

Learning the Value of Delayed Information

Wojciech Szpankowski
Purdue University
W. Layayette, U.S.A.

Abstract

Howard and DeGroot price a bit of information by how much it improves a decision the instant it is produced. But no report is instantaneous: a sensor reading or a market quote issued at time $t - d$ is acted on at t , after the world has moved on. How much is such a stale bit still worth? We answer by placing the delay inside the Howard–DeGroot value for a Markov world with transition matrix K . To our knowledge, this is the first formal theory of the value of delayed information. One idea organizes everything: delay acts as a low-pass filter on the world’s dynamics, damping each spectral mode geometrically, so only the slow modes survive, at a base rate set by the second eigenvalue of K . Three consequences follow. First, the world sets the base rate, but the decision sets its multiple, through a sharp trichotomy: a soft, hedged decision loses value at twice the base rate, a knife-edge decision between two actions at the base rate itself, and a committed decision is worth exactly nothing beyond a finite, endogenous deadline. At the knife edge, worth is the square root of Shannon information. Second, the best yes/no question about the past depends on how late the answer arrives: threshold the reward when it is fresh, but the slowest mode of K , the only signal a late answer still carries, when it is late. To leading order, this single bit is all a late report is worth. Third, delayed value is learnable from a length- n trajectory only up to a horizon: logarithmic in n for model-free plug-in estimators, but of order $\sqrt{n}$ for temporal information once K is fitted, as confirmed on a chain fitted to Beijing air-quality data.

1 Introduction

Ronald Howard’s value of information (Howard, 1966) asks how much a decision-maker’s expected reward improves by learning a bit of information before acting, $\text{VoI} = \mathbb{E}_x \max_a \mathbb{E}[u(Y, a) \mid x] - \max_a \mathbb{E}[u(Y, a)]$ (see also Raiffa and Schlaifer, 1961). DeGroot (1962) freed the same quantity from any particular reward by measuring it as the expected drop $U(\pi) - \mathbb{E}_x U(P(Y \mid x))$ of a concave uncertainty functional of the posterior, a reward-free view that Blackwell’s comparison of experiments (Blackwell, 1953) makes canonical. The Shannon instance is Lindley’s (Lindley, 1956), and Bernardo (1979) and Frankel and Kamenica (2019) characterize which uncertainty functionals arise from decisions. All are instantaneous: the signal concerns the very variable on which one acts. Yet on any real system (e.g., a sensor feed, a market, a delayed measurement) the world emits knowledge about its state at time $t - d$, and the decision is made at time t , on a state that has meanwhile evolved. Put simply, the problem is this. The world emits a piece of knowledge about its state at time $t - d$; the decision is made at time t , on a state that has meanwhile evolved. How much does the decision-maker gain from that delayed piece of knowledge? Throughout, a *stale* report or bit means one that is delayed in this sense. A primer on Howard and DeGroot instantaneous values is in Appendix B.

Delayed information has been studied for its freshness (age of information; Kaul et al., 2012; Sun et al., 2017; Yates et al., 2021) and for the mismatch it induces (age of incorrect information

(Maatouk et al., 2020), uncertainty of information (Chen and Liew, 2024)); these score a feed by time or by estimation error. Delayed-observation MDPs (Katsikopoulos and Engelbrecht, 2003; Walsh et al., 2009) optimize a policy given the delay, and Shaiderman and Solan (2026) study value of information in dynamic decision problems, but not as a function of a delay on a fixed Markov world. Our question is the one Howard asked, now as a function of delay: how much can a delayed bit be worth, and which bit of a stale report is worth reading when it will arrive late? Our object is the Howard/DeGroot value with a clock inserted—the signal concerns S_{t-d} but the action lands on S_t —so worth must flow through the world’s dynamics. We keep DeGroot’s reward-free functional (our $\Phi = -U$) and recover Howard’s and DeGroot’s quantities exactly at $d = 0$; monotonicity in d is a Blackwell garbling (Theorem 1). Everything past $d = 0$ (i.e., aging exponents, the information/worth decoupling, endogenous deadlines, rotating eigen-bits, and a learnability horizon) is what the delay adds.

Let the world be a finite, irreducible, aperiodic Markov chain S_t with kernel K , stationary law π , and eigenvalues $1 = \lambda_1 > |\lambda_2| \geq \dots$. A report of $S_{t-d} = s$ tells the decision-maker, about the acting state S_t , exactly the row $K^d(s, \cdot) = \pi + \varepsilon_s^{(d)}$, where $\varepsilon_s^{(d)} = \sum_{k \geq 2} \lambda_k^d \psi_k(s)(\pi \psi_k)$ scales the k -th eigenmode by λ_k^d . So *delay acts as a low-pass filter on the dynamics*: it passes the slow modes (large $|\lambda_k|$) and kills the fast ones, with a cutoff that sharpens as d grows. A decision is a direction in belief space (what to pay to know) and its delayed value is how much of that direction survives the filter. The filter also turns the direction itself: generically, as d grows, the most valuable bit drifts from the reward toward the world’s slowest mode.

Contributions. (i) A reward-free definition of delayed value (Section 2) and a tight DeGroot ceiling $\bar{\mathcal{V}}(d) = \Theta(|\lambda_2|^d)$ over all span-one (TV-Lipschitz) decisions (Section 3). (ii) The aging trichotomy and the temporal value spectrum (Section 4): a generically quantized exponent $\beta_\Phi = -\lim_d \frac{1}{d} \log \mathcal{V}_\Phi(d) \in \{2\mu, \mu, \infty\}$ with $\mu = \log(1/|\lambda_2|)$. (iii) The rotation of the optimal one-bit summary (Section 5): the most durable stale bit thresholds ψ_2 , and the optimal question rotates with d , driven by a backward recurrence that pulls the reward back. (iv) Learning the value curve (Section 6): a model-free horizon $d^* = \log n / 2\mu$ that no data-independent correction extends; a model-based horizon of order $\sqrt{n}$, sharp for temporal information; and $\mathcal{V}_\Phi(d)$ as the regret advantage of a learner who sees the world late. (v) Numerical confirmations on a chain fitted to Beijing air-quality data (Section 7).

2 Setup

Let $\mathcal{S} = \{1, \dots, N\}$, K reversible Markov (see Appendix F for an extension to general Markov) w.r.t. π , with π -orthonormal eigenfunctions $\psi_1 \equiv 1, \psi_2, \dots, \psi_N$ (i.e. $\langle \psi_j, \psi_k \rangle_\pi = \sum_s \pi(s) \psi_j(s) \psi_k(s) = \delta_{jk}$) and real eigenvalues $1 = \lambda_1 > |\lambda_2| \geq \dots$ (Levin and Peres, 2017). The spectral representation of the d -step kernel K^d is

$$K^d(s, s') = \pi(s') \left[1 + \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k(s') \right], \quad \varepsilon_s^{(d)}(s') := K^d(s, s') - \pi(s') = \pi(s') \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k(s'). \quad (1)$$

Note $\sum_{s'} \varepsilon_s^{(d)}(s') = 0$ (each $\psi_k \perp_\pi \psi_1$) and, crucially, $\sum_s \pi(s) \varepsilon_s^{(d)} = \pi K^d - \pi = 0$: the prior is preserved under the delayed experiment, which is what kills linear terms below (see Appendix A).

A *decision* is a bounded convex payoff $\Phi : \Delta(\mathcal{S}) \rightarrow \mathbb{R}$; the leading case is $\Phi(b) = \max_{a \in \mathcal{A}} \langle b, r_a \rangle$ with rewards $r_a \in \mathbb{R}^{\mathcal{S}}$ (Howard’s value of information). DeGroot uncertainty is $U = -\Phi$, concave.

A perfect report of S_{t-d} induces, on the acting state S_t , the posterior family $\{K^d(s, \cdot)\}_s$ with weights $\pi(s)$.

Definition 1 (Delayed value). For a convex payoff Φ , the value of d -delayed information is

$$\mathcal{V}_\Phi(d) := \mathbb{E}_{s \sim \pi}[\Phi(K^d(s, \cdot))] - \Phi(\pi) = \mathbb{E}_{s \sim \pi}[\Phi(\pi + \varepsilon_s^{(d)})] - \Phi(\pi). \quad (2)$$

We call $\Phi(\pi)$ the *blind payoff*: the value of acting on the prior alone, with no report; $\mathcal{A}^* = \arg \max_a \langle \pi, r_a \rangle$ is the set of blind-optimal actions.

The delay lives entirely in the belief argument of a fixed functional. Two canonical choices: $\Phi(b) = \max_a \langle b, r_a \rangle$ gives Howard's value of the delayed report; $\Phi(b) = -H(b)$ (neg-entropy) gives the *temporal information*

$$\mathcal{T}(d) = I(S_{t-d}; S_t) = H(\pi) - \mathbb{E}_s H(K^d(s, \cdot)) \quad (3)$$

of the state process (Szpankowski and Vaccaro, 2026).

Theorem 1 (Well-posedness). For every convex Φ : (i) $\mathcal{V}_\Phi(d) \geq 0$; (ii) $\mathcal{V}_\Phi(0) = \mathbb{E}_{s \sim \pi} \Phi(\delta_s) - \Phi(\pi)$ is the (undelayed) DeGroot value, and $\mathcal{V}_\Phi(\infty) = 0$; (iii) $\mathcal{V}_\Phi$ is non-increasing in d .

Proof. (i) $\mathbb{E}_s K^d(s, \cdot) = \pi K^d = \pi$, so by Jensen $\mathbb{E}_s \Phi(K^d(s, \cdot)) \geq \Phi(\pi)$. (ii) $K^0(s, \cdot) = \delta_s$; and $K^d(s, \cdot) \rightarrow \pi$ by ergodicity, whence $\Phi(K^d(s, \cdot)) \rightarrow \Phi(\pi)$ because a convex function is continuous at the interior point π (see Appendix B). (iii) By Chapman–Kolmogorov, $K^{d+1}(s, \cdot) = \sum_{s'} K(s, s') K^d(s', \cdot)$, so by Jensen and $\pi K = \pi$, $\mathbb{E}_s \Phi(K^{d+1}(s, \cdot)) \leq \sum_{s'} \pi(s') \Phi(K^d(s', \cdot))$; that is, the $(d+1)$ -report is a Blackwell garbling of the d -report (Blackwell, 1953). Subtract $\Phi(\pi)$. $\square$

Remark 1 (Monotonicity is a Markov property; a small delay can help). Only part (iii) uses that the world is Markov, through the conditional independence $S_{t-d-1} \perp S_t \mid S_{t-d}$, which makes the older report a garbling of the newer one. Parts (i)–(ii) need only stationarity and ergodicity, and the value itself is well-defined for any stationary source through the two-point law $Q_d(s, \cdot) = \mathcal{L}(S_t \mid S_{t-d} = s)$ (Markov is the special case $Q_d = K^d$) so that $\mathcal{V}_\Phi(d) = \mathbb{E}_{s \sim \pi}[\Phi(Q_d(s, \cdot))] - \Phi(\pi)$. When S_t is a function of a larger Markov chain—so Q_d is not the d -th power of any kernel on $\mathcal{S}$ —monotonicity can fail, and a staler report can be worth strictly more. By Blackwell's theorem (Blackwell, 1953), $\mathcal{V}_\Phi(d+1) \leq \mathcal{V}_\Phi(d)$ for every convex Φ if and only if the experiment Q_{d+1} is a garbling of Q_d .

Example 1 (Alternating channel). Let $X_t \in \{\pm 1\}$ be a symmetric two-state chain with $\mathbb{E}[X_t X_{t+m}] = \rho^m$, let B_t be i.i.d. fair ± 1 , and let a phase $B \in \{0, 1\}$ be uniform and independent of both. Observe $S_t = X_t$ for $t \equiv B \pmod{2}$ and $S_t = B_t$ otherwise, a stationary, uniform, non-Markov process (the phase never mixes, so no finite window of S is a sufficient statistic). Act on S_t from the report S_{t-d} . If d is odd the two slots have opposite parity, hence sit on opposite channels, so $S_{t-d} \perp S_t$ and $\mathcal{V}_\Phi(d) = 0$ for every Φ . If d is even they share a channel, and averaging over the unobserved phase gives $Q_d(s, \sigma) = \frac{1}{2}(1 + \frac{1}{2}\rho^d s \sigma)$ —a diluted bit of strength $\frac{1}{2}\rho^d$. For the boundary decision with contrast $v(s) = s$ ($\langle \pi, v \rangle = 0$), $\mathcal{V}_\Phi(d) = \frac{1}{4}\rho^d$ for d even and 0 when d is odd; its neg-entropy reading is $\mathcal{T}(d) = \frac{1}{2}[(1 + \varepsilon) \log(1 + \varepsilon) + (1 - \varepsilon) \log(1 - \varepsilon)]$ with $\varepsilon = \frac{1}{2}\rho^d$ for even d , and $\mathcal{T}(d) = 0$ for odd d . The freshest report ($d = 1$) lands on the noise tick and is worthless; stepping one further back ($d = 2$) returns to a world tick, so $\mathcal{V}_\Phi(1) = 0 < \mathcal{V}_\Phi(2) = \frac{1}{4}\rho^2$. For $\rho = 0.9$ we have $\mathcal{V}_\Phi = (0.500, 0, 0.203, 0, 0.164, \dots)$.

3 The DeGroot ceiling

How much value can any decision extract from a report about S_{t-d} ? Normalize decisions to be 1-Lipschitz in total variation—this contains every Howard decision whose rewards have span at most one. The envelope over this class is the total-variation (statistical) information, which is the DeGroot value of subset betting.

Theorem 2 (Ceiling). *Let $\bar{\mathcal{V}}(d) := \sum_s \pi(s) \|K^d(s, \cdot) - \pi\|_{\text{TV}}$. For every payoff Φ that is 1-Lipschitz in total variation,*

$$|\Phi(b) - \Phi(b')| \leq \|b - b'\|_{\text{TV}} \quad \text{for all } b, b' \in \Delta(\mathcal{S}), \quad (4)$$

and in particular for every Howard decision $\Phi(b) = \max_a \langle b, r_a \rangle$ whose rewards have span $\max_s r_a(s) - \min_s r_a(s) \leq 1$,

$$\mathcal{V}_\Phi(d) \leq \bar{\mathcal{V}}(d), \quad (5)$$

with equality for subset betting $\{r_A(s) = \mathbf{1}[s \in A] - \pi(A)\}_{A \subseteq \mathcal{S}}$ (so $\bar{\mathcal{V}}$ is a DeGroot value, not a bound). Moreover, if $|\lambda_2| > |\lambda_3|$ (the leading rate is simple),

$$\bar{\mathcal{V}}(d) = \frac{1}{2} \|\psi_2\|_{1,\pi}^2 |\lambda_2|^d + O(|\lambda_3|^d) = \Theta(|\lambda_2|^d), \quad (6)$$

where $\|\psi_2\|_{1,\pi} = \sum_s \pi(s) |\psi_2(s)|$.

Proof. By (4) with $b = K^d(s, \cdot)$ and $b' = \pi$, $\Phi(K^d(s, \cdot)) - \Phi(\pi) \leq \|K^d(s, \cdot) - \pi\|_{\text{TV}}$; averaging over $s \sim \pi$ gives (5). It remains to check that a Howard decision with span at most one satisfies (4). Fix b, b' and let a^* be optimal at b . Then $\Phi(b) - \Phi(b') \leq \langle b, r_{a^*} \rangle - \langle b', r_{a^*} \rangle = \langle b - b', r_{a^*} \rangle$. Since $\sum_s (b(s) - b'(s)) = 0$ we may recenter r_{a^*} by any constant; choosing $c = \frac{1}{2}(\max r_{a^*} + \min r_{a^*})$ gives $\|r_{a^*} - c\mathbf{1}\|_\infty \leq \frac{1}{2}$, hence $\langle b - b', r_{a^*} \rangle = \langle b - b', r_{a^*} - c\mathbf{1} \rangle \leq \|b - b'\|_1 \cdot \frac{1}{2} = \|b - b'\|_{\text{TV}}$. Exchanging b and b' gives (4). Equality: for subset betting $\Phi_{\text{TV}}(b) = \max_A (b(A) - \pi(A)) = \|b - \pi\|_{\text{TV}}$ and $\Phi_{\text{TV}}(\pi) = 0$; each r_A has span at most one. For (6), insert the leading ($k = 2$) term of (1): $\|\varepsilon_s^{(d)}\|_{\text{TV}} = \frac{1}{2} |\lambda_2|^d |\psi_2(s)| \|\psi_2\|_{1,\pi} + O(|\lambda_3|^d)$, and average using $\mathbb{E}_s |\psi_2(s)| = \|\psi_2\|_{1,\pi}$. $\square$

4 Aging and temporal value

How does $\mathcal{V}_\Phi(d)$ decay with d ? The ceiling is first order in $|\lambda_2|^d$; temporal information is second order; a generic hard decision has finite support in d . The exponent is set by the local shape of Φ at the prior: since $\mathbb{E}_s \varepsilon_s^{(d)} = 0$, the linear part of Φ averages out and the first surviving term decides. A Howard payoff $\max_a \langle \cdot, r_a \rangle$ splits $\Delta(\mathcal{S})$ into cells, one per action; it is linear inside a cell and kinked on the walls between cells. Three geometries occur. A *smooth* (hedged) payoff curves at π . In the *generic hard* case π lies inside one cell with a positive margin (true for Lebesgue-a.e. rewards). In the *boundary* case π lies on a wall, indifferent between two or more actions: non-generic, but decision-relevant, since a report is worth reading only where it can change the action, and a decision-maker who tunes rewards or thresholds to be indifferent at the prior sits exactly there.

Theorem 3 (Aging trichotomy). *Let $\lambda_2 \neq 0$, $\mu = \log(1/|\lambda_2|)$, and let $\underline{\beta}_\Phi, \bar{\beta}_\Phi$ be the $\liminf$ and $\limsup$ of $-\frac{1}{d} \log \mathcal{V}_\Phi(d)$ ($-\log 0 = \infty$), with common value β_Φ .*

(S) **Smooth.** *If $\Phi \in C^2$ near π , put $m_k(s') = \pi(s') \psi_k(s')$ and $w_k = m_k^\top \nabla^2 \Phi(\pi) m_k \geq 0$. Then*

$$\mathcal{V}_\Phi(d) = \frac{1}{2} \sum_{k \geq 2} w_k \lambda_k^{2d} + o(|\lambda_2|^{2d}) = \int x^d d\nu_\Phi(x) + o(|\lambda_2|^{2d}), \quad (7)$$

with temporal value spectrum $\nu_\Phi = \frac{1}{2} \sum_{k \geq 2} w_k \delta_{\lambda_k^2}$. If $w_2^* = \sum_{k: |\lambda_k| = |\lambda_2|} w_k > 0$, then $\mathcal{V}_\Phi(d) \sim \frac{1}{2} w_2^* |\lambda_2|^{2d}$ and $\beta_\Phi = 2\mu$; otherwise $\mathcal{V}_\Phi(d) = o(|\lambda_2|^{2d})$. For $\Phi = -H$ ($w_k \equiv 1$), $\mathcal{T}(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d} + O(|\lambda_2|^{3d})$.

(B) **Boundary.** Let $\Phi = \max_{a \in \mathcal{A}} \langle \cdot, r_a \rangle$, $\mathcal{A}$ finite, with $|\mathcal{A}^*| \geq 2$ for $\mathcal{A}^* = \arg \max_a \langle \pi, r_a \rangle$. Put $g(\varepsilon) = \max_{a \in \mathcal{A}^*} \langle \varepsilon, r_a \rangle$ and $\sigma_2 = \max_{\mathcal{A}^*} \langle \psi_2, r_a \rangle_\pi - \min_{\mathcal{A}^*} \langle \psi_2, r_a \rangle_\pi$. Then $\mathcal{V}_\Phi(d) = \mathbb{E}_s g(\varepsilon_s^{(d)})$ for $d \geq d_1$ (some $d_1 \geq 0$; $d_1 = 0$ if $\mathcal{A} = \mathcal{A}^*$), and if $|\lambda_2| > |\lambda_3|$,

$$\mathcal{V}_\Phi(d) = \frac{1}{2} |\lambda_2|^d \sigma_2 \|\psi_2\|_{1,\pi} + O(|\lambda_3|^d). \quad (8)$$

So $\beta_\Phi = \mu$ iff $\sigma_2 > 0$ (the blind optima disagree along ψ_2); otherwise $\beta_\Phi \geq \log(1/|\lambda_3|) > \mu$. For two optimizers, $v = r_{a_1} - r_{a_2}$ ($\langle \pi, v \rangle = 0$) gives $\sigma_2 = |\langle v, \psi_2 \rangle_\pi|$ and, for $d \geq d_1$,

$$\mathcal{V}_\Phi(d) = \frac{1}{2} \mathbb{E}_s |\langle \varepsilon_s^{(d)}, v \rangle|. \quad (9)$$

(Multiple $|\lambda_2|$: Appendix C.)

(H) **Generic hard.** If $\mathcal{A}^* = \{a^*\}$ with margin $\delta = \min_{a \neq a^*} \langle \pi, r_{a^*} - r_a \rangle > 0$, then $\mathcal{V}_\Phi(d) = 0$ for all $d \geq D_0$, where

$$D_0 \leq \max \left\{ 0, \left\lceil \frac{\log \left(\frac{\rho}{2\delta} \sqrt{\frac{1-\pi_{\min}}{\pi_{\min}}} \right)}{\log(1/|\lambda_2|)} \right\rceil \right\}, \quad (10)$$

$$\rho = \max_{a \neq a^*} \text{span}(r_{a^*} - r_a), \quad \pi_{\min} = \min_s \pi(s).$$

Proof sketch. (S) Taylor-expand $\Phi(\pi + \varepsilon)$ to second order; $\mathbb{E}_s \varepsilon_s^{(d)} = 0$ kills the linear term and $\mathbb{E}_s \psi_k \psi_l = \delta_{kl}$ diagonalizes the quadratic. (B) Near π only $\mathcal{A}^*$ competes, so $\Phi(\pi + \varepsilon) = \Phi(\pi) + g(\varepsilon)$; insert the leading term $\lambda_2^d \psi_2(s) m_2$ of (1) and use positive homogeneity of g with $\mathbb{E}_s \psi_2^\pm = \frac{1}{2} \|\psi_2\|_{1,\pi}$. (H) Beyond the right side of (10) every row $K^d(s, \cdot)$ lies in the cell of a^* , so $\mathbb{E}_s \Phi(K^d(s, \cdot)) = \langle \pi K^d, r_{a^*} \rangle = \Phi(\pi)$; monotonicity (Theorem 1) does the rest. Details in Appendix C. $\square$

The deadline in (H) is *endogenous*: no one imposed it. Its bound grows with $\log(\rho/\delta)$, the stakes relative to the margin; doubling δ lowers it by $\log 2/\mu$ steps, and it diverges in a slow world ($\mu \rightarrow 0$).

Reading the spectrum. In (7) a decision is a curvature $\nabla^2 \Phi(\pi)$, the world is a spectrum $\{\lambda_k^2\}$, and delayed value is their pairing ν_Φ , read out moment by moment as d grows: w_k is how much Φ bends when the belief moves from π along the k -th mode m_k . Temporal information bends equally in every direction ($w_k \equiv 1$). A finite-action decision (cases B, H) is not a moment sequence of ν_Φ : generically its leading rate is $|\lambda_2|$ (a kink at π), or its value is finitely supported in d (no kink at π)—that is the content of the trichotomy.

Example 2 (One family, three regimes). Let $K = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$, so $\pi = (\frac{1}{2}, \frac{1}{2})$, $\lambda_2 = \lambda = 1 - 2p$, $\psi_2 = (1, -1)$, and the report $S_{t-d} = s$ yields the belief $b_1 = \frac{1}{2} \pm \frac{1}{2} \lambda^d$. Let $\Phi_c(b) = \max(b_1, 1 - b_1, c)$: bet on state 1, bet on state 2, or take a safe payoff $c < 1$. (B) If $c \leq \frac{1}{2}$, the two bets tie at π , which sits on the wall between their cells; near π , $\Phi_c(b) - \Phi_c(\pi) = |b_1 - \frac{1}{2}|$, so $\mathcal{V}_\Phi(d) = \frac{1}{2} |\lambda|^d$ and $\beta_\Phi = \mu$. (H) If $c > \frac{1}{2}$, the hedge is blind-optimal with margin $\delta = c - \frac{1}{2}$ and $\mathcal{V}_\Phi(d) = (\frac{1}{2} |\lambda|^d - \delta)^+$, which vanishes for $d \geq D_0 = \lceil \log(1/2\delta)/\mu \rceil$; here $\rho = 1$, $\pi_{\min} = \frac{1}{2}$, so (10) is tight. (S) For the smooth $\Phi = -H$, $\mathcal{T}(d) = \log 2 - h(\frac{1}{2} + \frac{1}{2} |\lambda|^d) = \frac{1}{2} \lambda^{2d} + O(\lambda^{4d})$, with h the binary entropy, and $\beta_\Phi = 2\mu$. In all three, $\varepsilon_s^{(d)}$ has mean zero and size $|\lambda|^d$, so the first term of Φ that survives averaging sets the rate: a kink leaves $\mathbb{E}|\varepsilon| \sim |\lambda|^d$, curvature leaves $\mathbb{E}\varepsilon^2 \sim |\lambda|^{2d}$, a flat cell leaves 0.

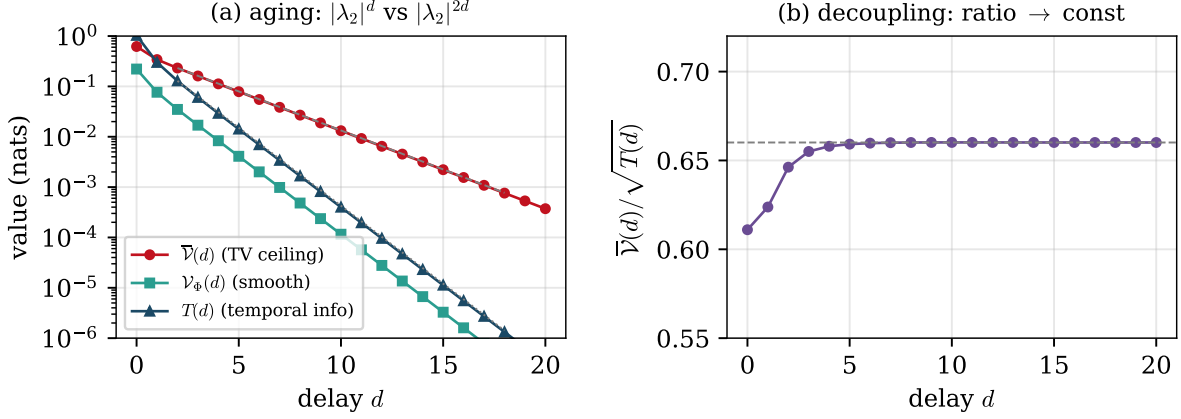


Figure 1: **Aging laws and decoupling.** (a) The ceiling $\bar{\mathcal{V}}(d)$ decays as $|\lambda_2|^d$ (fitted slope $-0.357 = -\mu$), while temporal information $\mathcal{T}(d)$ and a smooth soft-decision value decay as $|\lambda_2|^{2d}$ (fitted slope $-0.713 = -2\mu$). Grey guides have exact slopes. (b) The decoupling: $\bar{\mathcal{V}}(d)/\sqrt{\mathcal{T}(d)}$ converges to a chain constant ($\approx \|\psi_2\|_{1,\pi}^2/\sqrt{2}$), confirming $\mathcal{V} \asymp \sqrt{\mathcal{T}}$.

Corollary 1 (Information/worth decoupling). *Let $\beta_{\mathcal{T}} := \beta_{-H} = 2\mu$. In case (B) with $\sigma_2 > 0$ and $|\lambda_2| > |\lambda_3|$, $\mathcal{V}_{\Phi}(d) = \Theta(\sqrt{\mathcal{T}(d)})$: decision value is the square root of the temporal information. The conversion exponent $\kappa = \beta_{\Phi}/\beta_{\mathcal{T}}$ is $\frac{1}{2}, 1, \infty$ in the nondegenerate cases (B), (S) with $w_2^* > 0$, and (H).*

Remark 2 (Non-reversible chains). Reversibility is used only to make the eigenfunctions π -orthonormal. For any irreducible aperiodic K , take right eigenvectors ψ_k and left eigenvectors m_k with $m_j\psi_k = \delta_{jk}$ (Jordan blocks allowed). Theorem 1, the ceiling inequality (5), and the backward identity hold verbatim, and the exponents $\{2\mu, \mu, \infty\}$ survive up to a factor d^{m-1} from a Jordan block of size m at $|\lambda_2|$. Three things change (Appendix F). (i) Non-normality adds value: the χ^2 value is half the sum of the squared singular values of K^d , at least $\frac{1}{2} \sum_{k \geq 2} |\lambda_k|^{2d}$, with equality iff K is normal in $\ell^2(\pi)$. (ii) A complex pair $\lambda_2 = |\lambda_2|e^{i\theta}$ never cancels the leading term, so β_{Φ} exists; but the eigen-bit becomes a phase-bit $\mathbf{1}[\text{Re}(e^{i(d\theta+\phi)}\psi_2) > 0]$ that turns by θ per step and never freezes. (iii) $\mathcal{T}$ and $\bar{\mathcal{V}}$ are time-symmetric; Howard values are not; the estimator in Proposition 2 must be replaced by a one-step kernel.

5 What to ask when late

Suppose we may learn only one yes/no fact about the past state, $Y = f(S_{t-d})$ with $f: \mathcal{S} \rightarrow \{0, 1\}$ fixed in advance (possibly depending on d): a one-bit sensor, or a single question about S_{t-d} . The bit is a compression of the state, not the action; the receiver still chooses from all of $\mathcal{A}$. The question f splits $\mathcal{S}$ into cells C_0, C_1 . Which question is worth asking, and does the answer depend on how late it will be read? Having seen $Y = y$, the receiver's Bayes belief about the acting state S_t is

$$b_y(s') = \sum_{s \in C_y} \frac{\pi(s)}{\pi(C_y)} K^d(s, s'), \quad s' \in \mathcal{S}. \quad (11)$$

We treat a two-action boundary decision (several actions, Appendix C.7) with contrast $v = r_{a_1} - r_{a_2}$, $\langle \pi, v \rangle = 0$; the receiver plays a_1 iff $\langle b_y, v \rangle > 0$. The delay- d effective score

$$g_d(s) = \langle \varepsilon_s^{(d)}, v \rangle = \sum_{k \geq 2} \lambda_k^d \langle v, \psi_k \rangle_\pi \psi_k(s) \quad (12)$$

is the expected advantage of a_1 at time t given $S_{t-d} = s$, and $\mathbb{E}_s g_d = \langle \mathbb{E}_s \varepsilon_s^{(d)}, v \rangle = \langle \pi K^d - \pi, v \rangle = 0$. As a vector in $\ell^2(\pi)$ its k -th coordinate is damped by λ_k^d , so its direction turns from the reward contrast v ($d = 0$) toward the slowest mode ψ_2 ($d \rightarrow \infty$, when $\langle v, \psi_2 \rangle_\pi \neq 0$). Figure 2b traces this rotation through $|\cos_\pi(g_d, v)|$ and $|\cos_\pi(g_d, \psi_2)|$. See also Figure 5 in Appendix C.4.

Theorem 4 (Best one-bit question and its rotation). *For a two-action boundary decision,*

$$\mathcal{V}_f(d) = \left| \sum_{s \in C_0} \pi(s) g_d(s) \right| = \left| \sum_{k \geq 2} \lambda_k^d \langle v, \psi_k \rangle_\pi \langle \mathbf{1}_{C_0}, \psi_k \rangle_\pi \right|, \quad (13)$$

so the optimal question is $f_d^* = \mathbf{1}[g_d > 0]$, with value $\frac{1}{2} \|g_d\|_{1,\pi}$. Consequently the optimal question rotates with delay:

$$\begin{aligned} d = 0 : \quad & g_0 = v, \quad f_0^* = \mathbf{1}[v > 0] \quad (\text{threshold reward}); \\ d \rightarrow \infty : \quad & g_d = \lambda_2^d \langle v, \psi_2 \rangle_\pi (\psi_2 + O((|\lambda_3|/|\lambda_2|)^d)), \quad f_d^* = \mathbf{1}[\tau_d \psi_2 > 0], \quad \tau_d = \text{sign}(\lambda_2^d \langle v, \psi_2 \rangle_\pi), \end{aligned}$$

i.e. threshold the slow mode, when $\langle v, \psi_2 \rangle_\pi \neq 0$ and $|\lambda_2| > |\lambda_3|$ (τ_d only relabels the same partition). For every f and v ,

$$\mathcal{V}_f(d) \leq |\lambda_2|^d \|v\|_{2,\pi} \sqrt{\pi(C_0)\pi(C_1)} \leq \frac{1}{2} |\lambda_2|^d \|v\|_{2,\pi}, \quad (14)$$

and if $|\lambda_2| > |\lambda_3|$, any question with $\langle \mathbf{1}_{C_0}, \psi_2 \rangle_\pi = 0$ ages strictly faster, $\mathcal{V}_f(d) = O(|\lambda_3|^d)$. So the best delay-robustness of one bit is $\Theta(|\lambda_2|^d)$, achieved by the eigen-bit $\mathbf{1}[\psi_2 > 0]$.

Proof sketch. By Bayes plausibility the value is $\frac{1}{2} \sum_y |G_y|$ with $G_y = \sum_{s \in C_y} \pi(s) g_d(s)$; since $G_0 + G_1 = 0$, this is $|G_0|$, maximized by $\{g_d > 0\}$. Bound (14) is Cauchy-Schwarz and Parseval in (13). Details in Appendix C.4. $\square$

Since $g_d(s)$ is the expected advantage of a_1 given $S_{t-d} = s$, the best question is “is a_1 better?”: the answer bit is the recommended action itself. Delay changes only which states recommend a_1 : $\{v > 0\}$ when fresh, one side of $\{\psi_2 > 0\}$ when stale.

Several actions. With a finite $\mathcal{A}$, reporting the recommended action $a_d^*(s) = \arg \max_a (K^d r_a)(s)$ attains $\mathcal{V}_\Phi(d)$ and costs $\lceil \log_2 |\mathcal{A}_d| \rceil$ bits, where $\mathcal{A}_d$ is the set of actions recommended by some s . Delay shrinks $\mathcal{A}_d$: it can be all of $\mathcal{A}$ when fresh, lies in $\mathcal{A}^*$ for all large d , and is $\{a^*\}$, zero bits, for $d \geq D_0$ in case (H). Let $\mathcal{V}_1^*(d) = \max_f \mathcal{V}_f(d)$ be the best value of a one-bit report. One bit separates only two actions, so the best question thresholds $K^d(r_a - r_{a'})$ for the pair maximizing $\frac{1}{2} \|K^d(r_a - r_{a'})\|_{1,\pi} - \frac{1}{2} (\Delta_a + \Delta_{a'})$, where $\Delta_a = \Phi(\pi) - \langle \pi, r_a \rangle$ is the blind regret of a (Proposition 3, Appendix C.7). This pair can switch with d : a fresh question may involve a blind-suboptimal action whose reward varies strongly with the state, a stale one only actions in $\mathcal{A}^*$. The partition, however, stops depending on the rewards. In case (B) with $\sigma_2 > 0$ and $|\lambda_2| > |\lambda_3|$, every $a \in \mathcal{A}^*$ has $(K^d r_a)(s) = \Phi(\pi) + \lambda_2^d \psi_2(s) \langle r_a, \psi_2 \rangle_\pi + O(|\lambda_3|^d)$, so to leading order only the two actions with extreme $\langle r_a, \psi_2 \rangle_\pi$ are ever recommended, and the eigen-bit $\mathbf{1}[\psi_2 > 0]$ attains $\mathcal{V}_1^*(d) = \mathcal{V}_\Phi(d) + O(|\lambda_3|^d)$: to leading order, one bit recovers the full delayed value.

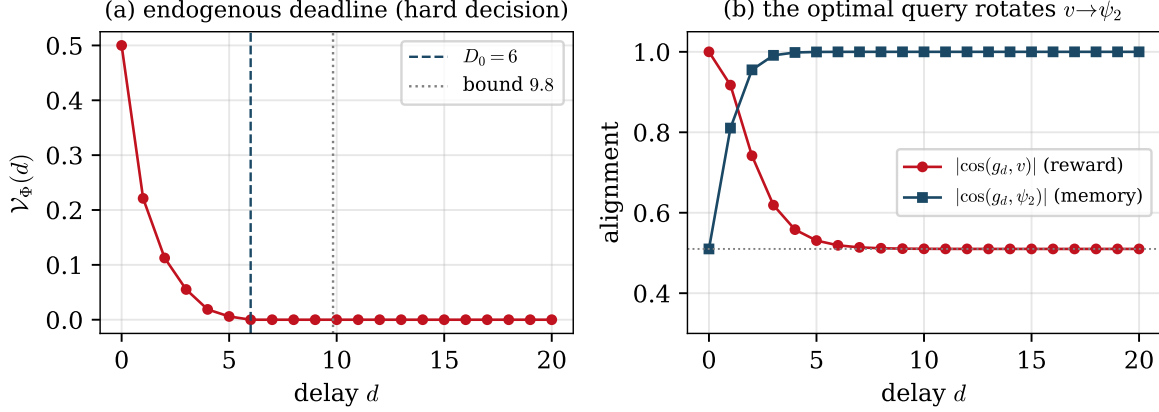


Figure 2: **Deadline and rotation.** (a) A hard decision with a comfortable blind optimum: value decays and then hits *exactly* zero at the endogenous deadline $D_0 = 6$, within the bound (10) ($D_0 \leq [9.8] = 10$). (b) The optimal one-bit question rotates from reward-aligned ($|\cos(g_d, v)| : 1 \rightarrow 0.51$) to memory-aligned ($|\cos(g_d, \psi_2)| : 0.51 \rightarrow 1$) as d grows; the residual $0.51 = |\cos(v, \psi_2)|$.

The backward reading. Rather than push the belief forward, pull the reward back from t to $t - d$: $\langle K^d(s, \cdot), r_a \rangle = \langle K^d r_a \rangle(s)$ by (19). For two actions, as $\langle \pi, v \rangle = 0$, this is $g_d = K^d v$, i.e. $g_d = K g_{d-1}$, $g_0 = v$: the Kolmogorov backward (one-shot Bellman) recurrence. For the full Howard payoff, with $\tilde{r}_a := K^d r_a$, (19) and $\langle \pi, \tilde{r}_a \rangle = \langle \pi, r_a \rangle$ give $V_\Phi(d) = \mathbb{E}_{s \sim \pi} \max_a \tilde{r}_a(s) - \max_a \langle \pi, \tilde{r}_a \rangle$: *the delayed value of the rewards $\{r_a\}$ is the undelayed DeGroot value of the pulled-back rewards $\{K^d r_a\}$, with the delay absorbed into the low-pass filter $K^d r_a = \sum_k \lambda_k^d \langle r_a, \psi_k \rangle \psi_k$.*

6 Learning delayed value

Given a trajectory $S_0, \dots, S_n$ of the chain (kernel unknown), up to what delay d can $\mathcal{T}(d)$ and the ceiling $\bar{V}(d)$ be estimated? The signals are exponentially small in d and estimation errors polynomial in n , so there is a horizon; it depends on whether the learner *measures* the signal at lag d or *computes* it from a one-step model. The online part at the end recovers any Howard value $V_\Phi(d)$. Throughout, $K > 0$ entrywise.

Model-free: measuring at lag d . A model-free learner uses only $\hat{P}_d(s, s') = \frac{1}{n-d} \sum_t \mathbf{1}[S_t = s, S_{t+d} = s']$; formally, a *model-free estimator* is any function of $\hat{P}_d$ (equivalently, invariant to reordering the lag- d pairs). The leading cases are the plug-ins $\hat{\mathcal{T}}(d) = I(\hat{P}_d)$ and $\hat{V}(d) = \frac{1}{2} \|\hat{P}_d - \hat{\pi}_0 \otimes \hat{\pi}_d\|_1$, with $\hat{\pi}_0, \hat{\pi}_d$ the marginals. Each cell of $\hat{P}_d$ is off by $\Theta(n^{-1/2})$ at every d , and both functionals are nonnegative, so this error cannot average out: it becomes an upward bias. Independence is a smooth minimum of MI ($I \approx \frac{1}{2} \chi^2$), so the bias is quadratic, $\Theta(n^{-1})$; TV has a kink there, so it is linear, $\Theta(n^{-1/2})$. The ratios signal/floor are $|\lambda_2|^{2d} n$ for $\mathcal{T}$ and its square root for $\bar{V}$, so both cross 1 for $d^* = \log n / 2\mu$.

Theorem 5 (Model-free learnability horizon). *Let $\mu = \log(1/|\lambda_2|)$ and $d^*(n) = \log n / 2\mu$. Achievability. If $d \leq (1 - \eta)d^*(n)$ for fixed $\eta > 0$, the plug-ins $\hat{\mathcal{T}}(d)$ and $\hat{V}(d)$ are consistent: their relative errors $\rightarrow 0$ in probability as $n \rightarrow \infty$. Floor. If $d \geq (1 + \eta)d^*(n)$, the true value falls below the noise floor (bias and fluctuation of the same order): $\mathbb{E}\hat{\mathcal{T}}(d) - \mathcal{T}(d) = \Theta(n^{-1})$ against*

$\mathcal{T}(d) = \Theta(|\lambda_2|^{2d})$, and $\mathbb{E}\widehat{\mathcal{V}}(d) - \overline{\mathcal{V}}(d) = \Theta(n^{-1/2})$ against $\overline{\mathcal{V}}(d) = \Theta(|\lambda_2|^d)$; neither plug-in is consistent. In the sampled-pairs model (n independent draws of (S_{t-d}, S_t)), no estimator is consistent (Proposition 4, Appendix D).

The $\sqrt{\cdot}$ -decoupling of the signals is offset exactly by the smoothness gap of the floors (Paninski, 2003). The proof uses McDiarmid-type concentration for Markov chains (Paulin, 2015) and a two-point argument (Tsybakov, 2009). Subtracting the bias does not help, because beyond d^* the fluctuation of $\widehat{\mathcal{T}}$ alone already exceeds $\mathcal{T}(d)$: Miller–Madow is of the form below, and Figure 4c shows the same for a data-driven jackknife.

Proposition 1 (No data-independent correction extends the horizon). *For any deterministic β_n and $d \geq (1 + \eta)d^*(n)$, $\text{sd}(\widehat{\mathcal{T}}(d)) \geq cn^\eta \mathcal{T}(d)$ and $\mathbb{P}(|(\widehat{\mathcal{T}}(d) - \beta_n)/\mathcal{T}(d) - 1| < \frac{1}{3}) \rightarrow 0$.*

Model-based: computing from one step. A model-based learner fits the one-step kernel $\widehat{K}$, where the signal is $\Theta(1)$, so each λ_k has relative error $\delta = O_P(n^{-1/2})$, and computes $\widehat{\lambda}_k^{2d} = \lambda_k^{2d}(1 + \delta)^{2d} \approx \lambda_k^{2d}(1 + 2d\delta)$. The exponentially small factor cancels in the ratio; the relative error grows only linearly in d .

Proposition 2 (Value of a model). *Let $\hat{\lambda}_k$ be the eigenvalues of the empirical kernel $\widehat{K}$ built from the symmetrized transition counts (so $\widehat{K}$ is reversible and the $\hat{\lambda}_k$ are real, estimated to $O_P(n^{-1/2})$ (Wolfer and Kontorovich, 2019)). The spectral estimator $\widetilde{\mathcal{T}}(d) = \frac{1}{2} \sum_{k \geq 2} \hat{\lambda}_k^{2d}$ satisfies*

$$\widetilde{\mathcal{T}}(d) = \mathcal{T}(d)(1 + O_P(dn^{-1/2}) + O(|\lambda_2|^d)),$$

so it is consistent for every $d = d_n \rightarrow \infty$ with $d_n = o(\sqrt{n})$, far beyond $d^* = \Theta(\log n)$.

Proof. $\max_k |\hat{\lambda}_k - \lambda_k| = O_P(n^{-1/2})$ by Weyl’s inequality, so $\sum_{k \geq 2} |\hat{\lambda}_k^{2d} - \lambda_k^{2d}| = |\lambda_2|^{2d} O_P(dn^{-1/2})$ for $d = o(\sqrt{n})$ by the mean-value theorem; the $O(|\lambda_2|^{3d})$ remainder of $\mathcal{T}(d)$ in Theorem 3(S) gives the $O(|\lambda_2|^d)$ term (Appendix D). $\square$

The $O(|\lambda_2|^d)$ term is a bias, not noise: $\widetilde{\mathcal{T}}$ estimates the quadratic proxy $\frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}$, exact only as $d \rightarrow \infty$. For small d the model-based plug-in $I(\hat{\pi} \otimes \widehat{K}^d)$, with $\hat{\pi}$ the stationary law of $\widehat{K}$, avoids it; it is used in Figure 4c.

Theorem 6 (Sharp model-based horizon). *Fix $\kappa \in (0, 1/(2N))$ and let $\mathcal{K}_\kappa$ be the reversible chains with $K \geq \kappa$ entrywise, $\lambda_2 \neq 0$ and $|\lambda_2| > |\lambda_3|$, started from their stationary laws. There are $c, p > 0$ depending only on κ such that, for every sequence $d_n \geq c\sqrt{n}$ and every estimator $\hat{\theta}_n$ of $\theta \in \{\mathcal{T}(d_n), \overline{\mathcal{V}}(d_n)\}$ measurable w.r.t. $S_0, \dots, S_n$,*

$$\liminf_{n \rightarrow \infty} \sup_{K \in \mathcal{K}_\kappa} \mathbb{P}_K \left(\left| \frac{\hat{\theta}_n}{\theta(K)} - 1 \right| \geq \frac{1}{3} \right) \geq p.$$

With Proposition 2, $\mathcal{T}(d)$ is learnable from a trajectory exactly up to $d = \Theta(\sqrt{n})$; for $\overline{\mathcal{V}}(d)$ the lower bound holds but a matching upper bound is not established here. The gap between $\log n$ and $\sqrt{n}$ is the price of refusing to model the world, exponential in reachable delay: a Markov analogue of the model-free/model-based gap in reinforcement learning (Azar et al., 2017; Jin et al., 2018; Tu and Recht, 2019), tight on the model-based side, and within the sampled-pairs model on the model-free side. The model must be right: for a non-Markov source $\widehat{K}^d$ is biased (Figure 4a).

Online recovery. Delayed value is also what an online learner earns, with nothing estimated. Take $\Phi = \max_a \langle \cdot, r_a \rangle$, rewards in $[0, 1]$. In the *stale-context protocol* the learner sees the chain d steps late: at round t it observes S_{t-d} , the context for round t and the feedback for round $t-d$; it commits to $p_t \in \Delta(\mathcal{A})$ and earns $\langle p_t, r.(S_t) \rangle$, revealed only at round $t+d$. A Bayesian knowing K would play $\arg \max_a \langle b_t, r_a \rangle$ with $b_t = K^d(S_{t-d}, \cdot)$; the learner knows neither K nor b_t . Feedback must be delayed as well: an immediate $r.(S_{t-1})$ could reveal S_{t-1} , and the learner would in effect see the world one step late. Let $\text{Adv}_n = \frac{1}{n} \sum_{t \leq n} (\langle p_t, r.(S_t) \rangle - r_{a^*}(S_t))$ be its per-round gain over the blind action $a^* = \arg \max_a \langle \pi, r_a \rangle$.

Algorithm. Run Hedge separately on each context s . The rounds with context s are dealt round-robin to $d+1$ copies, and the copy owning round t plays $p_t(a) \propto \exp(\eta \sum_{t'} r_a(S_{t'}))$, the sum over its own earlier rounds t' . These are at least $d+1$ rounds back, so their rewards are already revealed: each copy faces an undelayed problem.

Theorem 7 (Delayed value as an online advantage). *Let $d \leq n/2$. (i) Every policy in the stale-context protocol has $\text{Adv}_n \leq \mathcal{V}_\Phi(d) + O_P(\sqrt{(\tau_{\text{mix}} + d + 1)/n})$. (ii) The per-context Hedge above (Cesa-Bianchi and Lugosi, 2006; Freund and Schapire, 1997; Joulani et al., 2013) attains*

$$\text{Adv}_n = \mathcal{V}_\Phi(d) + O_P\left(\sqrt{(\tau_{\text{mix}} + N(d+1) \log |\mathcal{A}|)/n}\right).$$

Here $\tau_{\text{mix}} \leq \lceil \log(2/\pi_{\min})/\mu \rceil$ is the TV mixing time of K ; the O_P constants depend on K only through N and τ_{mix} .

So $\mathcal{V}_\Phi(d)$ is exactly the gain of seeing the world d steps late over seeing nothing: (i) is a ceiling for every stale learner, and (ii) reaches it without knowing K or π . The error has three sources: τ_{mix} (time averages approach π), d (the delay), and $N \log |\mathcal{A}|$ (one Hedge per context). Since $\mathcal{V}_\Phi(d) = \Theta(|\lambda_2|^d)$ for a boundary decision, the error resolves it for $d \leq (1 - \eta)d^*(n)$.

Proof sketch. Given $S_{\leq t-d}$, $S_t \sim b_t$, so no play earns more than $\Phi(b_t)$ in expectation; averaging over $S_{t-d} \sim \pi$ and subtracting a^* gives (i). Hedge on context s competes with the best fixed action for s , which is the Bayes action under $K^d(s, \cdot)$; this gives (ii). Details in Appendix D.6. $\square$

7 Numerical experiments

We use a reversible 3-state chain with $\pi = (0.2, 0.3, 0.5)$ and spectrum $\{1, 0.7, 0.3\}$ ($\mu = \log(1/0.7) = 0.357$), built from π -orthonormal eigenfunctions; reversibility ($\max_{s,s'} |\pi(s)K(s, s') - \pi(s')K(s', s)| < 10^{-16}$) and nonnegativity are checked. Figures 1 and 2 are exact computations from K^d ; Figure 3 uses simulated trajectories. Details are in Appendix E. In addition, in Appendix G we present four real-life examples illustrating our findings.

Aging trichotomy and decoupling (Figure 1). Fitted large- d slopes are -0.357 for $\bar{\mathcal{V}}$ and -0.713 for $\mathcal{T}$ and the smooth value, exactly $-\mu$ and -2μ (Theorem 3); $\bar{\mathcal{V}}(d)/\sqrt{\mathcal{T}(d)}$ flattens to 0.66 (Corollary 1), the limit $\|\psi_2\|_{1,\pi}^2/\sqrt{2}$ given by (6) and $\mathcal{T}(d) = \frac{1}{2}\lambda_2^{2d}(1 + o(1))$.

Endogenous deadline (Figure 2a). With actions “bet on state s ” plus a blind-optimal hedge, the value decays and is identically zero for $d \geq D_0 = 6$, inside the bound (10) ($D_0 \leq \lceil 9.8 \rceil = 10$): the deadline follows from mixing and margin, not an assumption.

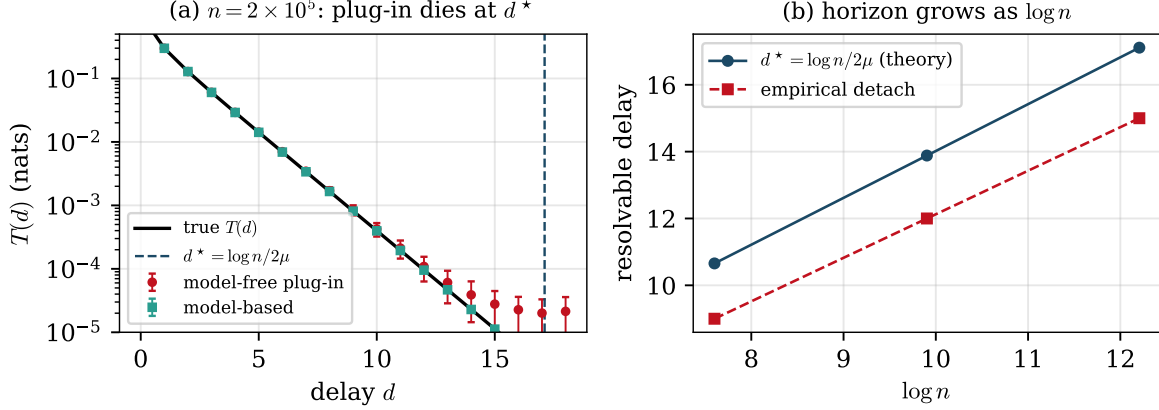


Figure 3: **Learnability horizon.** (a) With $n = 2 \times 10^5$ samples, the model-free plug-in estimate of $\mathcal{T}(d)$ tracks the truth until $d^* = \log n / 2\mu \approx 17$ and then saturates on its $\Theta(1/n)$ bias floor; the model-based (spectral) estimate tracks far longer. (b) The empirically resolvable delay grows linearly in $\log n$, matching $d^* = \log n / 2\mu$ ($n = 2 \times 10^3, 2 \times 10^4, 2 \times 10^5$).

Rotation of the optimal summary (Figure 2b). With v not aligned with ψ_2 , the optimal score g_d rotates from v toward ψ_2 ; the cosines cross and saturate at $|\cos(v, \psi_2)| = 0.51$ (see also Figure 5).

Learnability (Figure 3). The breakdown delay $\hat{d}_n$ of an estimator is the first d at which its median relative error exceeds 50%. The plug-in breaks down just below $d^* = \log n / 2\mu$, while the model-based estimate resolves far larger d (Proposition 2). Across $n \in \{2 \times 10^3, 2 \times 10^4, 2 \times 10^5\}$, $\hat{d}_n$ grows at the predicted rate $1/2\mu$ per e -fold of n (Theorem 5), offset by a constant set by the plug-in’s bias floor $\gamma_d/2n$. Here γ_d is the long-run variance constant of the lag- d pair counts (Appendix D): it tends to $\gamma_\infty = \sum_{k,l \geq 2} (1 + \lambda_k \lambda_l) / (1 - \lambda_k \lambda_l)$, exceeding the i.i.d.-pair value $(N - 1)^2$ through autocorrelation. Equating the floor to $\frac{1}{2}\mathcal{T}(d)$ gives $\hat{d}_n \approx d^* - \log(2\gamma_\infty)/2\mu$; here $\gamma_\infty = 7.18$, about 3.7 steps below d^* .

A chain fitted to real data (Figure 4). We quantize hourly Beijing PM_{2.5} readings (Liang et al., 2015) into five quintile states and fit a one-step kernel $\hat{K}$ by counting; it is nearly reversible (detailed-balance defect 4×10^{-3}) with real spectrum $\{1, 0.939, 0.828, 0.680, 0.535\}$, so $\mu = 0.063$ per hour: pollution regimes are a slow mode. Taking $\hat{K}$ as the world, (a)–(b) are exact; in (c) the truth is $\mathcal{T}(d)$ of $\hat{K}$, and estimators see only trajectories simulated from it. The only contact with the raw series is its lag- d mutual information (grey in (a)), where the Markov fit is visibly imperfect.

All three branches of the trichotomy appear. The ceiling ages at $|\lambda_2|^d$. A *soft* alert decision—the same two actions scored by $\Phi(b) = \tau \log \sum_a e^{(b, r_a) / \tau}$, $\tau = 0.5$ —ages at $|\lambda_2|^{2d}$ like temporal information (fitted slope $-0.1263 = -2\mu$). The *hard* version—alert (reward +1 in the top two quintiles, −1 otherwise) versus no alert, blind-optimal by margin $\delta = 0.2$, $\rho = 2$ —is worth exactly nothing once the report is $D_0 = 30$ hours old, inside the bound (10) ($D_0 \leq \lceil 36.6 \rceil = 37$): a report a day old is still worth acting on, one 30 hours old is not.

For the “high pollution” contrast $v = \mathbf{1}_{\{4,5\}} - \hat{\pi}(\{4,5\})\mathbf{1}$, g_d rotates from v toward ψ_2 within about ten hours (residual $|\cos(v, \psi_2)| = 0.87$). The rotation is in the score, not in the chosen bit:

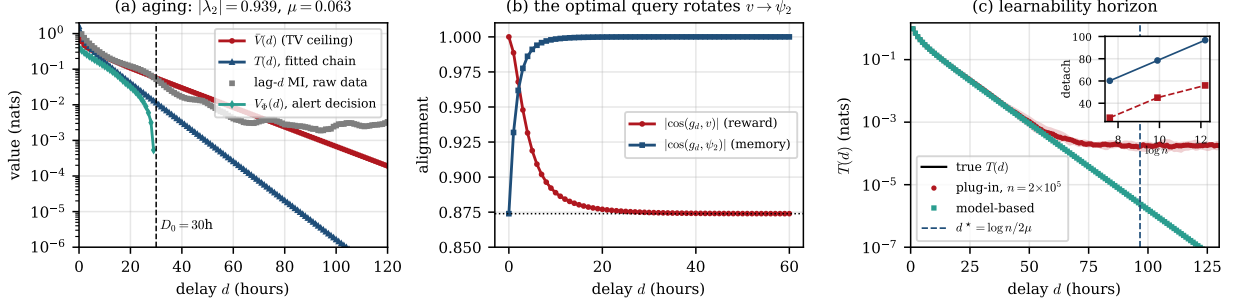


Figure 4: **A chain fitted to real data** (hourly Beijing PM_{2.5}, $n = 41,757$ valid hours, five quintile states). (a) The aging trichotomy in data: the ceiling decays as $|\lambda_2|^d$; temporal information and a soft alert decision decay as $|\lambda_2|^{2d}$ (fitted slopes $-0.1263 = -2\mu$); the hard alert decision is exactly zero after $D_0 = 30$ hours (bound (10): $\lceil 36.6 \rceil = 37$). The raw data’s lag- d mutual information (grey) decays more slowly than the fitted chain’s, a non-Markov tail. (b) The optimal one-bit question rotates from the reward contrast v (“high pollution”) toward ψ_2 ; residual $|\cos(v, \psi_2)| = 0.87$. (c) Learnability on trajectories simulated from the fitted chain, $n = 2 \times 10^5$: the plug-in floors out at $\gamma_d/2n$; a block-jackknife correction removes the mean floor but detaches at the same delay; the model-based plug-in $I(\hat{\pi} \otimes \hat{K}^d)$ tracks to $d = 130$. Inset: detach delay vs. $\log n$ (plug-in, squares; corrected, crosses) against d^* (circles).

$\psi_2 = (-1.36, -0.77, -0.01, 0.71, 1.44)$ already separates the top two quintiles from the rest, so the ψ_2 -threshold and the reward threshold coincide up to the middle quintile, where ψ_2 nearly vanishes.

In (c), the plug-in breaks down at $\hat{d}_n = 27, 45, 56$ h for $n = 2 \times 10^3, 2 \times 10^4, 2 \times 10^5$: the first increment (7.8 h per e -fold) matches $1/2\mu = 7.9$ h (least-squares slope 6.3 h). Beyond $\hat{d}_n$ it sits on its floor $\gamma_d/2n$, with $\gamma_d = 71$ –92 near $\gamma_\infty = 73.5$ (i.i.d. pairs: 16). The model-based estimate $I(\hat{\pi}_n \otimes \hat{K}_n^d)$ refits $\hat{K}_n$ from the same trajectory and errs only through it; its relative error grows like $d/\sqrt{n}$ (Proposition 2), so it drifts slowly from $T(d)$ instead of flooring, and does not break down by $d = 130$ for $n \geq 2 \times 10^4$.

Conclusion. Extending Howard and DeGroot to delayed information shows that delay is the operation that pulls value apart from information. The value $\mathcal{V}_\Phi(d)$ yields a quantized aging trichotomy with an endogenous deadline, $\sqrt{\cdot}$ -decoupling of worth from information, durable slow-mode bits as the best questions to ask late, and a $\log n$ versus sharp $\sqrt{n}$ horizon.

AI use statement

Generative AI tools were used in this work for the following purposes: writing and debugging experiment code, generating figures and experiments, drafting and editing parts of the text, and formulating and checking some proofs, especially in appendices. The overall problem formulation and the overall structure of the paper are the author’s own. Every proof was independently verified by the author, and all AI-assisted text, code, and figures were reviewed and edited. The author takes full responsibility for the content of this work, including any text, claims, or artifacts produced with the aid of generative AI.

References

- Mohammad Gheshlaghi Azar, Ian Osband, and Rémi Munos. Minimax regret bounds for reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pages 263–272, 2017.
- F. L. Bauer and C. T. Fike. Norms and exclusion theorems. *Numerische Mathematik*, 2:137–141, 1960.
- José M. Bernardo. Expected information as expected utility. *Annals of Statistics*, 7(3):686–690, 1979.
- David Blackwell. Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24(2):265–272, 1953.
- Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.
- Guoxin Chen and Soung Chang Liew. An index policy for minimizing the uncertainty-of-information of Markov sources. *IEEE Transactions on Information Theory*, 70(1):698–721, 2024.
- Yuxin Chen, Shervin Javdani, Amin Karbasi, J. Andrew Bagnell, Siddhartha Srinivasa, and Andreas Krause. Submodular surrogates for value of information. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3511–3518, 2015.
- Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2nd edition, 2006.
- Morris H. DeGroot. Uncertainty, information, and sequential experiments. *Annals of Mathematical Statistics*, 33(2):404–419, 1962.
- James Allen Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *The Annals of Applied Probability*, 1(1):62–87, 1991.
- Alexander Frankel and Emir Kamenica. Quantifying information and uncertainty. *American Economic Review*, 109(10):3650–3680, 2019.
- Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.
- Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2013.
- Ronald A. Howard. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1):22–26, 1966.
- Chi Jin, Zeyuan Allen-Zhu, Sébastien Bubeck, and Michael I. Jordan. Is Q-learning provably efficient? In *Advances in Neural Information Processing Systems 31*, pages 4863–4873, 2018.
- Galin L. Jones. On the Markov chain central limit theorem. *Probability Surveys*, 1:299–320, 2004.
- Pooria Joulani, András György, and Csaba Szepesvári. Online learning under delayed feedback. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 1453–1461, 2013.

- Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer, 1995.
- K. V. Katsikopoulos and S. E. Engelbrecht. Markov decision processes with delays and asynchronous cost collection. *IEEE Transactions on Automatic Control*, 48(4):568–574, 2003.
- Sanjit Kaul, Roy Yates, and Marco Gruteser. Real-time status: How often should one update? In *Proceedings of IEEE INFOCOM*, pages 2731–2735, 2012.
- David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2nd edition, 2017.
- Xuan Liang, Tao Zou, Bin Guo, Shuo Li, Haozhe Zhang, Shuyi Zhang, Hui Huang, and Song Xi Chen. Beijing PM2.5 data set. UCI Machine Learning Repository, <https://archive.ics.uci.edu/dataset/381>, 2015.
- Dennis V. Lindley. On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, 27(4):986–1005, 1956.
- Ali Maatouk, Saad Kriouile, Mohamad Assaad, and Anthony Ephremides. The age of incorrect information: A new performance metric for status updates. *IEEE/ACM Transactions on Networking*, 28(5):2215–2228, 2020.
- Julio Moro, James V. Burke, and Michael L. Overton. On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM Journal on Matrix Analysis and Applications*, 18(4):793–817, 1997.
- Liam Paninski. Estimation of entropy and mutual information. *Neural Computation*, 15(6):1191–1253, 2003.
- Daniel Paulin. Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electronic Journal of Probability*, 20:1–32, 2015. Paper no. 79.
- Howard Raiffa and Robert Schlaifer. *Applied Statistical Decision Theory*. Harvard Business School, 1961.
- Dmitry Shaiderman and Eilon Solan. Value of information in dynamic decision making. arXiv preprint arXiv:2608.29273, 2026.
- Bastian Steudel, Dominik Janzing, and Bernhard Schölkopf. Causal Markov condition for submodular information measures. In *Proceedings of the 23rd Conference on Learning Theory (COLT)*, pages 464–476, 2010.
- Yin Sun, Elif Uysal-Biyikoglu, Roy D. Yates, C. Emre Koksall, and Ness B. Shroff. Update or wait: How to keep your data fresh. *IEEE Transactions on Information Theory*, 63(11):7492–7508, 2017.
- Wojciech Szpankowski and Ugo Vaccaro. Temporal information: A modest proposal. preprint, 2026.
- Lloyd N. Trefethen and Mark Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, 2005.
- Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009.

- Stephen Tu and Benjamin Recht. The gap between model-based and model-free methods on the linear quadratic regulator: An asymptotic viewpoint. In *Proceedings of the 32nd Conference on Learning Theory (COLT)*, pages 3036–3083, 2019.
- T. J. Walsh, A. Nouri, L. Li, and M. L. Littman. Learning and planning in environments with delayed feedback. *Autonomous Agents and Multi-Agent Systems*, 18(1):83–105, 2009.
- Geoffrey Wolfer and Aryeh Kontorovich. Minimax learning of ergodic Markov chains. In *Proceedings of the 30th International Conference on Algorithmic Learning Theory (ALT)*, volume 98 of *PMLR*, pages 904–930, 2019.
- Roy D. Yates, Yin Sun, D. Richard Brown, Sanjit K. Kaul, Eytan Modiano, and Sennur Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, 2021.

Appendix

A Spectral preliminaries

Throughout, K is an $N \times N$ stochastic matrix, irreducible and aperiodic, reversible with respect to its stationary law π ($\pi(s)K(s, s') = \pi(s')K(s', s)$). Irreducibility gives $\pi > 0$. We write $\ell^2(\pi)$ for $\mathbb{R}^{\mathcal{S}}$ with the inner product $\langle f, g \rangle_\pi = \sum_s \pi(s) f(s) g(s)$ and norms $\|f\|_{p, \pi} = (\sum_s \pi(s) |f(s)|^p)^{1/p}$. We write $\mathbf{1} \in \mathbb{R}^{\mathcal{S}}$ for the all-ones vector ($\mathbf{1}(s) = 1$ for every s), and $\mathbf{1}_C$ or $\mathbf{1}[\cdot]$ for the indicator of a set $C \subseteq \mathcal{S}$ or an event.

Functions and measures. Two kinds of vectors in $\mathbb{R}^{\mathcal{S}}$ appear. *Functions* (rewards r_a , contrasts v , scores g_d) are acted on by K from the right, $(Kf)(s) = \sum_{s'} K(s, s') f(s')$. *Measures* (the prior π , beliefs b , rows $K^d(s, \cdot)$, perturbations $\varepsilon_s^{(d)}$) are acted on from the left, $(\mu K)(s') = \sum_s \mu(s) K(s, s')$. The eigenvectors of K come in matching pairs: for each $k \geq 1$, ψ_k below is a right eigenvector (a function), and

$$m_k(s') = \pi(s') \psi_k(s'), \quad s' \in \mathcal{S},$$

is the corresponding left eigenvector (a measure). Note that m_k is the entrywise product of π and ψ_k , a vector in $\mathbb{R}^{\mathcal{S}}$, not a matrix; and $m_1 = \pi$ since $\psi_1 \equiv 1$.

Lemma 1 (Spectral decomposition). *K is self-adjoint on $\ell^2(\pi)$, so it has real eigenvalues $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_N > -1$ and a π -orthonormal eigenbasis $\psi_1 \equiv 1, \psi_2, \dots, \psi_N$, $K\psi_k = \lambda_k \psi_k$. Moreover $m_k K = \lambda_k m_k$ and $\sum_{s'} m_j(s') \psi_k(s') = \langle \psi_j, \psi_k \rangle_\pi = \delta_{jk}$. For every $d \geq 0$ and $s, s' \in \mathcal{S}$,*

$$K^d(s, s') = \pi(s') \sum_{k \geq 1} \lambda_k^d \psi_k(s) \psi_k(s') = \sum_{k \geq 1} \lambda_k^d \psi_k(s) m_k(s'), \quad \sum_{k \geq 1} \psi_k(s) \psi_k(s') = \frac{\delta_{ss'}}{\pi(s)}. \quad (15)$$

In the first form of (15), $\psi_k(s) \psi_k(s')$ is the product of the s -th and s' -th entries of ψ_k ; the second form simply moves the factor $\pi(s')$ onto the second one. Read as matrices, the second form says $K^d = \sum_k \lambda_k^d \psi_k m_k^\top$: each term is a right eigenvector times a left eigenvector, and the $k = 1$ term $\mathbf{1} \pi^\top$ is the matrix whose rows all equal π .

Proof. Reversibility says $D^{1/2} K D^{-1/2}$ is symmetric, $D = \text{diag}(\pi)$; let $\{u_k\}$ be its orthonormal eigenvectors and set $\psi_k = D^{-1/2} u_k$, so $\langle \psi_j, \psi_k \rangle_\pi = u_j^\top u_k = \delta_{jk}$ and $K\psi_k = \lambda_k \psi_k$. Irreducibility and aperiodicity give $\lambda_1 = 1$ simple with $\psi_1 = \mathbf{1}$ and $|\lambda_k| < 1$ for $k \geq 2$ (Perron–Frobenius; [Levin and Peres, 2017](#), Ch. 12). Then $K^d = D^{-1/2} (\sum_k \lambda_k^d u_k u_k^\top) D^{1/2}$, i.e. $K^d = \sum_k \lambda_k^d (D^{-1/2} u_k) (D^{1/2} u_k)^\top$, where $D^{-1/2} u_k = \psi_k$ and $D^{1/2} u_k = D\psi_k = m_k$; entrywise this is (15). The same factorization gives $m_k K = \lambda_k m_k$, and $\sum_{s'} m_j(s') \psi_k(s') = u_j^\top u_k = \delta_{jk}$. Completeness $\sum_k u_k u_k^\top = I$ read entrywise is the last identity. $\square$

The lemma lists the eigenvalues in decreasing order; throughout the rest of the paper we re-index them by modulus, $|\lambda_2| \geq |\lambda_3| \geq \dots$, and “ $|\lambda_2| > |\lambda_3|$ ” means the leading rate is simple. Set $\mu = \log(1/|\lambda_2|) > 0$. Dropping the $k = 1$ term from (15) gives the deviation of the row from π ,

$$\varepsilon_s^{(d)} = K^d(s, \cdot) - \pi = \sum_{k \geq 2} \lambda_k^d \psi_k(s) m_k, \quad (16)$$

a measure in $\mathbb{R}^{\mathcal{S}}$: each term is the scalar $\lambda_k^d \psi_k(s)$ times the vector m_k . So $\psi_k(s)$ says how strongly state s loads on mode k , m_k is the direction in belief space that mode k moves, and λ_k^d is how much of it survives the delay.

Lemma 2 (Elementary facts). *For all $d \geq 0$:*

- (i) $\sum_{s'} \varepsilon_s^{(d)}(s') = 0$ for every s , and $\mathbb{E}_{s \sim \pi} \varepsilon_s^{(d)} = 0$.
- (ii) For every $v \in \mathbb{R}^{\mathcal{S}}$, $\langle \varepsilon_s^{(d)}, v \rangle = \sum_{k \geq 2} \lambda_k^d \psi_k(s) \langle v, \psi_k \rangle_{\pi}$.
- (iii) $\mathbb{E}_{s \sim \pi} [\psi_k(s) \psi_l(s)] = \delta_{kl}$ and $\sum_{k \geq 2} \psi_k(s)^2 = 1/\pi(s) - 1$.
- (iv) $\|\varepsilon_s^{(d)}\|_1 \leq (\sum_{k \geq 2} \lambda_k^{2d} \psi_k(s)^2)^{1/2} \leq |\lambda_2|^d \sqrt{1/\pi(s) - 1} \leq |\lambda_2|^d \sqrt{(1 - \pi_{\min})/\pi_{\min}}$.
- (v) $\max_s \|\varepsilon_s^{(d)}\|_1 \rightarrow 0$; in particular there is d_0 with $K^d(s, s') \geq \pi(s')/2$ for all s, s' and $d \geq d_0$.

Proof. (i) $\sum_{s'} m_k(s') = \langle \psi_k, \mathbf{1} \rangle_{\pi} = 0$ for $k \geq 2$; and $\mathbb{E}_s \psi_k(s) = \langle \psi_k, \mathbf{1} \rangle_{\pi} = 0$. (ii) The pairing of a measure with a function turns into the π -inner product: $\langle m_k, v \rangle = \sum_{s'} \pi(s') \psi_k(s') v(s') = \langle v, \psi_k \rangle_{\pi}$; apply this to each term of (16). (iii) The first is orthonormality; the second is (15) at $s = s'$ minus the $k = 1$ term. (iv) Divide the measure $\varepsilon_s^{(d)}$ by π to get the function $\varepsilon_s^{(d)}/\pi = \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k$ (its density with respect to π). Then $\|\varepsilon_s^{(d)}\|_1 = \|\varepsilon_s^{(d)}/\pi\|_{1,\pi} \leq \|\varepsilon_s^{(d)}/\pi\|_{2,\pi}$ by Jensen, and $\|\varepsilon_s^{(d)}/\pi\|_{2,\pi}^2 = \sum_{k \geq 2} \lambda_k^{2d} \psi_k(s)^2$ by orthonormality; then use (iii). (v) follows from (iv) since $|\lambda_2| < 1$. $\square$

B Howard and DeGroot values: a short primer

We recall the two classical notions of value of information that the paper builds on, and how the delayed value (2) fits into them. Throughout, θ is an unknown state in a finite set $\mathcal{S}$ with prior $\pi \in \Delta(\mathcal{S})$, and an experiment is a random variable X jointly distributed with θ . After observing $X = x$ the decision-maker holds the posterior $\pi_x = \mathbb{P}(\theta \in \cdot \mid X = x)$. Bayes plausibility, $\mathbb{E}_X \pi_X = \pi$, holds for every experiment. We write δ_s for the point mass at s , the posterior after a perfect observation $\theta = s$.

Howard: the value of information for a decision. A decision problem is a set of actions $\mathcal{A}$ with rewards $r_a \in \mathbb{R}^{\mathcal{S}}$, where $r_a(s)$ is the payoff of action a in state s . A decision-maker with belief b plays a Bayes action and earns $\Phi(b) = \max_{a \in \mathcal{A}} \langle b, r_a \rangle$. The value of information of X (Howard, 1966; Raiffa and Schlaifer, 1961) is the expected gain from acting on the posterior instead of the prior:

$$\text{VoI}(X) = \mathbb{E}_X \Phi(\pi_X) - \Phi(\pi) = \mathbb{E}_X \max_a \langle \pi_X, r_a \rangle - \max_a \langle \pi, r_a \rangle. \quad (17)$$

For a perfect experiment ($X = \theta$, so $\pi_X = \delta_{\theta}$) this is the expected value of perfect information, $\text{EVPI} = \mathbb{E}_{s \sim \pi} \max_a r_a(s) - \max_a \langle \pi, r_a \rangle$. Since Φ is a maximum of linear functions it is convex, so Jensen and Bayes plausibility give $0 \leq \text{VoI}(X) \leq \text{EVPI}$. Information is worth something only if it can change the action: $\text{VoI}(X) = 0$ whenever one action is optimal at every posterior π_x .

DeGroot: the value of information without a decision. DeGroot (1962) replaced the decision by an *uncertainty function*: any concave $U : \Delta(\mathcal{S}) \rightarrow \mathbb{R}$. The information of X about θ is the expected reduction of uncertainty,

$$I_U(X; \theta) = U(\pi) - \mathbb{E}_X U(\pi_X) \geq 0, \quad (18)$$

which we call the DeGroot value of X under U . For a perfect experiment it equals $U(\pi) - \mathbb{E}_{s \sim \pi} U(\delta_s)$. Standard choices:

- Shannon entropy $U = H$: $I_H(X; \theta) = H(\theta) - H(\theta | X)$ is the mutual information (Lindley, 1956), and $H(\delta_s) = 0$ gives the value $H(\pi)$ for a perfect experiment.
- Quadratic, $U(b) = 1 - \sum_s b(s)^2/\pi(s)$: $I_U(X; \theta) = \mathbb{E}_X \sum_s (\pi_X(s) - \pi(s))^2/\pi(s)$ is the χ^2 -information.
- Bayes risk of a decision problem, $U(b) = -\max_a \langle b, r_a \rangle$: I_U is exactly Howard's (17).

The two coincide. The last example shows that Howard's value is a DeGroot value. Conversely, every continuous concave U is a Bayes risk. The convex function $-U$ is the supremum of its affine minorants, and an affine function on the simplex has the form $b \mapsto \langle b, r \rangle$ (absorb the constant using $\sum_s b(s) = 1$). Hence $-U(b) = \sup_{r \in \mathcal{R}} \langle b, r \rangle$ for a (possibly infinite) set of reward vectors $\mathcal{R}$, one action per supporting hyperplane. For example, entropy is the Bayes risk of forecasting θ under the logarithmic score (Bernardo, 1979). So the two notions are the same object: DeGroot values are Howard values of possibly infinite decision problems, and a concave U is a decision problem with the rewards left implicit. Frankel and Kamenica (2019) characterize the measures of information that arise this way. Writing $\Phi = -U$ for the (convex) payoff, both read $\mathbb{E}_X \Phi(\pi_X) - \Phi(\pi)$; this is the form used in the paper.

Comparing experiments. X' is a *garbling* of X if X' is obtained from X through a channel that does not look at θ . Blackwell (1953) showed that X' is a garbling of X if and only if $I_U(X'; \theta) \leq I_U(X; \theta)$ for every concave U , or equivalently $\text{VoI}(X') \leq \text{VoI}(X)$ for every decision problem. The easy direction is Jensen: each posterior of X' is a mixture of posteriors of X . Different U can rank experiments that are not garblings of one another differently.

The delayed value in this language. Take the acting state $\theta = S_t$ and the stale report $X = S_{t-d}$. The posterior after $X = s$ is the row $K^d(s, \cdot)$, the law of X is π , and Bayes plausibility is stationarity, $\pi K^d = \pi$. Therefore the delayed value (2) is, for every d , the DeGroot value of the stale report, $\mathcal{V}_\Phi(d) = I_{-\Phi}(S_{t-d}; S_t) = \mathbb{E}_{s \sim \pi} \Phi(K^d(s, \cdot)) - \Phi(\pi)$. At $d = 0$ the report is perfect ($K^0(s, \cdot) = \delta_s$), so $\mathcal{V}_\Phi(0) = \mathbb{E}_{s \sim \pi} \Phi(\delta_s) - \Phi(\pi)$ is the value of perfect information: the EVPI for a Howard payoff and $H(\pi)$ for $\Phi = -H$. The monotonicity of Theorem 1(iii) is the easy direction of Blackwell's theorem, since S_{t-d-1} is a garbling of S_{t-d} as an experiment about S_t . The payoffs used in the paper are:

- Howard, $\Phi(b) = \max_a \langle b, r_a \rangle$: the value of a report that is d steps old.
- Temporal information, $\Phi = -H$: $\mathcal{T}(d) = I(S_{t-d}; S_t)$, the mutual information of (3).
- χ^2 , $\Phi(b) = \frac{1}{2} \sum_s b(s)^2/\pi(s)$: by (1), exactly $\mathcal{V}_\Phi(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}$.
- Subset betting, $\{r_A = \mathbf{1}_A - \pi(A)\}_{A \subseteq \mathcal{S}}$: the total-variation ceiling $\bar{\mathcal{V}}(d)$ of Theorem 2.

Most of the above is standard, perhaps except the third statement, which we prove next.

Lemma 3 (Exact χ^2 value). *Let $\Phi_\chi(b) = \frac{1}{2} \sum_s b(s)^2/\pi(s)$. Then for every $d \geq 0$, $\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}$.*

Proof. For $b = \pi + \varepsilon$ with $\sum_{s'} \varepsilon(s') = 0$, $\Phi_\chi(\pi + \varepsilon) = \frac{1}{2} \sum_{s'} \pi(s') + \sum_{s'} \varepsilon(s') + \frac{1}{2} \sum_{s'} \varepsilon(s')^2/\pi(s') = \Phi_\chi(\pi) + \frac{1}{2} \sum_{s'} \varepsilon(s')^2/\pi(s')$, with no remainder since Φ_χ is quadratic. By (1), $\varepsilon_s^{(d)}(s')/\pi(s') = \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k(s')$, so by π -orthonormality of $\{\psi_k\}$, $\sum_{s'} \varepsilon_s^{(d)}(s')^2/\pi(s') = \left\| \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k \right\|_{2,\pi}^2 =$

$\sum_{k \geq 2} \lambda_k^{2d} \psi_k(s)^2$. Averaging over $s \sim \pi$ and using $\mathbb{E}_s \psi_k(s)^2 = 1$ (Lemma 2(iii)) gives $\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}$. $\square$

What the delay adds to the classical picture is that the posteriors are not free: they are the rows of K^d , pinned to the dynamics. This is why the spectrum of K governs every value in the paper.

C Deferred proofs for the value laws

Remark 3 (Grounding in generic VoI). The payoff $\Phi(b) = \max_{a \in \mathcal{A}} \langle b, r_a \rangle$ is exactly the generic (Bayes-optimal) value-of-information functional $\text{VoI}(x) = \max_a \mathbb{E}_Y[u(Y, a) \mid x]$ of [Chen et al. \(2015\)](#); [Howard \(1966\)](#): identify the hidden target with the acting state $Y = S_t$, the observation with the stale report $x = \{S_{t-d} = s\}$ so that $\mathbb{P}(Y \mid x) = K^d(s, \cdot)$, and the reward with the utility $r_a = u(\cdot, a)$. Then $\text{VoI}(x) = \Phi(K^d(s, \cdot))$, and (2) is its stationary average net of the blind baseline $\Phi(\pi)$.

C.1 Proof of Theorem 2 (ceiling asymptotics)

The inequality and the equality case are proved in the body. For (6), write by (16)

$$\|\varepsilon_s^{(d)}\|_{\text{TV}} = \frac{1}{2} \sum_{s'} \pi(s') \left| \lambda_2^d \psi_2(s) \psi_2(s') + \sum_{k \geq 3} \lambda_k^d \psi_k(s) \psi_k(s') \right|.$$

Since $||a+b|-|a|| \leq |b|$ and $|\lambda_k| \leq |\lambda_3|$ for $k \geq 3$, the inner term equals $|\lambda_2|^d |\psi_2(s)| |\psi_2(s')| + O(|\lambda_3|^d)$ uniformly in s, s' (the constant is $\sum_{k \geq 3} \max_{s, s'} |\psi_k(s) \psi_k(s')|$). Multiplying by $\pi(s')$, summing over s' , and averaging $s \sim \pi$ gives $\bar{\mathcal{V}}(d) = \frac{1}{2} |\lambda_2|^d \mathbb{E}_s |\psi_2(s)| \|\psi_2\|_{1, \pi} + O(|\lambda_3|^d) = \frac{1}{2} \|\psi_2\|_{1, \pi}^2 |\lambda_2|^d + O(|\lambda_3|^d)$. Since $\psi_2 \neq 0$ we have $\|\psi_2\|_{1, \pi} > 0$, so the leading term dominates once $|\lambda_2| > |\lambda_3|$, giving $\Theta(|\lambda_2|^d)$.

C.2 Proof of Theorem 3 (aging trichotomy)

Throughout, $\lambda_2 \neq 0$, so $\mu \in (0, \infty)$, and $\underline{\beta}_\Phi, \bar{\beta}_\Phi$ denote the $\liminf$ and $\limsup$ of $-\frac{1}{d} \log \mathcal{V}_\Phi(d)$.

(S) Smooth. Let Φ be C^2 on a neighbourhood of π in the affine hull of $\Delta(\mathcal{S})$, with gradient $\nabla \Phi(\pi)$ and (symmetric) Hessian $\nabla^2 \Phi(\pi)$. Taylor's theorem with the mean-value remainder gives, for all ε in the tangent space $\{\sum_{s'} \varepsilon(s') = 0\}$ with $\|\varepsilon\|_1$ small,

$$\Phi(\pi + \varepsilon) = \Phi(\pi) + \langle \nabla \Phi(\pi), \varepsilon \rangle + \frac{1}{2} \varepsilon^\top \nabla^2 \Phi(\pi) \varepsilon + \varrho(\varepsilon), \quad |\varrho(\varepsilon)| \leq \omega(\|\varepsilon\|_1) \|\varepsilon\|_1^2,$$

where $\omega(r) \rightarrow 0$ as $r \rightarrow 0$ is the modulus of continuity of $\nabla^2 \Phi$ at π (times a constant). Insert $\varepsilon = \varepsilon_s^{(d)}$ and average over $s \sim \pi$. The linear term vanishes by Lemma 2(i). For the quadratic term, by (16) and Lemma 2(iii),

$$\mathbb{E}_s [\varepsilon_s^{(d)\top} \nabla^2 \Phi(\pi) \varepsilon_s^{(d)}] = \sum_{k, l \geq 2} \lambda_k^d \lambda_l^d \mathbb{E}_s [\psi_k(s) \psi_l(s)] m_k^\top \nabla^2 \Phi(\pi) m_l = \sum_{k \geq 2} \lambda_k^{2d} m_k^\top \nabla^2 \Phi(\pi) m_k.$$

For the remainder, Lemma 2(iv) gives $\|\varepsilon_s^{(d)}\|_1 \leq C |\lambda_2|^d$ uniformly in s , hence $\mathbb{E}_s |\varrho(\varepsilon_s^{(d)})| \leq \omega(C |\lambda_2|^d) C^2 |\lambda_2|^{2d} = o(|\lambda_2|^{2d})$. This proves (7) with $\nu_\Phi = \frac{1}{2} \sum_{k \geq 2} (m_k^\top \nabla^2 \Phi(\pi) m_k) \delta_{\lambda_k^2}$; each weight is ≥ 0 because a convex C^2 function has positive semidefinite Hessian on its domain. For k with $|\lambda_k| = |\lambda_2|$ we have $\lambda_k^{2d} = |\lambda_2|^{2d}$ (the eigenvalues are real), so with $w_2^* := \sum_{k: |\lambda_k| = |\lambda_2|} m_k^\top \nabla^2 \Phi(\pi) m_k$

the sum in (7) is $\frac{1}{2}w_2^*|\lambda_2|^{2d} + O(|\lambda'|^{2d})$, where $|\lambda'| < |\lambda_2|$ is the next distinct modulus. If $w_2^* > 0$, then $\mathcal{V}_\Phi(d) = \frac{1}{2}w_2^*|\lambda_2|^{2d}(1 + o(1))$, so $\beta_\Phi = 2\mu$. If $w_2^* = 0$, then $\mathcal{V}_\Phi(d) = o(|\lambda_2|^{2d})$ and $\underline{\beta}_\Phi \geq 2\mu$.

Temporal information. $\Phi(b) = \sum_s b(s) \log b(s)$ is C^∞ on the open simplex, with $\nabla^2 \Phi(\pi) = \text{diag}(1/\pi)$ and $\nabla^3 \Phi$ bounded near π ; then $m_k^\top \text{diag}(1/\pi) m_k = \sum_{s'} \pi(s') \psi_k(s')^2 = 1$ and the remainder is $O(\|\varepsilon\|_1^3) = O(|\lambda_2|^{3d})$, giving $\mathcal{T}(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d} + O(|\lambda_2|^{3d})$. The $k = 2$ term is positive, so $\mathcal{T}(d) = \Theta(|\lambda_2|^{2d})$ and $\beta_\mathcal{T} = 2\mu$ with no further condition. Moreover $\mathcal{T}(d) > 0$ for every d , since $\lambda_2 \neq 0$ implies $K^d \neq \mathbf{1}\pi$.

(B) Boundary. Let $\mathcal{A}$ be finite, $\mathcal{A}^* = \arg \max_a \langle \pi, r_a \rangle$ with $|\mathcal{A}^*| \geq 2$, so $\langle \pi, r_a \rangle = \Phi(\pi)$ for $a \in \mathcal{A}^*$. If $\mathcal{A} \neq \mathcal{A}^*$, put $\delta' = \min_{a \notin \mathcal{A}^*} (\Phi(\pi) - \langle \pi, r_a \rangle) > 0$ and $\rho' = \max_a \text{span}(r_a)$. Fix $a_1 \in \mathcal{A}^*$. For $a \notin \mathcal{A}^*$ and ε in the tangent space, recentering gives $|\langle \varepsilon, r_{a_1} - r_a \rangle| \leq \frac{1}{2} \|\varepsilon\|_1 \text{span}(r_{a_1} - r_a) \leq \rho' \|\varepsilon\|_1$, so $\langle \pi + \varepsilon, r_{a_1} - r_a \rangle \geq \delta' - \rho' \|\varepsilon\|_1 \geq 0$ whenever $\|\varepsilon\|_1 \leq \delta'/\rho'$. Hence

$$\Phi(\pi + \varepsilon) = \Phi(\pi) + g(\varepsilon), \quad g(\varepsilon) = \max_{a \in \mathcal{A}^*} \langle \varepsilon, r_a \rangle,$$

for $\|\varepsilon\|_1 \leq \delta'/\rho'$ (for all ε if $\mathcal{A} = \mathcal{A}^*$). By Lemma 2(iv) there is d_1 such that $\|\varepsilon_s^{(d)}\|_1 \leq \delta'/\rho'$ for all s and $d \geq d_1$ ($d_1 = 0$ if $\mathcal{A} = \mathcal{A}^*$); for such d , $\mathcal{V}_\Phi(d) = \mathbb{E}_s g(\varepsilon_s^{(d)})$.

Now let $|\lambda_2| > |\lambda_3|$ and write $\varepsilon_s^{(d)} = \lambda_2^d \psi_2(s) m_2 + \eta_s^{(d)}$ with $\eta_s^{(d)} = \sum_{k \geq 3} \lambda_k^d \psi_k(s) m_k$. As in Lemma 2(iv), $\|\eta_s^{(d)}\|_1 \leq (\sum_{k \geq 3} \lambda_k^{2d} \psi_k(s)^2)^{1/2} \leq |\lambda_3|^d \sqrt{(1 - \pi_{\min})/\pi_{\min}}$. The map g is Lipschitz on the tangent space, $|g(x) - g(y)| \leq \max_{a \in \mathcal{A}^*} |\langle x - y, r_a \rangle| \leq \frac{1}{2} \rho' \|x - y\|_1$, so $\mathbb{E}_s g(\varepsilon_s^{(d)}) = \mathbb{E}_s g(\lambda_2^d \psi_2(s) m_2) + O(|\lambda_3|^d)$. Let $\tau = \text{sign}(\lambda_2^d)$. Positive homogeneity of g gives

$$\mathbb{E}_s g(\lambda_2^d \psi_2(s) m_2) = |\lambda_2|^d (\mathbb{E}_s (\tau \psi_2)^+ g(m_2) + \mathbb{E}_s (\tau \psi_2)^- g(-m_2)).$$

Since $\mathbb{E}_s \psi_2 = 0$, $\mathbb{E}_s \psi_2^+ = \mathbb{E}_s \psi_2^- = \frac{1}{2} \|\psi_2\|_{1,\pi}$, so the bracket equals $\frac{1}{2} \|\psi_2\|_{1,\pi} (g(m_2) + g(-m_2))$ for either sign τ . Finally $\langle m_2, r_a \rangle = \langle \psi_2, r_a \rangle_\pi$, so $g(m_2) + g(-m_2) = \max_{a \in \mathcal{A}^*} \langle \psi_2, r_a \rangle_\pi - \min_{a \in \mathcal{A}^*} \langle \psi_2, r_a \rangle_\pi = \sigma_2$. This proves (8). If $\sigma_2 > 0$, $\mathcal{V}_\Phi(d) = \Theta(|\lambda_2|^d)$ and $\beta_\Phi = \mu$; if $\sigma_2 = 0$, $\mathcal{V}_\Phi(d) = O(|\lambda_3|^d)$ and $\underline{\beta}_\Phi \geq \log(1/|\lambda_3|) > \mu$.

For two optimizers a_1, a_2 , put $\bar{r} = \frac{1}{2}(r_{a_1} + r_{a_2})$ and $v = r_{a_1} - r_{a_2}$. Then $\max(x, y) = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|$ gives $g(\varepsilon) = \langle \varepsilon, \bar{r} \rangle + \frac{1}{2} |\langle \varepsilon, v \rangle|$; averaging and $\mathbb{E}_s \varepsilon_s^{(d)} = 0$ give $\mathcal{V}_\Phi(d) = \frac{1}{2} \mathbb{E}_s |\langle \varepsilon_s^{(d)}, v \rangle|$, the first line of (9), and $\sigma_2 = |\langle v, \psi_2 \rangle_\pi|$ gives the second.

When $|\lambda_2|$ is not simple, put $E = \{k : |\lambda_k| = |\lambda_2|\}$ and $\tau_k = \text{sign}(\lambda_k^d)$. The same argument gives the leading term $|\lambda_2|^d \mathbb{E}_s g(\sum_{k \in E} \tau_k \psi_k(s) m_k)$, up to $O(|\lambda'|^d)$ with $|\lambda'| < |\lambda_2|$ the next distinct modulus. If E contains eigenvalues of both signs, this coefficient depends on the parity of d and may vanish for one parity, so $\underline{\beta}_\Phi$ and $\bar{\beta}_\Phi$ may differ.

(H) Generic hard. Let $B_{a^*} = \{b \in \Delta(\mathcal{S}) : \langle b, r_{a^*} - r_a \rangle \geq 0 \ \forall a\}$ be the cell on which a^* is optimal. For $b = \pi + \varepsilon$ with $\sum_s \varepsilon(s) = 0$, recentering $u = r_{a^*} - r_a$ by $c_u = \frac{1}{2}(\max u + \min u)$ does not change $\langle \varepsilon, u \rangle$, and $\|u - c_u \mathbf{1}\|_\infty = \frac{1}{2} \text{span}(u) \leq \rho/2$, so

$$\langle \pi + \varepsilon, r_{a^*} - r_a \rangle \geq \delta - \|\varepsilon\|_1 \cdot \frac{\rho}{2} \geq 0 \quad \text{whenever } \|\varepsilon\|_1 \leq 2\delta/\rho.$$

By Lemma 2(iv), $\|\varepsilon_s^{(d)}\|_1 \leq |\lambda_2|^d \sqrt{(1 - \pi_{\min})/\pi_{\min}}$ for every s . Hence all rows $K^d(s, \cdot)$ lie in B_{a^*} as soon as $|\lambda_2|^d \sqrt{(1 - \pi_{\min})/\pi_{\min}} \leq 2\delta/\rho$, i.e. for every integer

$$d \geq \bar{D} := \max \left\{ 0, \left\lceil \frac{\log \left(\frac{\rho}{2\delta} \sqrt{(1 - \pi_{\min})/\pi_{\min}} \right)}{\log(1/|\lambda_2|)} \right\rceil \right\}.$$

For such d , $\Phi(K^d(s, \cdot)) = \langle K^d(s, \cdot), r_{a^*} \rangle$ for all s , so $\mathbb{E}_s \Phi(K^d(s, \cdot)) = \langle \pi K^d, r_{a^*} \rangle = \langle \pi, r_{a^*} \rangle = \Phi(\pi)$ and $\mathcal{V}_\Phi(d) = 0$. Define $D_0 = \min\{d : \mathcal{V}_\Phi(d) = 0\} \leq \bar{D}$; by Theorem 1(i),(iii) ($\mathcal{V}_\Phi \geq 0$ and non-increasing) $\mathcal{V}_\Phi(d) = 0$ for all $d \geq D_0$. This is (10) (note that $D_0 \leq \bar{D}$ requires the ceiling: the unrounded ratio need not bound D_0), and $\beta_\Phi = \infty$.

Remark 4 (Scope of the trichotomy). (i) For a finite-action Howard payoff, (B) and (H) are exhaustive: either the blind optimum is unique ((H)) or it is not ((B)). Such a payoff is locally linear at π in case (H), so it also satisfies (S) with $\nu_\Phi = 0$, consistent with $\mathcal{V}_\Phi \equiv 0$ eventually. Generically its exponent is μ or ∞ ; the 2μ branch belongs to soft decisions, whose payoff bends at π . (ii) The quantization $\beta_\Phi \in \{2\mu, \mu, \infty\}$ holds under the nondegeneracy conditions $w_2^* > 0$ in (S), and $\sigma_2 > 0$, $|\lambda_2| > |\lambda_3|$ in (B). When they fail, the value decays strictly faster, at a rate set by deeper modes of the chain. (iii) The quantization is a property of these two classes, not of convexity. For $1 < p < 2$, the bounded convex payoff $\Phi(b) = |\langle b - \pi, v \rangle|^p$ has $\mathcal{V}_\Phi(d) = \mathbb{E}_s |\langle \varepsilon_s^{(d)}, v \rangle|^p \sim |\lambda_2|^{pd} |\langle v, \psi_2 \rangle_\pi|^p \mathbb{E}_s |\psi_2(s)|^p$ when $\langle v, \psi_2 \rangle_\pi \neq 0$ and $|\lambda_2| > |\lambda_3|$, so $\beta_\Phi = p\mu$. (iv) $\underline{\beta}_\Phi$ and $\bar{\beta}_\Phi$ can differ, e.g. in (B) when $\lambda_3 = -\lambda_2$ and the leading term depends on the parity of d (it may vanish for one parity), or for complex λ_2 in non-reversible chains.

C.3 Proof of Corollary 1

By case (S) with $\Phi = -H$, $\mathcal{T}(d) = \Theta(|\lambda_2|^{2d})$, so $\beta_\mathcal{T} = 2\mu$. In case (B) with $\sigma_2 > 0$ and $|\lambda_2| > |\lambda_3|$, (8) gives $\mathcal{V}_\Phi(d) = \Theta(|\lambda_2|^d)$; hence $\mathcal{V}_\Phi(d) = \Theta(\sqrt{\mathcal{T}(d)})$ and $\kappa = \beta_\Phi/\beta_\mathcal{T} = \mu/2\mu = \frac{1}{2}$. In case (S) with $w_2^* > 0$, $\beta_\Phi = 2\mu$ and $\kappa = 1$. In case (H), $\mathcal{V}_\Phi(d) = 0$ for $d \geq D_0$, while $\mathcal{T}(d) > 0$ for all d because $\lambda_2 \neq 0$ implies $K^d \neq \mathbf{1}\pi$; so $\kappa = \infty$. When the nondegeneracy conditions fail, β_Φ is strictly larger than the values above and depends on deeper modes of K (Remark 4).

C.4 Proof of Theorem 4 (optimal one-bit question): the bound and the rotation

The exact formula (13) is proved in the body. By Cauchy–Schwarz in $\ell^2(\pi)$ applied to the second form of (13),

$$\mathcal{V}_f(d) \leq |\lambda_2|^d \sum_{k \geq 2} |\langle v, \psi_k \rangle_\pi| |\langle \mathbf{1}_{C_0}, \psi_k \rangle_\pi| \leq |\lambda_2|^d \|v\|_{2,\pi} \|\mathbf{1}_{C_0} - \pi(C_0)\mathbf{1}\|_{2,\pi} = |\lambda_2|^d \|v\|_{2,\pi} \sqrt{\pi(C_0)\pi(C_1)},$$

since $\sum_{k \geq 2} \langle \mathbf{1}_{C_0}, \psi_k \rangle_\pi^2 = \|\mathbf{1}_{C_0}\|_{2,\pi}^2 - \pi(C_0)^2 = \pi(C_0)(1 - \pi(C_0)) \leq \frac{1}{4}$. This is (14). If $\langle \mathbf{1}_{C_0}, \psi_2 \rangle_\pi = 0$ and $|\lambda_2| > |\lambda_3|$, the $k = 2$ term of (13) vanishes and the same argument over $k \geq 3$ gives $\mathcal{V}_f(d) \leq \frac{1}{2} |\lambda_3|^d \|v\|_{2,\pi} = O(|\lambda_3|^d)$. For $v = \psi_2$, $g_d = \lambda_2^d \psi_2$, so the optimal summary is the ψ_2 -threshold and its value is $\frac{1}{2} \|g_d\|_{1,\pi} = \frac{1}{2} |\lambda_2|^d \|\psi_2\|_{1,\pi}$, a fraction $\|\psi_2\|_{1,\pi} \leq 1$ of the bound $\frac{1}{2} |\lambda_2|^d \|\psi_2\|_{2,\pi}$: the upper bound is attained in rate. Finally, $g_0(s) = \langle \delta_s - \pi, v \rangle = v(s) - \langle \pi, v \rangle = v(s)$. If $\langle v, \psi_2 \rangle_\pi \neq 0$ and $|\lambda_2| > |\lambda_3|$, then $g_d = \lambda_2^d \langle v, \psi_2 \rangle_\pi (\psi_2 + O((|\lambda_3|/|\lambda_2|)^d))$ uniformly in s , so $g_d/\|g_d\|_{2,\pi} \rightarrow \pm \psi_2$ (the sign alternates with d if $\lambda_2 < 0$). On $\{s : \psi_2(s) \neq 0\}$ the sign of $g_d(s)$ is eventually that of $\pm \psi_2(s)$, so the optimal cell $\{g_d > 0\}$ eventually agrees there with $\{\psi_2 > 0\}$ or its complement, which define the same summary. States with $\psi_2(s) = 0$ are assigned by the deeper modes; they do not affect the leading-order value. If $\langle v, \psi_2 \rangle_\pi = 0$, the same statements hold with ψ_2 replaced by the first ψ_k with $\langle v, \psi_k \rangle_\pi \neq 0$ (when its eigenvalue modulus is simple).

In Figure 5 we illustrate the rotation for $N = 3$. The most durable one-bit summary is a threshold. For a centered $x \in \ell^2(\pi)$, the projection of x onto $\text{span}\{\mathbf{1}, \mathbf{1}_{C_0}\} = \text{span}\{\mathbf{1}, \hat{f}\}$ is $\langle \hat{f}, x \rangle_\pi \hat{f}$, so $\|x\|_{2,\pi}^2 - \langle \hat{f}, x \rangle_\pi^2$ is the error of the best $\ell^2(\pi)$ approximation of x by a function constant on C_0 and on C_1 . Maximizing $\langle \hat{f}, x \rangle_\pi^2$ over one-bit summaries is therefore optimal two-level quantization of x . If an optimal C_0 contained a state closer to the level of C_1 than to its own, moving it would

strictly decrease the error, since $\pi > 0$. So each optimal cell is a threshold set of x (up to ties at the midpoint), and with $x = \psi_2$ this identifies the most durable one-bit summary.

C.5 The pulled-back reward identity

For $\Phi(b) = \max_a \langle b, r_a \rangle$ and $b = K^d(s, \cdot)$,

$$\langle K^d(s, \cdot), r_a \rangle = \sum_{s'} K^d(s, s') r_a(s') = (K^d r_a)(s) = \tilde{r}_a(s), \quad (19)$$

so $\Phi(K^d(s, \cdot)) = \max_a \tilde{r}_a(s)$. For the blind term, by stationarity, $\langle \pi, r_a \rangle = \langle \pi K^d, r_a \rangle = \langle \pi, K^d r_a \rangle = \langle \pi, \tilde{r}_a \rangle$, so $\Phi(\pi) = \max_a \langle \pi, \tilde{r}_a \rangle$. The recurrence $\tilde{r}_a^{(h)} = K \tilde{r}_a^{(h-1)}$, $\tilde{r}_a^{(0)} = r_a$, reproduces $\tilde{r}_a^{(d)} = K^d r_a$, and by Lemma 1 $K^d r_a = \sum_{k \geq 1} \lambda_k^d \langle r_a, \psi_k \rangle_\pi \psi_k$; in the two-action case $v = r_{a_1} - r_{a_2}$ this is (12), since $\langle v, \psi_1 \rangle_\pi = \langle \pi, v \rangle = 0$.

C.6 Second-order value of one stale summary

Let $e = (f, d)$ be a one-bit summary f read at delay d , with cells C_0, C_1 , $\pi_y = \pi(C_y)$, and $\hat{f} = (\mathbf{1}_{C_0} - \pi_0 \mathbf{1}) / \sqrt{\pi_0 \pi_1}$ its centered, normalized score. Given $Y = y$ the belief about S_t is $b_y = \pi + \varepsilon_y$ with $\varepsilon_y = \frac{1}{\pi_y} \sum_{s \in C_y} \pi(s) \varepsilon_s^{(d)} = \sum_{k \geq 2} \lambda_k^d \frac{a_{y,k}}{\pi_y} m_k$, $a_{y,k} = \langle \mathbf{1}_{C_y}, \psi_k \rangle_\pi$. For smooth Φ , expanding to second order at π (the linear terms cancel since $\sum_y \pi_y \varepsilon_y = 0$), the value $F(\{e\}) = \sum_y \pi_y \Phi(b_y) - \Phi(\pi)$ is

$$F(\{e\}) = \frac{1}{2} \sum_y \pi_y \varepsilon_y^\top \nabla^2 \Phi(\pi) \varepsilon_y + o(|\lambda_2|^{2d}) = \frac{1}{2} \sum_{k,l \geq 2} \lambda_k^d \lambda_l^d m_k^\top \nabla^2 \Phi(\pi) m_l \sum_y \frac{a_{y,k} a_{y,l}}{\pi_y} + o(|\lambda_2|^{2d}).$$

Since $a_{0,k} + a_{1,k} = \langle \mathbf{1}, \psi_k \rangle_\pi = 0$, $\sum_y a_{y,k} a_{y,l} / \pi_y = a_{0,k} a_{0,l} (1/\pi_0 + 1/\pi_1) = a_{0,k} a_{0,l} / (\pi_0 \pi_1) = \langle \hat{f}, \psi_k \rangle_\pi \langle \hat{f}, \psi_l \rangle_\pi$, which gives

$$F(\{e\}) = \frac{1}{2} \sum_{k,l \geq 2} \lambda_k^d \lambda_l^d \langle \hat{f}, \psi_k \rangle_\pi \langle \hat{f}, \psi_l \rangle_\pi m_k^\top \nabla^2 \Phi(\pi) m_l + o(|\lambda_2|^{2d}). \quad (20)$$

When $m_k^\top \nabla^2 \Phi(\pi) m_l = 0$ for $k \neq l$ the double sum collapses to $\sum_{k \geq 2} w_k |\lambda_k|^{2d} \langle \hat{f}, \psi_k \rangle_\pi^2 = \sum_{k \geq 2} w_k r_{e,k}$ with $w_k = m_k^\top \nabla^2 \Phi(\pi) m_k$, i.e. (21); note $\sum_{k \geq 2} \langle \hat{f}, \psi_k \rangle_\pi^2 = \|\hat{f}\|_{2,\pi}^2 = 1$. For temporal information $\nabla^2 \Phi(\pi) = \text{diag}(1/\pi)$ and $m_k^\top \nabla^2 \Phi(\pi) m_l = \langle \psi_k, \psi_l \rangle_\pi = \delta_{kl}$, so the condition holds with $w_k \equiv 1$.

Multi-bit reports. For a smooth payoff the same filter governs any one-bit summary. Write $e = (f_e, d_e)$ for a summary f_e read at delay d_e , with π -centered, π -normalized score $\hat{f}_e$; to second order it is (Appendix C.6) $\frac{1}{2} \sum_{k,l \geq 2} \lambda_k^{d_e} \lambda_l^{d_e} \langle \hat{f}_e, \psi_k \rangle_\pi \langle \hat{f}_e, \psi_l \rangle_\pi m_k^\top \nabla^2 \Phi(\pi) m_l$; when $\nabla^2 \Phi(\pi)$ is diagonal in the eigenbasis—as for temporal information—this is $\frac{1}{2} \sum_{k \geq 2} w_k r_{e,k}$ with $w_k = m_k^\top \nabla^2 \Phi(\pi) m_k$ and the *delayed resolution*

$$r_{e,k} = |\lambda_k|^{2d_e} \langle \hat{f}_e, \psi_k \rangle_\pi^2 \in [0, 1] \quad \left(\sum_{k \geq 2} \langle \hat{f}_e, \psi_k \rangle_\pi^2 = 1 \right) : \quad (21)$$

a late summary resolves only the slow modes. Once the $k = 2$ term dominates (for $w_2 > 0$ and $|\lambda_2| > |\lambda_3|$), the best one-bit summary maximizes $\langle \hat{f}_e, \psi_2 \rangle_\pi^2$, and this maximizer is a threshold set

of ψ_2 .¹ The thresholds of $\psi_3, \psi_4, \dots$ are the next directions to survive. This suggests that the optimal m -bit stale experiment is the joint level set of $\psi_2, \dots, \psi_{m+1}$; we conjecture this and leave it, together with the generic-VoI selection problem over stale summaries (Chen et al., 2015), as an open problem.

C.7 Several actions: the best one-bit report

Let $\Phi(b) = \max_{a \in \mathcal{A}} \langle b, r_a \rangle$ with $\mathcal{A}$ finite, $\mathcal{A}^* = \arg \max_a \langle \pi, r_a \rangle$, and write $\Delta_a = \Phi(\pi) - \langle \pi, r_a \rangle \geq 0$ for the *blind regret* of a ($\Delta_a = 0$ iff $a \in \mathcal{A}^*$). A one-bit report f with cells C_0, C_1 is read at delay d (we allow a constant f); the receiver forms b_y by (11) and plays $\arg \max_a \langle b_y, r_a \rangle$, so $\mathcal{V}_f(d) = \sum_y \pi(C_y) \Phi(b_y) - \Phi(\pi)$. Let $\mathcal{V}_1^*(d) = \max_f \mathcal{V}_f(d)$ and, as in Section 5, $\tilde{r}_a = K^d r_a$.

Proposition 3 (Best one-bit report, finite $\mathcal{A}$). *For every $d \geq 0$,*

$$\mathcal{V}_1^*(d) = \max_{a, a' \in \mathcal{A}} \mathbb{E}_{s \sim \pi} \max(\tilde{r}_a(s), \tilde{r}_{a'}(s)) - \Phi(\pi) = \max_{a, a' \in \mathcal{A}} \left[\frac{1}{2} \|K^d(r_a - r_{a'})\|_{1, \pi} - \frac{1}{2}(\Delta_a + \Delta_{a'}) \right], \quad (22)$$

attained by the report $C_1 = \{s : \tilde{r}_a(s) > \tilde{r}_{a'}(s)\}$ for a maximizing pair (a, a') . Moreover:

- (i) *If $\mathcal{A} = \{a_1, a_2\}$ is a boundary decision, (22) is $\frac{1}{2} \|g_d\|_{1, \pi}$ and the report is f_d^* of Theorem 4.*
- (ii) *If $|\mathcal{A}^*| = 1$ (case (H) of Theorem 3), $\mathcal{V}_1^*(d) = 0$ for $d \geq D_0$.*
- (iii) *If $|\mathcal{A}^*| \geq 2$ (case (B)), there is d_1 such that for $d \geq d_1$ only pairs in $\mathcal{A}^*$ matter: $\mathcal{V}_1^*(d) = \frac{1}{2} \max_{a, a' \in \mathcal{A}^*} \|K^d(r_a - r_{a'})\|_{1, \pi}$. If moreover $|\lambda_2| > |\lambda_3|$, then with $c_a = \langle r_a, \psi_2 \rangle_\pi$,*

$$\mathcal{V}_1^*(d) = \frac{1}{2} |\lambda_2|^d \|\psi_2\|_{1, \pi} \left(\max_{a \in \mathcal{A}^*} c_a - \min_{a \in \mathcal{A}^*} c_a \right) + O(|\lambda_3|^d) = \mathcal{V}_\Phi(d) + O(|\lambda_3|^d), \quad (23)$$

and the eigen-bit $\mathbf{1}[\psi_2 > 0]$ attains (23).

Proof. Formula (22). Since $\pi(C_y) b_y = \sum_{s \in C_y} \pi(s) K^d(s, \cdot)$, we have $\pi(C_y) \langle b_y, r_a \rangle = \sum_{s \in C_y} \pi(s) \tilde{r}_a(s)$. If the receiver plays a_0 on $Y = 0$ and a_1 on $Y = 1$, the value is $\sum_{s \in C_0} \pi(s) \tilde{r}_{a_0}(s) + \sum_{s \in C_1} \pi(s) \tilde{r}_{a_1}(s) - \Phi(\pi) \leq \mathbb{E}_{s \sim \pi} \max(\tilde{r}_{a_0}(s), \tilde{r}_{a_1}(s)) - \Phi(\pi)$, which is at most the right side of the first equality. Conversely, for any pair (a, a') the report $C_1 = \{\tilde{r}_a > \tilde{r}_{a'}\}$ with the receiver playing a on C_1 and a' on C_0 achieves $\mathbb{E}_s \max(\tilde{r}_a, \tilde{r}_{a'}) - \Phi(\pi)$, and the receiver's best response can only do better. For the second equality use $\max(x, y) = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|$, $\tilde{r}_a - \tilde{r}_{a'} = K^d(r_a - r_{a'})$ and $\mathbb{E}_s \tilde{r}_a(s) = \langle \pi K^d, r_a \rangle = \langle \pi, r_a \rangle = \Phi(\pi) - \Delta_a$.

(i) Here $\Delta_{a_1} = \Delta_{a_2} = 0$ and $K^d(r_{a_1} - r_{a_2}) = K^d v = g_d$; the pairs (a, a) give 0.

(ii) Each b_y is a π -average of the rows $K^d(s, \cdot)$, $s \in C_y$, so by Jensen $0 \leq \mathcal{V}_f(d) \leq \mathcal{V}_\Phi(d)$; and $\mathcal{V}_\Phi(d) = 0$ for $d \geq D_0$ by Theorem 3(H).

(iii) By the proof of Theorem 3(B), there are $\delta' > 0$ and d_1 such that $\max_{a \in \mathcal{A}} \langle \pi + \varepsilon, r_a \rangle = \max_{a \in \mathcal{A}^*} \langle \pi + \varepsilon, r_a \rangle$ whenever $\|\varepsilon\|_1 \leq \delta'$, and $\|\varepsilon_s^{(d)}\|_1 \leq \delta'$ for all s and $d \geq d_1$. Each $b_y - \pi = \sum_{s \in C_y} \frac{\pi(s)}{\pi(C_y)} \varepsilon_s^{(d)}$ is a convex combination of the $\varepsilon_s^{(d)}$, so it lies in the same ℓ^1 -ball. Hence for $d \geq d_1$ every $\mathcal{V}_f(d)$, and so $\mathcal{V}_1^*(d)$, is unchanged when $\mathcal{A}$ is replaced by $\mathcal{A}^*$, and (22) with all $\Delta_a = 0$ gives the first claim. For $a, a' \in \mathcal{A}^*$, $v = r_a - r_{a'}$ has $\langle \pi, v \rangle = 0$, so by Lemma 2(ii), $K^d v = \lambda_2^d (c_a - c_{a'}) \psi_2 + e_d$ with $\|e_d\|_{1, \pi} \leq |\lambda_3|^d \sum_{k \geq 3} |\langle v, \psi_k \rangle_\pi| = O(|\lambda_3|^d)$. By the triangle inequality, $\frac{1}{2} \|K^d v\|_{1, \pi} = \frac{1}{2} |\lambda_2|^d |c_a - c_{a'}| \|\psi_2\|_{1, \pi} + O(|\lambda_3|^d)$; maximizing over the finitely many pairs gives the first equality in (23).

¹For centered x , $\|x\|_{2, \pi}^2 - \langle \hat{f}, x \rangle_\pi^2$ is the $\ell^2(\pi)$ error of the best approximation of x by a function constant on C_0 and on C_1 . Maximizing $\langle \hat{f}, x \rangle_\pi^2$ is therefore optimal two-level quantization of x , and Lloyd's nearest-level condition forces each cell to be a threshold set of x (up to ties), since $\pi > 0$.

For the second, for $d \geq d_1$, $\mathcal{V}_\Phi(d) = \mathbb{E}_s \max_{a \in \mathcal{A}^*} \langle \varepsilon_s^{(d)}, r_a \rangle$, and $\langle \varepsilon_s^{(d)}, r_a \rangle = \lambda_2^d \psi_2(s) c_a + O(|\lambda_3|^d)$ uniformly in s and a . With $\tau = \text{sign}(\lambda_2^d)$, $\max_a \tau \psi_2(s) c_a$ equals $\tau \psi_2(s) \max_a c_a$ if $\tau \psi_2(s) > 0$ and $\tau \psi_2(s) \min_a c_a$ otherwise. Since $\mathbb{E}_s \psi_2(s) = 0$, $\mathbb{E}_s (\tau \psi_2)^+ = \mathbb{E}_s (\tau \psi_2)^- = \frac{1}{2} \|\psi_2\|_{1,\pi}$, so $\mathcal{V}_\Phi(d) = \frac{1}{2} |\lambda_2|^d \|\psi_2\|_{1,\pi} (\max_a c_a - \min_a c_a) + O(|\lambda_3|^d)$. Finally, for the eigen-bit let $a_+ \in \arg \max_{\mathcal{A}^*} c_a$ and $a_- \in \arg \min_{\mathcal{A}^*} c_a$. Playing a_+ on $\{\tau \psi_2 > 0\}$ and a_- on its complement (a cell of $\mathbf{1}[\psi_2 > 0]$ for either sign of τ) gives, by the same expansion, $|\lambda_2|^d (\mathbb{E}_s (\tau \psi_2)^+ c_{a_+} - \mathbb{E}_s (\tau \psi_2)^- c_{a_-}) + O(|\lambda_3|^d)$, which is (23). The receiver's best response can only do better, and nothing exceeds $\mathcal{V}_1^*(d)$. $\square$

Reading. One bit can separate only two actions, so the best report is the best pairwise comparison: it asks whether a beats a' at time t , given S_{t-d} , and it thresholds the pulled-back contrast $K^d(r_a - r_{a'})$; Theorem 4 is the case $|\mathcal{A}| = 2$. The penalty $\frac{1}{2}(\Delta_a + \Delta_{a'})$ in (22) is new with $|\mathcal{A}| > 2$: a blind-suboptimal action can be worth reporting on when the report is fresh, because its reward varies strongly with the state, but once $d \geq d_1$ only blind-optimal actions are compared. So the optimal pair itself can switch with d . At small d it may involve an action outside $\mathcal{A}^*$; for $d \geq d_1$ it lies in $\mathcal{A}^*$; and within $\mathcal{A}^*$ it moves from the pair with the largest spread $\|K^d(r_a - r_{a'})\|_{1,\pi}$ to the pair with the largest slow-mode contrast $|c_a - c_{a'}|$. The partition, however, does not depend on the rewards in the limit: for every finite-action boundary decision the best stale report is the eigen-bit $\mathbf{1}[\psi_2 > 0]$, and only the two actions it recommends depend on $\{r_a\}$. By (23), this single bit already captures the leading term of the full delayed value, $\mathcal{V}_1^*(d)/\mathcal{V}_\Phi(d) \rightarrow 1$ when $\max_{\mathcal{A}^*} c_a > \min_{\mathcal{A}^*} c_a$: to leading order, a stale report of any size is worth no more than one well-chosen bit.

C.8 Random delay

Let D be independent of the chain. By conditioning, $\mathbb{E}_D \mathcal{V}_\Phi(D) = \sum_d \mathbb{P}(D = d) \mathcal{V}_\Phi(d)$. In case (H), $\mathcal{V}_\Phi(d) = 0$ for $d \geq D_0$, so $\mathbb{E}_D \mathcal{V}_\Phi(D) = \mathbb{E}[\mathcal{V}_\Phi(D) \mathbf{1}\{D < D_0\}]$ and the tail of D beyond D_0 is irrelevant.

For the χ^2 payoff $\Phi_\chi(b) = \frac{1}{2} \sum_s b(s)^2 / \pi(s)$ (convex and bounded on $\Delta(\mathcal{S})$), the expansion (7) is exact: the linear term $\sum_{s'} \varepsilon_s^{(d)}(s') = 0$ vanishes, and by orthonormality $\frac{1}{2} \mathbb{E}_s \sum_{s'} \varepsilon_s^{(d)}(s')^2 / \pi(s') = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}$ for every d . Fubini then gives $\mathbb{E}_D \mathcal{V}_\chi(D) = \frac{1}{2} \sum_{k \geq 2} \mathbb{E}[\lambda_k^{2D}] = \frac{1}{2} \sum_{k \geq 2} G_D(\lambda_k^2)$, with G_D the probability generating function of D .

For a general smooth Φ , $\mathcal{V}_\Phi(d) = \int x^d d\nu_\Phi(x) + \varrho_d$ with $\varrho_d = o(|\lambda_2|^{2d})$ only as $d \rightarrow \infty$, so $\mathbb{E}_D \mathcal{V}_\Phi(D) = \int G_D d\nu_\Phi + \mathbb{E} \varrho_D$. The remainder is negligible relative to the main term only if, for every fixed M , $\mathbb{P}(D \leq M)$ is small compared with $\mathbb{E}|\lambda_2|^{2D}$. This fails for $D \sim \text{Geom}(p)$: both quantities are $\Theta(p)$ as $p \rightarrow 0$, so $\int G_D d\nu_\Phi$ is a second-order proxy, not the value.

Remark 5 (Scope of the ceiling). (i) Convexity is not used in (5). It enters only through $\mathcal{V}_\Phi \geq 0$ (Theorem 1) and through the equality case, which makes $\bar{\mathcal{V}}$ itself a DeGroot value. (ii) The class is exactly the span-one Howard class. Conversely, every convex Φ satisfying (4) is a Howard payoff $\Phi(b) = \sup_a \langle b, r_a \rangle$ with (possibly infinitely many) actions of span at most one. Take the tangent affine minorants of Φ at relative-interior points of $\Delta(\mathcal{S})$, absorbing constants into r_a via $\langle b, \mathbf{1} \rangle = 1$. Their slopes have span at most one, because on $\{v : \sum_s v(s) = 0\}$ the dual norm of $\|\cdot\|_{\text{TV}}$ is the span. Continuity of Φ then extends the representation to the boundary. (iii) Some normalization is necessary. Replacing Φ by $c \Phi_{\text{TV}}$ with $c > 1$ gives $\mathcal{V}_\Phi = c \bar{\mathcal{V}} > \bar{\mathcal{V}}$.

D Deferred proofs for learnability

D.1 Setting and a concentration lemma

The learner observes $S_0, \dots, S_n$ from the stationary chain (the non-stationary case adds a bias $O(\tau_{\text{mix}}/n)$ in all statements below). Fix $d \geq 0$, $m = n - d$, and let

$$\hat{P}_d(s, s') = \frac{1}{m} \sum_{t=0}^{m-1} \mathbf{1}[S_t = s, S_{t+d} = s'], \quad P_d(s, s') = \pi(s)K^d(s, s'), \quad \Delta = \hat{P}_d - P_d.$$

Let $\hat{\pi}_0, \hat{\pi}_d$ be the two marginals of $\hat{P}_d$ (both estimate π) and $\Delta_1 = \hat{\pi}_0 - \pi$, $\Delta_2 = \hat{\pi}_d - \pi$. Let τ_{mix} be the total-variation mixing time of K ; under a spectral gap, $\tau_{\text{mix}} \leq \lceil \log(2/\pi_{\min})/\mu \rceil$ (Levin and Peres, 2017, Thm. 12.4).

Lemma 4 (Concentration of $\hat{P}_d$). *There is a constant C_K (depending on K only through τ_{mix} and N) such that for all $d \leq n/2$ and $u > 0$,*

$$\mathbb{P}(\|\Delta\|_1 \geq u) \leq 2N^2 \exp\left(-\frac{u^2 m}{C_K N^4}\right), \quad \text{hence} \quad \|\Delta\|_1 = O_P(\sqrt{\tau_{\text{mix}}/n}),$$

uniformly in $d \leq n/2$; the same holds for $\|\Delta_1\|_1, \|\Delta_2\|_1$.

Proof. Fix (s, s') . The function $F(S_0, \dots, S_n) = \hat{P}_d(s, s')$ satisfies the bounded-difference property with constants $c_t = 2/m$ (changing one S_t changes at most two indicator pairs). McDiarmid's inequality for Markov chains (Paulin, 2015, Thm. 2.1 and Cor. 2.10) gives $\mathbb{P}(|F - \mathbb{E}F| \geq \varepsilon) \leq 2 \exp(-\varepsilon^2/(c \tau_{\min} \sum_t c_t^2))$ with $\tau_{\min} \leq \tau_{\text{mix}}$ up to constants; here $\sum_t c_t^2 \leq 4(n+1)/m^2 \leq 12/m$, since $n+1 \leq 3m$ for $d \leq n/2$. Since the chain is stationary, $\mathbb{E}F = P_d(s, s')$. A union bound over N^2 entries and $\|\Delta\|_1 \leq N^2 \max_{s,s'} |\Delta(s, s')|$ give the display. The marginals are handled identically (or by $\|\Delta_i\|_1 \leq \|\Delta\|_1$). $\square$

Lemma 5 (Regularity of the two functionals). *Write $I(P)$ for the mutual information of a joint law P on $\mathcal{S}^2$ and $V(P) = \sum_s P_1(s) \|P(s, \cdot)/P_1(s) - P_2\|_{\text{TV}}$ for the TV functional (P_1, P_2 the marginals).*

- (i) *For all joint laws Q, P , $|V(Q) - V(P)| \leq C_1 \|Q - P\|_1$, with $C_1 = 3/2$.*
- (ii) *Let $d \geq d_0$ (Lemma 2(v)), so $P_d(s, s') \geq \pi(s)\pi(s')/2 \geq \pi_{\min}^2/2$. For $\|Q - P_d\|_1 \leq \pi_{\min}^2/4$, $I(Q) - I(P_d) = \langle L_d, Q - P_d \rangle + R(Q)$ with $L_d(s, s') = \log \frac{P_d(s, s')}{\pi(s)\pi(s')}$, $\|L_d\|_\infty \leq C_2 |\lambda_2|^d$, and $|R(Q)| \leq C_3 \|Q - P_d\|_1^2$.*

Proof. (i) $V(P) = \frac{1}{2} \sum_{s,s'} |P(s, s') - P_1(s)P_2(s')|$, and the map $P \mapsto P - P_1 \otimes P_2$ is Lipschitz in ℓ^1 with constant 3 (each marginal has ℓ^1 -Lipschitz constant 1, and $\|P_1 \otimes P_2 - Q_1 \otimes Q_2\|_1 \leq \|P_1 - Q_1\|_1 + \|P_2 - Q_2\|_1$), so $C_1 = 3/2$. (ii) $I(P) = \sum_{s,s'} P(s, s') \log P(s, s') - \sum_s P_1(s) \log P_1(s) - \sum_{s'} P_2(s') \log P_2(s')$ is C^∞ on $\{P \geq \pi_{\min}^2/4\}$, which contains the stated ball, with second derivatives bounded by $12/\pi_{\min}^2$ entrywise; since $|\Delta^\top H \Delta| \leq \max_{ij} |H_{ij}| \|\Delta\|_1^2$, the Taylor expansion with quadratic remainder holds with $C_3 = O(1/\pi_{\min}^2)$. The gradient in a direction Δ with $\sum \Delta = 0$ is $\sum_s \Delta(s, s') [\log P_d(s, s') + 1] - \sum_s \Delta_1(s) [\log \pi(s) + 1] - \sum_{s'} \Delta_2(s') [\log \pi(s') + 1] = \langle L_d, \Delta \rangle$, since $\sum_s \Delta_1(s) = \sum_{s'} \Delta_2(s') = 0$. Finally $L_d(s, s') = \log(1 + x)$ with $x = \sum_{k \geq 2} \lambda_k^d \psi_k(s) \psi_k(s') \geq -\frac{1}{2}$ for $d \geq d_0$, and $|\log(1 + x)| \leq 2|x|$ for $x \geq -\frac{1}{2}$, so $\|L_d\|_\infty \leq 2 \max_{s,s'} \sum_{k \geq 2} |\lambda_k|^d |\psi_k(s) \psi_k(s')| \leq C_2 |\lambda_2|^d$. $\square$

Lemma 6 (Two-sided size of the signals). *For every $d \geq 0$,*

$$\frac{\|\psi_2\|_{1,\pi}}{2\|\psi_2\|_\infty} |\lambda_2|^d \leq \bar{\mathcal{V}}(d) \leq \frac{1}{2} |\lambda_2|^d \sqrt{(1 - \pi_{\min})/\pi_{\min}},$$

and $\mathcal{T}(d) = \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d} (1 + O(|\lambda_2|^d))$, so $\frac{1}{4} |\lambda_2|^{2d} \leq \mathcal{T}(d) \leq C |\lambda_2|^{2d}$ for $d \geq d_1$. No assumption $|\lambda_2| > |\lambda_3|$ is needed.

Proof. By π -orthonormality, $\sum_{s'} \varepsilon_s^{(d)}(s') \psi_2(s') = \sum_{k \geq 2} \lambda_k^d \psi_k(s) \langle \psi_k, \psi_2 \rangle_\pi = \lambda_2^d \psi_2(s)$, so $\|\varepsilon_s^{(d)}\|_{\text{TV}} \geq |\lambda_2|^d |\psi_2(s)| / (2\|\psi_2\|_\infty)$; average over $s \sim \pi$. The upper bound is Lemma 2(iv). The statement for $\mathcal{T}$ is Theorem 3(S). $\square$

D.2 Proof of Theorem 5 (achievability)

Let $d \leq (1 - \eta)d^*(n)$, so $|\lambda_2|^d \geq n^{-(1-\eta)/2}$ and $|\lambda_2|^{2d} \geq n^{-(1-\eta)}$; also $d = O(\log n) \leq n/2$ for large n , and we may assume $d \geq \max(d_0, d_1)$ (for bounded d the signals are bounded below and consistency is the classical plug-in consistency). By Lemma 4, $\|\Delta\|_1 = O_P(n^{-1/2})$.

TV ceiling. $\hat{\mathcal{V}}(d) = V(\hat{P}_d)$ and $\bar{\mathcal{V}}(d) = V(P_d)$, so by Lemma 5(i), $|\hat{\mathcal{V}}(d) - \bar{\mathcal{V}}(d)| \leq C_1 \|\Delta\|_1 = O_P(n^{-1/2})$, while $\bar{\mathcal{V}}(d) \geq c |\lambda_2|^d \geq cn^{-(1-\eta)/2}$ by Lemma 6. The relative error is $O_P(n^{-\eta/2}) \rightarrow 0$.

Temporal information. On the event $\|\Delta\|_1 \leq \pi_{\min}^2/4$, whose probability tends to one, Lemma 5(ii) gives $|\hat{\mathcal{T}}(d) - \mathcal{T}(d)| \leq \|L_d\|_\infty \|\Delta\|_1 + C_3 \|\Delta\|_1^2 = O_P(|\lambda_2|^d n^{-1/2} + n^{-1})$, while $\mathcal{T}(d) \geq \frac{1}{4} |\lambda_2|^{2d}$. The relative error is $O_P(n^{-1/2} |\lambda_2|^{-d} + n^{-1} |\lambda_2|^{-2d}) = O_P(n^{-\eta/2} + n^{-\eta}) \rightarrow 0$.

D.3 Proof of Theorem 5 (bias floors of the plug-ins)

Temporal information. Expanding as in Lemma 5(ii) to third order (the third derivatives of I are bounded on the same ball),

$$\hat{\mathcal{T}}(d) - \mathcal{T}(d) = \langle L_d, \Delta \rangle + Q_d(\Delta) + O(\|\Delta\|_1^3),$$

$$Q_d(\Delta) = \frac{1}{2} \sum_{s,s'} \frac{\Delta(s, s')^2}{P_d(s, s')} - \frac{1}{2} \sum_s \frac{\Delta_1(s)^2}{\pi(s)} - \frac{1}{2} \sum_{s'} \frac{\Delta_2(s')^2}{\pi(s')},$$

valid on the event $\|\Delta\|_1 \leq \pi_{\min}^2/4$, whose complement has probability $\leq 2N^2 e^{-cn}$ by Lemma 4 (and $|\hat{\mathcal{T}} - \mathcal{T}| \leq 2 \log N$ always). Taking expectations, $\mathbb{E}\langle L_d, \Delta \rangle = 0$ (stationarity), $\mathbb{E}\|\Delta\|_1^3 = O(n^{-3/2})$, and $\mathbb{E}Q_d(\Delta) = \Theta(n^{-1})$ by the following computation.

Reduction to a doubly centered form. Write $Q_d(\Delta) = \frac{1}{2} \sum_{s,s'} \tilde{\Delta}(s, s')^2 / (\pi(s)\pi(s')) + O(|\lambda_2|^d \|\Delta\|_1^2)$ with $\tilde{\Delta} = \Delta - \Delta_1 \otimes \pi - \pi \otimes \Delta_2$. Indeed, $\sum \tilde{\Delta}^2 / (\pi \otimes \pi) = \sum \Delta^2 / (\pi \otimes \pi) - \sum \Delta_1^2 / \pi - \sum \Delta_2^2 / \pi$ by direct expansion using $\sum_{s'} \Delta(s, s') = \Delta_1(s)$, $\sum_s \Delta(s, s') = \Delta_2(s')$ and $\sum_s \Delta_1(s) = \sum_{s'} \Delta_2(s') = 0$; replacing P_d by $\pi \otimes \pi$ in the denominator costs the $O(|\lambda_2|^d)$ factor. With $A(x) = \mathbf{1}[x = s] - \pi(s)$ and $B(y) = \mathbf{1}[y = s'] - \pi(s')$,

$$\tilde{\Delta}(s, s') = \frac{1}{m} \sum_{t < m} \xi_t(s, s'), \quad \xi_t = A(S_t)B(S_{t+d}) - \mathbb{E}[A(S_0)B(S_d)],$$

a centered, bounded functional of (S_t, S_{t+d}) (and $\mathbb{E}[A(S_0)B(S_d)] = P_d(s, s') - \pi(s)\pi(s')$).

Covariance bounds. Write $\rho = |\lambda_2|$ and use $\|K^j A\|_\infty \leq C\rho^j$, $\|K^j B\|_\infty \leq C\rho^j$ (Lemma 2). For $0 < j \leq d$ the times are ordered $0 < j \leq d < j+d$; conditioning on (S_j, S_d) and using reversibility for the past, $\mathbb{E}[A(S_0)A(S_j)B(S_d)B(S_{j+d})] = \mathbb{E}[(K^j A)(S_j)A(S_j)B(S_d)(K^j B)(S_d)]$, so $|\text{Cov}(\xi_0, \xi_j)| \leq$

$C\rho^{2j} + C\rho^{2d} \leq 2C\rho^{2j}$. For $j > d$ the same conditioning gives $|\text{Cov}(\xi_0, \xi_j)| \leq C\rho^{2d}$, and the $(j - d)$ -step gap between the two pairs gives $|\text{Cov}(\xi_0, \xi_j)| \leq C\rho^{j-d}$. Hence $\sum_j |\text{Cov}(\xi_0, \xi_j)| \leq C'$ uniformly in d , and the tail $\sum_{|j|>d} O(d\rho^{2d}) \rightarrow 0$.

The limit. For fixed j , as $d \rightarrow \infty$ the blocks (S_0, S_j) and (S_d, S_{d+j}) decouple, so $\text{Cov}(\xi_0, \xi_j) \rightarrow \mathbb{E}[A(S_0)A(S_j)]\mathbb{E}[B(S_0)B(S_j)]$. Writing A_s for the function A built from state s , $\sum_s \mathbb{E}[A_s(S_0)A_s(S_j)]/\pi(s) = \sum_s (K^{|j|}(s, s) - \pi(s)) = \sum_{k \geq 2} \lambda_k^{|j|}$, and dominated convergence gives

$$\gamma_d := \sum_{s, s'} \frac{\sigma_d^2(s, s')}{\pi(s)\pi(s')} \rightarrow \gamma_\infty = \sum_{j \in \mathbb{Z}} \left(\sum_{k \geq 2} \lambda_k^{|j|} \right)^2 = \sum_{k, l \geq 2} \frac{1 + \lambda_k \lambda_l}{1 - \lambda_k \lambda_l} \geq (N - 1)^2,$$

where $\sigma_d^2(s, s') = \sum_{j \in \mathbb{Z}} \text{Cov}(\xi_0, \xi_j)$ is the Markov-chain asymptotic variance (Paulin, 2015, Prop. 3.1). The $j = 0$ term is the i.i.d.-pair value $(N - 1)^2$; the excess is the integrated autocorrelation of the pair counts. (For the Beijing chain of Appendix E, $\gamma_\infty \approx 73.5$ against $(N - 1)^2 = 16$.) By the uniform summability above, $m \mathbb{E} \tilde{\Delta}(s, s')^2 = \sum_{|j| < m} (1 - |j|/m) \text{Cov}(\xi_0, \xi_j) \rightarrow \sigma_d^2(s, s')$ uniformly in $d \leq n/2$. Hence, for $d \geq d_2$ and $m \geq m_0$, $\frac{1}{2}\gamma_\infty \cdot \frac{1}{2m} \leq \mathbb{E}Q_d(\Delta) \leq 2\gamma_\infty \cdot \frac{1}{2m}$, that is, $\mathbb{E}\hat{\mathcal{T}}(d) - \mathcal{T}(d) = \Theta(n^{-1})$ uniformly in $d \geq d_2$.

Conclusion. Since $\mathcal{T}(d) = \Theta(|\lambda_2|^{2d}) \leq Cn^{-(1+\eta)}$ for $d \geq (1 + \eta)d^*(n)$, the bias exceeds the signal by the factor $n^\eta \rightarrow \infty$; in particular $\mathbb{E}\hat{\mathcal{T}}(d)/\mathcal{T}(d) \rightarrow \infty$, and (as $\hat{\mathcal{T}} \geq 0$) the plug-in is not consistent in relative error. The fluctuation of the quadratic term is of the same order: $Q_d(\Delta)$ is a positive quadratic form in the centered sums, so $\text{sd}(Q_d(\Delta)) \asymp \mathbb{E}Q_d(\Delta) \asymp n^{-1}$; to leading order its law is a weighted sum of $(N - 1)^2$ independent χ_1^2 variables with total weight γ_d , scaled by $1/2m$ (Appendix D.4). Hence subtracting the bias—by Miller–Madow, which removes only the i.i.d. constant $(N - 1)^2/2n$ and so leaves most of the floor $\gamma_d/2n$ in place, or by a block jackknife, which removes the mean floor nonparametrically—leaves a centered estimator whose standard deviation still exceeds $\mathcal{T}(d)$ by the factor n^η ; this is what Figure 4(c) shows, and it is the finite-sample face of Proposition 4.

TV ceiling. Here the functional is non-smooth at the null. Set $\hat{\varepsilon}_s = \hat{P}_d(s, \cdot)/\hat{\pi}_0(s) - \hat{\pi}_d$ and $\varepsilon_s = \varepsilon_s^{(d)}$, so $\hat{\mathcal{V}}(d) = \sum_s \hat{\pi}_0(s) \|\hat{\varepsilon}_s\|_{\text{TV}}$. By Lemma 4, $\hat{\varepsilon}_s - \varepsilon_s = \zeta_s + O_P(n^{-1})$ where $\zeta_s = \frac{1}{\pi(s)} (\Delta(s, \cdot) - \Delta_1(s)K^d(s, \cdot)) - \Delta_2$ is a linear functional of the centered sums with $\sqrt{m}\zeta_s \Rightarrow \mathcal{N}(0, \Gamma_s)$ (Markov CLT), and Γ_s is nondegenerate on the sum-zero subspace for d large: by the covariance bounds above, as $d \rightarrow \infty$ the limiting covariance is a positive combination of $\text{diag}(\pi) - \pi\pi^\top$ (the $j = 0$ term, scaled by $1/\pi(s)$) and positive semidefinite autocorrelation terms, so it has rank $N - 1$. Then, for $d \geq (1 + \eta)d^*(n)$ (so $\|\varepsilon_s\|_1 \leq Cn^{-(1+\eta)/2} = o(n^{-1/2})$), using $|\|\varepsilon_s + \zeta_s\|_{\text{TV}} - \|\zeta_s\|_{\text{TV}}| \leq \|\varepsilon_s\|_{\text{TV}}$,

$$\mathbb{E}\hat{\mathcal{V}}(d) = \sum_s \pi(s) \mathbb{E}\|\zeta_s\|_{\text{TV}} + O(|\lambda_2|^d + n^{-1}) = \frac{c_V}{\sqrt{n}}(1 + o(1)),$$

$$c_V = \sum_s \pi(s) \mathbb{E}\|G_s\|_{\text{TV}} > 0, \quad G_s \sim \mathcal{N}(0, \Gamma_s),$$

by uniform integrability of $\sqrt{m}\|\zeta_s\|_{\text{TV}}$ (bounded by a concentrating sum). Since $\bar{\mathcal{V}}(d) = \Theta(|\lambda_2|^d) = O(n^{-(1+\eta)/2})$ by Lemma 6, again $\mathbb{E}\hat{\mathcal{V}}(d)/\bar{\mathcal{V}}(d) \rightarrow \infty$ and the plug-in is not consistent.

Proposition 4 (Converse in the sampled-pairs model). *Let the learner observe n independent pairs $(X_i, Y_i) \sim P_d = \pi \otimes K^d$ from an unknown chain in the class $\mathcal{K} = \{K > 0 \text{ entrywise, reversible}\}$, and let $\hat{\theta}_n$ be any estimator of $\theta(K) \in \{\mathcal{T}(d), \bar{\mathcal{V}}(d)\}$ that is a function of these pairs (equivalently,*

of $\hat{P}_d$). If $d = d_n \geq (1 + \eta)d^*(n)$, then

$$\liminf_{n \rightarrow \infty} \sup_{K \in \mathcal{K}} \mathbb{P}_K \left(\left| \frac{\hat{\theta}_n}{\theta(K)} - 1 \right| \geq \frac{1}{3} \right) \geq \frac{1}{4},$$

so no estimator in the sampled-pairs model has vanishing relative error beyond the horizon.

Proof. Fix $K \in \mathcal{K}$ with eigenpairs (λ_k, ψ_k) , $\lambda_2 \neq 0$ and $|\lambda_2| > |\lambda_3|$, and for $\alpha \in (0, 1)$ let K_α have the same π and ψ_k , eigenvalues $\lambda'_2 = \alpha^{1/(2d)}\lambda_2$ and $\lambda'_k = \lambda_k$ ($k \geq 3$). Since $K > 0$ and $|\lambda'_2 - \lambda_2| \leq |\lambda_2|(1 - \alpha^{1/(2d)}) \rightarrow 0$ as $d \rightarrow \infty$, $K_\alpha = K + (\lambda'_2 - \lambda_2)\psi_2 m_2^\top$ is entrywise positive and stochastic, reversible w.r.t. π , i.e. $K_\alpha \in \mathcal{K}$, for d large. Since $(\lambda'_2)^{2d} = \alpha\lambda_2^{2d}$ and $|\lambda'_2|^d = \sqrt{\alpha}|\lambda_2|^d$, Theorems 3(S) and 2 give $\mathcal{T}_\alpha(d)/\mathcal{T}(d) \rightarrow \alpha$ and $\bar{\mathcal{V}}_\alpha(d)/\bar{\mathcal{V}}(d) \rightarrow \alpha^{1/2}$. Take $\alpha = \frac{1}{3}$ for $\mathcal{T}$ and $\alpha = \frac{1}{9}$ for $\bar{\mathcal{V}}$, so that $\theta' = \theta(K_\alpha)$ satisfies $\theta'/\theta \rightarrow \frac{1}{3}$; for large d , $\theta' < \frac{1}{2}\theta$, and the relative-error balls $\{x : |x/\theta - 1| < \frac{1}{3}\} = (\frac{2}{3}\theta, \frac{4}{3}\theta)$ and $\{x : |x/\theta' - 1| < \frac{1}{3}\} \subset (0, \frac{2}{3}\theta)$ are disjoint. The two pair laws differ by $(P'_d - P_d)(s, s') = (\lambda_2'^d - \lambda_2^d)\pi(s)\pi(s')\psi_2(s)\psi_2(s')$, with $|\lambda_2'^d - \lambda_2^d| = (1 - \sqrt{\alpha})|\lambda_2|^d \leq |\lambda_2|^d$. Using $P_d \geq \pi \otimes \pi/2$ for $d \geq d_0$ and $\|\psi_2\|_{2,\pi} = 1$,

$$\chi^2(P'_d \| P_d) = \sum_{s, s'} \frac{(P'_d - P_d)^2}{P_d} \leq 2(\lambda_2^d - \lambda_2'^d)^2 \sum_{s, s'} \pi(s)\pi(s')\psi_2(s)^2\psi_2(s')^2 \leq 2|\lambda_2|^{2d},$$

so for $d \geq (1 + \eta)d^*(n)$, $\|P_d^{\otimes n} - P_d'^{\otimes n}\|_{\text{TV}}^2 \leq \frac{1}{2}\text{KL}(P_d'^{\otimes n} \| P_d^{\otimes n}) \leq \frac{n}{2}\chi^2(P'_d \| P_d) \leq n|\lambda_2|^{2d} \leq n^{-\eta} \rightarrow 0$. Since the two balls are disjoint, Le Cam's two-point lemma (Tsybakov, 2009, Thm. 2.2) gives $\max\{\mathbb{P}_K(|\hat{\theta}_n/\theta - 1| \geq \frac{1}{3}), \mathbb{P}_{K_\alpha}(|\hat{\theta}_n/\theta' - 1| \geq \frac{1}{3})\} \geq \frac{1}{2}(1 - \|P_d^{\otimes n} - P_d'^{\otimes n}\|_{\text{TV}}) \rightarrow \frac{1}{2} \geq \frac{1}{4}$. $\square$

Remark 6. The converse is stated for independent pairs because a trajectory carries more than its lag- d statistics: from $S_0, \dots, S_n$ the one-step kernel, and hence λ_2 , is estimable to $O(n^{-1/2})$, which is exactly the model-based route of Proposition 2. Theorem 6 shows that this route is optimal among all trajectory-based estimators.

D.4 Proof of Proposition 1 (no data-independent correction)

Let $d \geq (1 + \eta)d^*(n)$ and work on the event of Lemma 4. From the expansion above, $\hat{\mathcal{T}}(d) = \mathcal{T}(d) + \langle L_d, \Delta \rangle + Q_d(\Delta) + O(\|\Delta\|_1^3)$. The first term has standard deviation $O(\|L_d\|_\infty n^{-1/2}) = O(|\lambda_2|^d n^{-1/2}) = o(n^{-1})$ and the last is $O_P(n^{-3/2})$, so $n(\hat{\mathcal{T}}(d) - \mathbb{E}\hat{\mathcal{T}}(d)) = n(Q_d(\Delta) - \mathbb{E}Q_d(\Delta)) + o_P(1)$.

Limit law. $m\tilde{\Delta} = \sum_t \xi_t$ with ξ_t a centered bounded functional of (S_t, S_{t+d}) ; the Markov-chain CLT (Jones, 2004) gives $\sqrt{m}\tilde{\Delta} \Rightarrow \mathcal{N}(0, \Sigma_d)$ jointly over (s, s') , and $mQ_d(\Delta) \Rightarrow W_d = \frac{1}{2}G^\top \text{diag}(1/\pi \otimes \pi)G$, $G \sim \mathcal{N}(0, \Sigma_d)$. By the covariance computation in the previous subsection, $\Sigma_d \rightarrow \Sigma_\infty = \sum_{j \in \mathbb{Z}} C_j \otimes C_j$ with $C_j = \text{Cov}(e_{S_0}, e_{S_j}) = D(K^{|j|} - \mathbf{1}\pi)$, $D = \text{diag}(\pi)$. In the eigenbasis, $C_j \otimes C_j = \sum_{k, l \geq 2} (\lambda_k \lambda_l)^{|j|} (m_k \otimes m_l)(m_k \otimes m_l)^\top$, so $\Sigma_\infty = \sum_{k, l \geq 2} \frac{1 + \lambda_k \lambda_l}{1 - \lambda_k \lambda_l} (m_k \otimes m_l)(m_k \otimes m_l)^\top$ is positive semidefinite of rank $(N - 1)^2$. Hence W_∞ is a weighted sum of $(N - 1)^2$ independent χ_1^2 variables with positive weights: it has a continuous law and variance $v_\infty > 0$. Uniform integrability of $(mQ_d(\Delta))^2$ follows from $mQ_d(\Delta) \leq Cm\|\Delta\|_1^2$ and the sub-Gaussian tails of Lemma 4, so $\text{Var}(mQ_d(\Delta)) \rightarrow v_\infty$ and $\text{sd}(\hat{\mathcal{T}}(d)) = \sqrt{v_\infty}/n(1 + o(1))$. Since $\mathcal{T}(d) \leq Cn^{-(1+\eta)}$, $\text{sd}(\hat{\mathcal{T}}(d))/\mathcal{T}(d) \geq cn^\eta \rightarrow \infty$.

Any deterministic correction fails. The event $|\hat{\mathcal{T}}(d) - \beta_n|/\mathcal{T}(d) < \frac{1}{3}$ is $\{n\hat{\mathcal{T}}(d) \in J_n\}$ with $J_n = (n\beta_n + \frac{2}{3}n\mathcal{T}(d), n\beta_n + \frac{4}{3}n\mathcal{T}(d))$, an interval of length $\frac{2}{3}n\mathcal{T}(d) \leq Cn^{-\eta} \rightarrow 0$. Since $n(\hat{\mathcal{T}}(d) - \mathbb{E}\hat{\mathcal{T}}(d))$ converges in law to $W_\infty - \mathbb{E}W_\infty$, whose law is continuous, the convergence of distribution functions is uniform (Pólya), and the probability of an interval of vanishing length tends to zero, whatever its location.

D.5 Proof of Theorem 6 (sharp model-based horizon)

Let $\kappa < 1/(2N)$ and fix $K \in \mathcal{K}_{2\kappa}$ with eigenpairs (λ_k, ψ_k) . (The class is nonempty: with π uniform, $K = \mathbf{1}\pi^\top + \eta\psi m^\top$ for any $\psi \perp_\pi \mathbf{1}$, $\|\psi\|_{2,\pi} = 1$, $m = \pi\psi$, has $\lambda_2 = \eta$, $\lambda_3 = 0$, and $K \geq 2\kappa$ for η small, since $1/N > 2\kappa$.) Put $\varepsilon_n = c_0/\sqrt{n}$ with $c_0 = \sqrt{\kappa}/2$, and $K' = K + \varepsilon_n \lambda_2 \psi_2 m_2^\top$: the same π and eigenvectors, with $\lambda'_2 = \lambda_2(1 + \varepsilon_n)$ and $\lambda'_k = \lambda_k$ otherwise. Rows of K' sum to one ($\sum_{s'} m_2(s') = 0$), K' is reversible w.r.t. π (same eigenbasis), and, since $\pi(s) = \sum_x \pi(x)K(x, s) \geq 2\kappa$ gives $|\psi_2(s)m_2(s')| \leq \pi_{\min}^{-1/2} \leq (2\kappa)^{-1/2}$, we have $K' \geq 2\kappa - \varepsilon_n(2\kappa)^{-1/2} \geq \kappa$ for $n \geq n_0(\kappa)$; also $|\lambda_3| < |\lambda_2| < |\lambda'_2| \leq \bar{\rho} := \frac{1}{2}(1 + |\lambda_2|) < 1$ for $n \geq n_0$. So $K, K' \in \mathcal{K}_\kappa$.

Indistinguishability. Both chains are started from π . By the chain rule, $\text{KL}(\mathbb{P}_K^{(n)} \|\mathbb{P}_{K'}^{(n)}) = n \sum_s \pi(s) \text{KL}(K(s, \cdot) \| K'(s, \cdot))$. Using $\text{KL} \leq \chi^2$, $K' \geq \kappa$ and $\pi(s') \leq 1$,

$$\text{KL}(K(s, \cdot) \| K'(s, \cdot)) \leq \sum_{s'} \frac{(K(s, s') - K'(s, s'))^2}{K'(s, s')} \leq \frac{\varepsilon_n^2 \lambda_2^2 \psi_2(s)^2}{\kappa} \sum_{s'} \pi(s')^2 \psi_2(s')^2 \leq \frac{\varepsilon_n^2 \psi_2(s)^2}{\kappa},$$

where the last step uses $|\lambda_2| \leq 1$ and $\sum_{s'} \pi(s')^2 \psi_2(s')^2 \leq \|\psi_2\|_{2,\pi}^2 = 1$. Averaging over $s \sim \pi$ and using $\|\psi_2\|_{2,\pi} = 1$ again, $\text{KL}(\mathbb{P}_K^{(n)} \|\mathbb{P}_{K'}^{(n)}) \leq n\varepsilon_n^2/\kappa = c_0^2/\kappa = \frac{1}{2}$, so $\|\mathbb{P}_K^{(n)} - \mathbb{P}_{K'}^{(n)}\|_{\text{TV}} \leq \sqrt{\frac{1}{2}\text{KL}} \leq \frac{1}{2}$ by Pinsker's inequality, for every $n \geq n_0$ and every d .

Separation. Put $c = 2 \log 3/c_0 = 2 \log 3 \sqrt{2/\kappa}$ and let $d = d_n \geq c\sqrt{n}$. Since $\log(1+x) \geq x/2$ for $x \in [0, 1]$, $(|\lambda'_2|/|\lambda_2|)^d = (1 + \varepsilon_n)^d \geq e^{d\varepsilon_n/2} \geq e^{cc_0/2} = 3$ and $(\lambda'_2/\lambda_2)^{2d} \geq 9$. K and K' share π and the eigenbasis, and differ only in λ_2 . By Theorem 3(S) and (6), with constants depending only on π and the common eigenbasis,

$$\mathcal{T}_K(d) = \frac{1}{2}\lambda_2^{2d}(1 + r_d), \quad \mathcal{T}_{K'}(d) = \frac{1}{2}\lambda_2'^{2d}(1 + r'_d),$$

$$\bar{\mathcal{V}}_K(d) = \frac{1}{2}\|\psi_2\|_{1,\pi}^2 |\lambda_2|^d (1 + q_d), \quad \bar{\mathcal{V}}_{K'}(d) = \frac{1}{2}\|\psi_2\|_{1,\pi}^2 |\lambda_2'|^d (1 + q'_d),$$

where $|r_d|, |r'_d|, |q_d|, |q'_d| \leq C((|\lambda_3|/|\lambda_2|)^d + \bar{\rho}^d) \rightarrow 0$ as $n \rightarrow \infty$ (because $d_n \rightarrow \infty$, $|\lambda_3|/|\lambda_2| \leq |\lambda_3|/|\lambda_2| < 1$ and $|\lambda_2| \leq \bar{\rho}$). Hence, for $\theta \in \{\mathcal{T}(d), \bar{\mathcal{V}}(d)\}$ and θ' its value under K' , $\theta'/\theta \geq 3(1 + o(1)) > 2$ for $n \geq n_1(\kappa, K)$. Then the relative-error balls $B = \{x : |x/\theta - 1| < \frac{1}{3}\} = (\frac{2}{3}\theta, \frac{4}{3}\theta)$ and $B' = (\frac{2}{3}\theta', \frac{4}{3}\theta')$ are disjoint, since $\frac{2}{3}\theta' > \frac{4}{3}\theta$.

Le Cam. Let $A = \{\hat{\theta}_n \in B\}$. As $A \cap \{\hat{\theta}_n \in B'\} = \emptyset$, $\mathbb{P}_K(|\hat{\theta}_n/\theta - 1| \geq \frac{1}{3}) + \mathbb{P}_{K'}(|\hat{\theta}_n/\theta' - 1| \geq \frac{1}{3}) \geq \mathbb{P}_K(A^c) + \mathbb{P}_{K'}(A) \geq 1 - \|\mathbb{P}_K^{(n)} - \mathbb{P}_{K'}^{(n)}\|_{\text{TV}} \geq \frac{1}{2}$ (Tsybakov, 2009, Thm. 2.2). So for $n \geq \max(n_0, n_1)$ one of $K, K' \in \mathcal{K}_\kappa$ has failure probability at least $\frac{1}{4}$, and $\liminf_n \sup_{K \in \mathcal{K}_\kappa} \mathbb{P}_K(|\hat{\theta}_n/\theta(K) - 1| \geq \frac{1}{3}) \geq p := \frac{1}{4}$, with $c = 2 \log 3 \sqrt{2/\kappa}$ and p depending only on κ . Together with Proposition 2, which gives consistency for $d_n = o(\sqrt{n})$, $\mathcal{T}(d)$ is learnable from a trajectory exactly up to $d = \Theta(\sqrt{n})$. $\square$

D.6 Proof of Theorem 7 (online recovery)

Notation. For a state s write $r.(s) = (r_a(s))_{a \in \mathcal{A}} \in [0, 1]^{\mathcal{A}}$ for the rewards of all actions in state s ; for a randomized action $p \in \Delta(\mathcal{A})$, $\langle p, r.(s) \rangle = \sum_a p(a)r_a(s)$ is its payoff in state s . The pulled-back reward of action a (Section 5, Appendix C) is the function $(K^d r_a)(s) = \sum_{s'} K^d(s, s') r_a(s') = \mathbb{E}[r_a(S_t) \mid S_{t-d} = s] = \langle K^d(s, \cdot), r_a \rangle$, the expected reward of a now, given that the state was s at d steps earlier. Accordingly $(K^d r.)(s) \in [0, 1]^{\mathcal{A}}$ is the vector of these expected rewards, and $\langle p, (K^d r.)(s) \rangle = \sum_a p(a)(K^d r_a)(s)$ is the expected payoff of p in context s .

Setup. Index the stationary chain from time $1 - d$. The learner observes this single state process with lag d : at round $t = 1, \dots, n$ the newest state it has observed is S_{t-d} . This observation plays two roles. It is the *context* for round t , and it is the *feedback* for round $t - d$, since it reveals $r.(S_{t-d})$, the reward vector that determined the payoff of round $t - d$. The protocol at round t is:

1. observe S_{t-d} (context for round t ; feedback for round $t - d$),
2. play $p_t \in \Delta(\mathcal{A})$ (a randomized action),
3. earn $X_t = \langle p_t, r.(S_t) \rangle$ (realized now, not observed),
4. $r.(S_t)$ is revealed at round $t + d$ (when S_t is observed).

The learner commits to p_t without seeing S_t ; the world then pays according to S_t , and the learner learns the amount d rounds later. (For example, with hourly rounds and $d = 2$, an alert decided at 3 pm uses the 1 pm reading, and whether 3 pm was polluted becomes known at 5 pm.)

Two objects should be kept apart. The learner plays $p_t \in \Delta(\mathcal{A})$, a distribution over actions. A Bayesian who knew K would hold the belief $b_t := K^d(S_{t-d}, \cdot) \in \Delta(\mathcal{S})$ about the current state, and would play a Bayes action $\arg \max_a \langle b_t, r_a \rangle$, earning $\Phi(b_t)$ in expectation. The learner knows neither K nor b_t ; it uses only the observed contexts and the revealed reward vectors, so the maps r_a may be known or unknown. Theorem 7 states that the learner's average gain over the blind action nevertheless approaches $\mathcal{V}_\Phi(d) = \mathbb{E}_{s \sim \pi} \Phi(K^d(s, \cdot)) - \Phi(\pi)$, the value of acting on b_t .

Let U be the learner's internal randomness and $\mathcal{F}_t = \sigma(S_{\leq t-d}, U)$. At round t the learner has seen the contexts $S_{\leq t-d}$ and the reward vectors $r.(S_{t'})$ for $t' \leq t - d$; the latter are functions of $S_{\leq t-d}$, so p_t is $\mathcal{F}_t$ -measurable. Rewards lie in $[0, 1]$, so $0 \leq \Phi \leq 1$ and $|\Phi(b) - \Phi(b')| \leq \|b - b'\|_{\text{TV}}$ (Theorem 2). Write $X_t := \langle p_t, r.(S_t) \rangle$, $M_t := X_t - \mathbb{E}[X_t | \mathcal{F}_t]$, and let $\hat{\pi}$ be the empirical law of $S_{1-d}, \dots, S_{n-d}$ and $\hat{P}_d$ the empirical law of the n pairs (S_{t-d}, S_t) , $t = 1, \dots, n$. The proof combines four lemmas: a Bayes bound per round (Lemma 7), lag- d noise (Lemma 8), concentration of time averages (Lemma 9), and a regret bound for delayed per-context Hedge (Lemma 10).

Lemma 7 (Stale Bayes bound). *For every policy and every t , with $b_t = K^d(S_{t-d}, \cdot)$, $\mathbb{E}[X_t | \mathcal{F}_t] = \langle p_t, (K^d r.) (S_{t-d}) \rangle = \sum_a p_t(a) \langle b_t, r_a \rangle \leq \Phi(b_t)$, with equality iff p_t is supported on $\arg \max_a \langle b_t, r_a \rangle$.*

Proof. U is independent of the chain, and by the Markov property S_t depends on $S_{\leq t-d}$ only through S_{t-d} , with conditional law b_t . Hence $\mathbb{E}[r_a(S_t) | \mathcal{F}_t] = (K^d r_a)(S_{t-d}) = \langle b_t, r_a \rangle$. Since p_t is $\mathcal{F}_t$ -measurable, $\mathbb{E}[X_t | \mathcal{F}_t] = \sum_a p_t(a) \langle b_t, r_a \rangle \leq \max_a \langle b_t, r_a \rangle = \Phi(b_t)$, with equality exactly when p_t puts no mass on a non-maximizer. $\square$

Lemma 8 (Lag- d noise). *For every policy, $\frac{1}{n} \sum_{t \leq n} M_t = O_P(\sqrt{(d+1)/n})$.*

Proof. M_t is a function of $S_{\leq t}$ and U , hence $\mathcal{F}_{t+d}$ -measurable, and $\mathbb{E}[M_{t'} | \mathcal{F}_{t'}] = 0$. If $t' \geq t + d$, then M_t is $\mathcal{F}_{t'}$ -measurable and $\mathbb{E}[M_t M_{t'}] = \mathbb{E}[M_t \mathbb{E}[M_{t'} | \mathcal{F}_{t'}]] = 0$. So only pairs with $|t - t'| < d$ are correlated, and $|M_t| \leq 1$ gives $\text{Var}(\sum_{t \leq n} M_t) \leq (2d+1)n$. Apply Chebyshev. $\square$

Lemma 9 (Time averages). *Uniformly in $d \leq n/2$: (a) $\|\hat{\pi} - \pi\|_1 = O_P(\sqrt{\tau_{\text{mix}}/n})$ and $\|\hat{P}_d - P_d\|_1 = O_P(\sqrt{\tau_{\text{mix}}/n})$; (b) for every fixed $f : \mathcal{S} \rightarrow [0, 1]$, both $\frac{1}{n} \sum_{t \leq n} f(S_{t-d})$ and $\frac{1}{n} \sum_{t \leq n} f(S_t)$ equal $\mathbb{E}_\pi f + O_P(\sqrt{\tau_{\text{mix}}/n})$.*

Proof. (a) is Lemma 4 applied to the stationary path $S_{1-d}, \dots, S_n$. This path has length $n + d$ and n lag- d pairs, and $d \leq n/2 \leq (n + d)/2$. $\hat{\pi}$ is the first marginal of $\hat{P}_d$. (b) The first average is $\sum_s \hat{\pi}(s) f(s)$, and $|\sum_s (\hat{\pi} - \pi)(s) f(s)| \leq \|\hat{\pi} - \pi\|_1$. The second average is the same with $\hat{\pi}$ replaced by the second marginal of $\hat{P}_d$, which is within $\|\hat{P}_d - P_d\|_1$ of π . $\square$

Lemma 10 (Delayed per-context Hedge). *For each s let $B_s = \{t \leq n : S_{t-d} = s\}$ be the rounds with context s , and $n_s = |B_s|$. Within B_s , run $d + 1$ copies of Hedge in round-robin; copy j plays the rounds of B_s whose rank in B_s is $\equiv j \pmod{d+1}$. Then, for a universal constant C , $\sum_{t \in B_s} X_t \geq \max_a \sum_{t \in B_s} r_a(S_t) - C\sqrt{(d+1)n_s \log |\mathcal{A}|}$.*

Proof. Two consecutive rounds $t < t'$ of the same copy have d other rounds of B_s between them, so $t' \geq t + d + 1$. The reward vector $r_{\cdot}(S_t)$ of round t is revealed at round $t + d < t'$. Each copy has therefore seen all its feedback before it plays again (Joulani et al., 2013), and faces an undelayed full-information problem with reward vectors $r_{\cdot}(S_t) \in [0, 1]^{\mathcal{A}}$. These vectors, and the sets B_s , depend only on the chain and not on the learner's play, so the adversary is oblivious and the regret bound holds pathwise. Copy j plays $n_{s,j}$ rounds. The anytime Hedge bound (Cesa-Bianchi and Lugosi, 2006, Thm. 2.3) gives regret at most $C\sqrt{n_{s,j} \log |\mathcal{A}|}$; its additive $O(\sqrt{\log |\mathcal{A}|})$ term is absorbed because $n_{s,j} \geq 1$ for every copy that plays. The best action of copy j may differ across copies, so the sum of the copies' comparators is at least $\max_a \sum_{t \in B_s} r_a(S_t)$. Summing over j and using $\sum_j \sqrt{n_{s,j}} \leq \sqrt{(d+1)n_s}$ (Cauchy-Schwarz) gives the display. $\square$

Remark 7 (Why $d + 1$ copies). Within a round, observation precedes play, so $r_{\cdot}(S_t)$ is already available when round $t + d$ is played. Hence d copies suffice when $d \geq 1$ (consecutive rounds of a copy then satisfy $t' \geq t + d$). We use $d + 1$ copies so that the case $d = 0$ (one copy, undelayed Hedge) is covered by the same statement; this affects only the constant.

Proof of (i). For every policy,

$$\begin{aligned} \text{Adv}_n &= \underbrace{\frac{1}{n} \sum_{t \leq n} \mathbb{E}[X_t \mid \mathcal{F}_t]}_{\leq \frac{1}{n} \sum_t \Phi(K^d(S_{t-d}, \cdot))} + \frac{1}{n} \sum_{t \leq n} M_t - \frac{1}{n} \sum_{t \leq n} r_{a^*}(S_t), \end{aligned}$$

where the brace is Lemma 7. Apply Lemma 9(b) with $f(s) = \Phi(K^d(s, \cdot))$ to the first term, giving $\mathbb{E}_{s \sim \pi} \Phi(K^d(s, \cdot)) + O_P(\sqrt{\tau_{\text{mix}}/n})$. Apply Lemma 8 to the second term, giving $O_P(\sqrt{(d+1)/n})$. Apply Lemma 9(b) with $f = r_{a^*}$ to the third term, giving $\langle \pi, r_{a^*} \rangle + O_P(\sqrt{\tau_{\text{mix}}/n}) = \Phi(\pi) + O_P(\sqrt{\tau_{\text{mix}}/n})$. Hence $\text{Adv}_n \leq \mathcal{V}_{\Phi}(d) + O_P(\sqrt{(\tau_{\text{mix}} + d + 1)/n})$.

Proof of (ii). Sum Lemma 10 over s and use $\sum_s \sqrt{n_s} \leq \sqrt{Nn}$:

$$\frac{1}{n} \sum_{t \leq n} X_t \geq \frac{1}{n} \sum_s \max_a \sum_{t \in B_s} r_a(S_t) - C\sqrt{N(d+1) \log |\mathcal{A}|/n}.$$

Let $\hat{q}_s$ be the empirical law of $\{S_t\}_{t \in B_s}$, the empirical counterpart of $b_t = K^d(s, \cdot)$ on the rounds with context s . Then

$$\max_a \sum_{t \in B_s} r_a(S_t) = n_s \Phi(\hat{q}_s) \geq n_s \Phi(K^d(s, \cdot)) - n_s \|\hat{q}_s - K^d(s, \cdot)\|_{\text{TV}}.$$

Since $n_s \hat{q}_s(s') = n \hat{P}_d(s, s')$, $n_s = n \hat{\pi}(s)$ and $P_d(s, s') = \pi(s) K^d(s, s')$,

$$\begin{aligned} \frac{1}{n} \sum_s n_s \|\hat{q}_s - K^d(s, \cdot)\|_{\text{TV}} &= \frac{1}{2} \sum_{s, s'} |\hat{P}_d(s, s') - \hat{\pi}(s) K^d(s, s')| \\ &\leq \frac{1}{2} (\|\hat{P}_d - P_d\|_1 + \|\hat{\pi} - \pi\|_1) = O_P(\sqrt{\tau_{\text{mix}}/n}) \end{aligned}$$

by Lemma 9(a). Finally,

$$\frac{1}{n} \sum_s n_s \Phi(K^d(s, \cdot)) = \sum_s \hat{\pi}(s) \Phi(K^d(s, \cdot)) = \mathbb{E}_{s \sim \pi} \Phi(K^d(s, \cdot)) + O_P(\sqrt{\tau_{\text{mix}}/n}),$$

and the blind term is handled as in (i). Hence $\text{Adv}_n \geq \mathcal{V}_\Phi(d) - O_P(\sqrt{(\tau_{\text{mix}} + N(d+1) \log |\mathcal{A}|)/n})$. Combined with (i), this proves the display.

Horizon. For a nondegenerate boundary decision $\mathcal{V}_\Phi(d) \geq c|\lambda_2|^d \geq cn^{-(1-\eta)/2}$ when $d \leq (1-\eta)d^*(n)$, while $d = O(\log n)$ makes the error $O_P(\sqrt{\log n/n})$. The relative error is therefore $O_P(\sqrt{\log n} n^{-\eta/2}) \rightarrow 0$.

D.7 Proof of Proposition 2 (value of a model)

The estimator. Let $n_{ss'}$ be the number of transitions $s \rightarrow s'$ in $S_0, \dots, S_n$ and set

$$\hat{Q}(s, s') = \frac{n_{ss'} + n_{s's}}{2n}, \quad \hat{\pi}(s) = \sum_u \hat{Q}(s, u), \quad \hat{K}(s, s') = \frac{\hat{Q}(s, s')}{\hat{\pi}(s)}.$$

$\hat{Q}$ is symmetric by construction, so $\hat{\pi}(s)\hat{K}(s, s') = \hat{Q}(s, s')$ is symmetric: $\hat{K}$ is reversible w.r.t. $\hat{\pi}$, and summing over s shows that $\hat{\pi}$ is exactly its stationary law. Since $K > 0$ entrywise, every transition is observed with probability tending to one; on that event $\hat{\pi} > 0$ and $\hat{K} > 0$, so $\hat{K}$ is irreducible and aperiodic and its top eigenvalue 1 is simple. Write $D = \text{diag}(\pi)$, $\hat{D} = \text{diag}(\hat{\pi})$, $Q(s, s') = \pi(s)K(s, s')$, and $A = D^{-1/2}QD^{-1/2} = D^{1/2}KD^{-1/2}$, $\hat{A} = \hat{D}^{-1/2}\hat{Q}\hat{D}^{-1/2} = \hat{D}^{1/2}\hat{K}\hat{D}^{-1/2}$. Both are real symmetric (Q is symmetric by reversibility of K) and similar to K and $\hat{K}$, so their eigenvalues are those of K and $\hat{K}$: real, with top eigenvalue exactly 1.

The two symmetric matrices are close. With $\hat{P}_1$ the empirical law of the n consecutive pairs and $P_1 = Q$, we have $\hat{Q} = \frac{1}{2}(\hat{P}_1 + \hat{P}_1^\top)$ and, by reversibility, $Q = \frac{1}{2}(P_1 + P_1^\top)$. Hence $\max_{s, s'} |\hat{Q}(s, s') - Q(s, s')| \leq \max_{s, s'} |\hat{P}_1(s, s') - P_1(s, s')| = O_P(n^{-1/2})$ by Lemma 4 with $d = 1$, and $\|\hat{\pi} - \pi\|_1 \leq \|\hat{P}_1 - P_1\|_1 = O_P(n^{-1/2})$ (this is also the minimax rate of [Wolfer and Kontorovich, 2019](#)). The map $(Q, \pi) \mapsto D^{-1/2}QD^{-1/2}$ is Lipschitz on $\{\pi \geq \pi_{\min}/2\}$, with a constant depending only on $\pi_{\min}$, and $\hat{\pi} \geq \pi_{\min}/2$ with probability tending to one. Therefore every entry of $\hat{A} - A$ is $O_P(n^{-1/2})$, and, N being fixed, $\varepsilon_n := \|\hat{A} - A\|_{\text{op}} \leq \|\hat{A} - A\|_F \leq N \max_{s, s'} |(\hat{A} - A)(s, s')| = O_P(n^{-1/2})$.

Weyl's inequality. Order the eigenvalues of A and of $\hat{A}$ algebraically, $1 = \nu_1 > \nu_2 \geq \dots \geq \nu_N$ and $1 = \hat{\nu}_1 > \hat{\nu}_2 \geq \dots \geq \hat{\nu}_N$. For real symmetric matrices Weyl's inequality gives $|\hat{\nu}_j - \nu_j| \leq \varepsilon_n$ for every j ; no eigen-gap is needed. The paper indexes eigenvalues by modulus instead, but $\{\nu_j\}_{j \geq 2} = \{\lambda_k\}_{k \geq 2}$ as multisets (the top eigenvalue 1 is simple in both matrices), and $\sum_{k \geq 2} \lambda_k^{2d}$ is a symmetric function of this multiset. We may therefore pair the eigenvalues in algebraic order, and relabel $\hat{\lambda}_k$ so that $\max_{k \geq 2} |\hat{\lambda}_k - \lambda_k| \leq \varepsilon_n = O_P(n^{-1/2})$.

From eigenvalues to their powers. For every $k \geq 2$, $|\lambda_k| \leq |\lambda_2|$ and $|\hat{\lambda}_k| \leq |\lambda_2| + \varepsilon_n$, so the mean-value theorem gives

$$|\hat{\lambda}_k^{2d} - \lambda_k^{2d}| \leq 2d \varepsilon_n (|\lambda_2| + \varepsilon_n)^{2d-1} = \frac{2d \varepsilon_n}{|\lambda_2|} |\lambda_2|^{2d} \left(1 + \frac{\varepsilon_n}{|\lambda_2|}\right)^{2d-1}$$

($\lambda_2 \neq 0$). For $d = o(\sqrt{n})$, $(1 + \varepsilon_n/|\lambda_2|)^{2d} = \exp(O_P(dn^{-1/2})) = 1 + o_P(1)$. Summing over the $N-1$ indices, $|\frac{1}{2} \sum_{k \geq 2} \hat{\lambda}_k^{2d} - \frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d}| = |\lambda_2|^{2d} O_P(dn^{-1/2})$.

Conclusion. By Lemma 6, $\frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d} \geq \frac{1}{2} |\lambda_2|^{2d}$ and $\frac{1}{2} \sum_{k \geq 2} \lambda_k^{2d} = \mathcal{T}(d)(1 + O(|\lambda_2|^d))$. Dividing, $\tilde{\mathcal{T}}(d) = \mathcal{T}(d)(1 + O_P(dn^{-1/2}) + O(|\lambda_2|^d))$. The relative error vanishes when $d = d_n \rightarrow \infty$ with $d_n = o(\sqrt{n})$; for bounded d the $O(|\lambda_2|^d)$ bias of the quadratic proxy remains. No simplicity of $|\lambda_2|$ is needed. The constants in O_P depend on K only through N , $\pi_{\min}$, $|\lambda_2|$ and τ_{mix} . The condition $d = o(\sqrt{n})$ enters only through the factor $(1 + \varepsilon_n/|\lambda_2|)^{2d}$; at $d \asymp \sqrt{n}$ it is a constant multiplicative error, which is where Theorem 6 applies.

E Experimental details

Chain construction. Given a target stationary law π and a target spectrum $\{1, \lambda_2, \dots, \lambda_N\}$, choose a π -orthonormal basis $\psi_1 = \mathbf{1}, \psi_2, \dots, \psi_N$ (Gram–Schmidt in $\ell^2(\pi)$ from a random start), and set $K(s, s') = \pi(s')[1 + \sum_{k \geq 2} \lambda_k \psi_k(s) \psi_k(s')]$. Reversibility holds by construction (Lemma 1); nonnegativity is checked and the basis resampled otherwise. For the 3-state chain of Figures 1 and 2: $\pi = (0.2, 0.3, 0.5)$, $\{\lambda_k\} = \{1, 0.7, 0.3\}$, $\mu = 0.357$, $\max_{s, s'} |\pi(s)K(s, s') - \pi(s')K(s', s)| < 10^{-16}$.

Decisions. Smooth: $\Phi(b) = -H(b)$ (temporal information) and a soft-max payoff $\Phi(b) = \tau \log \sum_a e^{(b, r_a)/\tau}$. Boundary: two actions with $v = r_{a_1} - r_{a_2}$ chosen so that $\langle \pi, v \rangle = 0$ and $|\cos_\pi(v, \psi_2)| = 0.51$. Hard: actions “bet on state s ” ($r_s = e_s$) plus a hedge $r_h = (\pi(s) + \gamma)_s$ with γ chosen so the hedge is blind-optimal with margin $\delta > 0$; the observed deadline is $D_0 = 6$ and the bound (10) gives 9.8.

Learnability. Trajectories of length $n \in \{2 \times 10^3, 2 \times 10^4, 2 \times 10^5\}$ from the 3-state chain, 50 independent replicates; the plug-in $\tilde{\mathcal{T}}(d)$ is the empirical MI of (S_t, S_{t+d}) ; the model-based $\tilde{\mathcal{T}}(d) = \frac{1}{2} \sum_{k \geq 2} \hat{\lambda}_k^{2d}$ uses eigenvalues of the empirical one-step kernel. The “empirical detach” point in Figure 3b is the smallest d at which the plug-in’s relative error exceeds 50%.

Beijing chain. The UCI Beijing PM_{2.5} data set (Liang et al., 2015) has 43,824 hourly rows (2010–2014), 41,757 with a valid PM_{2.5} reading; missing hours break the sample path and are excluded from the transition counts. States are the quintiles of the valid readings (cut points 23, 54, 93, 157 $\mu\text{g}/\text{m}^3$); $\hat{K}$ is the row-normalized transition count matrix. Its eigenvalues are real, $\{1, 0.939, 0.828, 0.680, 0.535\}$, the stationary law is $(0.20, 0.20, 0.20, 0.20, 0.20)$ up to 3×10^{-3} , and $\max_{s, s'} |\hat{\pi}(s)\hat{K}(s, s') - \hat{\pi}(s')\hat{K}(s', s)| = 3.6 \times 10^{-3}$; we use $\hat{K}$ as the world and its right eigenvector $\psi_2 = (-1.36, -0.77, -0.01, 0.71, 1.44)$, normalized in $\ell^2(\hat{\pi})$. Panel (a): $\mathcal{T}(d)$, $\bar{\mathcal{V}}(d)$, the hard-alert value and the soft-alert value $\mathbb{E}_s \Phi(\hat{K}^d(s, \cdot)) - \Phi(\hat{\pi})$ with $\Phi(b) = \tau \log \sum_a e^{(b, r_a)/\tau}$, $\tau = 0.5$ (slope fitted on $d \in [40, 100]$: -0.1263 ; $-2\mu = -0.1263$), are computed exactly from $\hat{K}^d$; the grey points are the empirical mutual information of (S_t, S_{t+d}) over all valid pairs in the raw series. Alert decision: $r_{\text{alert}} = +1$ on quintiles 4, 5 and -1 on 1–3, $r_{\text{no}} = 0$; blind margin $\delta = 0.2$, $\rho = 2$, so (10) gives $\bar{D} = \lceil \log(5\sqrt{0.8/0.2})/0.0631 \rceil = 37$ (the value 36.6 before rounding); the exact deadline is $D_0 = 30$. Panel (b): $v = \mathbf{1}_{\{4,5\}} - \hat{\pi}(\{4,5\})\mathbf{1}$, $g_d = (\hat{K}^d - \mathbf{1}\hat{\pi}^\top)v$, cosines in $\ell^2(\hat{\pi})$. Panel (c): 20 trajectories of length $n \in \{2 \times 10^3, 2 \times 10^4, 2 \times 10^5\}$ simulated from $\hat{K}$ (stationary start); the plug-in is the empirical mutual information of (S_t, S_{t+d}) ; the model-based estimate is $I(\hat{\pi}_n \otimes \hat{K}_n^d)$ with $\hat{K}_n$ the one-step kernel estimated from the same trajectory; “detach” is the smallest d at which the median relative error exceeds 50%: 27, 45, 56 hours for the plug-in. The plug-in floor (median over $d \in [111, 130]$) is 2.3×10^{-2} , 2.1×10^{-3} , 1.8×10^{-4} nats for the three n , i.e. $\gamma_d/2n$ with $\gamma_d = 92, 83, 71$; the i.i.d.-pair value would be $(N-1)^2 = 16$, and the excess is the integrated autocorrelation of the pair counts (compare $(1+\lambda_2)/(1-\lambda_2) = 32$). Bias corrections: Miller–Madow,

$\widehat{\mathcal{T}}_{\text{MM}} = \widehat{\mathcal{T}} - (m_{xy} - m_x - m_y + 1)/2n$ with m the numbers of nonzero cells, detaches at 29, 40, 57 h; the delete-one-block jackknife with $B = 10$ contiguous blocks of pairs, $\widehat{\mathcal{T}}_{\text{JK}} = B\widehat{\mathcal{T}} - (B-1)\widehat{\mathcal{T}}^{(-i)}$, lowers the mean floor to 2.8×10^{-3} , 3.8×10^{-4} , 5.7×10^{-6} nats but detaches at 22, 40, 55 h. The model-based estimate detaches at 60 h for $n = 2 \times 10^3$ and not within $d \leq 130$ for the larger n .

F Non-reversible chains

Throughout this section K is an irreducible, aperiodic stochastic matrix with stationary law π , not assumed reversible. Reversibility entered the paper only through Lemma 1 and parts (iii)–(iv) of Lemma 2: a π -orthonormal eigenbasis, $m_k = \pi\psi_k$, and the row bound $\|\varepsilon_s^{(d)}\|_1 \leq |\lambda_2|^d \sqrt{1/\pi(s) - 1}$. We replace these, list what survives unchanged, and then record what is genuinely new: a cost of non-normality (Appendix F.3), a sharpened trichotomy in which complex eigenvalues cannot cancel the leading term (Appendix F.4), a rotating optimal summary (Appendix F.5), a forward/backward asymmetry (Appendix F.6), and consequences for learning (Appendix F.7). Appendix F.8 works out a drifting walk on a cycle.

F.1 Spectral calculus without self-adjointness

Adjoint, normality. The time reversal $K^*(s, s') = \pi(s')K(s', s)/\pi(s)$ is the adjoint of K in $\ell^2(\pi)$ and the kernel of the reversed stationary chain. K is reversible iff $K = K^*$ and (π) -normal iff $KK^* = K^*K$. Write $\Pi = \mathbf{1}\pi^\top$; then $K\Pi = \Pi K = \Pi^2 = \Pi$, so

$$K^d - \Pi = (K - \Pi)^d, \quad d \geq 1, \quad (24)$$

and $K - \Pi$ has spectral radius $|\lambda_2| < 1$, with eigenvalues indexed by modulus, $1 = \lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \dots$, now complex in general (non-real ones in conjugate pairs).

Jordan form. Let $K = \Pi + \sum_j (\lambda_j P_j + N_j)$ over the eigenvalues $\lambda_j \neq 1$, with spectral projectors P_j ($P_j P_l = \delta_{jl} P_j$, $P_j \Pi = \Pi P_j = 0$) and nilpotents $N_j = P_j N_j = N_j P_j$ of index m_j (the largest Jordan block of λ_j). By (24),

$$\varepsilon_s^{(d)} = e_s^\top (K^d - \Pi) = \sum_j \sum_{i=0}^{m_j-1} \binom{d}{i} \lambda_j^{d-i} e_s^\top N_j^i P_j. \quad (25)$$

Diagonalizable case. Take right eigenvectors $K\psi_k = \lambda_k \psi_k$ and left eigenvectors $m_k K = \lambda_k m_k$, biorthonormal: $\sum_{s'} m_j(s') \psi_k(s') = \delta_{jk}$, $\psi_1 = \mathbf{1}$, $m_1 = \pi$. Then (1) and (16) hold verbatim in the form $K^d(s, s') = \sum_k \lambda_k^d \psi_k(s) m_k(s')$ and $\varepsilon_s^{(d)} = \sum_{k \geq 2} \lambda_k^d \psi_k(s) m_k$, the non-real terms coming in conjugate pairs. The normalization fixes each product $\psi_k m_k$ but not ψ_k alone; every formula below is invariant under $\psi_k \mapsto c\psi_k$, $m_k \mapsto m_k/c$.

What survives, what fails. For $k \geq 2$, $\mathbb{E}_\pi \psi_k = \pi \psi_k = 0$ (as $\pi \psi_k = \pi K \psi_k = \lambda_k \pi \psi_k$) and $\sum_{s'} m_k(s') = m_k \mathbf{1} = 0$. Hence Lemma 2(i)–(ii) survive: $\sum_{s'} \varepsilon_s^{(d)}(s') = 0$, $\mathbb{E}_s \varepsilon_s^{(d)} = 0$, $\langle \varepsilon_s^{(d)}, v \rangle = ((K^d - \Pi)v)(s)$. What fails is $\mathbb{E}_\pi[\psi_k \psi_l] = \delta_{kl}$ and $m_k = \pi \psi_k$. Their failure is measured by the Gram matrices and the eigenvalue condition numbers

$$R_{kl} = \langle \psi_k, \psi_l \rangle_\pi, \quad L_{kl} = \left\langle \frac{m_k}{\pi}, \frac{m_l}{\pi} \right\rangle_\pi, \quad \kappa_k = \|\psi_k\|_{2,\pi} \|m_k/\pi\|_{2,\pi} \geq |m_k \psi_k| = 1, \quad (26)$$

with $\langle f, g \rangle_\pi = \sum_s \pi(s) f(s) \overline{g(s)}$. κ_k is Wilkinson's condition number of λ_k in $\ell^2(\pi)$; for diagonalizable K , $\kappa_k \equiv 1$ iff the eigenbasis is π -orthogonal iff K is π -normal. Reversible chains are normal; so are, e.g., all random walks on groups with uniform π (Appendix F.8).

Lemma 11 (Row bounds without reversibility). *Let $s_k = \|K^k - \Pi\|_\pi$, the operator norm on $\ell^2(\pi)$ (the largest non-trivial singular value of K^k). For all $d \geq 0$, $s \in \mathcal{S}$, $k \geq 1$,*

$$\|\varepsilon_s^{(d)}\|_1 \leq \|K^d - \Pi\|_\pi \sqrt{1/\pi(s) - 1} \leq s_k^{\lfloor d/k \rfloor} \sqrt{1/\pi(s) - 1}. \quad (27)$$

Moreover $s_k^{1/k} \rightarrow |\lambda_2|$, and with m the largest Jordan block among eigenvalues of modulus $|\lambda_2|$, there are $0 < c_K \leq C_K$ with

$$c_K d^{m-1} |\lambda_2|^d \leq \|K^d - \Pi\|_\pi \leq C_K d^{m-1} |\lambda_2|^d, \quad d \geq 1. \quad (28)$$

If K is normal, $s_1 = |\lambda_2|$ and (27) is Lemma 2(iv) verbatim.

Proof. Let $f = \varepsilon_s^{(d)}/\pi$. For $g \in \ell^2(\pi)$, $\langle f, g \rangle_\pi = ((K^d - \Pi)g)(s) = \langle \delta_s/\pi, (K^d - \Pi)g \rangle_\pi = \langle (K^{*d} - \Pi)(\delta_s/\pi - \mathbf{1}), g \rangle_\pi$, using $(K^{*d} - \Pi)\mathbf{1} = 0$. So $\|\varepsilon_s^{(d)}\|_1 = \|f\|_{1,\pi} \leq \|f\|_{2,\pi} \leq \|K^{*d} - \Pi\|_\pi \|\delta_s/\pi - \mathbf{1}\|_{2,\pi}$, and $\|A^*\| = \|A\|$, $\|\delta_s/\pi - \mathbf{1}\|_{2,\pi}^2 = 1/\pi(s) - 1$. Submultiplicativity in (24) gives the second inequality; Gelfand's formula gives $s_k^{1/k} \rightarrow \rho(K - \Pi) = |\lambda_2|$. The upper bound in (28) is (25). For the lower bound, if λ with $|\lambda| = |\lambda_2|$ has a block of size m and $K - \Pi = VJV^{-1}$, then $\|(K - \Pi)^d\| \geq \|J^d\|/\text{cond}(V) \geq \binom{d}{m-1} |\lambda_2|^{d-m+1}/\text{cond}(V)$, reading off the corner entry of the Jordan block. If K is normal, $\|K - \Pi\|_\pi = \rho(K - \Pi)$. $\square$

s_1^2 is the second eigenvalue of Fill's multiplicative reversibilization KK^* (Fill, 1991). It can equal 1 for an ergodic non-reversible chain (e.g. one with a deterministic step), which is why (27) is stated for k -step blocks; $\max_k (1 - s_k^2)/k$ is Paulin's pseudo-spectral gap (Paulin, 2015).

Lemma 12 (Non-real eigenvectors are not collinear). *If $K\psi = \lambda\psi$ with $\lambda \notin \mathbb{R}$, then for every $\alpha \in \mathbb{R}$ the real vector $\text{Re}(e^{i\alpha}\psi)$ is non-zero; equivalently, the points $\psi(s) \in \mathbb{C}$ do not lie on one line through 0. Consequently $\mathbb{E}_\pi |\psi|^2 > |\mathbb{E}_\pi \psi^2|$, and for any $m \neq 0$ the real matrix $\text{Re}(e^{i\alpha}\psi m^\top)$ is non-zero for every α .*

Proof. If $\text{Re}(e^{i\alpha}\psi) = 0$ then $e^{i\alpha}\psi = iu$ with u real and non-zero, and $Ku = \lambda u$ forces $\lambda \in \mathbb{R}$. Equality in $|\mathbb{E}\psi|^2 \leq \mathbb{E}|\psi|^2$ (with $\pi > 0$) means $\psi(s)^2$ has constant phase, i.e. $\psi = e^{i\gamma}u$ with u real: excluded. For the matrix, pick s' with $m(s') \neq 0$; vanishing would put all $\psi(s)$ on one line through 0. $\square$

F.2 What carries over verbatim

The following use only stationarity, Jensen, Chapman–Kolmogorov, or $\mathbb{E}_s \varepsilon_s^{(d)} = 0$, and hold for every irreducible aperiodic K without change: Theorem 1 and Remark 1 (Blackwell monotonicity); the inequality (5) and its equality case in Theorem 2; the local reduction $\mathcal{V}_\Phi(d) = \mathbb{E}_s g(\varepsilon_s^{(d)})$ for $d \geq d_1$ in Theorem 3(B), with d_1 now taken from Lemma 11; existence of the endogenous deadline in Theorem 3(H) (its bound changes, Theorem 8); the exact part of Theorem 4 (first equality of (13), optimality of $f_d^* = \mathbf{1}[g_d > 0]$, value $\frac{1}{2}\|g_d\|_{1,\pi}$, with $g_d = (K^d - \Pi)v$); Proposition 3, formula (22) and parts (i)–(ii); the pulled-back reward identity (19); Lemma 4 (Paulin's mixing-time McDiarmid bound does not require reversibility) and Lemma 5; and Theorem 7, whose constants depend on K only through N and τ_{mix} , now bounded via (27) (Appendix F.7).

F.3 The χ^2 value and the cost of non-normality

Proposition 5 (Exact χ^2 value). *For $\Phi_\chi(b) = \frac{1}{2} \sum_s b(s)^2 / \pi(s)$ and every $d \geq 0$,*

$$\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \|K^d - \Pi\|_{HS(\pi)}^2 = \frac{1}{2} \sum_{k \geq 2} \sigma_k(K^d)^2 \geq \frac{1}{2} \sum_{k \geq 2} |\lambda_k|^{2d}, \quad (29)$$

where $\sigma_1(K^d) = 1 \geq \sigma_2(K^d) \geq \dots$ are the singular values of K^d in $\ell^2(\pi)$ (of $D^{1/2}K^dD^{-1/2}$, $D = \text{diag } \pi$). Equality holds at $d = 1$ iff it holds for all d iff K is π -normal. If K is diagonalizable,

$$\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \sum_{k, l \geq 2} \overline{\lambda_k^d} \lambda_l^d R_{kl} L_{kl}, \quad (30)$$

a quadratic form in $(\lambda_k^d)_{k \geq 2}$ with the positive semidefinite matrix $R \circ L$ (Hadamard product), whose diagonal is $(\kappa_k^2)_{k \geq 2}$.

Proof. The computation of Lemma 3 gives $\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \sum_s \pi(s) \sum_{s'} \varepsilon_s^{(d)}(s')^2 / \pi(s') = \frac{1}{2} \|A\|_F^2$ with $A = D^{1/2}(K^d - \Pi)D^{-1/2}$. With $u = \sqrt{\pi}$, $B = D^{1/2}K^dD^{-1/2}$ satisfies $Bu = B^\top u = u$, so $B = uu^\top + A$ with $Au = A^\top u = 0$, and the singular values of A are those of B with one 1 removed. A has eigenvalues λ_k^d ($k \geq 2$) and 0, and Schur's inequality $\sum |\lambda_i(A)|^2 \leq \|A\|_F^2$ holds with equality iff A is normal (Horn and Johnson, 2013, Thm. 2.5.4). Since $uu^\top$ commutes with B and $B^\top$, A is normal iff B is, iff K^d is π -normal; K normal implies all powers normal. For (30), ε is real, so $\varepsilon^2 = |\varepsilon|^2$; expand $|\sum_k \lambda_k^d \psi_k(s) m_k(s')|^2$ and sum. R and L are Gram matrices, so $R \circ L \succeq 0$ by the Schur product theorem. $\square$

In a reversible chain $R = L = I$ and (30) is Lemma 3. In general the spectrum alone underestimates delayed information: the excess is the transient growth of non-normal powers (Trefethen and Embree, 2005), invisible to the eigenvalues. For a simple real λ_2 it is a constant factor, $\mathcal{V}_{\Phi_\chi}(d) \sim \frac{1}{2} \kappa_2^2 \lambda_2^{2d}$; with a Jordan block it is polynomial, $\asymp d^{2(m-1)} |\lambda_2|^{2d}$.

F.4 The aging trichotomy

Let $E = \{k \geq 2 : |\lambda_k| = |\lambda_2|\}$ (the peripheral eigenvalues), m the largest Jordan block in E , and $|\lambda'| < |\lambda_2|$ the next modulus, m' its largest block. We say λ_2 is a *simple real eigenvalue* if $E = \{2\}$, $\lambda_2 \in \mathbb{R}$, $m = 1$; and a *simple complex pair* if $E = \{2, \bar{2}\}$, $\lambda_2 = |\lambda_2|e^{i\theta}$ with $\theta \in (0, \pi)$, $m = 1$. In the latter case $\varepsilon_s^{(d)} = 2|\lambda_2|^d \text{Re}(e^{id\theta} \psi_2(s) m_2) + \eta_s^{(d)}$, $\|\eta_s^{(d)}\|_1 = O(d^{m'-1} |\lambda'|^d)$. For rewards put $c_a = m_2 r_a = \sum_{s'} m_2(s') r_a(s') \in \mathbb{C}$ (in a reversible chain $c_a = \langle \psi_2, r_a \rangle_\pi$).

Theorem 8 (Aging, non-reversible). *Let $\lambda_2 \neq 0$, $\mu = \log(1/|\lambda_2|)$. Constants below depend on K .*

(C) **Ceiling.** $c d^{m-1} |\lambda_2|^d \leq \bar{\mathcal{V}}(d) \leq C d^{m-1} |\lambda_2|^d$. For a simple real λ_2 , $\bar{\mathcal{V}}(d) = \frac{1}{2} \|\psi_2\|_{1,\pi} \|m_2\|_1 |\lambda_2|^d + O(d^{m'-1} |\lambda'|^d)$. For a simple complex pair, $\bar{\mathcal{V}}(d) = |\lambda_2|^d \kappa_C(d\theta) + O(d^{m'-1} |\lambda'|^d)$ with

$$\kappa_C(\alpha) = \mathbb{E}_s \sum_{s'} |\text{Re}(e^{i\alpha} \psi_2(s) m_2(s'))|$$

continuous, 2π -periodic and $\min_\alpha \kappa_C > 0$.

(S) **Smooth.** Let $\Phi \in C^2$ near π , $H = \nabla^2 \Phi(\pi)$. Then

$$\mathcal{V}_\Phi(d) = \frac{1}{2} \mathbb{E}_s [\varepsilon_s^{(d)\top} H \varepsilon_s^{(d)}] + o(d^{2(m-1)} |\lambda_2|^{2d}) \leq C d^{2(m-1)} |\lambda_2|^{2d},$$

if K is diagonalizable the leading term is $\frac{1}{2} \sum_{k,l \geq 2} (\lambda_k \lambda_l)^d \mathbb{E}_\pi[\psi_k \psi_l] m_k^\top H m_l$. For a simple complex pair,

$$\mathcal{V}_\Phi(d) = |\lambda_2|^{2d} (A_\Phi + \operatorname{Re}(e^{2id\theta} B_\Phi)) + o(|\lambda_2|^{2d}), \quad (31)$$

$A_\Phi = \mathbb{E}_\pi |\psi_2|^2 m_2^* H m_2$, $B_\Phi = \mathbb{E}_\pi [\psi_2^2] m_2^\top H m_2$, and $A_\Phi - |B_\Phi| \geq (\mathbb{E}_\pi |\psi_2|^2 - |\mathbb{E}_\pi \psi_2^2|) m_2^* H m_2$, which is > 0 iff $m_2^* H m_2 > 0$; then $\beta_\Phi = 2\mu$. Temporal information satisfies, with no condition on the spectrum,

$$c d^{2(m-1)} |\lambda_2|^{2d} \leq \mathcal{T}(d) \leq C d^{2(m-1)} |\lambda_2|^{2d}, \quad \mathcal{T}(d) = \mathcal{V}_{\Phi_\chi}(d) (1 + O(d^{m-1} |\lambda_2|^d)), \quad (32)$$

so $\beta_\mathcal{T} = 2\mu$; for a simple real λ_2 , $\mathcal{T}(d) \sim \frac{1}{2} \kappa_2^2 \lambda_2^{2d}$; for a simple complex pair (31) holds with $A = \kappa_2^2$ and $|B| < \kappa_2^2$.

(B) **Boundary.** For $d \geq d_1$, $\mathcal{V}_\Phi(d) = \mathbb{E}_s g(\varepsilon_s^{(d)})$ and $\mathcal{V}_\Phi(d) \leq C d^{m-1} |\lambda_2|^d$. For a simple real λ_2 , (9) holds verbatim with $\sigma_2 = \max_{\mathcal{A}^*} c_a - \min_{\mathcal{A}^*} c_a$. For a simple complex pair,

$$\mathcal{V}_\Phi(d) = |\lambda_2|^d \kappa_B(d\theta) + O(d^{m-1} |\lambda'|^d), \quad \kappa_B(\alpha) = 2 \mathbb{E}_s \max_{a \in \mathcal{A}^*} \operatorname{Re}(e^{i\alpha} \psi_2(s) c_a), \quad (33)$$

and either $\kappa_B \equiv 0$ (iff c_a is constant on $\mathcal{A}^*$) or $\min_\alpha \kappa_B > 0$, in which case $\beta_\Phi = \mu$ (the limit exists). For two optimal actions with contrast v , $\kappa_B(\alpha) = \mathbb{E}_s |\operatorname{Re}(e^{i\alpha} c_v \psi_2(s))|$, $c_v = m_2 v$.

(H) **Generic hard.** $\mathcal{V}_\Phi(d) = 0$ for $d \geq D_0$, where, for every k with $s_k < 1$,

$$D_0 \leq k \left\lceil \frac{\log \left(\frac{\rho}{2\delta} \sqrt{(1 - \pi_{\min})/\pi_{\min}} \right)}{\log(1/s_k)} \right\rceil_+. \quad (34)$$

For normal K ($s_1 = |\lambda_2|$) this is (10); as $k \rightarrow \infty$ the rate $\log(1/s_k)/k \rightarrow \mu$ recovers its exponent.

Proof. (C) Upper bound: (27)–(28). Lower bound: for $\|f\|_\infty \leq 1$, $\|\varepsilon_s^{(d)}\|_{\text{TV}} \geq \frac{1}{2} |((K^d - \Pi)f)(s)|$, so $\bar{\mathcal{V}}(d) \geq \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \|(K^d - \Pi)f\|_{1,\pi} \geq c_{N,\pi_{\min}} \|K^d - \Pi\|_\pi$ by equivalence of norms in dimension N ; apply (28). The asymptotics follow as in Appendix C from the leading term of (25), using $\|a+b\| - \|a\| \leq \|b\|$; for the pair, $\|2\operatorname{Re}(e^{i\alpha} \psi_2(s) m_2)\|_{\text{TV}} = \sum_{s'} |\operatorname{Re}(e^{i\alpha} \psi_2(s) m_2(s'))|$. κ_C is continuous on the circle and vanishes nowhere by Lemma 12.

(S) The Taylor argument of Appendix C uses only $\mathbb{E}_s \varepsilon_s^{(d)} = 0$ and $\|\varepsilon_s^{(d)}\|_1 \leq C d^{m-1} |\lambda_2|^d$ (Lemma 11); without orthonormality the quadratic term keeps its cross terms, giving the bilinear form stated. For the pair, write $\varepsilon = z + \bar{z} + \eta$, $z = \lambda_2^d \psi_2(s) m_2$; then $\frac{1}{2} \mathbb{E}[(z + \bar{z})^\top H (z + \bar{z})] = \operatorname{Re} \mathbb{E}[z^\top H z] + \mathbb{E}[z^\top H \eta]$, which is (31). Since $H \succeq 0$ on the tangent space (which contains m_2 , as $m_2 \mathbf{1} = 0$), $|m_2^\top H m_2| = |\langle \bar{m}_2, m_2 \rangle_H| \leq m_2^* H m_2$, hence $|B_\Phi| \leq |\mathbb{E} \psi_2^2| m_2^* H m_2$, and Lemma 12 gives the strict gap. For $\mathcal{T}$, $H = \operatorname{diag}(1/\pi)$ is positive definite on the tangent space and the cubic remainder is $O(d^{3(m-1)} |\lambda_2|^{3d})$, so $\mathcal{T}(d) = \frac{1}{2} \|K^d - \Pi\|_{HS(\pi)}^2 (1 + O(d^{m-1} |\lambda_2|^d))$; the two-sided bound follows from (28) and $\|\cdot\|_\pi \leq \|\cdot\|_{HS(\pi)} \leq \sqrt{N} \|\cdot\|_\pi$. The simple-eigenvalue constants are the diagonal of (30).

(B) The reduction and the Lipschitz bound on g are as in Appendix C. For a simple real λ_2 , $\langle \lambda_2^d \psi_2(s) m_2, r_a \rangle = \lambda_2^d \psi_2(s) c_a$, and $\mathbb{E}_\pi \psi_2^\pm = \frac{1}{2} \|\psi_2\|_{1,\pi}$ since $\mathbb{E}_\pi \psi_2 = 0$; the rest is verbatim. For the pair, $\langle 2\operatorname{Re}(e^{i\alpha} \psi_2(s) m_2), r_a \rangle = 2\operatorname{Re}(e^{i\alpha} \psi_2(s) c_a)$, giving (33). Fix α and put $X_a(s) = \operatorname{Re}(e^{i\alpha} \psi_2(s) c_a)$, so $\mathbb{E}_\pi X_a = 0$ and $\kappa_B(\alpha) = 2 \mathbb{E} \max_a X_a \geq 2 \max_a \mathbb{E} X_a = 0$. Equality forces some a_0 with $\max_a X_a = X_{a_0}$ π -a.s.; then $X_{a_0} - X_b \geq 0$ has mean zero, so $X_{a_0} \equiv X_b$, i.e. $\operatorname{Re}(e^{i\alpha} (c_{a_0} - c_b) \psi_2(s)) = 0$ for all s , and Lemma 12 gives $c_b = c_{a_0}$. So if c_a is not constant on $\mathcal{A}^*$, $\kappa_B > 0$ everywhere, and by continuity on the circle $\min \kappa_B > 0$. The two-action form is $\max(x, y) = \frac{x+y}{2} + \frac{|x-y|}{2}$.

(H) As in Appendix C, all rows lie in the cell of a^* once $\|\varepsilon_s^{(d)}\|_1 \leq 2\delta/\rho$ for all s ; by (27) this holds when $s_k^{\lfloor d/k \rfloor} \sqrt{(1 - \pi_{\min})/\pi_{\min}} \leq 2\delta/\rho$. $\square$

Corollary 2 (Decoupling survives). *For every irreducible aperiodic K with $\lambda_2 \neq 0$, $\bar{\mathcal{V}}(d) \asymp \sqrt{\mathcal{T}(d)}$, with no simplicity assumption (both are $\asymp d^{m-1}|\lambda_2|^d$). For a boundary decision and a simple real eigenvalue or simple complex pair with $\kappa_B \neq 0$, $\mathcal{V}_\Phi(d) \asymp \sqrt{\mathcal{T}(d)}$. The conversion exponents $\kappa \in \{\frac{1}{2}, 1, \infty\}$ of Corollary 1 are unchanged.*

Complex eigenvalues sharpen the trichotomy. In a reversible chain the leading coefficient can vanish on a parity class when $\lambda_3 = -\lambda_2$, so that $\underline{\beta}_\Phi \neq \bar{\beta}_\Phi$. A simple complex pair cannot do this: by Lemma 12 the rotating mode $\text{Re}(e^{id\theta}\psi_2)$ never vanishes, so its coefficient is bounded away from zero and the limit β_Φ exists in (C), (S) (when $m_2^*Hm_2 > 0$) and (B) (when c_a is non-constant on $\mathcal{A}^*$). A gap $\underline{\beta}_\Phi < \bar{\beta}_\Phi$ requires two or more peripheral eigenvalues with different arguments (two pairs, or a pair together with $\pm|\lambda_2|$), where the coefficient is an almost periodic function of d that may approach 0. Even then $\mathcal{T}$ and $\bar{\mathcal{V}}$ have $\beta = 2\mu$ and μ by (32) and (C).

F.5 Which one-bit summary to read: the rotating phase-bit

The exact part of Theorem 4 holds (Appendix F.2). The bound (14) becomes, by Cauchy–Schwarz in $\mathcal{V}_f(d) = |\langle \mathbf{1}_{C_0} - \pi(C_0)\mathbf{1}, g_d \rangle_\pi|$,

$$\mathcal{V}_f(d) \leq \sqrt{\pi(C_0)\pi(C_1)} \|K^d - \Pi\|_\pi \|v\|_{2,\pi} \leq C d^{m-1} |\lambda_2|^d \|v\|_{2,\pi}. \quad (35)$$

For a simple real λ_2 the rotation statement of Theorem 4 holds with ψ_2 the right eigenvector and the condition $\langle v, \psi_2 \rangle_\pi \neq 0$ replaced by $c_v = m_2 v \neq 0$; a summary with $\sum_{s \in C_0} \pi(s) \psi_2(s) = 0$ still ages at $O(d^{m'-1}|\lambda'|^d)$. For a complex pair the picture changes.

Proposition 6 (Phase-bit). *Let $\lambda_2 = |\lambda_2|e^{i\theta}$ be a simple complex pair, v a two-action boundary contrast with $c_v = m_2 v = |c_v|e^{i\phi} \neq 0$, and $\alpha_d = d\theta + \phi$. Then $g_d = 2|\lambda_2|^d |c_v| \text{Re}(e^{i\alpha_d}\psi_2) + O(d^{m'-1}|\lambda'|^d)$, and the optimal summary $f_d^* = \mathbf{1}[g_d > 0]$ agrees with the phase-bit $\mathbf{1}[\text{Re}(e^{i\alpha_d}\psi_2) > 0]$ on every s with $|\text{Re}(e^{i\alpha_d}\psi_2(s))| > Cd^{m'-1}(|\lambda'|/|\lambda_2|)^d$. Its value is $\mathcal{V}_{f_d^*}(d) = |\lambda_2|^d |c_v| \mathbb{E}_\pi |\text{Re}(e^{i\alpha_d}\psi_2)| + O(d^{m'-1}|\lambda'|^d)$, bounded below by a positive multiple of $|\lambda_2|^d$. The phase-bits $\{\mathbf{1}[\text{Re}(e^{i\alpha}\psi_2) > 0] : \alpha \in [0, 2\pi)\}$ form at most $2N$ labelled partitions (N up to complement), changing only when α crosses $\frac{\pi}{2} - \arg \psi_2(s) \pmod{\pi}$. Hence f_d^* does not converge: if $\theta/2\pi = p/q$ it is eventually q -periodic; if $\theta/\pi \notin \mathbb{Q}$, each phase-bit is optimal for an asymptotic fraction of delays equal to the length of its α -arc over 2π (Weyl equidistribution).*

Proof. $g_d = (K^d - \Pi)v = \sum_{k \geq 2} \lambda_k^d (m_k v) \psi_k$; keep the pair and bound the rest by (25). The sign of $g_d(s)$ is that of the leading term whenever the latter exceeds the remainder. The value is $\frac{1}{2}\|g_d\|_{1,\pi}$; the lower bound is Lemma 12 and continuity in α . State s changes side exactly when $\text{Re}(e^{i\alpha}\psi_2(s))$ changes sign, at two antipodal angles; the $2N$ angles cut the circle into at most $2N$ arcs. $\square$

Geometrically: plot the N points $\psi_2(s) \in \mathbb{C}$. The best stale bit asks on which side of a line through the origin $\psi_2(S_{t-d})$ lies, and the line turns by θ per step of delay. Compare Figure 2: in a reversible chain the score g_d converges to $\pm\psi_2$ and the chosen bit freezes after finitely many steps; here the score keeps rotating in the real plane $\text{span}\{\text{Re}\psi_2, \text{Im}\psi_2\}$ and so does the bit. What the slow mode retains is a phase, and a stale report is useful only if the question tracks that phase.

Several actions: one bit may not suffice. Formula (22) and Proposition 3(i)–(ii) hold. For $d \geq d_1$ and a simple complex pair, the action recommended at s is, up to the same switching neighborhoods, $\arg \max_{a \in \mathcal{A}^*} \text{Re}(e^{i\alpha_d}\psi_2(s)c_a)$: a vertex of the polygon $P = \text{conv}\{c_a : a \in \mathcal{A}^*\} \subset \mathbb{C}$ exposed by the direction $e^{i\alpha_d}\psi_2(s)$. If the c_a are collinear, $\max_a \text{Re}(zc_a)$ is attained at the two

ends of the segment for every z , and the proof of Proposition 3(iii) goes through: $\mathcal{V}_1^*(d) = \mathcal{V}_\Phi(d) + O(d^{m'-1}|\lambda'|^d)$ (in a reversible chain the c_a are real, hence always collinear). If P has three or more vertices and the directions $\{e^{i\alpha}\psi_2(s)\}_s$ expose three of them, one bit recovers only a fraction of $\mathcal{V}_\Phi(d)$ at those phases, and the recommended-action report needs up to $\lceil \log_2 |\text{ext } P| \rceil$ bits. So the slogan “a late report is worth one well-chosen bit” is a property of real slow modes.

F.6 Forward and backward delayed value

Running the chain backwards replaces K by K^* . Under K^* the pair law of (S_{t-d}, S_t) is the transpose of that under K , so $\mathcal{V}_\Phi^{K^*}(d)$ is the value, for an action on S_t , of the *future* state S_{t+d} (hindsight, or smoothing). K^* has the same eigenvalues as K , right eigenvectors m_k/π and left eigenvectors $\pi\psi_k$.

Proposition 7 (Time symmetry). *$\mathcal{T}$, $\bar{\mathcal{V}}$ and $\mathcal{V}_{\Phi_\chi}$ are invariant under $K \mapsto K^*$. Howard values are not in general: for a simple real λ_2 , the forward boundary constant in (9) is $\sigma_2 \|\psi_2\|_{1,\pi}$ with σ_2 built from $c_a = m_2 r_a$, the backward one is $\sigma_2^* \|m_2\|_1$ with $c_a^* = \langle \psi_2, r_a \rangle_\pi$; the forward eigen-bit thresholds the right eigenvector ψ_2 , the backward one the left density m_2/π . The deadlines D_0 differ as well.*

Proof. $I(S_{t-d}; S_t)$ is symmetric in its arguments; $\bar{\mathcal{V}}(d) = \frac{1}{2} \sum_{s,s'} |P_d(s, s') - \pi(s)\pi(s')|$ is transpose-invariant; $\|A^*\|_{HS} = \|A\|_{HS}$. For the eigenvectors: $(K^*(m_k/\pi))(s) = (m_k K)(s)/\pi(s) = \lambda_k m_k(s)/\pi(s)$, and similarly on the left. Apply Theorem 8(B) to K^* . $\square$

In a reversible chain $m_2/\pi = \psi_2$ and the two readings coincide. Information about the world is time-symmetric; the worth of a report depends on whether it comes from the past or the future.

F.7 Learning delayed value

Model-free. Theorem 5 holds with the same $\mu = \log(1/|\lambda_2|)$: the proofs use only Lemma 4, Lemma 5 (with $C_2|\lambda_2|^d$ replaced by $C_2 d^{m-1}|\lambda_2|^d$ via (27)) and two-sided signal bounds, now (C) and (32). A Jordan block shifts the horizon by $d^*(n) = \frac{\log n}{2\mu} + O(\frac{(m-1)\log \log n}{\mu})$. The floor constant is unchanged: $\gamma_\infty = \sum_{j \in \mathbb{Z}} (\text{tr } K^{|j|} - 1)^2 = \sum_{k,l \geq 2} \frac{1+\lambda_k \lambda_l}{1-\lambda_k \lambda_l} \geq (N-1)^2$, since its derivation in Appendix D uses only $\sum_s \mathbb{E}[A_s(S_0)A_s(S_j)]/\pi(s) = \text{tr } K^{|j|} - 1 = \sum_{k \geq 2} \lambda_k^{|j|}$ (trace, with algebraic multiplicities), which is real; the covariance bounds use K^* in place of reversibility. Proposition 4 holds: for a simple real λ_2 use the perturbation $K_\alpha = K + (\lambda'_2 - \lambda_2)\psi_2 m_2$, which preserves row sums ($m_2 \mathbf{1} = 0$) and π ($\pi\psi_2 = 0$); for a pair perturb $2\text{Re}((\lambda'_2 - \lambda_2)\psi_2 m_2)$ with $\lambda'_2 = \alpha^{1/(2d)}\lambda_2$. The χ^2 bound picks up a factor κ_2^2 .

Model-based: the symmetrized estimator is inconsistent. The estimator of Proposition 2 symmetrizes the counts, so $\hat{Q} \rightarrow \frac{1}{2}(Q + Q^\top)$ and $\hat{K} \rightarrow K_+ = \frac{1}{2}(K + K^*)$, the additive reversibilization. Hence $\tilde{\mathcal{T}}(d) \rightarrow \frac{1}{2} \sum_{k \geq 2} \lambda_k (K_+)^{2d}$, which has the wrong exponent in general, in either direction: for normal K the eigenvalues of K_+ are $\text{Re } \lambda_k$, so a non-real λ_2 makes $\tilde{\mathcal{T}}$ decay too fast (Appendix F.8); for strongly non-normal K the numerical range, hence $\lambda_2(K_+)$, can exceed $|\lambda_2|$. The fix is to keep the raw kernel $\hat{K}(s, s') = n_{ss'}/n_s$ and estimate by matrix powers, avoiding any eigenvalue matching: $\hat{\mathcal{T}}_{\text{MB}}(d) = I(\hat{\pi} \otimes \hat{K}^d)$ or $\tilde{\mathcal{T}}(d) = \frac{1}{2} \|\hat{K}^d - \hat{\Pi}\|_{HS(\hat{\pi})}^2$. (The first is already used in Figure 4(c).)

Proposition 8 (Value of a model, non-reversible). *Let $K > 0$ be diagonalizable with simple peripheral spectrum (a simple real λ_2 or a simple complex pair). Then $\tilde{\mathcal{T}}(d) = \mathcal{T}(d)(1 + O_P(dn^{-1/2}) + O(|\lambda_2|^d))$, so $\tilde{\mathcal{T}}$ is consistent for every $d = d_n \rightarrow \infty$ with $d_n = o(\sqrt{n})$. The constants depend on N , $\pi_{\min}$, $|\lambda_2|$, τ_{mix} , the condition number of the eigenbasis and the minimal eigenvalue separation.*

Proof sketch. $\|\hat{K} - K\| = O_P(n^{-1/2})$ (Wolfer and Kontorovich, 2019) and $\|\hat{\pi} - \pi\|_1 = O_P(n^{-1/2})$ (Lemma 4, $d = 1$). By Bauer–Fike (Bauer and Fike, 1960), each eigenvalue of $\hat{K}$ is within $\text{cond}(V)\varepsilon_n$ of one of K ; simple eigenvalues and their spectral projectors are analytic in the perturbation (Kato, 1995), so $\hat{\lambda}_k = \lambda_k + O_P(n^{-1/2})$ and $\hat{P}_k = P_k + O_P(n^{-1/2})$. Then $\hat{K}^d - \hat{\Pi} = \sum_{k \geq 2} \hat{\lambda}_k^d \hat{P}_k$ with $\hat{\lambda}_k^d = \lambda_k^d(1 + O_P(dn^{-1/2}))$ for $d = o(\sqrt{n})$. The peripheral part dominates with relative remainder $O((|\lambda'|/|\lambda_2|)^d)$; its HS norm is bounded below by a multiple of $|\lambda_2|^d$ uniformly in the phase (Lemma 12), so its relative error is $O_P(dn^{-1/2})$. Finally $\mathcal{T} = \mathcal{V}_{\Phi_\chi}(1 + O(|\lambda_2|^d))$ by (32). $\square$

The two-point construction of Theorem 6 perturbs λ_2 along $\psi_2 m_2$ (or along the pair) and goes through for this class, with κ also bounding the eigenbasis condition number; so $\Theta(\sqrt{n})$ remains sharp for diagonalizable chains.

Remark 8 (Defective chains). The constants in Proposition 8 blow up near a chain whose λ_2 has a Jordan block of size $m \geq 2$. There an $O(\varepsilon)$ perturbation moves the eigenvalue by $\Theta(\varepsilon^{1/m})$ (Lidskii–Vishik–Lyusternik; Moro et al., 1997), so $\hat{\lambda}_2^d$ is off by a factor $\exp(\Theta(dn^{-1/(2m)}))$. This suggests a model-based horizon of order $n^{1/(2m)}$ uniformly over a neighborhood of a defective chain. A matching lower bound would perturb K by $cn^{-1/2}$ in a direction that splits the block (keeping the KL divergence $O(1)$) while moving the top modulus by $\asymp n^{-1/(2m)}$; we expect this to give $\Theta(n^{1/(2m)})$ but do not prove it. The gap with the model-free $\log n$ stays exponential.

Online recovery. Theorem 7 is unchanged, with the reversible bound on τ_{mix} replaced by the consequence of (27), $\tau_{\text{mix}} \leq \min_{k:s_k < 1} k \lceil \log(2\sqrt{(1 - \pi_{\min})/\pi_{\min}}) / \log(1/s_k) \rceil$, or by Paulin’s bound in terms of the pseudo-spectral gap (Paulin, 2015).

F.8 Example: a drifting walk on a cycle

On $\mathcal{S} = \mathbb{Z}_N$, $N \geq 3$, let $K = \alpha I + \beta S + \gamma S^{-1}$, $K(s, s+1) = \beta$, $K(s, s-1) = \gamma$, with $\alpha \geq \frac{1}{2}$, $\beta > \gamma \geq 0$, $\alpha + \beta + \gamma = 1$. Then π is uniform and K is circulant, hence π -normal; it is reversible iff $\beta = \gamma$. Its eigenvectors $\varphi_j(s) = \omega^{js}$, $\omega = e^{2\pi i/N}$, are orthonormal in $\ell^2(\pi)$, with $\lambda_j = \alpha + \beta\omega^j + \gamma\omega^{-j}$ and $|\lambda_j|^2 = (\alpha + (\beta + \gamma)\cos x_j)^2 + (\beta - \gamma)^2 \sin^2 x_j$, $x_j = 2\pi j/N$. Since $\alpha \geq \frac{1}{2} \geq \beta + \gamma$ and $(\beta + \gamma)^2 \geq 4\beta\gamma$, $|\lambda_j|^2$ is non-decreasing in $\cos x_j$, so $\lambda_2 = \lambda_1 = |\lambda_2|e^{i\theta}$ and $\bar{\lambda}_2 = \lambda_{N-1}$ form a simple complex pair with $\theta \in (0, \frac{\pi}{2})$ (here the cycle index $j = 1$ plays the role of the mode 2).

Value laws. Normality gives $\mathcal{V}_{\Phi_\chi}(d) = \frac{1}{2} \sum_{j \neq 0} |\lambda_j|^{2d}$ exactly (Proposition 5), $\kappa \equiv 1$, and $B = \mathbb{E}_\pi[\varphi_1^2] \mathbb{E}_\pi[\varphi_1^2] = 0$ (as $\mathbb{E}_\pi \omega^{2s} = 0$): $\mathcal{T}(d) \sim |\lambda_2|^{2d}$ with no oscillation. The rotation lives entirely in the decision.

The phase-bit. For the contrast $v(s) = \cos(2\pi s/N)$ (“is the state in the right half of the cycle?”), exactly and for every d , $g_d(s) = |\lambda_2|^d \cos(\frac{2\pi s}{N} + d\theta)$, $f_d^* = \mathbf{1}[\cos(\frac{2\pi s}{N} + d\theta) > 0]$. The best stale bit asks whether the state was in the half-cycle that the drift will have carried into the decision-relevant half by time t ; the half advances by $d\theta N/2\pi \approx d(\beta - \gamma)$ sites, the mean displacement. It never freezes; its value is $\frac{1}{2}|\lambda_2|^d \frac{1}{N} \sum_s |\cos(2\pi s/N + d\theta)| \approx |\lambda_2|^d/\pi$ for large N . Under K^* (drift reversed) the bit turns the other way (Proposition 7).

Transport preserves information. The additive reversibilization $K_+ = \alpha I + \frac{\beta + \gamma}{2}(S + S^{-1})$ has the same π and holding probability but eigenvalues $\text{Re } \lambda_j$, so $|\lambda_2(K_+)| < |\lambda_2(K)|$. For large N , with $x = 2\pi/N$, $\mu(K) \approx \frac{1}{2}(\beta + \gamma - (\beta - \gamma)^2)x^2$ against $\mu(K_+) \approx \frac{1}{2}(\beta + \gamma)x^2$: deterministic drift moves information around the cycle without destroying it, so the drifting chain’s stale reports age

more slowly than its reversibilization's. The symmetrized estimator of Proposition 2 would report the faster rate $2\mu(K_+)$.

G Four small examples

Throughout, $\Phi(b) = \max_{a \in \mathcal{A}} \langle b, r_a \rangle$ (Howard), so by the pulled-back identity $\mathcal{V}_\Phi(d) = \mathbb{E}_{s \sim \pi} \max_a (K^d r_a)(s) - \max_a \langle \pi, r_a \rangle$, and the policy at delay d is the best action given the stale report $S_{t-d} = s$, i.e. $\arg \max_a (K^d r_a)(s)$. One step of the chain is the unit of delay. All numbers are exact.

Example A: umbrella and an old weather report ($N = 2$, case (H)). Weather is Sunny or Rainy, one step = one day: $K = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$, $\pi = (\frac{2}{3}, \frac{1}{3})$, $\lambda_2 = 0.7$. Umbrella costs 1 (nuisance) if sunny and 0 if rainy; None costs 0 if sunny and 3 if rainy: $r_U = (-1, 0)$, $r_N = (0, -3)$. Blind, $\langle \pi, r_U \rangle = -\frac{2}{3} > \langle \pi, r_N \rangle = -1$, so $\mathcal{A}^* = \{U\}$ with margin $\delta = \frac{1}{3}$: case (H) of Theorem 3. After a report “sunny d days ago”, $\mathbb{P}(\text{rain}) = \frac{1}{3}(1 - 0.7^d)$, and None is optimal iff $0.7^d > \frac{1}{4}$, i.e. $d \leq 3$.

d (days)	$\mathcal{V}_\Phi(d)$	report Sunny	report Rainy
0	0.667	None	Umbrella
1	0.400	None	Umbrella
2	0.213	None	Umbrella
3	0.083	None	Umbrella
≥ 4	0	Umbrella	Umbrella

The report changes the action only while it is at most 3 days old, and from day 4 on it is worth exactly zero: $D_0 = 4$. With $\rho = \text{span}(r_U - r_N) = 4$ and $\pi_{\min} = \frac{1}{3}$, the bound (10) gives $D_0 \leq \lceil \log(4\sqrt{2}/(2\delta)) / \log(1/0.7) \rceil = 6$.

Example B: same weather, indifferent decision ($N = 2$, case (B)). Lower the rain cost from 3 to 2, $r_N = (0, -2)$. Now $\langle \pi, r_U \rangle = \langle \pi, r_N \rangle = -\frac{2}{3}$, so $\mathcal{A}^* = \{U, N\}$: case (B). The policy is the same at every delay (None after Sunny, Umbrella after Rainy), and since $N = 2$ the remainder in (9) vanishes: $\mathcal{V}_\Phi(d) = \frac{1}{2}|\lambda_2|^d \sigma_2 \|\psi_2\|_{1,\pi} = \frac{2}{3}(0.7)^d$, with $\sigma_2 = |\langle v, \psi_2 \rangle_\pi| = \sqrt{2}$, $\|\psi_2\|_{1,\pi} = \frac{2\sqrt{2}}{3}$. Changing one cost from 3 to 2 turns a report that is worthless after 4 days into one that is valuable forever, decaying at exactly $|\lambda_2|^d$: the trichotomy (H) vs. (B) on two states.

Example C: a stale traffic snapshot ($N = 3$, three actions, rotation). A commuter sees the highway state d steps ago, one step = 10 minutes; states Free, Busy, Jam:

$$K = \begin{pmatrix} 0.85 & 0.15 & 0 \\ 0.20 & 0.50 & 0.30 \\ 0 & 0.25 & 0.75 \end{pmatrix}, \quad \pi = (0.377, 0.283, 0.340), \quad \lambda_2 = 0.815, \quad \lambda_3 = 0.285.$$

Rewards are minus travel times in minutes, with times Highway (22, 35, 60), Side streets (28, 30, 44), Train (34, 34, 34) in states (Free, Busy, Jam). Blind expected times are 38.6, 34.0, 34.0, so $\mathcal{A}^* = \{\text{Side, Train}\}$: case (B).

d (minutes)	$\mathcal{V}_\Phi(d)$ (min saved)	report Free	report Busy	report Jam
0	5.66	Highway	Side	Train
10	3.85	Highway	Side	Train
20	2.89	Highway	Train	Train
30	2.07	Highway	Train	Train
40	1.37	Highway	Train	Train
50	1.02	Side	Train	Train
60	0.83	Side	Train	Train
120	0.24	Side	Train	Train

- *The same report calls for different actions at different ages.* “Busy” means Side streets when fresh but Train once it is 20 minutes old: Busy tends to become Jam, so an old “Busy” is a jam warning. “Free” means Highway for up to 40 minutes, then Side streets.
- *Only $\mathcal{A}^*$ survives.* Highway is not blind-optimal and is used only while the report is fresh. From $d = 5$ on only Side and Train are chosen, as in the proof of Theorem 3(B), and $\mathcal{V}_\Phi(d) = \frac{1}{2}|\lambda_2|^d \sigma_2 \|\psi_2\|_{1,\pi} + O(|\lambda_3|^d)$ by (9), with $\sigma_2 = 6.20$ and $\|\psi_2\|_{1,\pi} = 0.917$. The leading term reproduces $\mathcal{V}_\Phi(d)$ to four digits from $d = 8$ on.
- *Rotation of the best bit* (Theorem 4). For the Side/Train contrast $v = r_{\text{Side}} - r_{\text{Train}} = (6, 4, -10)$ ($\langle \pi, v \rangle = 0$), the best one-bit report thresholds $g_d = K^d v$: {Free, Busy} vs {Jam} for $d = 0, 1$, and {Free} vs {Busy, Jam} for $d \geq 2$. Fresh, the bit asks “is it jammed?” (the sign pattern of v); stale, it asks “is it free?”, the sign pattern of the slow mode $\psi_2 = (1.22, -0.29, -1.11)$. The cosine $|\cos_\pi(g_d, \psi_2)|$ is 0.86, 0.98, 0.997, 1.000 for $d = 0, 1, 2, 3$. From $d = 5$ on, this single bit is worth all of $\mathcal{V}_\Phi(d)$.

Example D: a stale market regime, two bettors ($N = 2$, cases (B) and (S)). The market is in a Bull or Bear regime, one step = one month: $K = \begin{pmatrix} 0.95 & 0.05 \\ 0.10 & 0.90 \end{pmatrix}$, $\pi = (\frac{2}{3}, \frac{1}{3})$, $\lambda_2 = 0.85$. Regime information is naturally stale: business-cycle dates are announced many months after the fact, and output data arrive a quarter late. Consider two bettors who see the regime of d months ago.

- *Risk-neutral (Howard).* Actions Long, Cash, Short with monthly returns $r_L = (2, -4)\%$, $r_C = 0$, $r_S = -r_L$. Blind, all three earn 0, so $\mathcal{A}^* = \mathcal{A}$: case (B). The policy never changes: Long after a Bull report, Short after a Bear report, at every d . By (9), exactly, $\mathcal{V}_\Phi(d) = \frac{1}{2}|\lambda_2|^d \sigma_2 \|\psi_2\|_{1,\pi} = \frac{8}{3}(0.85)^d\%$ per month, with $\sigma_2 = \langle r_L - r_S, \psi_2 \rangle_\pi = 4\sqrt{2}$ and $\|\psi_2\|_{1,\pi} = \frac{2\sqrt{2}}{3}$.
- *Kelly (log-utility).* Each month the bettor splits wealth as $q \in \Delta(\mathcal{S})$ across the two regimes at odds $o(s)$, so a belief b yields growth rate $\Phi(b) = \max_q \sum_s b(s) \log(q(s)o(s)) = \langle b, \log o \rangle - H(b)$, attained at $q = b$. This is a Howard payoff with action set $\Delta(\mathcal{S})$. Its linear part $\langle b, \log o \rangle$ averages out by $\mathbb{E}_s \varepsilon_s^{(d)} = 0$, so for any odds $\mathcal{V}_\Phi(d) = H(\pi) - \mathbb{E}_s H(K^d(s, \cdot)) = \mathcal{T}(d)$, the temporal information (3): *the value of a stale report to a Kelly bettor is exactly the temporal information*, in nats of log-wealth growth per month (cf. Cover and Thomas, 2006, Ch. 6). The payoff is smooth, so case (S) gives $\mathcal{T}(d) \sim \frac{1}{2}|\lambda_2|^{2d}$. Here the policy does change with d : the bettor stakes the posterior $q = K^d(s, \cdot)$, which drifts from the report toward π .

d (months)	risk-neutral $\mathcal{V}_\Phi(d)$ (%/mo)	Kelly $\mathcal{T}(d)$ (nats/mo)	Kelly stake on Bull (Bull / Bear report)	ratio $\mathcal{V}_\Phi/\sqrt{\mathcal{T}}$
0	2.667	0.637	1.00 / 0.00	3.34
1	2.267	0.396	0.95 / 0.10	3.60
3	1.638	0.191	0.87 / 0.26	3.75
6	1.006	0.070	0.79 / 0.42	3.81
12	0.379	0.010	0.71 / 0.57	3.80
24	0.054	0.0002	0.67 / 0.65	3.78

The same stale signal ages at two rates: like $|\lambda_2|^d$ for the bettor who goes all-in (a kink) and like $|\lambda_2|^{2d}$ for the Kelly bettor (curvature). By $d = 12$ the Kelly value has fallen to 1.6% of its fresh level, while the risk-neutral value retains 14%. The ratio $\mathcal{V}_\Phi(d)/\sqrt{\mathcal{T}(d)}$ settles at $\sigma_2\|\psi_2\|_{1,\pi}/\sqrt{2} = 3.77$, as Corollary 1 predicts: decision value is the square root of temporal information, $\kappa = \frac{1}{2}$.

H Concluding remarks and open problems

Delay is the operation that pulls value apart from information. One object—the value orbit $\mathcal{V}_\Phi(d)$ —produces an aging trichotomy with a quantized exponent, the $\sqrt{\cdot}$ -decoupling of worth from information, eigen-bits as the durable ones, and a $\log n$ model-free horizon against a sharp $\sqrt{n}$ model-based one. Howard supplies the $d = 0$ anchor and DeGroot the reward-free functional; the delay d and the spectrum of K are the new content.

Open problems. (1) The model-free horizon over all trajectory-based estimators: Proposition 4 proves d^* is the limit when only lag- d pairs are seen, and Proposition 1 rules out data-independent corrections; whether any estimator that does not build a model can pass $\log n$ from a trajectory is open. (Theorem 6 settles the model-based side.) (2) The optimal m -bit stale experiment: the delayed resolution (21) singles out the level sets of the top slow modes $\psi_2, \dots, \psi_{m+1}$, and we conjecture their joint level set is exactly optimal; proving this, and solving the generic-VoI selection problem of Chen et al. (2015) over stale summaries at cost (value of information is not submodular for reports of one chain, cf. Steudel et al., 2010), remain open. (3) Noisy stale reports—a POMDP with delayed observations, where the backward recurrence $g_k = Kg_{k-1}$ acquires a maximum at every step and D_0 becomes a stopping rule. (4) Non-Markov worlds: under strong mixing the two-point law Q_d of Remark 1 still drives the value, but $\mathcal{T}$ can decay sub-exponentially and—as Example 1 shows—the value need no longer be monotone in d ; characterizing the envelope and its non-monotonicity is open.

[figure placeholder: fig_rotation3d.pdf]

Figure 5: **The rotation of the optimal one-bit summary, seen in three dimensions** ($N = 3$). *Setting.* A reversible chain on $\mathcal{S} = \{1, 2, 3\}$ with $\pi = (0.2, 0.3, 0.5)$ and spectrum $\{1, \lambda_2, \lambda_3\} = \{1, 0.7, 0.3\}$, one-step kernel $K = \begin{pmatrix} 0.501 & 0.323 & 0.175 \\ 0.216 & 0.649 & 0.135 \\ 0.070 & 0.081 & 0.849 \end{pmatrix}$, and π -orthonormal eigenfunctions $\psi_2 = (-0.88, -1.08, 1.00)$, $\psi_3 = (-1.80, 1.08, 0.07)$; and a two-action boundary decision with contrast $v = 0.51\psi_2 - 0.86\psi_3 = (1.10, -1.48, 0.45)$, so that $\langle \pi, v \rangle = 0$, $\|v\|_{2,\pi} = 1$ and $|\cos_\pi(v, \psi_2)| = 0.51$. We draw functions $f : \mathcal{S} \rightarrow \mathbb{R}$ in the coordinates $x_s = \sqrt{\pi(s)}f(s)$, which map $\ell^2(\pi)$ isometrically onto Euclidean $\mathbb{R}^3$, so angles in the picture are π -angles. Since $\mathbb{E}_s g_d(s) = 0$, every delay- d score $g_d = K^d v = \lambda_2^d \langle v, \psi_2 \rangle_\pi \psi_2 + \lambda_3^d \langle v, \psi_3 \rangle_\pi \psi_3$ of (12) lies in the plane $\mathbf{1}^\perp = \text{span}\{\psi_2, \psi_3\}$. (a) *The plane inside $\mathbb{R}^3$.* The grey arrow is the constant direction $\mathbf{1}$ and the tilted disk is $\mathbf{1}^\perp$. By Theorem 4 the optimal summary is $f_d^* = \mathbf{1}[g_d > 0]$, i.e. $C_0 = \{s : g_d(s) > 0\}$; it depends only on the sign pattern of g_d , which is constant on each of the six sectors into which the three coordinate planes $\{x_s = 0\}$ cut the disk. Only one of these planes matters here: the sheet $\{x_1 = 0\}$, which meets the disk along the dashed line $\{g(1) = 0\}$. The red sector is the cell $C_0 = \{1, 3\}$ (the reward threshold, $\mathbf{1}[v > 0]$); the blue sector is $C_0 = \{3\}$ (the eigen-bit $\mathbf{1}[\psi_2 > 0]$). The dots are the normalized scores $g_d/\|g_d\|_{2,\pi}$ for $d = 0, 1, 2, \dots$; they slide along the rim from v to ψ_2 . (b) *Delay as height.* The same plane, in the (ψ_2, ψ_3) frame, with the delay d stacked upward. The direction of g_d is the angle α_d with $\tan \alpha_d = \frac{\langle v, \psi_3 \rangle_\pi}{\langle v, \psi_2 \rangle_\pi} (\lambda_3/\lambda_2)^d$, so the curve climbs the cylinder and bends toward the ψ_2 line at the geometric rate $(\lambda_3/\lambda_2)^d = (3/7)^d$. The boundary $\{g(1) = 0\}$ becomes a vertical wall; the curve starts on the red side and crosses the wall at the (continuous) delay $d^\dagger = \log\left(\frac{-\langle v, \psi_3 \rangle_\pi \psi_3(1)}{\langle v, \psi_2 \rangle_\pi \psi_2(1)}\right) / \log(\lambda_2/\lambda_3) \approx 1.46$, where $g_d(1)$ changes sign ($g_1(1) = 0.15 > 0$, $g_2(1) = -0.08 < 0$). Hence the most useful bit of the stale report is the reward threshold $C_0 = \{1, 3\}$ for $d \leq 1$ and the eigen-bit $C_0 = \{3\}$ for every $d \geq 2$. *Reading.* (i) Delay acts on the score, not on the decision: $g_d = K g_{d-1}$ is the backward recurrence of Section 5, and each application of K shrinks the ψ_3 -coordinate by 0.3 but the ψ_2 -coordinate only by 0.7, which is the low-pass filter in geometric form. (ii) The *direction* of g_d converges to $\pm\psi_2$, so the chosen bit freezes after finitely many steps; its *value* $\mathcal{V}_{f_d^*}(d) = \frac{1}{2}\|g_d\|_{1,\pi}$ keeps decaying as $\Theta(|\lambda_2|^d)$ (Theorem 4). (iii) A switch of the bit happens only if v and ψ_2 lie in different sectors; when they share a sector, the score still rotates but the chosen partition never changes (as for the Beijing alert decision in Section 7, where the rotation is in the score, not in the chosen bit). (iv) For general N the picture is the same with $\mathbf{1}^\perp$ of dimension $N - 1$, cut into cones by the hyperplanes $\{f(s) = 0\}$; the trajectory of $g_d/\|g_d\|$ is drawn toward ψ_2 at rate $(|\lambda_3|/|\lambda_2|)^d$.