When Checking Takes Time: An Exact Two-Probe Stopping Frontier

Technical addendum to B-Theory's self-modeling studies
Version 1.0 | 9 November 2026 | Local release candidate

ABSTRACT

We solve a finite Bayesian prediction problem in which each diagnostic probe reads a hidden binary mechanism and then permits a random transition before delivering its reading. A point forecast may subsequently be disclosed to the mechanism, which responds through a specified COPY/NOT evaluator. For initial fault probability and observation error both 1/4, at most two probes, transition probability h in [0,1/2], disclosure probability lambda in [0,1], and probe cost c >= 0, we obtain the exact minimum expected prediction error plus probe cost. A branch-dominance proof reduces all policies to immediate stopping or a single contingent two-probe plan. The latter can improve the objective even though its first probe alone makes prediction worse. We give exact examples of beneficial contingent diagnosis and harmful mandatory continuation, and distinguish executing a check from having the option to check. This is a worked specialization of classical Bayesian filtering and optimal stopping, with a reproducible finite certificate. Historical priority, architectural advantage, and consciousness conclusions are not claimed.

1. PURPOSE AND RELATION TO PRIOR WORK

This addendum isolates one small result: the complete parameter frontier for the observation, transition, and disclosure contract defined below. The contribution is its explicit solution and proof certificate. Neither costly information acquisition nor planning beyond the immediate value of a measurement is proposed as new.

Monahan [1] studies stopping in a hidden Markov system with costly information and potentially nonmonotone policies. This establishes a direct classical antecedent at the problem-class level. Only the publisher abstract was accessible in this review; no theorem-by-theorem exclusion of overlap with its full text is asserted.

Krishnamurthy [2, Section II, equations (10) and (17)] treats noisy Markov-chain sampling with controllable observation intervals through Bayesian filtering and a Bellman equation. The displayed filter propagates the state and then incorporates the next measurement. Here a probe measures first and delivers that measurement after a transition; stopping forecasts a binary response rather than declaring a target-state arrival. These distinctions specify the present calculation, not a new alternative to partially observable decision theory.

Ahmad and Yu [3] optimize sensing with response-error, temporal, and sensor-switching costs and compare with one-step criteria. In a distinct subsequent paper, Ahmad, Huang, and Yu [7, Section 2.6] explain how a one-step approximation can stop too readily. These are antecedents for the importance of a contingent sensing plan. Bayesian sequential change diagnosis [4] likewise supplies established inference and stopping machinery for changing latent regimes. Our symmetric reversible transitions and two-probe terminal objective are a deliberately small different fixture.

Within B-Theory, the reflective self-model manuscript [5] already distinguishes physical mechanism from represented model, propagates known hazards, and compares explicit with compiled implementations. The acquisition manuscript [6] treats the availability and cost of predictive evidence. This addendum adds a solved case in which every executed probe exposes the mechanism to another transition before its evidence can be used. It does not re-present the earlier fixed-point or leakage results as contributions. A separate integration

note identifies overlap with the self-diagnosis draft without changing that author's files.

2. MODEL AND INFORMATION CONTRACT

The hidden physical mechanism is indexed by theta in {0,1}. At a fixed task input 0, theta=0 denotes COPY and theta=1 denotes NOT. The prior is Pr(theta=1)=1/4. A policy knows all model parameters but does not observe theta, noise, or transitions directly. At each decision it can stop or execute a probe, with at most two probes total. There are no other observations, free waiting actions, repairs, or interventions.

Each executed probe has this order: (i) read Y = theta XOR Z, with independent Pr(Z=1)=1/4; (ii) during the probe interval, replace theta by theta XOR D, where Pr(D=1)=h; (iii) deliver Y and choose whether to stop or probe again. Every executed probe uses fresh independent Z and D. Thus the read itself is non-mutating, but the complete probe consumes a transition interval. A second probe reads the then-current mechanism. We assume 0 <= h <= 1/2. No extra transition or actuation delay follows stopping. Immediate stopping incurs no probe transition. Outcomes under different stopping rules therefore concern their respective stopping times, not a common fixed deadline.

At stopping the policy selects a private point forecast P in {0,1} from its observation history and any independent private randomization. An independent E with Pr(E=1)=lambda is then drawn. The evaluator sees P only if E=1 and produces the action

```
A = theta XOR (E P). (1)
```

Here 0 <= lambda <= 1. With E=0 the input is the fixed task bit 0; with E=1 the input is P. This is a stipulated response mechanism, not an unrestricted adversarial responder. The objective is

```
R = Pr(P != A) + c E[N], c >= 0, N <= 2. (2)
```

All policies share this prior, noise law, evaluator, timing and budget. The formula covers ideal probability laws for real parameters. It does not assert that arbitrary real probabilities can be sampled with a bounded number of fair bits. Worked examples below use dyadic probabilities.

3. FILTERING AND TERMINAL LOSS

Write b for the conditional probability that the current theta equals 1. For a probe observation y, the reading-time posterior B_y(b) and its delivered-time propagation are

```
B_0(b) = b/(3-2b), B_1(b) = 3b/(1+2b),
T_h(b) = h + (1-2h)b, b_next = T_h(B_y(b)). (3)
```

The probability of y=1 is (1+2b)/4. Equation (3) follows directly by Bayes' rule and the independent symmetric flip. Updating a fixed-mechanism likelihood without T_h would estimate the wrong time-indexed variable.

Lemma 1 (terminal risk). At belief b the minimum conditional prediction error is

```
r_lambda(b) = (1-lambda) min(b,1-b) + lambda b. (4)
```

Proof. In the sealed branch A=theta, so a maximum-posterior forecast minimizes error. In the disclosed branch P != (theta XOR P) exactly when theta=1, regardless of P. Averaging the branches proves (4). For lambda<1 a maximum-posterior forecast is optimal; at lambda=1 every forecast ties. We choose forecast 0 when b=1/2. Randomized forecasts cannot improve this minimum. End proof.

For reference the ordinary finite-horizon Bellman recursion is

```
V_0(b) = r_lambda(b),
V_n(b) = min{r_lambda(b), c + sum_y Pr(y|b) V_(n-1)(T_h(B_y(b)))}. (5)
```

The theorem below evaluates V_2(1/4) explicitly. Equation (5) is classical dynamic programming, not a separate novelty claim.

4. EXACT FRONTIER

Theorem 1. Under precisely the contract in Section 2, for every h in [0,1/2], lambda in [0,1], and c >= 0,

```
R* = 1/4 - max{0, 3(1-lambda)(1-2h)/32 - h/2 - 11c/8}. (6)
```

One optimal policy stops immediately with P=0 unless the expression inside the maximum is positive. When positive, it probes once, stops with P=0 after Y_1=0, and after Y_1=1 probes again and forecasts the second reading. Strict improvement occurs exactly when

```
c < [3(1-lambda) - (22-6lambda)h]/44. (7)
```

At equality immediate stopping and the contingent policy tie. Other redundant policies may tie at degenerate endpoints. The statement specifies an optimal policy, not a unique optimizer.

Proof. The first reading has probabilities and delivered-time posteriors

```
Pr(Y_1=0)=5/8, b_0=1/10+4h/5;
Pr(Y_1=1)=3/8, b_1=1/2. (8)
```

Both beliefs are at most 1/2. Stopping after exactly one probe therefore has error E[b_Y]=1/4+h/2, independently of lambda. With its cost c it cannot improve on immediate risk 1/4.

We next show that continuation after Y_1=0 is never helpful. For any b <= 1/2 the expected reading-time Bayes error after a binary noisy probe with error 1/4 is min(b,1/4). To verify this identity, sum the smaller of the two joint state/observation masses for each observation: the y=0 term is b/4 and the y=1 term is min(3b/4,(1-b)/4). Their sum equals b for b <= 1/4 and 1/4 otherwise. Because h <= 1/2,

```
min(T_h(u),1-T_h(u)) = h+(1-2h)min(u,1-u). (9)
```

Thus one further probe has sealed expected error h+(1-2h)min(b,1/4) and disclosed expected error T_h(b). Apply this at b=b_0. If b_0 <= 1/4 (h <= 3/16), both terms equal T_h(b_0) >= b_0. If b_0 >= 1/4, the sealed term minus b_0 is 3/20-3h/10 >= 0, and the disclosed term minus b_0 is h(1-2b_0) >= 0. Their weighted average cannot improve on stopping risk b_0; adding c preserves dominance. Hence continuation after 0 can be removed from any policy without increasing its objective.

After Y_1=1 the belief is 1/2. The next reading gives posteriors 1/4 and 3/4 with equal probabilities, then T_h propagates them. Its sealed expected error is 1/4+h/2; its disclosed expected error is 1/2. Compared with stopping error 1/2, the conditional improvement before price is (1-lambda)(1/4-h/2).

The only possibly useful nontrivial tree therefore probes once and repeats only after 1. It has expected probe count 1+3/8=11/8 and gross error

```
R_A = 1/4+h/2 - (3/32)(1-lambda)(1-2h). (10)
```

There are five observation-ordered stopping patterns before assigning terminal forecasts: immediate stop, one probe, continuation only after 0, continuation only after 1, and continuation after both. The dominance arguments remove all except immediate stopping and the last candidate in (10). Conditional minimization handles all terminal forecasts. Any randomized history policy is a mixture of deterministic contingent plans with independent random seeds and cannot outperform the best plan. Comparing R_A+11c/8 with 1/4 proves (6), and rearrangement gives (7). If the contingent tree improves at the root, its second probe after 1 is also locally beneficial: with g=(1-lambda)(1/4-h/2), root gain is (3/8)g-h/2-11c/8 > 0, which implies g>c. Its terminal rule is valid for h<1/2; at h=1/2 choose the stated tie convention. End proof.

Endpoints. At h=0, (6) reduces to 1/4-max{0,3(1-lambda)/32-11c/8}. At h=1/2 every delivered reading concerns a state fully randomized by the last transition, so immediate stopping is optimal. At lambda=1 point forecasts cannot alter disclosed error; (6) again selects immediate stopping. With c=0, positive benefit requires h<3(1-lambda)/(22-6lambda); for lambda=1 this condition is empty. The zero-cost critical h is 3/22 when sealed and 9/82 when lambda=1/4. Negative right sides in (7) admit no nonnegative cost.

5. BENEFIT AND HARM IN EXACT EXAMPLES

Take h=1/16, lambda=1/4, c=1/128. Immediate prediction has error and total risk 1/4. A single mandatory probe has gross error 9/32, already worse before price. The contingent plan has gross error 225/1024, expected probe cost 11/1024, and total risk 59/256. Its net improvement is 5/256. A policy evaluating the first probe only by its immediate improvement declines it and misses this positive value. The example illustrates the theorem; it is not evidence that nonmyopic sensing is historically new.

The contingent plan's terminal current-state beliefs are 3/20 after history 0, 9/32 after 10, and 23/32 after 11. It forecasts 1 only after 11. If a second probe is instead forced after history 0, the current error 3/20 rises to T_(1/16)(3/20)=31/160, an increase of 7/160 even at zero price. The second reading cannot reverse the maximum-posterior decision on that branch, and the added transition interval worsens expected accuracy. This comparison holds for every lambda because all terminal forecasts remain 0.

For an unconditional example, take h=1/4, lambda=0, c=0 and require exactly two probes. After the first delivered reading the beliefs are 3/10 and 1/2. From either belief the optimal second-probe error is 1/4+h/2=3/8, whereas immediate stopping has error 1/4. These are comparisons of mandatory execution with stopping earlier. Offering an optional probe cannot worsen the optimized objective: the larger policy class contains immediate stopping. Nor does the example show that additional evidence alone is harmful about a fixed target; the target mechanism evolves during the executed probe.

6. INTERPRETATION AND LIMITS

The result is about a finite prediction interface. Learning the mechanism, selecting a point forecast, and predicting the distribution of the evaluator's eventual action are distinct tasks. Conditional on terminal belief b and chosen P, before E is drawn the action law is Pr(A=1|history,P)=b+lambda P(1-2b). A correct probability report is not certainty about a realized action. If such a law is reported, it must use a separate field not read by this evaluator; a reactive probability port would be a different problem.

An explicit Bayesian controller and a compiled controller selecting the same tree and leaf forecasts agree on every coupled outcome. This does not measure implementation costs or exclude advantages under different resource constraints, but it establishes no advantage for naming an internal representation a self-model. The mechanism can equally well be interpreted as an external device. No claim about subjective awareness follows.

Two probes, prior and observation error 1/4, known symmetric independent transitions, fixed transition exposure per executed probe, and no final delay are essential to the stated formula. Unknown hazards, asymmetric transitions, more probes, repair actions, different disclosure channels, or a fixed common deadline require fresh analysis. The present solution does not establish an optimal policy structure for those problems. In particular it cannot use the attenuation calculation for a single stopping-independent late shock: here stopping controls how many transitions occur.

7. REPRODUCIBLE CERTIFICATE AND STATUS

The accompanying standard-library Python program enumerates every deterministic depth-two contingent policy. If $S_0=\{0,1\}$ is the set of terminal forecasts, recursively $S_n=\{0,1\} \cup (S_{n-1} \times S_{n-1})$. Hence $|S_1|=6$ and $|S_2|=38$. Each pair branches on the delivered observation. The 64 elementary outcomes are the six bits (initial theta, first noise, first flip, second noise, second flip, disclosure). They are independent under the declared product law. At endpoints some outcomes have zero mass. A probe's noise and flip are applied only when that probe is executed; unused latent bits are summed out and never observed. This proves coverage for the declared finite policy class. Independent randomized policies cannot improve by the mixture argument above.

Exact rational comparisons corroborate (6) at 160 parameter/cost cases, including endpoints and boundary ties. The continuum proof is the argument in Section 4, not the finite test grid. Fresh proof and hostile-review reports, their dispositions, source-reading scopes, and reproduction commands accompany this manuscript. These are internal AI reviews, not human peer review. The release recommendation is for a modest technical addendum with classical attribution; no exhaustive historical-priority claim or physical validation is made.

REFERENCES

[1] George E. Monahan (1980). Optimal Stopping in a Partially Observable Markov Process with Costly Information. Operations Research 28(6), 1319-1334. https://doi.org/10.1287/opre.28.6.1319 . Scope inspected: publisher abstract only; full-text theorem overlap unresolved.

[2] Vikram Krishnamurthy (2012). When to look at a noisy Markov chain in sequential decision making if measurements are costly? arXiv:1208.3291v1. https://arxiv.org/abs/1208.3291 . Scope inspected: Sections I-III, especially II-A equation (10), II-B equation (17), and the assumptions of Theorem 1.

[3] Sheeraz Ahmad and Angela J. Yu (2013). Active Sensing as Bayes-Optimal Sequential Decision Making. UAI 2013, 15-24; PMLR R11. https://proceedings.mlr.press/r11/ahmad13a.html . Scope inspected: publisher abstract and bibliographic record only. The linked PDF returned an unsupported-content-type error; no 2013 full-text comparison is claimed.

[4] Savas Dayanik, Christian Goulding, and H. Vincent Poor (2008). Bayesian Sequential Change Diagnosis. Mathematics of Operations Research 33(2), 475-496. https://arxiv.org/abs/0710.4847 . Scope inspected in the underlying research: model and posterior/stopping formulation, Sections 2-3.

[5] B-Theory. When Does an Explicit Self-Model Help? Reflective self-model manuscript (R4), Sections 3-4. https://www.b-theory.com/manuscripts-html/reflective-self-models/ . Local research snapshot accessed 4 November 2026; current-site refresh failed in the web reader on 9 November 2026.

[6] B-Theory. From Observable States to Acquirable Predictive Models. https://www.b-theory.com/manuscripts-html/acquirable-predictive-models/ . Local research snapshot accessed 4 November 2026; overlap mapping appears in the source ledger.

[7] Sheeraz Ahmad, He Huang, and Angela J. Yu (2014). Cost-sensitive Bayesian control policy in human active sensing. Frontiers in Human Neuroscience 8:955. https://doi.org/10.3389/fnhum.2014.00955 . Open full text: https://pmc.ncbi.nlm.nih.gov/articles/PMC4253738/ . Scope inspected: Sections 2.2 and 2.6, equations (3)-(7), and adjoining Section 2.7 theorem statement. This is a distinct publication from [3].