Acquiring a Predictive Capability through Restricted Noisy Experiments

Exact frontiers and certified asymmetric irreversible acquisition

Brian Theory
AI-assisted research draft

13 November 2026

D02 — Classification: technical methods note. The results below are proved within the declared models. They are elementary specializations of statistical testing and experiment comparison; historical novelty is not claimed. This is an internally reviewed draft, not an externally peer-reviewed article.

Abstract

A predictive representation may exist without being identifiable through a permitted sensing interface. We formulate finite-horizon acquisition as a minimax decision problem, distinguish parameter identification from predicting a declared target, and derive a two-model risk obstruction directly from transcript laws. In a binary family with a forbidden exact probe and available binary symmetric channels, we solve the minimax absolute prediction risk exactly. For a fixed channel library an optimal policy repeatedly uses the least noisy permitted channel; the risk is the target separation times a binomial majority error, including the correct randomized tie term. The conclusion remains valid against adaptive policies because every weaker allowed experiment can be simulated from the best channel. A future-target example shows that prediction can become easier while historical identification remains equally difficult. Exact transcript enumeration and rational dynamic programming corroborate the formulas. We further solve an irreversible scout-and-commit protocol, give an exact five-probe adaptive allocation advantage over every fixed allocation, and derive a priced stopping specialization with a strict finite example. We then remove the symmetry shortcut: finite legal-policy mixtures and prior-based lower bounds certify asymmetric minimax risks, with exact six-candidate examples and a common-interface model-mismatch bound. These methods are standard decision-theoretic tools; historical novelty is not established. The analysis does not solve general constrained stochastic experiment design, infer finite physical storage from real-dimensional predictive states, or extend a homotopy-based admissibility theorem beyond its original deterministic assumptions.

1 Problem, baseline, and actual increment

The question is operational: given a library of executable experiments, what prediction error can an agent attain after a specified number of observations? Computing a sufficient predictive state from a known model is a different problem from acquiring that state through the actual interface. It is also different from identifying every parameter of the model. A parameter can be unidentifiable and irrelevant to a requested prediction; conversely, an apparently accurate predictor can fail on an intervention outside its acquisition policy.

The maintained B-Theory report *From Observable States to Acquirable Predictive Models*, revised 7 October 2026, studies noiseless queries on a particular finite ordered carrier and known deterministic

actions. Its Theorem 2 and Corollary 3 characterize reportable targets and the finest attainable transcript partition, while Section 7 changes the timing contract when actions alternate with queries [13]. The present model adds stochastic readout and an explicit bounded prediction loss but does not carry over that report's structural permission rule. It is a specialization with a different experiment library, not a generalization of its entire classification.

There is further overlap with *Causal Self-Diagnosis with Noisy Tests*, version 0.3 of 5 November 2026. Its finite-law framework, oracle reductions, repetition arguments, and finite decision calculations already address substantial portions of stochastic acquisition [10, Sections 5–9]. Thus the broad thesis that noisy permitted observations determine attainable diagnosis is not new in the corpus. The useful increment of this note is a completely solved benchmark that separates acquisition error from the separation of the evaluation targets and includes a finite-budget optimality proof against all allowed adaptive policies.

Reading the continuation. Sections 7–8 give symmetric adaptive and stopping controls. Sections 10–12 remove the symmetry shortcut, give executable asymmetric certificates, and state what model mismatch does and does not protect. The general methods are standard; exact finite evidence and the physical permission contract are the focus.

Nearest external work. Predictive-state representations encode action-conditional test predictions and can have finitely many real coordinates even when they attain infinitely many values; the linear representation theorem of Littman et al. [6, Theorem 1, Sections 1–2] is an existence/construction result for a supplied finite POMDP. Causal-state equivalence groups histories with the same future distributions [11, Definition 5, Theorem 1 and Corollary 1]. Neither fact by itself supplies a feasible experiment protocol. Controlled sensing for hypothesis testing directly optimizes observation choices; Nitinawarat et al. [9, Section II and Proposition 1] treat stationary conditionally independent observation laws and establish a binary fixed-sample asymptotic error-exponent result. Our binary calculation is finite-sample and exact because the channels are ordered by explicit simulation. It is much narrower than that literature.

2 Operational contract

2.1 General finite formulation

Let Θ be a finite set of known candidate models, and let $\mathcal{A}, \mathcal{Y}$ be finite action and observation alphabets. Nature fixes an unknown $\theta \in \Theta$. At acquisition step t, the controller sees its transcript

$$H_t = (A_1, Y_1, \ldots, A_t, Y_t)$$

and chooses A_{t+1} from a specified nonempty set $\mathcal{A}(H_t)$. The same legality function is supplied under every model; it cannot reveal θ. A randomized policy uses private randomness independent of nature. The system then makes its transition and returns its observation. A finite kernel

$$K_\theta(s', y \mid s, a)$$

therefore specifies transition-before-readout timing. Initial state laws and all candidate kernels are known. The latent state and the true model label are hidden. Actions may alter state in this general notation, but the solved example below has a fixed acquisition state.

After exactly n probes, the controller outputs $\widehat{v}(H_n)$ in a specified decision space. The target v_θ is a model-dependent test prediction, fixed by the evaluation protocol. If a target instead depends

on the acquisition trajectory, that dependence must be included in its definition and in the loss; the lower bound below is stated for fixed v_θ. For a metric $d \le 1$ the criterion is

$$R_n^* = \inf_{\pi,\widehat{v}} \sup_{\theta\in\Theta} \mathbb{E}_\theta^\pi d(\widehat{v}(H_n), v_\theta). \tag{1}$$

The infimum includes all history-based randomized policies obeying the same interface. There is no computational or memory constraint in (1). A lossless fixed-length encoding of the complete transcript needs at most $\lceil n \log_2(|\mathcal{A}||\mathcal{Y}|)\rceil$ bits, plus fixed code and numeric precision. We do not silently replace those costs by the dimension of a probability vector. Every comparator receives the same candidate family, legality rule, horizon, and cost accounting. An unrestricted oracle that can use a forbidden sensor is explicitly a different interface.

2.2 Equivalence is indexed by the permitted experiment

For a fixed policy π, write P_θ^π for its law on the finite transcript H_n. Terminal randomization is a common Markov kernel applied afterward; reports need not have a finite alphabet. Two parameters are n-experiment-equivalent if these laws agree for every permitted policy. This is weaker than equality of their hidden transition tables and stronger than equality under one selected policy. Equality of observations under a passive policy need not imply equality under interventions. Histories that have probability zero under every candidate can be assigned arbitrary actions; they have no effect on the risks. Conditional expressions below are used only on histories of positive probability under the relevant law.

3 A transcript obstruction with a complete proof

For finite distributions P, Q, let

$$\mathrm{TV}(P,Q) = \frac{1}{2}\sum_z |P(z) - Q(z)| = 1 - \sum_z \min\{P(z), Q(z)\}.$$

Proposition 3.1 (Two-model metric bound; standard two-point argument)**.** *Fix θ_0, θ_1 with target distance $\Delta = d(v_0, v_1)$. Every policy and estimator satisfy*

$$\max_{i\in\{0,1\}} \mathbb{E}_i^\pi d(\widehat{v}, v_i) \;\ge\; \frac{\Delta}{2}\left(1 - \mathrm{TV}(P_0^\pi, P_1^\pi)\right).$$

Consequently, identical transcript laws for every permitted policy imply $R_n^* \ge \Delta/2$.

Proof. Condition on a transcript z and on any estimator randomization. The triangle inequality gives $d(\widehat{v}(z), v_0) + d(\widehat{v}(z), v_1) \ge \Delta$. Multiply by $\min(P_0^\pi(z), P_1^\pi(z))$, sum over transcripts, and average over randomization. The sum of the two risks is at least $\Delta \sum_z \min(P_0^\pi(z), P_1^\pi(z))$. Their maximum is at least half their sum. The infimum over policies preserves the inequality when all transcript distances vanish. □

This proof establishes a lower bound, not a general exact minimax formula. With more than two parameters, the hardest pair need not characterize the whole decision problem. Even for two parameters, unequal costs and asymmetric channels can require a nonuniform least favorable prior.

Proposition 3.2 (Adaptive information budget in the memoryless case)**.** *Suppose that, conditional on θ and the selected action a, each new observation has mass function k_θ^a, independently of the preceding transcript. Assume $k_0^a \ll k_1^a$ for every allowed action and*

$$\kappa = \max_a \mathrm{KL}(k_0^a \| k_1^a) < \infty,$$

where KL uses natural logarithms. Then every admissible policy satisfies

$$\mathrm{KL}(P_0^\pi \| P_1^\pi) \le n\kappa, \qquad R_n^* \ge \frac{\Delta}{2}\left(1 - \min\{1, \sqrt{n\kappa/2}\}\right).$$

Proof. The likelihood of a transcript is the product of the policy's conditional action probabilities and the selected observation probabilities. In the likelihood ratio, the action factors cancel at each common history: the policy cannot depend on the unknown parameter. Taking the logarithm, expectation under model zero, and conditioning successively on histories gives

$$\mathrm{KL}(P_0^\pi \| P_1^\pi) = \sum_{t=1}^{n} \mathbb{E}_0^\pi \, \mathrm{KL}(k_0^{A_t} \| k_1^{A_t}) \le n\kappa.$$

For completeness, the finite-distribution Pinsker bound follows by taking the event $B = \{z : P(z) > Q(z)\}$. The log-sum inequality yields $\mathrm{KL}(P\|Q) \ge \mathrm{kl}(P(B)\|Q(B))$. For fixed $q \in (0,1)$, the binary KL as a function of p has value and derivative zero at $p = q$, and second derivative $1/[p(1-p)] \ge 4$. Integrating twice gives $\mathrm{kl}(p\|q) \ge 2(p-q)^2$; endpoints follow by limits. Since $P(B) - Q(B) = \mathrm{TV}(P, Q)$, the asserted variation bound follows. Apply Proposition 3.1 and take the infimum over policies. □

The factorization assumption is substantive. A persistent common sensor fault need not yield independent evidence when the same probe is repeated. Nor does this proposition apply without modification to an unknown changing state. The displayed lower bound is often loose and becomes zero when its information budget is large; it should not replace the exact calculation when that calculation is available.

4 An exactly solved restricted experiment

4.1 Fixed acquisition, declared future target

The acquisition state is a bit $\theta \in \{0, 1\}$. A probe of type a returns

$$Y_t = \theta \oplus N_t, \qquad N_t \sim \mathrm{Bernoulli}(\eta_a),$$

with independent fresh noise across steps, conditional on the chosen actions. All probes have unit cost and preserve θ. The allowed channel set is finite, with

$$0 < \eta_* := \min_{a \in \mathcal{A}} \eta_a < 1/2, \qquad \eta_a \le 1/2.$$

A perfect readout is available only to an explicitly unrestricted comparator. The number n of acquisition probes is fixed before observing data. There is no early-stop benefit or state change during acquisition.

To evaluate a future prediction, start an independent test trajectory at $X_0 = \theta$ after acquisition and flip the bit independently with known probability $\rho \in [0, 1/2]$ on each test step. Its horizon is h, held equal across all policies. The requested probability is

$$v_\theta = \mathbb{P}(X_h = 1 \mid X_0 = \theta) = \frac{1}{2} + (\theta - \tfrac{1}{2})(1 - 2\rho)^h. \tag{2}$$

Indeed, $v_\theta(h+1) = \rho + (1-2\rho)v_\theta(h)$ with $v_\theta(0) = \theta$; subtracting $1/2$ and iterating proves (2). Thus $\Delta_h = |v_1 - v_0| = (1-2\rho)^h$. When $h = 0$, define $\Delta_0 = 1$, including $\rho = 1/2$. The target is a probability, not the realized future bit. Loss is $|\widehat{v} - v_\theta|$. This excludes the irreducible noise of a realized future outcome and directly measures acquisition of the model's requested forecast.

4.2 Every weaker channel can be simulated

Lemma 4.1 (Explicit channel simulation)**.** *If $0 \le \eta_* < 1/2$ and $\eta \in [\eta_*, 1/2]$, a BSC(η) observation is obtained from a BSC(η_*) observation by an independent additional flip of probability*

$$q = \frac{\eta - \eta_*}{1 - 2\eta_*} \in [0, 1/2].$$

An adaptive n-probe policy using any of the allowed channels can therefore be simulated using n independent BSC(η_*) observations and independent random bits.

Proof. The composition has crossover probability $\eta_*(1-q) + (1-\eta_*)q = \eta$. To simulate an adaptive policy, feed it the previously simulated observations, use its prescribed next action, and garble the next fresh strongest-channel reading by the corresponding q. Conditional on every previous simulated history, the next observation has exactly the required law. Induction proves equality of the full transcript laws under each value of θ. The simulator has not inspected θ. □

This is the sufficiency direction of Blackwell comparison by randomization, here proved explicitly. A primary modern proof is de Oliveira [1, Theorem 1 and Section 4]; the historical comparison-of-experiments literature has not been exhaustively audited here. No priority claim follows from this binary specialization. The conclusion relies on equal costs, fresh observations, and no action-dependent state effects. A weaker but much cheaper channel can have better value per resource unit; the lemma alone says nothing about that changed problem.

Theorem 4.2 (Exact minimax predictive-acquisition frontier)**.** *Under the binary contract above, for every $n \ge 0$,*

$$R^*_{n,h} = \Delta_h \, e_n(\eta_*), \tag{3}$$

$$e_n(\eta) = \sum_{k=\lfloor n/2 \rfloor + 1}^{n} \binom{n}{k} \eta^k (1-\eta)^{n-k} + \begin{cases} \frac{1}{2}\binom{n}{n/2}\eta^{n/2}(1-\eta)^{n/2}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases} \tag{4}$$

For $n = 0$, the formula gives $e_0 = 1/2$. An optimal policy always uses a channel with noise η_*, reports v_1 if a strict majority of observations are one and v_0 if a strict majority are zero, and uses an unbiased tie decision. Reporting $(v_0 + v_1)/2$ on a tie has the same risk.

Proof. First consider the strongest channel and a uniform prior on θ. The likelihood ratio for k ones among n observations is $[(1-\eta_*)/\eta_*]^{2k-n}$. Conditional absolute loss for a report $v \in [v_0, v_1]$ is

$$(1-p)(v - v_0) + p(v_1 - v),$$

where p is the posterior probability of $\theta = 1$. This affine expression is minimized at v_1 when $p > 1/2$, at v_0 when $p < 1/2$, and by any point of the interval on a tie. Reports outside the interval cannot improve loss. Hence majority is Bayes optimal. Under either parameter, an error has loss Δ_h and occurs when more than half of the independent noise bits flip, with half loss on ties. This gives the stated value and equal risks for both parameters.

For any rule, maximum risk is at least its uniform-prior Bayes risk. Since majority attains the optimal Bayes risk with equal risks, it is minimax for the strongest-channel experiment. Lemma 4.1 shows that every allowed adaptive policy is a randomized postprocessing of that experiment, so none has a lower Bayes risk. The same lower bound therefore applies to all permitted policies. The always-strongest majority policy is allowed and attains it. □

The frontier depends on the loss. Absolute error in the model probability selects a posterior median, which is an endpoint in this two-model problem. It is not the squared-error or Brier frontier. For one probe with $h = 0$ and $\eta = 0.2$, the posterior mean under the uniform prior is 0.2 or 0.8; its squared-error risk under either parameter is

$$(1-\eta)\eta^2 + \eta(1-\eta)^2 = \eta(1-\eta) = 0.16,$$

whereas the proved absolute-risk optimum is 0.2. In this $h = 0$ example the target is itself a bit, so squared error is the Brier loss. No claim about other scoring rules is inferred from (3).

Boundary cases and finite versus asymptotic statements. If $\eta_* = 0$, one observation identifies θ exactly. If every channel has $\eta_a = 1/2$, observations are independent of θ and risk is $\Delta_h/2$ for all n. If $0 < \eta_* < 1/2$ and $\Delta_h > 0$, every finite transcript has positive probability under both models, so the risk is strictly positive at every finite n. Nevertheless it tends to zero. Hoeffding's bounded-sum inequality [3, Theorem 2, equation (2.6), p. 16] gives

$$e_n(\eta_*) \le \exp\{-2n(1/2-\eta_*)^2\}.$$

Thus a sufficient condition for $R^*_{n,h} \le \varepsilon$ with $0 < \varepsilon < \Delta_h$ is

$$n \ge \frac{\log(\Delta_h/\varepsilon)}{2(1/2-\eta_*)^2}.$$

This is a sufficient sample count, not an exact inverse frontier. When $\Delta_h = 0$, the target is the same under both models and risk is zero without acquisition, although θ remains unknown.

5 Examples, resource accounting, and failed generalizations

At noise $\eta_* = 0.2$, the historical-target risk is 0.2 after one observation, 0.05792 after five, and 0.01958144 after ten. An exact sensor would give zero after one probe, but it is forbidden in the restricted experiment. A controller using the permitted but uninformative channel has risk 1/2. These comparators have different declared channel rights; no architecture advantage is inferred.

For $\rho = 0.1$ and $h = 5$, the two future probabilities are separated by $0.8^5 = 0.32768$. With five acquisition probes the risk is therefore 0.0189792256. Historical-identification error is still 0.05792. The reduction is caused by the evaluation target losing sensitivity to the historical bit, not by acquiring more information. If all test trajectories mix to the same distribution, long-horizon probability prediction may be easy while origin identification remains difficult.

The majority implementation needs the count of ones and the probe index: at most $2\lceil \log_2(n+1) \rceil$ mutable bits, plus finite program and parameter storage. The probability pair in (2) requires an agreed precision if physically represented. The complete transcript is unnecessary for this symmetric example; a count is sufficient because the channel laws are known and stationary. This does not prove a universal logarithmic-memory requirement for stochastic prediction.

Three changes defeat the exact proof. First, if every repetition shares one unknown flip bit, majority repeats the same mistake and e_n is no longer binomial. Second, if probes perturb the target, channel simulation need not preserve the evaluation law. Third, if a sensor has unknown polarity, a nominally low error rate does not fix its orientation; the pair $(\theta, \text{polarity})$ can have an exact label symmetry. That is a D03 calibration problem. Finally, under history-dependent permission constraints a simulated weaker query might change which later actions are allowed. Lemma 4.1 here uses a fixed channel library with no cumulative permissions.

6 Reproducible verification

The file `code/reproduce.py` computes (3), saves `data/frontier.csv`, and regenerates Figure 1. It also enumerates every binary transcript for $n \leq 10$ at five noise values and verifies $e_n = (1 - \mathrm{TV})/2$. A separate exact-rational Bayesian dynamic program optimizes over the three allowed noise levels $1/5, 7/20, 1/2$ for horizons 0 through 8. All nine values agree with the majority expression. These are bounded implementation checks; the channel-simulation proof supplies the all-horizon claim. No

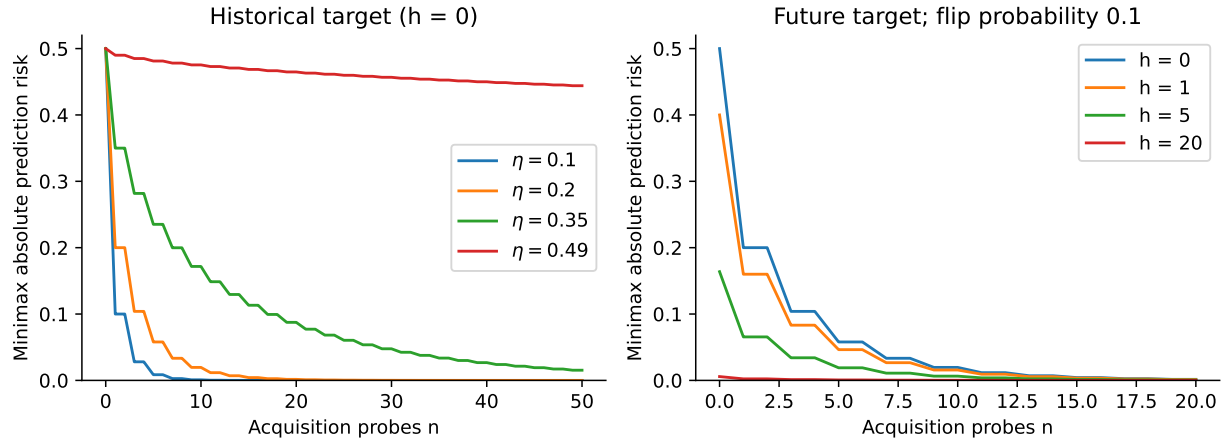


Figure 1: Analytical minimax risks, not empirical estimates. Left: historical target. Right: fixed acquisition noise 0.2 with future-target horizon h and flip probability 0.1. The different right-hand curves use different declared targets; they do not show improved acquisition of the same historical bit. Editable SVG and generating Python source are included.

random seed is needed for these deterministic calculations. The rational dynamic program verifies Bayesian values, and symmetry is what identifies them with minimax values here. That baseline script does not establish minimax equality for asymmetric models; Section 10 supplies a separate asymmetric certification method.

7 Irreversible sensing modes: an adaptive finite-budget frontier

The dominant-channel benchmark removes the reason to adapt. We now resolve one specific instance of the previously proposed mutually exclusive sensing problem. The increment is an explicit finite contract and its exact comparisons; neither adaptive sensing nor belief-state dynamic programming is new.

7.1 Executable protocol and matched comparators

Nature fixes two hidden bits $(M, B) \in \{0,1\}^2$. The mode M determines which sensor is useful; the task is to predict B, with absolute loss $|\widehat{b} - B|$. A controller starts uncommitted with total budget $N \in \mathbb{Z}_{\geq 0}$. While uncommitted it may pay one unit for a scout

$$Z = M \oplus U, \qquad U \sim \text{Bernoulli}(\alpha), \quad 0 \leq \alpha \leq 1/2.$$

Alternatively it irrevocably commits to mode $a \in \{0,1\}$. Commitment is a decision with no extra charge. Every remaining unit buys a target reading: BSC(η) of B if $a = M$, or a fresh fair bit if $a \neq M$, where $0 \leq \eta \leq 1/2$. No further scouts or mode changes are allowed. All noises are fresh and independent conditional on (M, B) and the actions. Exactly N paid observations are used. Zero remaining observations permit an arbitrary report. No counterfactual readout from the unchosen mode is exposed.

The model label now has four possibilities although the target has two. The good/bad mode is fixed across target readings: marginalizing it produces a mixture of product laws, *not* an iid BSC with averaged crossover probability. A uniform prior on (M, B) is used to prove a lower bound; the performance criterion remains the maximum risk over the four fixed parameter values.

We compare: (i) fully adaptive scouting duration and commitment; (ii) a scout count s fixed independently of observations, followed by data-dependent commitment and $N - s$ target readings; and (iii) open-loop acquisition actions, including commitment independent of all observations. Terminal prediction may use all acquired data in every class. Private randomization is allowed in every class. These are identical sensor rights and total budgets with different feedback restrictions.

Lemma 7.1 (Risk after commitment)**.** *Under the uniform prior, before commitment let* $p = \mathbb{P}(M = 1 \mid Z_{1:s})$*. With* r *readings left, the minimum conditional Bayes risk after choosing a mode is*

$$g_r(p) = \frac{1}{2} - \max\{p, 1-p\} \left(\frac{1}{2} - e_r(\eta) \right).$$

An optimizer selects a posterior-most-likely mode and reports the majority of the target readings, randomizing unbiasedly at ties.

Proof. Scouting leaves B fair and independent of M. For a selected mode with posterior probability w of being correct, the target transcript has laws

$$Q_b(y) = w P_b^\eta(y) + (1-w) 2^{-r}.$$

Their signed difference is $w(P_1^\eta - P_0^\eta)$, so majority remains Bayes-optimal, including $w = 0$ where every report ties. Averaging errors gives $we_r + (1-w)/2$, minimized by the largest available w. This argument retains the persistent mixture and does not replace it by independent effective noise. □

Theorem 7.2 (Fixed allocation, adaptive allocation, and minimax)**.** *For a fixed integer scout count* $0 \leq s \leq N$ *and* $r = N - s$*, the exact minimax risk is*

$$F_{N,s} = e_r(\eta) + \left(\frac{1}{2} - e_r(\eta) \right) e_s(\alpha). \tag{5}$$

The best open-loop acquisition risk is

$$O_N = \frac{1}{4} + \frac{1}{2} e_N(\eta). \tag{6}$$

Define $V_0(p) = 1/2$ *and, for* $r \geq 1$,

$$\begin{aligned} q(p) &= \alpha + (1-2\alpha)p, \\ p_1 &= \frac{(1-\alpha)p}{q(p)}, \qquad p_0 = \frac{\alpha p}{1-q(p)}, \\ V_r(p) &= \min\left\{g_r(p),\ q(p)V_{r-1}(p_1) + (1-q(p))V_{r-1}(p_0)\right\}. \end{aligned} \tag{7}$$

Zero-probability branches are omitted, avoiding division by zero. The fully adaptive minimax risk is $V_N(1/2)$, *and*

$$V_N(1/2) \leq \min_{0\leq s\leq N} F_{N,s} \leq O_N.$$

Each V_r *is symmetric about* $1/2$, *concave and piecewise affine. No single posterior threshold structure is asserted.*

Proof. For fixed s, posterior-MAP estimation of M from scouts has error $e_s(\alpha)$. Conditional on selecting the correct mode the target-majority error is $e_r(\eta)$; conditional on an incorrect choice it is $1/2$. Lemma 7.1 proves Bayes optimality, yielding (5). These rules have equal risks under all four parameter values, by complementing scout bits and mode labels or complementing target bits and the final report. Uniform-prior Bayes risk is a lower bound on maximum risk, so the equal-risk rule is minimax.

For an open-loop fixed mode and any fixed allocation, under the uniform prior the mode is correct with probability $1/2$. Scouts can change that posterior probability but cannot change which target-majority decision is optimal: the signed likelihood difference in Lemma 7.1 has the same sign. Averaging therefore gives $1/4 + e_{N-s}/2$. Since additional observations cannot worsen optimal Bayes risk, e_n is nonincreasing; hence skipping scouts gives the smallest lower bound. A uniformly randomized mode and majority attain it with equal risks. Open-loop random mixtures of allocations cannot improve an average below its smallest constituent value. This proves (6) for the entire declared class.

Before commitment, the scout posterior p, remaining budget r, and the known uncommitted permission state are sufficient. Commitment incurs $g_r(p)$. A scout incurs the second term of (7) by Bayes' rule and induction. These exhaust the permitted next decisions. Finite horizons and alphabets ensure the minima are attained. Randomization cannot lower a conditional scalar minimum. For each r, the expected loss of any fixed finite deterministic policy tree is affine in the initial mode probability p; there are finitely many trees after restricting terminal predictions to endpoint majority rules. Their minimum is piecewise affine and concave. Relabeling the modes proves symmetry.

Finally symmetrize a Bayes-optimal tree over the four transformations generated by mode-complementation and target-bit complementation, with independent uniform private randomness. The protocol and loss are invariant, and this finite group acts transitively on (M, B). The transformed rule has the original uniform-prior Bayes risk at each fixed parameter. Thus $V_N(1/2)$ is attained as maximum risk and equals minimax risk. The inclusions of policy classes give the inequalities. □

7.2 A strict gap and a counterexample to cost-unmatched claims

For $N = 5$ and $\alpha = \eta = 1/5$, the exact values are

$$V_5(1/2) = \frac{1061}{6250} = 0.16976 < \min_s F_{5,s} = \frac{229}{1250} = 0.1832 < O_5 = \frac{3487}{12500} = 0.27896.$$

The fixed-allocation risks for $s = 0, \dots, 5$ are respectively

$$\left(\frac{3487}{12500}, \frac{229}{1250}, \frac{229}{1250}, \frac{289}{1250}, \frac{289}{1250}, \frac{1}{2}\right).$$

An optimal adaptive rule takes two scouts. If they agree, it commits to their mode and takes three target readings. If they disagree, it takes a third scout, commits to its indicated mode, and takes two target readings. All branches use five paid observations. Under each fixed (M, B), its error is

$$(1-\alpha)^2 e_3(\eta) + \frac{\alpha^2}{2} + 2\alpha(1-\alpha)\left[(1-\alpha)e_2(\eta) + \frac{\alpha}{2}\right] = \frac{1061}{6250}. \tag{8}$$

The following backward values certify optimality, as distinct from merely constructing an upper bound. Symmetry supplies the omitted complement posteriors; the columns compare immediate commitment with one scout followed by optimal continuation.

Remaining r	Mode posterior p	Commit $g_r(p)$	Scout continuation
1	1/2	7/20	1/2
1	16/17	37/170	1/2
1	256/257	1034/5140	1/2
2	4/5	13/50	13/50
2	64/65	133/650	133/650
3	1/2	151/500	13/50
3	16/17	541/4250	37/170
4	4/5	229/1250	1061/6250
5	1/2	3487/12500	1061/6250

The additional precision of adaptive scouting is small but strict: its advantage over the best fixed allocation is $84/6250 = 0.01344$. Under the uniform prior every scout has zero mutual information about B alone, yet scouting can improve subsequent acquisition of B. A rule maximizing only immediate information about the target would overlook this instrumental value. This is not a universal failure theorem for information-based design objectives.

Paying for a scout is not always worthwhile. With $N = 3$, $\eta = 1/5$ and $\alpha = 49/100$, one scout followed by two target readings has minimax risk $347/1000$, whereas skipping the scout and using three target readings has risk $151/500$. Holding the number of target readings fixed while adding a free scout would hide this counterexample. At $\alpha = 1/2$, scouts are wholly uninformative about mode; at $\eta = 1/2$, no policy predicts B better than $1/2$.

The stopping in this section concerns the *scouting phase*; total sample size is fixed. In the terminology of Naghshvar and Javidi [8, Section II-B], this is fixed-size adaptive acquisition, not a variable-total-sample sequential test. The next section treats actual stopping of acquisition.

8 Stopping acquisition when probes have a price

Return to the original binary fixed library of BSCs, now permit a stopping time $0 \le T \le N$ and charge $c \ge 0$ for each executed probe. The criterion is

$$\inf_{\pi, T, \widehat{v}} \max_{b \in \{0,1\}} \mathbb{E}_b[|\widehat{v} - v_b| + cT].$$

The probability target separation $\Delta = |v_1 - v_0|$ remains fixed, independent of stopping and acquisition history. This is an additive cost objective, not a hard expected-budget constraint or a theorem about a moving target. Define for the best channel η_*

$$q_\eta(p) = \eta_* + (1 - 2\eta_*)p, \qquad u_1(p) = \frac{(1-\eta_*)p}{q_\eta(p)}, \quad u_0(p) = \frac{\eta_* p}{1 - q_\eta(p)}.$$

Proposition 8.1 (Standard finite-horizon optimal stopping specialization)**.** *Put* $J_0(p) = \Delta \min(p, 1-p)$ *and*

$$J_r(p) = \min\left\{\Delta \min(p, 1-p),\ c + q_\eta(p) J_{r-1}(u_1(p)) + (1 - q_\eta(p)) J_{r-1}(u_0(p))\right\},$$

again omitting impossible branches. Then the minimax additive objective equals $J_N(1/2)$*. There is an optimal policy that uses only the strongest channel whenever it continues.*

Proof. The stopped-policy simulation uses a fresh strongest-channel observation only when the simulated policy continues, then garbles it as in Lemma 4.1. It reproduces the entire stopped transcript and its stopping-time law under both parameters, including the number of paid probes. Thus weaker channels cannot improve the objective. For the strongest channel, the posterior suffices. Stopping costs $\Delta \min(p, 1-p)$ and continuing costs the second expression, proving the recursion by backward induction. Reflection $b \mapsto 1 - b$ and complementing every reading symmetrizes any Bayes-optimal rule to equal total risks under the two parameters. The uniform-prior lower bound is therefore attained as minimax risk. □

For $N = 3$, $\eta_* = 1/5$, $\Delta = 1$ and $c = 1/50$, an optimum takes two readings, stops if they agree, and otherwise takes a third. Its error is $e_3 = 13/125$, its expected sample size under each parameter is $2 + 2\eta_*(1 - \eta_*) = 58/25$, and its objective is $94/625 = 0.1504$. Backward checks are: at $p = 16/17$ and one reading left, stop for $1/17$; at $p = 1/2$ and one left, continue for $11/50$; at $p = 4/5$ and two left, continue for $163/1250$; at the initial prior continue for $94/625$. The best fixed sample count $n \in \{0, 1, 2, 3\}$ has objective

$$\min_n [e_n(1/5) + n/50] = \min\{1/2, 11/50, 6/25, 41/250\} = 41/250 = 0.164.$$

Randomizing a fixed count independently of observations cannot beat its smallest constituent objective. This exact small example demonstrates a sequentiality gain under a stated price and cap; it does not introduce a new stopping theory.

9 Prior art, evidence, and the D04 boundary

The comparison categories and the optimal-stopping machinery have direct antecedents. Naghshvar and Javidi [7, Section 3, Fact 1] give a belief-state stop-or-sense Bellman equation for active hypothesis testing. Naghshvar and Javidi [8, Sections II and IV] distinguish sequentiality from adaptivity and analyze their gains. Li et al. [5, Section III-A, equations (13)–(15)] study finite-horizon stopping and online sensor selection with usage constraints and sampling prices. Our irreversible two-stage permission state and target-only loss differ from their contracts, but a new restricted instance is not a new dynamic-programming principle.

The current search also found the September 2026 revision of Vershinin et al. [15], which addresses unequal and random sensing costs and action cancellation, and Vijayan and Mandal [16], which studies fixed-budget active-testing exponents. These are relevant continuing research lines, not support for a historical-priority claim here. Source editions and inspection depths are recorded in the new source ledger. The search was targeted and is not exhaustive; the precise toy formula's historical novelty remains unverified.

The original computations were rerun from isolated copies: all 55 transcript checks and nine rational Bayesian-DP checks passed, and both original data files were byte-identical. The extension script `../code/verify_extensions.py` independently optimizes a full posterior on all four (M, B) states, without invoking Lemma 7.1 inside that oracle, and agrees exactly with the scalar recurrence

on 200 parameter/horizon combinations. It independently checks the fixed-allocation formula against a constrained full-state oracle in 900 cases. It also directly enumerates paths of the displayed five-probe policy under every parameter, verifies all fixed allocations, the budget counterexample, and the total-stopping example. A separate specialist wrote a second exact rational recursion and statewise policy calculation. These are bounded computational checks, not formal proof verification. The printed arguments establish the model-level results; no empirical sensor data were used.

D04's finite-library theorem assumes paid common validation episodes that can acquire the union of candidate tests. The mutually exclusive sensors here violate that assumption; observing a scout and one selected sensor does not produce the counterfactual outputs of all modes. This gives D04 an executable boundary example, not a coupled representation-selection theorem. The bits, candidate modes and noise laws are supplied and fixed here. Nothing learns a new representation language, estimates the channel laws from paid labels, prices its code, or optimizes representation-revision costs. The useful handoff is to retain D02's acquisition/permission state and D04's validation/revision charges in a future joint model. Merely renaming a mode a concept does not prove that joint result.

10 Beyond symmetry: executable minimax certificates

The preceding exact frontiers exploit symmetric channels and equalizer policies. An asymmetric nuisance parameter removes that shortcut. This section supplies a general finite certification contract and a larger asymmetric benchmark. The underlying methods are finite zero-sum game duality and belief-state control, not new minimax principles. Their purpose is to make the computational claim checkable without assuming a uniform prior is least favorable.

10.1 Finite permissions, hidden dynamics, and a fixed unknown model

Let $i \in \mathcal{I}$ be a static unknown candidate index, $x \in \mathcal{X}$ a finite hidden controlled state, and $q \in \mathcal{Q}$ an observed permission state. Candidate i supplies initial law $\mu_i(x)$ and kernel

$$K_i(x', q', y \mid x, q, a), \qquad a \in \mathcal{A}(q).$$

All alphabets, the report set $\mathcal{D}$, and horizon $N \in \mathbb{Z}_{\geq 0}$ are finite. The same nonempty legality map $\mathcal{A}(q)$ applies under every candidate. The controller observes (q', y) after each action, not i or x'. The initial q_0 is fixed and known. Irreversibility is encoded by the permission transitions: lost actions need not return. A bounded stopping rule can instead be padded to N using a shared absorbing stop action. The terminal loss $\ell(i, d) \in [0, 1]$ and target are fixed across comparisons. This section has no running cost; importing the priced objective of Section 8 requires a bound on total loss range or a different accounting argument.

A legal deterministic tree maps each available observable history to an action and its terminal report. Write r_i^j for tree j's risk under candidate i. A tree never selects an action using the true candidate. Perfect recall and private randomization are allowed. There are finitely many legal trees, albeit potentially an enormous number.

Proposition 10.1 (Standard finite-game certificate)**.** *For a distribution w on executable legal trees and a prior λ on $\mathcal{I}$, set*

$$U(w) = \max_i \sum_j w_j r_i^j, \qquad L(\lambda) = \min_j \sum_i \lambda_i r_i^j.$$

Then $L(\lambda) \leq R_N^* \leq U(w)$. The exact minimax value satisfies

$$R_N^* = \min_w U(w) = \max_\lambda L(\lambda).$$

Consequently a matching rational pair $L(\lambda) = U(w)$ certifies optimality. A positive difference gives a certified enclosure, not a numerical declaration of equality.

Proof. Pre-sample the policy's private random choices at every node of its finite history tree. This gives a random deterministic legal tree with the same candidatewise risks. Thus all admissible randomized policies are included in the simplex of mixtures. Any prior-weighted risk is at most that policy's maximum risk; minimizing proves the lower inequality. The displayed mixture is executable, proving the upper inequality. The optimization of U is the finite linear program minimizing u subject to $\sum_j w_j r_i^j \le u$ for every i, $w_j \ge 0$ and $\sum_j w_j = 1$. Its dual maximizes t subject to $t \le \sum_i \lambda_i r_i^j$ for every j, $\lambda_i \ge 0$ and $\sum_i \lambda_i = 1$. Finite linear-program duality gives equality; feasible bounded finite simplices ensure attainment. □

The lower certificate does not require listing all trees. Starting from $b_0(i,x) = \lambda_i \mu_i(x)$, define

$$B_0(b,q) = \min_{d \in \mathcal{D}} \sum_{i,x} b(i,x)\ell(i,d),$$
$$B_r(b,q) = \min_{a \in \mathcal{A}(q)} \sum_{q',y} p_b(q', y \mid q, a) B_{r-1}(b_{a,q',y}, q'),$$

where

$$p_b(q', y \mid q, a) b_{a,q',y}(i,x') = \sum_x b(i,x) K_i(x', q', y \mid x, q, a).$$

Ignore zero-probability branches. Backward induction proves $L(\lambda) = B_N(b_0, q_0)$. The belief retains correlations between target, nuisance, and hidden state, in addition to remaining permissions. This is the established POMDP information-state method [7, 5]; finite-horizon piecewise-linear value methods go back at least to Smallwood and Sondik [12]. We checked the latter's publisher metadata/abstract, not its entire proof. No efficiency claim follows: reachable beliefs and policy trees can grow exponentially.

Two errors that the certificate prevents. First, replacing the static candidate by a new worst-case choice after every observation gives a different, stronger adversarial model. The prior is selected once. Second, a Bayes-optimal tree at a least-favorable prior need not itself be minimax. For one observation with $P_0(Y=0) = 1$ and $P_1(Y=0) = 1/2$, reporting Y has uniform-prior error $1/4$ but worst error $1/2$. At least-favorable prior $\lambda_1 = 2/3$, both decisions after $Y = 0$ are Bayes ties. Their risk vectors are $(0, 1/2)$ and $(1, 0)$, neither minimax. Mixing them with weights $2/3, 1/3$ gives $(1/3, 1/3)$ and the matching certificate $R^* = 1/3$.

Even both separate target and nuisance marginals can be insufficient. On $(M, B) \in \{0,1\}^2$, beliefs $(1/4, 1/4, 1/4, 1/4)$ and $(1/2, 0, 0, 1/2)$ have identical marginals. A noiseless observation of M leaves optimal target error $1/2$ in the first belief and makes it zero in the second. This is a state-sufficiency counterexample; both beliefs are not claimed reachable from the original independent scout-only prior.

10.2 An asymmetric example with a complete small certificate

Use the two-mode irreversible protocol with $N = 2$, but let

$$\mathbb{P}(Z = 1 \mid M = 0) = 1/10, \quad \mathbb{P}(Z = 0 \mid M = 1) = 3/10,$$

and give the correct sensors crossover probabilities $1/10$ in mode zero and $3/10$ in mode one. A wrong-mode sensor remains fair. A scout followed by a target reading, with $x = \mathbb{P}(a = 0 \mid Z = 0)$

and $y = \mathbb{P}(a = 0 \mid Z = 1)$, has bit-symmetric mode risks

$$r_0 = \frac{1}{2} - \frac{2}{5}\left(\frac{9}{10}x + \frac{1}{10}y\right), \qquad r_1 = \frac{1}{2} - \frac{1}{5}\left[\frac{3}{10}(1-x) + \frac{7}{10}(1-y)\right].$$

Choosing $x = 10/21$, $y = 0$ makes both risks $23/70$. Execute this by mixing "choose mode Z" with weight $10/21$ and "always choose mode one" with weight $11/21$, after the scout. This is a feasible upper certificate.

For the lower certificate use prior $(1/14, 1/14, 3/7, 3/7)$ on $(M, B) = (00, 01, 10, 11)$. If the first action commits to a target sensor, the second must use that sensor. The two-read mixture retains the majority sign and $e_2 = e_1$, so prior risks for fixed modes zero and one are $31/70$ and $23/70$. If the first action scouts, the posterior mode-zero mass after $Z = 0$ is $1/3$, making the two modes tied at target risk $11/30$; after $Z = 1$ it is $1/43$, favoring mode one. Averaging yields $23/70$. A second scout without any target reading leaves error $1/2$. These exhaust the legal first-step choices and their optimal continuations, proving the lower bound and hence the exact minimax value $23/70$.

By contrast, the uniform-prior Bayes policy selects mode Z and has average error $1/4$, but worst error $9/25 > 23/70$. The best open-loop policy has minimax risk $11/30$. The latter follows by mixing fixed modes zero and one to equalize $1/2 - (2/5)t$ and $3/10 + (1/5)t$ at $t = 1/3$; independently sampled scouts do not improve the optimal majority decision for a fixed mode. The independently written specialist script also exhaustively enumerates all 176 deterministic legal two-step trees, including every terminal decoder, and verifies both lower certificates. This small result is a regression test for the general method, not a priority claim.

11 An asymmetric six-candidate reproducible study

To avoid simply increasing the size of the symmetric example, let $M \in \{0, 1, 2\}$, $B \in \{0, 1\}$, with six fixed candidates and target B. Scout S has three outputs, with rows indexed by M:

$$P(Z = \cdot \mid M) = \begin{pmatrix} .7 & .2 & .1 \\ .1 & .6 & .3 \\ .2 & .1 & .7 \end{pmatrix}.$$

The three binary target sensors have probabilities of output one as follows. Every printed decimal is an exact rational hundredth or tenth.

Sensor	00	01	10	11	20	21
A_0	.08	.88	.30	.55	.42	.60
A_1	.25	.60	.18	.82	.35	.52
A_2	.33	.65	.40	.58	.12	.75

All outcomes are fresh conditional on (M, B) and actions. Cross-mode sensors are informative, no bit-reflection symmetry is assumed, and the channels cannot be replaced by the good/fair mixture formula. Before commitment the learner may scout or take a target reading. Taking the first target reading commits to its sensor, after which only that sensor remains available. Every observation costs one and total count is fixed at N; commitment has no separate free observation. The tables are invented synthetic laws, not empirical sensor calibration, biomedical validation, or evidence of practical savings.

The *open-loop* comparator randomizes over legal schedules independently of outcomes and may decode using the full observed transcript. Legal schedules are $S^s A_a^{N-s}$, including all scouts. The

reversible comparator may use any sensor or scout at any step and has explicitly greater action rights; it isolates the value of relaxing irreversibility, not an architecture advantage. All classes use the same terminal target, sample count, known laws and private randomization.

11.1 Certificates, executable policies, and negative controls

The solver repeatedly asks a Bayesian dynamic program for a best-response tree to a candidate prior and adds its candidatewise risk vector to a finite-game LP. Floating-point LPs propose priors and mixtures; their active systems are reconstructed over the rationals. An exact Bayes recursion then computes the lower bound and an executable mixture's risk gives the upper bound. If reconstruction fails, rationalized feasible candidates still give an enclosure; equality is never inferred from solver tolerance. The saved policy is a directed acyclic graph of actions, observation branches and terminal reports. It receives observations and its own random mixture choice, never a hidden state label.

All 17 reported cases have *exactly zero rational certificate gap*. The independent verifier reads only the saved model/policy/certificate, checks legal action sequences and horizon, evaluates candidatewise policy risks, and independently recomputes the Bayes lower bound without the optimizer or its oracle. The small asymmetric example supplies a separately enumerated control. The certificate checks are algorithmic verification under supplied kernels, not formal proof-assistant verification.

N	Irreversible adaptive	Open-loop	Reversible adaptive
0	.500000000	.500000000	.500000000
1	.330151462	.330151462	.330151462
2	.268760234	.306356429	.268619242
3	.231259259	.257860485	.224222209
4	.176811713	.235287802	.176125933
5	.157907608	.213936487	not computed

These are rounded exact minimax values for the declared finite instances. At $N = 4$ the irreversible value is

$$\frac{187750191664730450608334 27}{106186512543671256255000000}.$$

Its stored optimal mixture has six trees, with a combined 34-node graph including the separate uniform-prior baseline. All six candidatewise risks equal this fraction. The selected deterministic uniform-prior Bayes rule has worst risk $1133883/5000000 = 0.2267766$. This reports that specific tie-broken rule, not a claim that every possible uniform-Bayes tie resolution has that worst risk. The exact primal/dual certificate is the reason to call the returned mixture minimax.

The adaptive advantage over open-loop at $N = 4$ is about 0.0584761; the extra advantage of reversible rights is only 0.000685779. The latter small positive gap should not be advertised as a robust practical effect. A gap's magnitude, not only its sign, matters when the model is approximate.

12 Model mismatch and the limits of the numerical advantage

Proposition 12.1 (Common-contract finite-horizon coupling bound)**.** *Let K and $\widehat{K}$ have the same candidates, initial laws, finite alphabets, initial permission, legality map, horizon and terminal loss in $[0, 1]$. Suppose for every candidate and legal (x, q, a),*

$$\mathrm{TV}\big(K_i(\cdot \mid x, q, a), \widehat{K}_i(\cdot \mid x, q, a)\big) \le \delta, \qquad 0 \le \delta \le 1,$$

where the compared outcome includes (x', q', y). *Put* $\epsilon_N = 1 - (1-\delta)^N$. *Every fixed policy's candidatewise risks change by at most* ϵ_N, *and*

$$|R^*_N(K) - R^*_N(\widehat{K})| \le \epsilon_N.$$

A policy certified on $\widehat{K}$ *by bounds* $[L, U]$ *has worst risk at most* $U + \epsilon_N$ *on* K *and regret at most* $(U - L) + 2\epsilon_N$. *The true minimax value lies in* $[L - \epsilon_N, U + \epsilon_N]$, *clipped to* $[0, 1]$.

Proof. Couple equal initial latent states and the same private policy randomness. While latent states and complete observed histories agree, the selected actions agree and are legal. Maximal coupling of the next joint outcome preserves equality with probability at least $1 - \delta$. Induction gives probability at least $(1-\delta)^N$ of agreement through the horizon. On that event the report and loss agree; elsewhere their difference is at most one. This proves the policywise risk bound. Maximum and infimum preserve uniform bounds, proving value sensitivity. For the transferred policy, its risk is at most $U + \epsilon_N$, while the true optimum is at least $L - \epsilon_N$; subtract. The bounds concern a fixed perturbed candidate family, not a reselected hidden parameter at every history. □

This is a standard finite coupling/simulation bound. Explicit POMDP model and transferred-policy sensitivity have substantial primary antecedents [2]; recent policy-observation robustness work explicitly separates fixed from history-dependent model deviations [4]. Costed switching [14] and usage-constrained sensing [5] also predate this note. Our result does not replace their richer models or establish a new general robustness theory.

The common-contract condition is necessary. If a first action has identical readout under two systems but irreversibly disables a later informative sensor in only one, readout-only total variation at that step is zero while the future executable policy sets differ. Comparing the joint permission successor exposes that difference; silently changing legality makes policy transfer undefined. Changes in target loss need a separate discrepancy term. The terminal-loss theorem does not retain its constant for unbounded cumulative costs.

For the six-candidate study, suppose *externally justified* common-contract kernel error is at most $\delta = 0.005$. At $N = 4$, $2\epsilon_N = 0.03970099875$. Applying the theorem separately to the two policy classes preserves an adaptive-over-open-loop value gap of at least 0.01877509. By contrast, even $\delta = 0.0001$ gives $2\epsilon_N \approx 0.00079988$, larger than the reversible-rights gap; this certificate is inconclusive about its persistence. This does not prove that the small gap reverses. At $\delta = 0.02$, the bound also becomes inconclusive for the larger adaptivity gap.

We additionally perturb all binary target probabilities by $\pm\delta$, toward or away from ambiguous target responses, and evaluate the *frozen nominal policy* exactly; the scout kernel is unchanged. These are eight structured perturbations per certificate, not optimization over every allowed perturbed family. For irreversible $N = 4$, moving probabilities toward ambiguity raises worst risk from 0.1768117 to 0.1831660 at $\delta = 0.005$, 0.2023407 at 0.02, and 0.2411337 at 0.05. All 136 saved stress checks obey the coupling bound. No observed data establish that any real sensor satisfies these error radii. Obtaining such a bound requires a separately specified calibration experiment and its costs; a deployment learner cannot inspect (M, B) for free.

13 What improved, and what still needs human judgment

The larger contribution is not the mere use of more states. The revision removes the symmetry-based identification of Bayes with minimax, supplies executable policy/lower-bound certificates, proves the required common-interface transfer bound, and records when the apparent advantage is too small for that bound to protect. It retains the simpler exact examples as transparent controls. The

prior-art audit strengthens attribution and may reinforce, rather than improve, the weak evidence for historical novelty.

For human feedback, the useful questions are specific: does the irreversible-permission certificate/benchmark illuminate an existing sequential-design problem; which historical experiment-selection or nuisance-testing literature subsumes the construction; and what physically measured irreversible interface would justify an empirical study? The next mathematical work should treat learned channel laws or a coupled paid-validation contract, with confidence sets and a fixed versus adaptive uncertainty convention stated explicitly. The present methods note supplies neither a scalable general algorithm nor empirical validation, and no theorem about D04's endogenous representation learning follows.

14 Unresolved investigation and portfolio relation

Precisely unresolved problem. The new scout-and-commit family resolves one specific fixed-mode, fresh-noise case, not general constrained stochastic acquisition. The finite permission-state certificate now covers supplied finite hidden controlled models, without providing a closed frontier or scalable algorithm. Unknown channel reliability, statistically learned uncertainty sets, arbitrary unbounded histories, and coupled paid-validation/representation revision remain unresolved. Any state reduction or approximation still requires a certificate.

D02 versus D04. This note holds the evaluation target fixed and asks what an interface can acquire. D04 chooses among representations using task performance and paid distinctions. Its selection theorem is not the same theorem as the channel frontier, although a representation-selection experiment can use this frontier as a lower-bound subproblem. If a future D04 manuscript merely selects among these three channels, it should be merged into this note. The present drafts retain distinct decision variables.

Being, Becoming, and The MIND. The operational content of Being is the predictive capability available under a present experiment budget. Becoming is an alteration of that capability through observations or changes to the interface. A model-based agent can also recognize a task-relevant ambiguity and refrain from claiming identification. These are useful functional targets. They do not establish phenomenal consciousness or a discovered self/world boundary.

Publication assessment. The solved benchmark, explicit timing, and failure cases suit a methods appendix to the acquisition/diagnosis line. The new exact examples strengthen the methods material but do not verify standalone novelty. A separate submission would need a broader priority audit and assessment of whether the restricted family adds a substantive result beyond established controlled sensing. The existing acquisition and diagnosis notes remain the closest internal antecedents.

References

[1] Henrique de Oliveira. Blackwell's informativeness theorem using diagrams. *Games and Economic Behavior*, 109:126–131, 2018. doi: 10.1016/j.geb.2017.12.008. URL `https://henriquedeoliveira.com/wp-content/uploads/2020/12/blackwells-informativeness-theorem-using-diagrams.pdf`. Author preprint dated 30 November 2017; relevant Theorem 1 and Section 4 read. Journal metadata verified against publisher and institutional record. Accessed 13 November 2026.

[2] Yunus Emre Demirci, Ali Devran Kara, and Serdar Yüksel. Sensitivity of filter kernels and robustness bounds to transition and measurement kernel perturbations in partially observable stochastic control, 2025. URL https://arxiv.org/abs/2508.10658v3. v3, 12 September 2025. Latest abstract and selected passages inspected. Earlier v1 Section 6.2 was read separately under its earlier title; no full-proof audit.

[3] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830. URL https://www.cs.rpi.edu/academics/courses/spring06/random/hoefding.pdf.

[4] Benjamin Kraske, Qi Heng Ho, Federico Rossi, Morteza Lahijanian, and Zachary Sunberg. Robustness analysis of POMDP policies to observation perturbations, 2026. URL https://arxiv.org/abs/2604.21256v1. v1, 23 April 2026. Abstract and Sections 2.3 and 3 inspected; fixed/history-dependent uncertainty distinction.

[5] Shang Li, Xiaoou Li, Xiaodong Wang, and Jingchen Liu. Sequential hypothesis test with online usage-constrained sensor selection. *IEEE Transactions on Information Theory*, 65(7):4392–4410, 2019. doi: 10.1109/TIT.2019.2910730. URL https://arxiv.org/abs/1601.06447v1. Read 2016 primary preprint Sections I–III-A, especially equations (13)–(15).

[6] Michael L. Littman, Richard S. Sutton, and Satinder Singh. Predictive representations of state. In *Advances in Neural Information Processing Systems*, volume 14, pages 1555–1561, 2001. URL https://papers.neurips.cc/paper_files/paper/2001/file/1e4d36177d71bbb3558e43af9577d70e-Paper.pdf.

[7] Mohammad Naghshvar and Tara Javidi. Active sequential hypothesis testing. *The Annals of Statistics*, 41(6):2703–2738, 2013. doi: 10.1214/13-AOS1144. URL https://arxiv.org/abs/1203.4626v4. Section 3, Fact 1 read in the primary arXiv reprint; not a full proof audit.

[8] Mohammad Naghshvar and Tara Javidi. Sequentiality and adaptivity gains in active hypothesis testing. *IEEE Journal of Selected Topics in Signal Processing*, 7(5):768–782, 2013. URL https://arxiv.org/abs/1211.2291v1. Read primary 10 November 2012 preprint, especially policy categories in Section II-B.

[9] Sirin Nitinawarat, George K. Atia, and Venugopal V. Veeravalli. Controlled sensing for multihypothesis testing. *IEEE Transactions on Automatic Control*, 58(10):2451–2464, 2013. doi: 10.1109/TAC.2013.2261188. URL https://vvv.ece.illinois.edu/papers/journal/niti-atia-veer-tac-2013.pdf.

[10] Research development draft prepared for Brian Theory. Causal self-diagnosis with noisy tests: Observation laws, irreversible access, and task recovery, 2026. URL https://www.b-theory.com/manuscripts/causal-self-diagnosis-r1-v0.3-2026-11-05.pdf. Version 0.3, 5 November 2026. Original submission authorship unresolved; fixed 38-page reading copy. Accessed 13 November 2026.

[11] Cosma Rohilla Shalizi and James P. Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *Journal of Statistical Physics*, 104:817–879, 2001. doi: 10.1023/A:1010388907793. URL https://arxiv.org/abs/cond-mat/9907176. Read in arXiv:cond-mat/9907176v2; relevant source locators are Definition 5 and Section V.

[12] Richard D. Smallwood and Edward J. Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973. doi: 10.1287/opre.21.5.1071. Publisher metadata and abstract checked; full original proof not re-audited.

[13] Brian Theory. From observable states to acquirable predictive models, 2026. URL `https://www.b-theory.com/manuscripts-html/acquirable-predictive-models/`. Revised technical report, 7 October 2026; site reading copy 1. Accessed 13 November 2026.

[14] Nidhin Koshy Vaidhiyan and Rajesh Sundaresan. Active search with a cost for switching actions, 2015. URL `https://arxiv.org/abs/1505.02358v1`. Read primary v1, 10 May 2015; abstract, introduction and model context. Asymptotic proofs not audited.

[15] George Vershinin, Asaf Cohen, and Omer Gurewitz. On cost-aware designs for sequential hypothesis testing, 2026. URL `https://arxiv.org/abs/2512.19067v3`. arXiv:2512.19067v3, 24 September 2026. Abstract, introduction and model scope inspected; proofs not audited.

[16] Sushant Vijayan and Himadri Mandal. PDE-based bounds on optimal exponents in active simple hypothesis testing, 2026. URL `https://arxiv.org/abs/2610.36657v1`. arXiv:2610.36657v1, 27 October 2026. Abstract and Sections I–II inspected by the independent reviewer; proof audit not claimed.