{G}(Q)$ are in bijection with linear extensions of Q .*

Proof. A maximal path has the form

$$\emptyset = I_0 \subset I_1 \subset \dots \subset I_n = V, \quad |I_t| = t.$$

Let v_t be the unique element of $I_t \setminus I_{t-1}$. Condition (7) implies that every predecessor of v_t appears among $v_1, \dots, v_{t-1}$, so $(v_1, \dots, v_n)$ is a linear extension. Conversely, each prefix of a linear extension contains every predecessor of each of its elements and is therefore an ideal. Its successive prefixes form a path. The two constructions are inverse. $\square$

This is the hinge of the paper: an observed line is a maximal path in the lattice of a graph we do not completely know.

4 A HINDSIGHT ORACLE ON A CANDIDATE GRAPH

We first state the traversal problem on an arbitrary finite layered graph G_Q indexed by a compatible geometry Q . The canonical choice will later be $G_Q = \mathcal{G}(Q)$.

Let the layers be $\mathcal{X}_0^Q, \dots, \mathcal{X}_T^Q$, with root x_0^Q , and assume every vertex is reachable. For $u \in \mathcal{X}_{t-1}^Q$ define

$$R_t^Q(u) = \{v \in \mathcal{X}_t^Q : (u, v) \in E_Q\}.$$

A feasible path $x = (x_1, \dots, x_T)$ receives positive vertex rewards $r_t^Q(x_t)$. Each traversed edge has retention $a_t^Q(x_{t-1}, x_t) \in (0, 1]$. The path's multiplicative value is

$$J_Q(x; Y) = \prod_{t=1}^T a_t^Q(x_{t-1}, x_t) r_t^Q(x_t), \quad (8)$$

where Y is the realized environment. Define movement cost

$$D_Q(x) = - \sum_{t=1}^T \log a_t^Q(x_{t-1}, x_t). \quad (9)$$

The hindsight oracle ranks layer t with a positive surrogate s_t^Q . It knows the realized scores, but follows a local feasible rule. Its path g^Q satisfies

$$s_t^Q(g_t^Q) \geq \eta_t^Q \max_{v \in R_t^Q(g_{t-1}^Q)} s_t^Q(v), \quad 0 < \eta_t^Q \leq 1. \quad (10)$$

The factor η_t^Q absorbs approximate search, rounding, and pruning.

Two different distortions matter. Put

$$h_t^Q(x) = \frac{r_t^Q(x)}{s_t^Q(x)}, \quad \Delta_t^Q = \frac{\max_{x \in \mathcal{X}_t^Q} h_t^Q(x)}{\min_{x \in \mathcal{X}_t^Q} h_t^Q(x)} \geq 1, \quad (11)$$

and

$$\chi_t^Q = \frac{\max_{x \in \mathcal{X}_t^Q} s_t^Q(x)}{\max_{x \in R_t^Q(g_{t-1}^Q)} s_t^Q(x)} \geq 1. \quad (12)$$

The first asks whether the surrogate preserves relative reward. The second asks whether the best-looking state is still reachable.

This distinction is central. The factor Δ_t^Q measures an error in how states are scored; χ_t^Q measures an error created by the path itself. Even a perfect surrogate has $\Delta_t^Q = 1$ but may have $\chi_t^Q > 1$ if an earlier choice has closed the branch containing the layer optimum.

For the statistical statement, let Π_Q be a finite or countable class of feasible causal policies. A fixed $\pi \in \Pi_Q$ induces a path $y^{Q,\pi}(Y)$ and log value

$$L_Q(\pi; Y) = \log J_Q(y^{Q,\pi}(Y); Y).$$

After seeing Σ_k but before seeing the remaining randomness in Y , assign a geometry prior ρ_{Σ_k} on $\mathcal{F}(\Sigma_k)$ and a conditional policy prior q_Q on Π_Q . Set

$$K_Q(\pi) = -\log q_Q(\pi).$$

Assume the conditional lower tail

$$\mathbb{P}(\mathbb{E}[L_Q(\pi; Y) \mid \Sigma_k] - L_Q(\pi; Y) \geq u \mid \Sigma_k) \leq \exp\left(-\frac{u^2}{2V_{Q,\pi}}\right) \quad (13)$$

for every fixed pair (Q, π) and $u \geq 0$.

5 THE HIDDEN-GEOMETRY ORACLE INEQUALITY

Theorem 6 (Serialized hindsight-oracle theorem). *Under the preceding definitions:*

(a) Pathwise certificate. *For every realization, every $Q \in \mathcal{F}(\Sigma_k)$, and every feasible comparator path y ,*

$$\log \frac{J_Q(y; Y)}{J_Q(g^Q; Y)} \leq D_Q(g^Q) - D_Q(y) + \sum_{t=1}^T (\log \Delta_t^Q + \log \chi_t^Q - \log \eta_t^Q). \quad (14)$$

The four coefficients are jointly sharp: for arbitrary $d_t \geq 0$, $\Delta_t \geq 1$, $\eta_t \in (0, 1]$, $\chi_1 = 1$, and $\chi_t \geq 1$ for $t \geq 2$, an instance attains equality with $D(g) = \sum_t d_t$ and $D(y) = 0$.

(b) Uniform geometry-policy inequality. *For every $\delta \in (0, 1)$, with conditional probability at least $1 - \delta$, simultaneously for every $Q \in \mathcal{F}(\Sigma_k)$ and $\pi \in \Pi_Q$,*

$$\begin{aligned} \log J_Q(g^Q; Y) &\geq \mathbb{E}[L_Q(\pi; Y) \mid \Sigma_k] - B_Q(\pi; Y) \\ &\quad - \sqrt{2V_{Q,\pi} \left(\log \frac{1}{\rho_{\Sigma_k}(Q)} + K_Q(\pi) + \log \frac{1}{\delta} \right)}, \end{aligned} \quad (15)$$

where

$$B_Q(\pi; Y) = D_Q(g^Q) - D_Q(y^{Q,\pi}) + \sum_{t=1}^T (\log \Delta_t^Q + \log \chi_t^Q - \log \eta_t^Q). \quad (16)$$

The pair (Q, π) may be selected after observing Y .

(c) Necessity of geometry complexity. *For a uniform geometry prior, the last term contains $\log N_k$. For every $N_k \geq 16$, there is a family satisfying (13), with one policy per geometry and all deterministic terms zero, such that any simultaneous lower guarantee with probability at least $3/4$ requires penalty at least*

$$\sqrt{\frac{V}{2}} \log N_k. \quad (17)$$

Part (a) is a deterministic accounting identity: every unit of regret is assigned to movement, score distortion, approximate maximization, or lost reachability. Part (b) makes that accounting valid even when the analyst chooses both the dependency graph and the comparator after seeing the outcome. Part (c) shows why uncertainty about the graph must then carry a statistical price; it is not removable bookkeeping.

5.1 Proof of the pathwise statement

Fix Q and Y and suppress superscripts. For any comparator choice y_t ,

$$\frac{h_t(g_t)}{h_t(y_t)} \geq \frac{\min_x h_t(x)}{\max_x h_t(x)} = \frac{1}{\Delta_t}.$$

Approximate greediness and accessibility give

$$s_t(g_t) \geq \eta_t \max_{x \in R_t(g_{t-1})} s_t(x) = \frac{\eta_t}{\chi_t} \max_{x \in \mathcal{X}_t} s_t(x) \geq \frac{\eta_t}{\chi_t} s_t(y_t).$$

Since $r_t = s_t h_t$,

$$\log \frac{r_t(y_t)}{r_t(g_t)} \leq \log \Delta_t + \log \chi_t - \log \eta_t. \quad (18)$$

Expanding the path ratio yields

$$\begin{aligned} \log \frac{J(y)}{J(g)} &= \sum_{t=1}^T \log \frac{r_t(y_t)}{r_t(g_t)} + \sum_{t=1}^T \log \frac{a_t(y_{t-1}, y_t)}{a_t(g_{t-1}, g_t)} \\ &= \sum_{t=1}^T \log \frac{r_t(y_t)}{r_t(g_t)} - D(y) + D(g). \end{aligned}$$

Substitution of (18) proves (14).

For joint sharpness, create states g_1, y_1 at the first layer, both reachable from the root. At each $t \geq 2$, create g_t, y_t, u_t and edges

$$g_{t-1} \rightarrow g_t, \quad g_{t-1} \rightarrow u_t, \quad y_{t-1} \rightarrow y_t,$$

with $u_{t-1} \rightarrow u_t$ when defined; omit $g_{t-1} \rightarrow y_t$. Set

$$s_1(g_1) = \eta_1, \quad s_1(y_1) = 1,$$

and for $t \geq 2$ set

$$s_t(g_t) = \eta_t, \quad s_t(u_t) = 1, \quad s_t(y_t) = \chi_t.$$

Thus the reachable maximum from the greedy prefix is 1, while the layer maximum is χ_t . Define

$$h_t(g_t) = h_t(u_t) = 1, \quad h_t(y_t) = \Delta_t, \quad r_t = s_t h_t.$$

Give the greedy edge retention e^{-d_t} and every edge on the y - and u -chains retention 1. The y -chain is globally optimal and

$$J(g) = \prod_t e^{-d_t} \eta_t, \quad J(y) = \prod_t \chi_t \Delta_t.$$

Taking the log ratio gives equality in (14) for every prescribed sequence. This proves part (a).

5.2 Proof of the uniform statement

Condition on Σ_k and give pair (Q, π) joint prior mass

$$w(Q, \pi) = \rho_{\Sigma_k}(Q) q_Q(\pi).$$

These masses sum to one. For a fixed pair set

$$u_{Q, \pi} = \sqrt{2V_{Q, \pi} \log \frac{1}{\delta w(Q, \pi)}}.$$

By (13), the probability that

$$L_Q(\pi; Y) < \mathbb{E}[L_Q(\pi; Y) \mid \Sigma_k] - u_{Q,\pi}$$

is at most $\delta w(Q, \pi)$. Summing over the countable collection of pairs bounds the probability of any failure by δ . Hence, on one simultaneous event,

$$L_Q(\pi; Y) \geq \mathbb{E}[L_Q(\pi; Y) \mid \Sigma_k] - u_{Q,\pi} \quad (19)$$

for every pair. Rearranging (14) with comparator $y^{Q,\pi}$ gives

$$\log J_Q(g^Q; Y) \geq L_Q(\pi; Y) - B_Q(\pi; Y).$$

Insert (19) and use

$$\log \frac{1}{w(Q, \pi)} = \log \frac{1}{\rho_{\Sigma_k}(Q)} + K_Q(\pi).$$

This proves (15). Because the event is simultaneous, later selection of (Q, π) does not invalidate it. The proof is the discrete Occam form underlying broader PAC-Bayesian and description-length principles [22, 23].

5.3 Proof of necessity

Fix $N = N_k \geq 16$. Give every geometry one policy and a one-step graph with one state. Let

$$Z_Q \stackrel{\text{iid}}{\sim} N(0, 1), \quad L_Q = \sqrt{V} Z_Q.$$

Equivalently $J_Q = e^{\sqrt{V} Z_Q} > 0$. Greedy and comparator coincide, and $D = 0$, $\Delta = \chi = \eta = 1$. Gaussian tails satisfy (13).

Put

$$b_0 = \sqrt{\frac{V}{2} \log N}, \quad z_0 = \frac{b_0}{\sqrt{V}} = \sqrt{\frac{1}{2} \log N}.$$

Independence and $1 - x \leq e^{-x}$ imply

$$\mathbb{P}\left(\min_Q L_Q \geq -b_0\right) = (1 - \Phi(-z_0))^N \leq e^{-N\Phi(-z_0)}.$$

The Mills lower bound

$$\Phi(-z) \geq \frac{z}{1+z^2} \frac{e^{-z^2/2}}{\sqrt{2\pi}}, \quad z > 0,$$

and $e^{-z_0^2/2} = N^{-1/4}$ give

$$N\Phi(-z_0) \geq \frac{z_0}{1+z_0^2} \frac{N^{3/4}}{\sqrt{2\pi}}.$$

At $N = 16$ the right side exceeds 1. It increases thereafter because, writing $N = e^{2z_0^2}$, the logarithmic derivative of

$$\frac{z_0}{1+z_0^2} e^{3z_0^2/2}$$

is $z_0^{-1} - 2z_0/(1+z_0^2) + 3z_0 > 0$. Therefore

$$\mathbb{P}(\min_Q L_Q \geq -b_0) < e^{-1} < \frac{3}{4}.$$

A smaller penalty only decreases this probability. Any simultaneous 3/4-confidence lower guarantee must therefore pay at least b_0 , proving part (c).

For a recovered height-two order with m comparisons, Proposition 3 gives $N = 2^m$ and hence a necessary penalty $\sqrt{Vm \log 2/2}$. A complete balanced height-two order has $m = \lfloor n^2/4 \rfloor$, producing order $n\sqrt{V}$ even after its agreement order has been recovered.

6 THE GEOMETRY OF ACCESSIBILITY

We now specialize to $G_Q = \mathcal{G}(Q)$. For equal-rank ideals $I, J \in \mathcal{I}_t(Q)$ define exchange distance

$$d_t(I, J) = \frac{1}{2}|I \Delta J| = t - |I \cap J|. \quad (20)$$

For a prefix $G \in \mathcal{I}_{t-1}(Q)$, its one-step frontier is

$$\text{Succ}_Q(G) = \{G \cup \{v\} : \text{pred}_Q(v) \subseteq G, v \notin G\}.$$

Define the rank- t displacement radius

$$\kappa_t(Q, G) = \max_{I \in \mathcal{I}_t(Q)} \min_{J \in \text{Succ}_Q(G)} d_t(I, J). \quad (21)$$

To see what this radius measures, stand at a prefix G containing $t - 1$ completed operations and compare it with a desirable rank- t ideal I . Every next move must retain all of G and add one operation. Consequently, each element of $G \setminus I$ is a commitment that cannot be repaired in the next step; the projection theorem says that the number of such commitments is not just a lower bound, but the exact distance to the closest reachable state.

Theorem 7 (Ideal projection). *Let Q be a finite poset, $G \in \mathcal{I}_{t-1}(Q)$, and $I \in \mathcal{I}_t(Q)$.*

1. *Any two ideals in $\mathcal{I}_t(Q)$ are connected by exactly their exchange distance in valid remove-and-add exchanges.*
2. *The projection of I onto the successor frontier is exact:*

$$\min_{J \in \text{Succ}_Q(G)} d_t(I, J) = |G \setminus I|. \quad (22)$$

3. *Consequently,*

$$\kappa_t(Q, G) = \max_{I \in \mathcal{I}_t(Q)} |G \setminus I| \leq \min\{t - 1, n - t\}. \quad (23)$$

Proof. Take $I, J \in \mathcal{I}_t(Q)$ with $I \neq J$. Choose y minimal in $J \setminus I$. Every predecessor of y lies in J ; by minimality none lies in $J \setminus I$. Thus

$$\text{pred}_Q(y) \subseteq I \cap J.$$

Choose x maximal in $I \setminus J$. Removing x from I preserves the ideal property. Indeed, if a retained element z depended on x , then either $z \in I \setminus J$, contradicting maximality of x , or $z \in I \cap J$, in which case the ideal J would contain its predecessor x , a contradiction. Therefore

$$I' = (I \setminus \{x\}) \cup \{y\}$$

is an ideal of size t . Moreover $|I' \cap J| = |I \cap J| + 1$, so

$$d_t(I', J) = d_t(I, J) - 1.$$

Iteration produces a path of $d_t(I, J)$ valid exchanges. No exchange can repair more than one element of $I \setminus J$, so fewer are impossible. This proves part 1.

For part 2, $I \setminus G$ is nonempty because $|I| = t > |G| = t - 1$. Choose y minimal in $I \setminus G$. Every predecessor of y lies in I , and minimality forces it into G . Hence $J = G \cup \{y\}$ is a valid successor. Since $y \in I \setminus G$,

$$\begin{aligned} d_t(I, J) &= t - |I \cap J| \\ &= t - (|I \cap G| + 1) \\ &= (t - 1) - |I \cap G| = |G \setminus I|. \end{aligned}$$

Any successor adds only one element to G and can increase overlap with I by at most one, so no successor is closer. This proves (22).

Maximizing (22) gives the equality in (23). Finally, $|G \setminus I| \leq |G| = t - 1$ and $G \setminus I \subseteq V \setminus I$, whose size is $n - t$. The upper bound follows. $\square$

The nested optimization in (21) has disappeared. The radius counts how many elements of the chosen prefix can be absent from a competing ideal at the next rank.

Theorem 8 (Structural accessibility). *Suppose the rank- t log score is L_t -Lipschitz:*

$$|\log s_t(I) - \log s_t(J)| \leq L_t d_t(I, J) \quad (I, J \in \mathcal{I}_t(Q)). \quad (24)$$

If $I_t^* \in \arg \max_{I \in \mathcal{I}_t(Q)} s_t(I)$, then

$$\boxed{\log \chi_t^Q(G) \leq L_t |G \setminus I_t^*| \leq L_t \kappa_t(Q, G).} \quad (25)$$

The last inequality is sharp for every Q, G, L_t .

Proof. By Theorem 7, some $J^* \in \text{Succ}_Q(G)$ satisfies

$$d_t(I_t^*, J^*) = |G \setminus I_t^*|.$$

Therefore

$$\begin{aligned} \log \chi_t^Q(G) &= \log s_t(I_t^*) - \log \max_{J \in \text{Succ}_Q(G)} s_t(J) \\ &\leq \log s_t(I_t^*) - \log s_t(J^*) \\ &\leq L_t d_t(I_t^*, J^*) = L_t |G \setminus I_t^*| \\ &\leq L_t \kappa_t(Q, G). \end{aligned}$$

For sharpness, choose I_t^* maximizing $|G \setminus I|$ and define

$$s_t(I) = \exp[-L_t d_t(I, I_t^*)].$$

The reverse triangle inequality makes $\log s_t$ L_t -Lipschitz. Its global maximum is 1, and exact projection gives

$$\max_{J \in \text{Succ}_Q(G)} s_t(J) = e^{-L_t \kappa_t(Q, G)}.$$

Thus $\log \chi_t^Q(G) = L_t \kappa_t(Q, G)$. $\square$

Combining Theorems 6 and 8 gives the structural certificate

$$\boxed{\log \frac{J_Q(y; Y)}{J_Q(g^Q; Y)} \leq D_Q(g^Q) - D_Q(y) + \sum_{t=1}^T (\log \Delta_t^Q + L_t \kappa_t(Q, G_{t-1}^Q) - \log \eta_t^Q).} \quad (26)$$

The certificate is useful because it replaces the score-dependent reachability ratio χ_t^Q with a geometric quantity determined by the dependency order and the current prefix. Under Lipschitz scores, one unit of exchange distance can lose at most L_t units of log score, and at most κ_t such units separate the next frontier from any rank- t alternative. The product $L_t \kappa_t$ is therefore exactly the worst loss that geometry alone can certify.

Proposition 9 (Monotonicity and the trace-only barrier). *If $Q_1 \subseteq Q_2$ and G is an ideal of Q_2 , then*

$$\kappa_t(Q_1, G) \geq \kappa_t(Q_2, G). \quad (27)$$

Moreover, the empty order $Q_\emptyset$ belongs to every fiber, and for any $(t - 1)$ -element set G ,

$$\kappa_t(Q_\emptyset, G) = \min\{t - 1, n - t\}. \quad (28)$$

Hence for a prefix G of any observed trace,

$$\sup_{Q \in \mathcal{F}(\Sigma_k)} \kappa_t(Q, G) = \min\{t - 1, n - t\}. \quad (29)$$

Proof. Removing comparisons creates ideals, so $\mathcal{I}_t(Q_2) \subseteq \mathcal{I}_t(Q_1)$. The closed form (23) then gives

$$\max_{I \in \mathcal{I}_t(Q_1)} |G \setminus I| \geq \max_{I \in \mathcal{I}_t(Q_2)} |G \setminus I|.$$

This proves (27).

Every permutation extends the empty order, so $Q_\emptyset$ is always compatible and every t -element subset is an ideal. If $n - t \geq t - 1$, choose I disjoint from G , giving $|G \setminus I| = t - 1$. Otherwise choose $V \setminus I$ as any $(n - t)$ -element subset of G , giving $|G \setminus I| = n - t$. The universal upper bound in (23) proves (28). An observed prefix is an ideal of every compatible order, so taking the supremum proves (29). $\square$

The direction of monotonicity matters. Since the true order satisfies $P \subseteq A_k$,

$$\kappa_t(P, G) \geq \kappa_t(A_k, G). \quad (30)$$

The observed agreement radius is a lower bound on hidden-geometry accessibility until enough structure has been recovered. No deterministic manipulation of order-only traces can improve the worst case because the empty order never leaves the legality fiber. Side information, a restricted model class, or a scheduler likelihood is indispensable.

Nor can combinatorics alone control accessibility for arbitrary scores. If $\kappa_t(Q, G) > 0$, choose a nonsuccessor ideal I^* and set $s_t(I^*) = M$, $s_t(I) = 1$ otherwise. Then $\chi_t = M$, which is unbounded as $M \rightarrow \infty$. The Lipschitz condition is not cosmetic; it couples value to geometry.

7 RECOVERY FROM REPEATED TRACES

Let the true order be P . Suppose

$$\sigma_1, \dots, \sigma_k \stackrel{\text{iid}}{\sim} \mu_P, \quad \text{supp}(\mu_P) \subseteq \mathcal{L}(P). \quad (31)$$

Uniform sampling from $\mathcal{L}(P)$ is the reference case, but no uniformity is needed below.

For an unordered incomparable pair $\{u, v\} \in \text{Inc}(P)$, fix either labeling and define

$$p_{uv} = \mathbb{P}_{\sigma \sim \mu_P}(u <_\sigma v). \quad (32)$$

Relabeling replaces p_{uv} by $1 - p_{uv}$ and leaves every expression invariant. Poset dimension is the minimum number of linear extensions whose intersection equals P [5, 6]; it gives a deterministic floor, while the rarest orientation controls random waiting time.

The recovery mechanism is simple. An incomparable pair is mistakenly retained in A_k only while all observed traces happen to orient it the same way. Seeing the pair once in each direction removes that false comparison forever, so the last rare reversal determines the waiting time for full recovery.

Proposition 10 (Dimension floor). *If $\tau(P) = \min\{k : A_k = P\}$, then $\tau(P) \geq \dim(P)$ on every sample path.*

Proof. At recovery, the first $\tau(P)$ traces have intersection P , so they form a realizer of P . Dimension is the smallest cardinality of any realizer. $\square$

Theorem 11 (Repeated-trace transfer). *Define*

$$\varepsilon_k(P, \mu_P) = \sum_{\{u,v\} \in \text{Inc}(P)} [p_{uv}^k + (1 - p_{uv})^k]. \quad (33)$$

Then:

1. *Recovery satisfies*

$$\max_{\{u,v\} \in \text{Inc}(P)} [p_{uv}^k + (1 - p_{uv})^k] \leq \mathbb{P}(A_k \neq P) \leq \varepsilon_k(P, \mu_P). \quad (34)$$

2. *If P is not a chain and*

$$q_{\min} = \min_{\{u,v\} \in \text{Inc}(P)} \min\{p_{uv}, 1 - p_{uv}\} > 0,$$

then

$$\varepsilon_k(P, \mu_P) \leq 2|\text{Inc}(P)|e^{-kq_{\min}}. \quad (35)$$

Thus $\mathbb{P}(A_k = P) \geq 1 - \alpha$ whenever

$$k \geq \frac{1}{q_{\min}} \log \frac{2|\text{Inc}(P)|}{\alpha}. \quad (36)$$

3. *On $A_k = P$,*

$$\mathcal{G}(A_k) = \mathcal{G}(P), \quad \kappa_t(A_k, G) = \kappa_t(P, G) \quad (37)$$

for every valid rank and prefix.

4. *Build $\mathcal{G}(A_k)$ after observing the traces and run the surrogate-greedy oracle there. If (13) and (24) hold, then with joint probability at least $1 - \varepsilon_k(P, \mu_P) - \delta$, the graph is the true graph and, simultaneously for all $\pi \in \Pi_P$,*

$$\begin{aligned} \log J_P(g^P; Y) &\geq \mathbb{E}[L_P(\pi; Y) \mid \Sigma_k] \\ &\quad - \left[D_P(g^P) - D_P(y^{P,\pi}) + \sum_t (\log \Delta_t^P + L_t \hat{\kappa}_{t,k} - \log \eta_t^P) \right] \\ &\quad - \sqrt{2V_{P,\pi} \left(K_P(\pi) + \log \frac{1}{\delta} \right)}, \end{aligned} \quad (38)$$

where the certificate uses the observed quantity

$$\hat{\kappa}_{t,k} = \kappa_t(A_k, G_{t-1}^P) = \kappa_t(P, G_{t-1}^P) \quad \text{on } A_k = P. \quad (39)$$

Proof. Every trace extends P , so $P \subseteq A_k$ on every sample path. Recovery fails exactly when at least one incomparable pair has displayed only one orientation. For fixed $\{u, v\}$, independence makes this event have probability

$$p_{uv}^k + (1 - p_{uv})^k.$$

Keeping one event gives the lower bound in (34); the union bound gives the upper bound.

Put $q_{uv} = \min\{p_{uv}, 1 - p_{uv}\} \leq 1/2$. Then

$$\begin{aligned} p_{uv}^k + (1 - p_{uv})^k &= q_{uv}^k + (1 - q_{uv})^k \\ &\leq 2(1 - q_{uv})^k \leq 2e^{-kq_{uv}} \leq 2e^{-kq_{\min}}. \end{aligned}$$

Summing proves (35); solving its right side $\leq \alpha$ proves (36). A chain has no incomparable pairs and is recovered immediately.

If $A_k = P$, their ideals, ranks, successor relations, and exchange metrics coincide. Definition (21) proves (37).

For the final claim, condition on each realized Σ_k . Assign failure mass $\delta q_{A_k}(\pi)$ to each policy and repeat the weighted-tail argument from Theorem 6. It produces an event $\mathcal{E}_{\text{pol}}$ with

$$\mathbb{P}(\mathcal{E}_{\text{pol}}^c \mid \Sigma_k) \leq \delta, \quad \mathbb{P}(\mathcal{E}_{\text{pol}}^c) \leq \delta.$$

Let $\mathcal{R}_k = \{A_k = P\}$. The first part gives

$$\mathbb{P}(\mathcal{R}_k^c) \leq \varepsilon_k(P, \mu_P).$$

Hence

$$\mathbb{P}(\mathcal{R}_k \cap \mathcal{E}_{\text{pol}}) \geq 1 - \varepsilon_k(P, \mu_P) - \delta.$$

On $\mathcal{R}_k$, the plug-in graph, greedy path, policy paths, and radii are those of P . On $\mathcal{E}_{\text{pol}}$, simultaneously for all policies,

$$L_P(\pi; Y) \geq \mathbb{E}[L_P(\pi; Y) \mid \Sigma_k] - \sqrt{2V_{P,\pi}(K_P(\pi) + \log(1/\delta))}.$$

Apply structural certificate (26) and substitute (39). This proves (38). $\square$

The timing of selection explains why (38) has no geometry-description term. The agreement graph is chosen from traces before the residual environment Y is examined. If an analyst instead selects another $Q \in \mathcal{F}(\Sigma_k)$ after seeing Y , the term $-\log \rho_{\Sigma_k}(Q)$ from Theorem 6 returns.

The complete positive chain is

repeated traces $\longrightarrow$ agreement recovery $\longrightarrow$ true ideal lattice $\longrightarrow$ observed $\kappa_t \longrightarrow$ certified hindsight traversal
--

8 EXACT FINITE EXPERIMENTS

The theorems are finite and exact; the experiment asks whether their intermediate quantities behave nontrivially before the asymptotic recovery event. In particular: does $\kappa_t(A_k, G)$ approach $\kappa_t(P, G)$ smoothly, can the two coincide before $A_k = P$, and how much slack separates the actual greedy gap, exact accessibility, and structural radius?

8.1 Design and exact computation

For even n , split $V = B \cup U$ into equal lower and upper blocks. Include each possible comparison $b < u$, $(b, u) \in B \times U$, independently with probability p . The result has height at most two and is automatically transitively closed.

The main bridge experiment uses

$$n \in \{12, 14, 16, 18, 20\}, \quad p \in \{0.10, 0.50, 0.90\},$$

with 20 independently seeded posets per cell, for 300 posets total. Each produces 512 independent uniform linear extensions, evaluated at

$$k \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512\}.$$

Sampling is exactly uniform. For ideal I , let $C(I)$ be the number of extensions completing the residual order. The down-set recurrence is

$$C(V) = 1, \quad C(I) = \sum_{v: I \cup \{v\} \in \text{Succ}_P(I)} C(I \cup \{v\}). \quad (40)$$

At state I , choose available v with probability $C(I \cup \{v\})/C(I)$. Multiplying these conditional probabilities telescopes to $1/C(\emptyset)$ for every extension. The same ideal enumeration computes κ_t and the global optimum.

The controlled score field isolates accessibility. For each rank, draw a target ideal Z_t as a prefix of an independent uniform linear extension and set

$$\log s_t(I) = -Ld_t(I, Z_t), \quad L = 0.7. \quad (41)$$

The targets are fixed across trace checkpoints. By the reverse triangle inequality, (41) is L -Lipschitz. We take $r_t = s_t$, $a_t = 1$, and $\eta_t = 1$, so movement, surrogate distortion, and approximate optimization vanish. At every checkpoint the experiment computes the exact greedy path, the globally optimal path by dynamic programming, the realized accessibility sum $\sum_t \log \chi_t$, and the structural sum $\sum_t L\kappa_t$.

This score field gives each layer a preferred ideal and makes value decay exponentially with exchange distance from it. It is deliberately transparent: if the greedy prefix closes the preferred branch, the experiment records exactly how much of the loss comes from inaccessibility, without confounding it with an imperfect surrogate or costly movement.

For geometric recovery, $\kappa_t(A_k, G)$ is compared with $\kappa_t(P, G)$ along the prefixes of the first observed trace, a path valid in every compatible geometry. A post-run audit also recomputed the radii along the actual greedy prefixes of every unrecovered final instance. The implementation uses integer bit masks, exact extension counts, and deterministic seed 20260901. A separate pilot used 192 posets through $n = 14$ and seed 20260831.

8.2 Results

All 3,000 poset/checkpoint evaluations satisfy

$$\text{actual log optimality gap} \leq \sum_t \log \chi_t \leq \sum_t L\kappa_t. \quad (42)$$

Thus the experiment exercises both nontrivial inequalities, not only degenerate equality cases. The actual greedy gap is positive in 83.8% of rows and accessibility loss in 89.5%.

Table 1: Recovery of the full agreement order and convergence of the prefixwise displacement radius.

k	$\mathbb{P}(A_k = P)$	mean κ error	k	$\mathbb{P}(A_k = P)$	mean κ error
1	.000	2.4223	32	.270	.0445
2	.000	1.1735	64	.477	.0177
4	.000	.5246	128	.727	.0046
8	.003	.2145	256	.910	.0006
16	.120	.1068	512	.990	.0000

Full recovery reaches 99% at 512 traces. The three remaining cases, two at $n = 18$ and one at $n = 20$, each contain exactly one false comparability. Yet their reference-path radius is exact at every rank. The direct greedy-path audit gives the stronger operational check: in all three cases,

$$\kappa_t(A_{512}, G_{t-1}^{A_{512}}) = \kappa_t(P, G_{t-1}^{A_{512}}) \quad \text{for every } t.$$

The graph is not fully known, but the geometric quantity used by the structural certificate is already right along the path the oracle actually takes.

Conditional on a nonzero denominator, the mean ratio of actual gap to exact accessibility is .270 (median .257), and the mean ratio of accessibility to the structural bound is .450 (median .471). Both inequalities are sometimes exact, while their average slack reflects two different protections: χ_t compares with the score-maximizing rank ideal; κ_t protects against the farthest ideal whether or not it carries the maximum score.

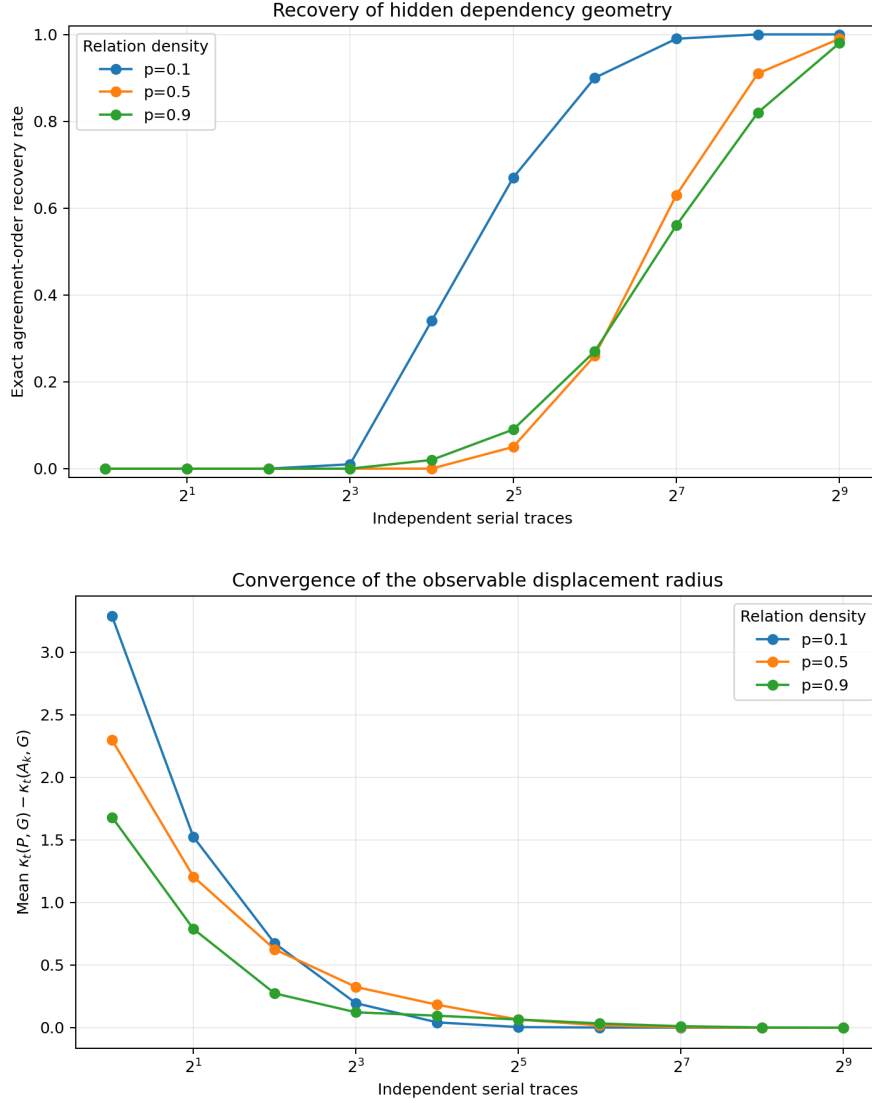


Figure 1: Recovery of hidden dependency geometry. Top: exact agreement-order recovery; sparse orders recover sooner, while medium and dense orders retain rare orientations longer. Bottom: mean difference $\kappa_t(P, G) - \kappa_t(A_k, G)$ along a fixed observed path; monotonicity requires nonnegativity, and the experiment shows convergence to zero. Both panels aggregate over the five sizes.

The exact run took 3,333.5 seconds on an eight-core laptop while using a sequential implementation. Computation grows rapidly because the dynamic program ranges over ideals, but the mathematical checks are auditable: the source records every row, and focused tests compare extension counts, sampling probabilities, projection identities, and oracle values with literal enumeration on small orders.

The earlier 192-poset pilot independently found the same qualitative pattern: all 1,536 inequalities passed, recovery reached 98.4% at 128 traces, and mean radius error was zero even in the three unrecovered cases. The larger bridge experiment therefore confirms rather than discovers the phenomenon.

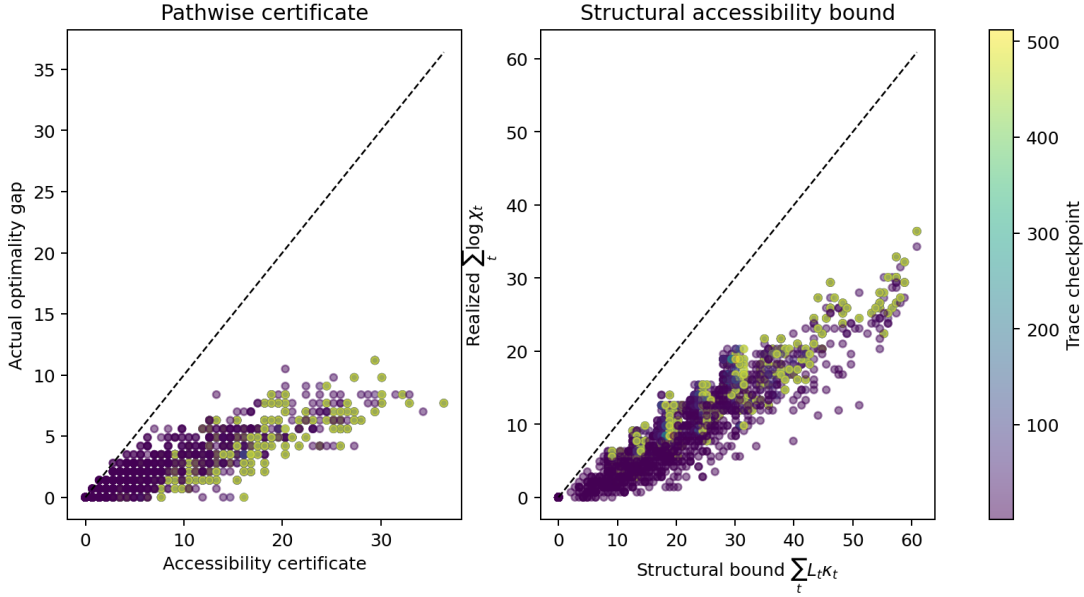


Figure 2: Both stages of the certificate. Every point lies on or below the dashed equality line. Color records the trace checkpoint.

9 WHERE THE RESULT FITS

Concurrency and process discovery. Lamport’s happened-before relation and Mazurkiewicz trace theory make partial order primary and a linearization secondary [11, 12]. Process mining infers workflow structure from event logs and has long confronted concurrency hidden by sequential records [13]. Our observation model is intentionally austere: only labeled order is retained. This permits a prior-free inverse fiber and exact impossibility statement, then asks a downstream decision question rather than only a reconstruction question.

Posets, rankings, and computation. Linear extensions and order ideals are classical, as are dimension and the hardness of counting extensions [7, 6, 8]. Statistical work treats partial orders as latent parameters behind queue-like rankings and introduces structured priors or restricted order classes to make inference possible [14]. We instead keep every deterministically compatible order until a scheduler model is introduced. The rare-orientation theorem makes the scheduler dependence explicit through p_{uv} .

Greedy and online traversal. Precedence scheduling studies policies on a known order [15, 16]; metrical task systems and smoothed online optimization study service and movement on a known metric or convex state space [18, 19, 20]. Learning-augmented variants quantify trust in predictions [21]. Our oracle is an offline diagnostic rather than an online competitive algorithm: it may see the realized reward field, yet its local rule can close future branches. The new term is χ_t , and the ideal theorem turns that term into $L_t \kappa_t$.

Selection after hindsight. The prior-weighted union argument is a classical Occam device, closely related to PAC-Bayesian and MDL reasoning [22, 23]. Its role here is precise. A path selected because it won on the realized environment is not automatically a credible policy target; a geometry selected because it supports that winner is not automatically a credible explanation of the trace. Their description costs add inside one square-root penalty. The Gaussian construction shows that this is not merely an artifact of the proof when candidate geometries can behave independently.

10 LIMITS AND OPEN EDGES

The theory isolates a clean mechanism, and its boundaries are visible.

First, exact graph recovery is sufficient, not necessary, for exact κ along a particular path. The experiment exposes this gap, but a finite-sample theorem for *functional recovery* of κ without full recovery remains open. Such a theorem should depend only on incomparable pairs capable of changing the maximizing ideals in (23), a smaller and path-dependent witness set.

Second, the scheduler model assumes independent traces. Real executions may share caches, priorities, or adaptive state. Mixing or martingale analogues of the orientation bound would widen the model without changing the geometric core.

Third, Lipschitz regularity is essential for a score-independent structural bound. Other regularities, including discrete convexity, submodularity, or curvature on distributive lattices, may replace the worst-radius certificate with a tighter objective-aware quantity.

Fourth, exact ideal enumeration is exponential in the worst case, as expected from the complexity of counting linear extensions [7]. Approximation, structured poset classes, and antimatroid algorithms are natural computational directions [8, 9, 14]. The theorem itself does not require exact enumeration; only the reported experiment does.

Finally, the controlled reward field validates mechanism rather than domain performance. Weighted scheduling objectives, nonuniform schedulers, and natural workflow rewards are the next empirical tests. Movement, surrogate distortion, and approximate maximization were deliberately set to zero here so that geometric inaccessibility could be seen without camouflage.

11 CONCLUSION

The trace begins as evidence and ends as a graph only after assumptions have done their work. Between those points lies a family of possible dependency orders, and each order gives the hindsight rule a different frontier of reachable decisions. The oracle inequality separates the costs incurred along that frontier; the ideal-projection theorem reduces its geometric cost to the explicit radius κ_t ; and the repeated-trace theorem gives a finite-sample condition under which the observed lattice may safely stand in for the hidden one.

The empirical result adds a quieter possibility. A final false comparison need not matter. The order can remain wrong in one place while the relevant displacement, and therefore the certificate along the chosen path, is already exact. Complete reconstruction is a sufficient stopping rule; decision-complete geometry may arrive earlier.

The paper in two sentences. A sequential trace tells us what happened but not what had to happen, so it hides the decision graph faced by a hindsight rule. We prove that one geometric distance controls this loss of reachability, and repeated traces can recover that distance and certify the rule, sometimes before they recover the entire graph.

REPRODUCIBILITY

The theorem note, exact runner, tests, event-level results, and figures are retained in the public repository [24]. The main experiment is in `results/serialized_decision_graphs/bridge_v1`; the independent pilot is in `results/serialized_decision_graphs/pilot_v1`. No stochastic approximation is used in the linear-extension sampler.

REFERENCES

- [1] J. R. Landers, “From dependency geometry to linear traces: Serialization as an information-losing realization of computation,” 2026. Online manuscript.
- [2] J. R. Landers, “An analysis of graph traversal with a hindsight oracle,” 2026. Online manuscript.
- [3] G. Birkhoff, *Lattice Theory*, 3rd ed., American Mathematical Society, 1967.
- [4] R. P. Stanley, *Enumerative Combinatorics*, vol. 1, 2nd ed., Cambridge University Press, 2012.

- [5] B. Dushnik and E. W. Miller, “Partially ordered sets,” *Amer. J. Math.*, 63(3):600–610, 1941.
- [6] W. T. Trotter, *Combinatorics and Partially Ordered Sets: Dimension Theory*, Johns Hopkins University Press, 1992.
- [7] G. Brightwell and P. Winkler, “Counting linear extensions,” *Order*, 8:225–242, 1991.
- [8] G. Pruesse and F. Ruskey, “Generating linear extensions fast,” *SIAM J. Comput.*, 23(2):373–386, 1994.
- [9] G. Pruesse and F. Ruskey, “Gray codes from antimatroids,” *Order*, 10:239–252, 1993.
- [10] B. Korte, L. Lovász, and R. Schrader, *Greedoids*, Springer, 1991.
- [11] L. Lamport, “Time, clocks, and the ordering of events in a distributed system,” *Commun. ACM*, 21(7):558–565, 1978.
- [12] V. Diekert and G. Rozenberg, eds., *The Book of Traces*, World Scientific, 1995.
- [13] W. M. P. van der Aalst, A. J. M. M. Weijters, and L. Mărușter, “Workflow mining: Discovering process models from event logs,” *IEEE Trans. Knowl. Data Eng.*, 16(9):1128–1142, 2004.
- [14] C. Jiang, G. K. Nicholls, and J. E. Lee, “Bayesian inference for vertex-series-parallel partial orders,” in *Proc. UAI*, PMLR 216:995–1004, 2023.
- [15] R. L. Graham, “Bounds for certain multiprocessing anomalies,” *Bell Syst. Tech. J.*, 45(9):1563–1581, 1966.
- [16] R. P. Brent, “The parallel evaluation of general arithmetic expressions,” *J. ACM*, 21(2):201–206, 1974.
- [17] R. D. Blumofe and C. E. Leiserson, “Scheduling multithreaded computations by work stealing,” *J. ACM*, 46(5):720–748, 1999.
- [18] A. Borodin, N. Linial, and M. E. Saks, “An optimal on-line algorithm for metrical task systems,” *J. ACM*, 39(4):745–763, 1992.
- [19] N. Chen, G. Goel, and A. Wierman, “Smoothed online convex optimization in high dimensions via online balanced descent,” in *Proc. COLT*, PMLR 75:1574–1594, 2018.
- [20] G. Goel and A. Wierman, “An online algorithm for smoothed online convex optimization,” *SIGMETRICS Perform. Eval. Rev.*, 47(2):6–8, 2019.
- [21] N. Christianson, J. Shen, and A. Wierman, “Optimal robustness–consistency tradeoffs for learning-augmented metrical task systems,” in *Proc. AISTATS*, PMLR 206:9377–9399, 2023.
- [22] D. A. McAllester, “Some PAC-Bayesian theorems,” *Machine Learning*, 37:355–363, 1999.
- [23] J. Rissanen, “Modeling by shortest data description,” *Automatica*, 14(5):465–471, 1978.
- [24] J. R. Landers, “Trade Deck Research: Serialized decision graph artifacts,” 2026. Public repository.