WHEN IS CALIBRATED CONFIDENCE SUFFICIENT FOR OPTIMAL STOPPING IN SEQUENTIAL PREDICTION?

Anonymous authors
Paper under double-blind review

ABSTRACT

Sequential predictors use confidence to decide whether to answer, abstain, or spend more computation. Calibration at each fixed time is insufficient for this task: it pools histories with the same score even when they differ in current error risk or future informativeness. We separate these two roles. For the terminal answer/abstain choice, the relevant quantity is the probability that the current answer is correct conditional on the observed history. We study scores equal to this conditional probability, the binary form of an ideal forecast relative to the available information. This condition holds at every fixed time exactly when it holds at every bounded stopping time, and it characterizes calibrated scores that support Bayes-optimal terminal decisions for every reject cost. For continuation, however, even an ideal confidence forecast need not be a sufficient statistic for control: equal scores can require opposite actions and leave a constant-order risk gap. A Bellman recursion based only on confidence is valid uniformly over bounded horizons, costs, and terminal losses exactly when the next-score law depends on history only through the current score. Calibration at each fixed time can also leave a constant-order gap. If the score is uniformly within ε of the history-conditional probability, optimizing the surrogate stopping problem on the same filtration incurs at most $2c_{\mathrm{err}}\varepsilon$ regret; a tie-free construction makes the constant tight. Numerical examples recover the theoretical gaps. A real-data diagnostic reveals residual history dependence but no decision-cost gap under the tested costs.

1 INTRODUCTION

Early-exit classifiers, adaptive-computation models, and reasoning agents produce a sequence of candidate answers, often with a confidence score at each step (Graves, 2016; Teerapittayanon et al., 2016; Kaya et al., 2019; Schwartz et al., 2020; Xin et al., 2020; Elbayad et al., 2020; Kadavath et al., 2022; Kuhn et al., 2023). When that score determines whether the system answers, abstains, or computes for another step, it becomes part of the decision rule rather than merely a summary of predictive performance.

We study a finite-label system that proposes an answer $\hat{Y}_t$ to a target Y and reports confidence p_t after observing the history available at time t. Each additional step costs c_{step}, a wrong answer costs c_{err}, and abstention costs c_{abs}. The controller must decide both *when to stop* and *what to do after stopping*. These decisions require different information. The terminal answer/abstain choice depends on the probability that the current answer is correct. The choice to continue also depends on what the next computation is likely to reveal. In sequential-control terminology, p_t is a sufficient statistic of the observed history for this stopping problem only when both quantities depend on the history through p_t.

A simple example shows why the distinction matters. Let $c_{\mathrm{err}} = 1$, $c_{\mathrm{step}} = 0.10$, and disallow abstention. Suppose two histories both have a current correctness probability of $p_1 = 0.75$, so stopping incurs expected error cost 0.25. On the first history, the next step reveals the label, and continuing costs only 0.10. On the second, the next step reveals nothing, and continuing costs 0.35. Thus the same perfectly valid current confidence leads to opposite optimal actions. Knowing the cost of stopping does not determine the value of waiting.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Ordinary calibration does not even guarantee the first quantity. It asks that predictions assigned score p be correct a p fraction of the time (Dawid, 1982; 1985); applying this identity separately at each fixed time does not strengthen its conditioning information. Histories pooled at the same score can still have different conditional probabilities of correctness. For example, two histories may both report 0.75 while their actual probabilities of correctness are 0.60 and 0.90. Pooling is harmless only when the omitted history carries no information relevant to the decision.

This observation leads to the stronger identity $p_t = \mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t)$, where $\mathcal{F}_t$ denotes everything available to the controller by time t. In forecasting terminology, this makes p_t an ideal forecast of the correctness event relative to the information set $\mathcal{F}_t$ (Gneiting & Ranjan, 2013). It is a correct-specification condition for a conditional probability, not ordinary score-level calibration and not a distribution-free guarantee supplied by recalibration. For calibrated scores, this identity characterizes Bayes-optimal terminal decisions across all reject costs. It still does not generally support the decision to continue.

The paper follows this distinction from terminal choice to sequential control. First we characterize the conditional target required when the stopping time itself depends on the history. We then show that even a score equal to this conditional probability can leave an order-c_{err} gap when treated as a sufficient statistic of history, and identify the additional Markov-kernel condition that closes this gap. Finally we quantify what can go wrong under calibration at each fixed time and how uniform errors relative to the history-conditional probability affect the full-information problem.

Contributions.

- **Terminal decision.** We show that confidence equals the history-conditional probability of correctness at every fixed time if and only if the same equality holds at every bounded stopping time. More substantively, a score supports Bayes-optimal answer/abstain decisions for every reject cost exactly when it determines the full-history correctness probability; ordinary score calibration then forces the two to be equal (Theorems 1 and 2).
- **Continuation decision.** We construct a process whose confidence is ideal relative to the history but for which confidence-only policies lose $\Theta(c_{\mathrm{err}})$. We then prove that stopping values factor through p_t for every bounded horizon, terminal loss, and positive step cost exactly when the next-score law depends on history only through the current score (Theorems 4 and 5 and Proposition 2). We also give concrete binary and multiclass sufficient models.
- **Weaker and approximate reports.** We show that confidence calibration at each fixed time can leave a $\Theta(c_{\mathrm{err}})$ gap between report-based and full-information policies. If p_t is uniformly within ε of the history-conditional probability, the full-information surrogate policy has regret at most $2c_{\mathrm{err}}\varepsilon$; tie-free constructions make the constant tight (Theorems 7 and 8 and Propositions 6 and 8).

The reject rule and fixed-coverage ranking are classical (Chow, 1970; El-Yaniv & Wiener, 2010), as are the Snell recursion and Markov dynamic programming given a kernel (Snell, 1952; Chow et al., 1971; Peskir & Shiryaev, 2006; Åström, 1965; Smallwood & Sondik, 1973; Bertsekas, 2017). Theorems 1, 3(b,c), and 6, along with Proposition 1, organize these standard facts. The paper-specific results are Theorem 2's stopping-time converse, the separations in Theorems 4 and 7, Theorem 5's characterization for reported confidence, and the tight bound in Theorem 8. Section 4 gives the detailed positioning.

2 Setup: Confidence Calibration and Ideal Forecasts

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let X denote an input query and let $Y \in \mathcal{Y}$ be a ground-truth label, where $\mathcal{Y}$ is finite. A sequential prediction system reveals signals $Z_1, Z_2, \ldots$ and generates the natural filtration

$$\mathcal{F}_t := \sigma(X, Z_1, \ldots, Z_t), \qquad \mathcal{F}_0 := \sigma(X).$$

At each time $t \geq 1$, the system outputs a candidate answer $\hat{Y}_t \in \mathcal{Y}$ and a confidence $p_t \in [0, 1]$, both $\mathcal{F}_t$-measurable. Write $\mathcal{G}_t := \sigma((\hat{Y}_s, p_s) : 1 \leq s \leq t)$ for the filtration generated by the reported outputs; generally $\mathcal{G}_t \subseteq \mathcal{F}_t$, and the inclusion can be strict. Throughout, $\mathbb{P}(A \mid Z)$ is shorthand for $\mathbb{E}[\mathbf{1}_A \mid \sigma(Z)]$.

A stopping time τ takes values in $\{1, 2, \ldots\}$ and is *bounded* if $\tau \leq T$ almost surely for some deterministic $T < \infty$. For a stopping time τ, we write

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for every } t \geq 0\}.$$

When τ is bounded, this is equivalent to requiring $A \cap \{\tau = t\} \in \mathcal{F}_t$ for every t. At time τ, the controller chooses $a_\tau \in \{\text{ANSWER}, \text{ABSTAIN}\}$, with total cost

$$\mathcal{C}(\tau, a_\tau) = c_{\text{step}}\tau + c_{\text{err}}\mathbf{1}_{\{a_\tau = \text{ANSWER}\}}\mathbf{1}_{\{Y \neq \hat{Y}_\tau\}} + c_{\text{abs}}\mathbf{1}_{\{a_\tau = \text{ABSTAIN}\}}, \tag{1}$$

For a policy $\pi = (\tau, a_\tau)$, we also write $\mathcal{C}(\pi) := \mathcal{C}(\tau, a_\tau)$. Here τ counts realized sequential prediction steps, including the mandatory first step: the system must reveal $(\hat{Y}_1, p_1)$ before any stopping choice is available, so c_{step} is always incurred once. In the dynamic program, costs are evaluated relative to the current time t, yielding the increment $c_{\text{step}}(\tau - t)$. We assume $c_{\text{step}} > 0$, $c_{\text{err}} > 0$, and $c_{\text{abs}} \in (0, c_{\text{err}}) \cup \{+\infty\}$. The case $c_{\text{abs}} = +\infty$ is the no-abstention variant. When the conditional probability that the current answer is correct equals p, the Bayes-optimal terminal loss is

$$g(p) := \min\{c_{\text{err}}(1-p), c_{\text{abs}}\}. \tag{2}$$

When $c_{\text{abs}} < \infty$, we define the boundary threshold $p^\dagger := 1 - \frac{c_{\text{abs}}}{c_{\text{err}}}$. At $p = p^\dagger$, answering, abstaining, and every mixture have the same conditional Bayes loss, so any deterministic or randomized tie rule is Bayes-optimal. For definiteness we use the canonical measurable selector

$$a^\dagger(p) := \begin{cases} \text{ANSWER}, & c_{\text{err}}(1-p) \leq c_{\text{abs}}, \\ \text{ABSTAIN}, & c_{\text{err}}(1-p) > c_{\text{abs}}. \end{cases} \tag{3}$$

Thus $a^\dagger(p) = \text{ANSWER}$ iff $p \geq p^\dagger$, with answer chosen on the boundary. None of the regret bounds below depends on this choice (Remark 1 and Section A.12).

Definition 1 (Calibration and ideal confidence)**.** *For a predictor $(\hat{Y}_t, p_t)_{t \geq 1}$:*

(i) ***Confidence calibration at each fixed time*** *holds if, for every fixed $t \geq 1$,*

$$\mathbb{P}(Y = \hat{Y}_t \mid p_t) = p_t \qquad \textit{a.s.}$$

(ii) ***Confidence calibration at a stopping time*** *τ holds if*

$$\mathbb{P}(Y = \hat{Y}_\tau \mid p_\tau) = p_\tau \qquad \textit{a.s.}$$

(iii) ***Ideal confidence relative to the history*** *holds if, for every $t \geq 1$,*

$$\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t) = p_t \qquad \textit{a.s.} \tag{4}$$

Condition (iii) is the binary specialization of an ideal forecast relative to an information set (Gneiting & Ranjan, 2013). It is a correct-specification condition: the reported score must equal $\mathbb{E}[\mathbf{1}_{\{Y=\hat{Y}_t\}} \mid \mathcal{F}_t]$, not merely its average given the score. Ordinary confidence calibration conditions on $\sigma(p_t)$; condition (iii) conditions on all observed information in $\mathcal{F}_t$. Below, "ideal confidence" is shorthand for condition (iii). Proposition 9 makes the distinction exact: condition (iii) is confidence calibration at each fixed time plus $\sigma(p_t)$-measurability of the conditional probability of correctness. Next-token calibration of a generative model does not imply this answer-level identity.

The strength of the condition depends on the prediction problem. For binary classification, the pair $(\hat{Y}_t, p_t)$ determines the entire conditional label distribution: p_t is the probability of the announced label and $1 - p_t$ the probability of the other label. With three or more labels, condition (iii) fixes only the probability assigned to the announced answer; it does not determine how the remaining probability is distributed. This distinction becomes important for continuation in Proposition 4.

3 Confidence Calibration and Optimal Stopping

We begin with the terminal decision and ask what information a score must contain for the conditional cost of stopping to be determined, even when the stopping time depends on the observed history. We then turn to continuation and ask when that same score contains enough information to value another computation step. The last two parts show what ordinary fixed-time confidence calibration can miss and how errors in the history-conditional probability translate into regret.

3.1 Calibration at History-Dependent Stopping Times

The first result says that no additional condition appears merely because the evaluation time is chosen adaptively: equality with the history-conditional probability at every fixed time is equivalent to the same equality at every bounded stopping time. This elementary stopping-time identity connects the forecasting condition to the terminal decision below.

Theorem 1 (Characterization by bounded stopping times)**.** *Let $(\hat{Y}_t, p_t)_{t\geq 1}$ be adapted to $(\mathcal{F}_t)_{t\geq 0}$. The following are equivalent:*

(a) for every $t \geq 1$, $\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t) = p_t$ almost surely (equivalently, p_t is a version of $\mathbb{E}[\mathbf{1}_{\{Y=\hat{Y}_t\}} \mid \mathcal{F}_t]$);[1]

(b) for every bounded stopping time τ, $\mathbb{P}(Y = \hat{Y}_\tau \mid \mathcal{F}_\tau) = p_\tau$ almost surely.

Moreover, condition (a) implies confidence calibration at every fixed time and at every bounded stopping time τ.

Proof sketch. The stopped identity decomposes over $\{\tau = t\}$. The reverse direction tests deterministic times. Both score-conditional calibration statements follow by the tower property. The ideal-forecast condition is equivalent to fixed-time confidence calibration plus $\sigma(p_t)$-measurability of the conditional probability of correctness (Proposition 9). Details: Section A.1.

For a given stopping time, requiring Bayes-optimal answer-versus-abstain decisions for *every* reject cost is much stronger than ordinary calibration: the reported score must determine the full-information conditional probability of correctness itself. The thresholding direction is classical (Chow, 1970); the converse, and the equality clause under calibration, are the content here. Fixed-time relatives include threshold and decision calibration (Sahoo et al., 2021; Zhao et al., 2021; Franc et al., 2023; Corvelo Benz & Gomez Rodriguez, 2023).

Theorem 2 (When a score suffices for every reject cost)**.** *Fix a bounded stopping time τ and define the conditional probability that the current answer is correct given all information available at stopping,*

$$r_\tau := \mathbb{P}(Y = \hat{Y}_\tau \mid \mathcal{F}_\tau).$$

Let $s_\tau \in [0,1]$ be any $\mathcal{F}_\tau$-measurable score. For $\lambda \in [0,1]$, consider the normalized answer-vs-abstain loss (corresponding to reject cost $c_{\mathrm{abs}}/c_{\mathrm{err}} = 1-\lambda$):

$$\ell_\lambda(a) := \mathbf{1}_{\{a=\textsc{answer}\}}\mathbf{1}_{\{Y\neq\hat{Y}_\tau\}} + (1-\lambda)\mathbf{1}_{\{a=\textsc{abstain}\}}.$$

The following are equivalent:

(a) for every rational $\lambda \in [0,1]$ (a countable threshold class suffices), there exists a $\sigma(s_\tau)$-measurable answer/abstain rule that is Bayes-optimal for ℓ_λ among all $\mathcal{F}_\tau$-measurable rules;

(b) r_τ admits a $\sigma(s_\tau)$-measurable version.

If, in addition, $\mathbb{P}(Y = \hat{Y}_\tau \mid s_\tau) = s_\tau$ almost surely, then these conditions imply $s_\tau = r_\tau$ almost surely. Rational thresholds are used only to obtain one common null set; once r_τ is $\sigma(s_\tau)$-measurable, Bayes score-based rules exist for every real reject cost.

Proof sketch. Optimal answer sets for threshold λ are the superlevel sets of r_τ. Rational λ are countable and dense, so one null set suffices and the cuts generate $\sigma(r_\tau)$. Calibration plus the tower property then forces $s_\tau = r_\tau$. Recoverability is weaker than equality: $s_\tau = r_\tau^2$ still supports every reject cost (Section A.15). Full proof: Section A.2.

Once the stopping rule is chosen, an ideal confidence score can be thresholded exactly as in selective prediction.

Proposition 1 (Selective prediction at a stopping time)**.** *Assume condition (iii) of Definition 1 and let $\tau \leq T$ be a bounded stopping time. Among all $\mathcal{F}_\tau$-measurable acceptance probabilities $\alpha_\tau : \Omega \to [0,1]$ with $\mathbb{E}[\alpha_\tau] = \gamma > 0$, the conditional error*

$$\frac{\mathbb{E}[(1-p_\tau)\alpha_\tau]}{\gamma}$$

[1] Conditional expectations are defined up to almost-sure equality. A version is one representative of that class.

is minimized by thresholding p_τ:

$$\alpha_\tau^\star = \mathbf{1}_{\{p_\tau > \theta\}} + \rho\, \mathbf{1}_{\{p_\tau = \theta\}}$$

for some $\theta \in [0,1]$ and $\rho \in [0,1]$ chosen so that $\mathbb{E}[\alpha_\tau^\star] = \gamma$. Here α_τ is the probability, given the history at τ, of answering rather than abstaining; γ is coverage; and the displayed ratio is the error probability among accepted instances.

Proof sketch. This is the classical Neyman–Pearson thresholding of p_τ at fixed coverage (Chow, 1970; El-Yaniv & Wiener, 2010); Section A.3.

3.2 Why the Current Score Need Not Value Another Step

Ideal confidence settles the loss of stopping given the observed history. It says nothing about how informative the next step will be, and therefore does not by itself make confidence a sufficient statistic for the dynamic program.

Theorem 3 (Full-information optimal stopping with ideal confidence)**.** *Assume a finite horizon $T < \infty$ and condition (iii) of Definition 1. Then:*

(a) for every stopping time τ with $1 \le \tau \le T$, the minimum conditional Bayes loss at time τ is $g(p_\tau)$ and is attained by the canonical selector $a^\dagger(p_\tau)$; when $c_{\mathrm{abs}} = +\infty$ this selector always answers, and when $c_{\mathrm{abs}} < \infty$ it answers iff $p_\tau \ge p^\dagger$, with abstention also Bayes-optimal on the boundary $p_\tau = p^\dagger$;

(b) if we define

$$L_t := g(p_t), \qquad \mathcal{T}_t := \{\tau : \tau \text{ is a stopping time and } t \le \tau \le T\},$$

together with the recursion

$$J_T := L_T, \qquad J_t := \min\{L_t,\ c_{\mathrm{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\} \quad \text{for } t = T-1, \dots, 1,$$

then

$$J_t = \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_t} \mathbb{E}[c_{\mathrm{step}}(\tau - t) + L_\tau \mid \mathcal{F}_t] \qquad \text{a.s.}$$

for every t;[2]

(c) the stopping time

$$\tau_t^* := \inf\{s \in \{t, \dots, T\} : J_s = L_s\}$$

is optimal for the continuation problem started at time t. In particular, τ_1^* together with the terminal action rule from part (a) is Bayes-optimal for the original finite-horizon problem.

Proof sketch. Part (a) is the stopped identity from Theorem 1. Parts (b)–(c) are the discrete-time Snell envelope (Snell, 1952; Chow et al., 1971; Peskir & Shiryaev, 2006). The unconditional Bayes risk is $c_{\mathrm{step}} + \mathbb{E}[J_1]$. Details: Section A.4.

Theorem 4 (Ideal confidence is not sufficient for optimal stopping)**.** *By cost rescaling, normalize to $c_{\mathrm{err}} = 1$, fix $c_{\mathrm{abs}} = +\infty$, and let $c_{\mathrm{step}} \in (0, 1/4)$. There exists a two-step sequential prediction problem whose confidence is ideal relative to the history and such that:*

(a) $p_1 = 3/4$ almost surely;
(b) the Bayes-optimal decision at time 1 is to continue on one $\mathcal{F}_1$-event and to stop on another; and
(c) the full-information value J_1 from Theorem 3 is not $\sigma(p_1)$-measurable. In particular, no representation $J_1 = V_1(p_1)$ is possible.

Proof sketch. With $H \in \{a, b\}$ equiprobable and $Y \sim \mathrm{Bernoulli}(3/4)$ independent of H, reveal H at $t = 1$ and Y only on $\{H = a\}$ at $t = 2$. Reports $p_1 = 3/4$ and $p_2 = 1$ on a, $p_2 = 3/4$ on b satisfy condition (iii), yet continuing is strictly optimal on a and strictly suboptimal on b. Section A.5 contains the full proof and Section A.15 the costs for $c_{\mathrm{step}} = 0.1$.

The following quantitative version shows that the gap can be of order c_{err}.

[2] The infimum runs over an uncountable family of stopping times. A pointwise infimum of uncountably many random variables need not be measurable, so the standard object is the essential infimum (Chow et al., 1971; Peskir & Shiryaev, 2006). The display is the assertion that the backward recursion computes that value.

Proposition 2 (Suboptimality of confidence-only policies). *For every $c_{\mathrm{err}} > 0$ and every $\delta \in (0, c_{\mathrm{err}}/4)$, there is a 2-step problem satisfying condition (iii) with $c_{\mathrm{abs}} = +\infty$ for which $R^*_{\mathrm{scalar}} - R^*_{\mathrm{full}} \geq c_{\mathrm{err}}/4 - \delta$, where R^*_{full} is the Bayes risk over all $(\mathcal{F}_t)$-adapted policies and R^*_{scalar} over policies adapted to $(\sigma(p_1, \ldots, p_t))_{t\geq 1}$. The gap approaches $c_{\mathrm{err}}/4$ in the parametrized family of Section A.6.*

Ideal confidence identifies the Bayes loss of stopping given the history, but it does not identify the continuation value. A Bellman recursion based only on confidence is valid uniformly over the class in Theorem 5 exactly when the future confidence path depends on the past only through the current confidence, as formalized next.

Because the reduced state space is $[0, 1]$, a standard Borel space, the conditional-independence and kernel formulations below are equivalent. We state both explicitly so the measurability requirements for scalar reduction are explicit.

Theorem 5 (Characterization of confidence as a Markov state). *Fix a finite horizon $T < \infty$ and an adapted confidence process $(p_t)_{t=1}^T$ with values in $[0, 1]$. For $1 \leq t \leq u \leq T$, let*

$$\mathcal{T}_t^u := \{\tau : \tau \text{ is a stopping time and } t \leq \tau \leq u\}.$$

For any bounded Borel loss $\ell : [0, 1] \to \mathbb{R}$ and step cost $c > 0$, define the truncated value process

$$J_t^{u,\ell,c} := \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_t^u} \mathbb{E}[c(\tau - t) + \ell(p_\tau) \mid \mathcal{F}_t].$$

The following are equivalent:

(a) for every $u \leq T$, bounded Borel ℓ, and $c > 0$, there exist Borel functions $V_t^{u,\ell,c}$ such that $J_t^{u,\ell,c} = V_t^{u,\ell,c}(p_t)$ almost surely for all $t \leq u$;
(b) for every $t < T$, the future confidence path $(p_{t+1}, \ldots, p_T)$ is conditionally independent of $\mathcal{F}_t$ given p_t;
(c) for every $t < T$, there exists a Borel kernel K_t on $[0, 1]$ such that

$$\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = K_t(p_t, B) \qquad \text{a.s.}$$

for every Borel set $B \subseteq [0, 1]$.

Under these equivalent conditions, the scalar Bellman recursion from Theorem 6 holds for every choice of u, ℓ, c.

Proof sketch. On a standard Borel space, (b) and (c) are equivalent, and one-step kernels compose to a product kernel for the full future path. Backward induction gives (a). Conversely, applying (a) to $\ell = \pm h$ forces $\mathbb{E}[h(p_{t+1}) \mid \mathcal{F}_t]$ to be $\sigma(p_t)$-measurable. Full proof: Section A.7.

Assumption 1 (Markov confidence model). *For this subsection we assume:*

(i) a finite horizon $T < \infty$;
(ii) ideal confidence in the sense of Definition 1(iii);
(iii) there exist Borel kernels K_t on $[0, 1]$ such that, for every Borel set $B \subseteq [0, 1]$,

$$\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = K_t(p_t, B) \qquad \text{a.s. for } t = 1, \ldots, T - 1.$$

Condition (iii) is exactly the Markov property that Theorem 5 identifies as necessary and sufficient for value factorization over every subhorizon, bounded Borel loss, and positive step cost.

Theorem 6 (Bellman recursion for a Markov confidence process). *Under Assumption 1, the value process from Theorem 3 factors through the current confidence. Given the kernel, this is standard dynamic programming with a sufficient statistic of history (Bertsekas, 2017; Åström, 1965; Smallwood & Sondik, 1973). There exist Borel functions $V_t : [0, 1] \to \mathbb{R}_+$ such that $J_t = V_t(p_t)$ almost surely and, with $\mathcal{T}_t$ as in Theorem 3,*

$$V_T(p) := g(p), \qquad V_t(p) := \min\left\{g(p),\ c_{\mathrm{step}} + \int_0^1 V_{t+1}(p')\, K_t(p, dp')\right\}$$

for $t = T - 1, \ldots, 1$. Consequently,

$$V_t(p_t) = \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_t} \mathbb{E}[c_{\mathrm{step}}(\tau - t) + g(p_\tau) \mid \mathcal{F}_t] \qquad \text{a.s.} \tag{5}$$

In particular:

(a) the stopping time

$$\tau^* := \inf\{t \le T : V_t(p_t) = g(p_t)\}$$

is optimal for the original stopping problem;

(b) the canonical terminal action at τ^* is $a^\dagger(p_{\tau^*})$; when $c_{\mathrm{abs}} < \infty$ this answers iff $p_{\tau^*} \ge p^\dagger$, with abstention also Bayes-optimal on the boundary; and

(c) the total Bayes risk is $c_{\mathrm{step}} + \mathbb{E}[V_1(p_1)]$: the first sequential step is mandatory before any stopping decision is available, so its cost c_{step} is always incurred.

Proof sketch. The kernel assumption identifies $\mathbb{E}[J_{t+1} \mid \mathcal{F}_t]$ with a Borel function of p_t, reducing the recursion of Theorem 3 to one on $[0, 1]$. Full details: Section A.10.

Interpretation. Ideal confidence identifies $g(p_t)$; the kernel reduces the full-history problem to a Bellman recursion on $[0, 1]$. For this g, the kernel is sufficient, not claimed necessary; without it, scalar policies can still lose $\Theta(c_{\mathrm{err}})$ (Theorem 4 and Proposition 2). A reduction that holds uniformly over losses, costs, and subhorizons is equivalent to the kernel. That equivalence is a Markov function (Rogers & Pitman, 1981) and a lumpable state abstraction (Kemeny & Snell, 1960; Givan et al., 2003; Li et al., 2006) on a standard Borel space, so it also holds for π_t on $\Delta^{|\mathcal{Y}|-1}$ (Section A.7). Under maximum a posteriori (MAP) prediction, $p_t = \max_y \pi_t(y)$ is a bounded submartingale (Proposition 10).

Proposition 3 (Conditionally independent evidence yields a Markov confidence process)**.** *If $|\mathcal{Y}| = 2$, $\hat{Y}_t \equiv 1$, $p_t := \mathbb{P}(Y = 1 \mid \mathcal{F}_t)$, and, for every $t \ge 1$, $Z_{t+1} \perp \mathcal{F}_t \mid Y$ with $Z_{t+1} \mid (Y = y) \sim \mu_y$, then confidence is ideal relative to the history and Assumption 1(iii) is satisfied by the stationary kernel induced by Bayes updating (Section A.8). Moreover, the Bayes kernel preserves concavity, so Corollary 1 applies whenever $c_{\mathrm{abs}} = +\infty$.*

Proposition 4 (When confidence is a Markov state with multiple classes)**.** *Let $\mathcal{Y}$ be finite, and write $\pi_t := (\mathbb{P}(Y = y \mid \mathcal{F}_t))_{y \in \mathcal{Y}}$.*

(i) ***Terminal answer versus abstain.*** *For any finite label set, if $p_t = \mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t)$, then answering costs $c_{\mathrm{err}}(1 - p_t)$ and abstaining costs c_{abs}. Unselected-class probabilities do not enter, so the scalar always suffices for the terminal decision (Theorem 3(a)).*

(ii) ***Signals about current-answer correctness.*** *Suppose the candidate is frozen after the first step, $\hat{Y}_t = \hat{Y}_1$ for all $t \ge 1$. Let $C := \mathbf{1}_{\{Y = \hat{Y}_1\}}$ and $p_t := \mathbb{P}(C = 1 \mid \mathcal{F}_t)$. If there are probability measures ν_0, ν_1 such that, conditionally on $C = c$, the signals $Z_2, Z_3, \ldots$ are i.i.d. with law ν_c and independent of $\mathcal{F}_1$, then confidence is ideal relative to the history and Assumption 1(iii) holds with the stationary Borel kernel induced by Bayes updating of C. Consequently Theorem 6 applies with Markov state p_t for any $|\mathcal{Y}|$, and Corollary 1 applies whenever $c_{\mathrm{abs}} = +\infty$.*

(iii) ***General multiclass continuation.*** *Suppose π_t is Markov relative to $\mathcal{F}_t$ with Borel kernel Q_t, and $p_t = m_t(\pi_t)$ for a Borel map m_t. If a Borel kernel K_t satisfies*

$$Q_t\big(\pi, m_{t+1}^{-1}(B)\big) = K_t(m_t(\pi), B)$$

for every reachable π and Borel $B \subseteq [0, 1]$, then Assumption 1(iii) holds. Thus posterior states with the same current confidence may be merged only when they induce the same next-confidence law.

Freezing $\hat{Y}_t \equiv \hat{Y}_1$ is essential in (ii): if the candidate switches, the correctness event changes. More generally, π_t is a valid reduced state only under a model, such as the Markov assumption in (iii), that makes its transition law independent of the remaining history.

Proof sketch. Part (i) is Theorem 3(a). Part (ii) substitutes C for Y in Proposition 3; see Section A.9. Part (iii) pushes Q_t forward through m_{t+1} and invokes Theorem 5. Adapted step costs are recorded in Section A.13.

Corollary 1 (Upper-threshold stopping without abstention)**.** *Assume Assumption 1 and $c_{\mathrm{abs}} = +\infty$, so $g(p) = c_{\mathrm{err}}(1 - p)$. Suppose further that each operator*

$$(\mathcal{K}_t f)(p) := \int_0^1 f(p')\, K_t(p, dp')$$

maps concave functions on $[0,1]$ *to concave functions, and that* $K_t(1,\{1\}) = 1$ *for all* $t < T$*. Then each* V_t *is concave and there exist deterministic thresholds* $\beta_t \in [0,1]$ *such that stopping is optimal whenever* $p_t > \beta_t$*, continuation is optimal whenever* $p_t < \beta_t$*, and only boundary ties can occur at* $p_t = \beta_t$*.*

Assumption 1(iii) is exact. The scalar Bellman recursion is robust to small total-variation deviations from this Markov reduction, with linear ηT degradation.

Proposition 5 (Robustness to an approximate Markov kernel)**.** *Suppose condition (iii) of Definition 1 holds and, for some Borel kernels* $\widetilde{K}_t$*, every* $t < T$ *satisfies* $\|\mathbb{P}(p_{t+1} \in \cdot \mid \mathcal{F}_t) - \widetilde{K}_t(p_t, \cdot)\|_{\mathrm{TV}} \le \eta$ *almost surely. Let* $\widetilde{V}_t$ *be the scalar Bellman recursion of Theorem 6 computed with* $\widetilde{K}_t$*. Then* $|J_t - \widetilde{V}_t(p_t)| \le 2(T-t)\,\eta\, c_{\mathrm{err}}$ *a.s., and the total Bayes-risk gap is at most* $2(T-1)\,\eta\, c_{\mathrm{err}}$*.*

3.3 Limits of Calibration at Each Fixed Time

The preceding subsection held the current correctness probability fixed and showed that future informativeness can vary across histories. Ordinary confidence calibration applied separately at each fixed time permits a more basic failure: histories with the same reported score may already have different current correctness probabilities. Restricting the policy to the reported outputs can then increase Bayes risk by $\Theta(c_{\mathrm{err}})$.

Theorem 7 (Fixed-time calibration does not imply optimal stopping from reports)**.** *There exists a two-step sequential prediction problem with* $c_{\mathrm{abs}} = +\infty$ *and a predictor such that:*

- *(i) the reported pair* $(\hat{Y}_1, p_1)$ *is constant almost surely, and confidence is calibrated at* $t = 1$ *and* $t = 2$*;*
- *(ii) condition (iii) of Definition 1 fails at* $t = 1$*; and*
- *(iii) every* $\mathcal{G}_t$*-adapted stopping policy has expected cost at least* 0.5*, whereas the* $\mathcal{F}_t$*-optimal stopping policy attains risk* 0.425*.*

Proof sketch. A hidden state $H \in \{a, b\}$ is equiprobable, with $\mathbb{P}(Y = 1 \mid H = a) = 0.6$ and $\mathbb{P}(Y = 1 \mid H = b) = 0.9$. The predictor always reports $\hat{Y}_1 = 1$ and $p_1 = 0.75$: calibrated at that time, but not ideal relative to $\mathcal{F}_1$. With $c_{\mathrm{step}} = 0.25$ and $c_{\mathrm{err}} = 1$, stopping now costs 0.65 on a and 0.35 on b, while the report assigns expected loss 0.50. The full-information policy continues on a and stops on b; any $\mathcal{G}_t$-policy must pick one action and costs 0.5. This fails Theorem 2; Section A.11.

Proposition 6 (Suboptimality of report-based policies)**.** *For every* $c_{\mathrm{err}} > 0$ *and every* $\delta \in (0, 1/4)$*, there exists a 2-step problem calibrated at each fixed time, with* $c_{\mathrm{abs}} = +\infty$ *and reported pair* $(\hat{Y}_1, p_1)$ *constant a.s. such that the Bayes-risk gap between the* $\mathcal{F}_t$*-optimal policy and the best* $\mathcal{G}_t$*-adapted policy equals* $\delta\, c_{\mathrm{err}}/2$*. The construction parametrizes Theorem 7 by* $q_a = 3/4 - \delta,\ q_b = 3/4 + \delta$*; the gap is shown empirically in Figure 1 (left).*

3.4 Robustness to Approximate Conditional Probabilities

The exact identity (4) is strong. Because Theorem 4 shows that confidence-only stopping can fail even when the score equals the history-conditional probability, the relevant robustness question is how approximation error perturbs the full-information stopping problem on the same filtration. Accordingly, both the true and surrogate stopping problems below are optimized over $(\mathcal{F}_t)$.

Theorem 8 (Regret from approximate conditional probabilities)**.** *Assume a finite horizon* T *and define the history-conditional probability that the current answer is correct,*

$$p_t^* := \mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t).$$

Suppose $\sup_{1 \le t \le T} |p_t - p_t^*| \le \varepsilon$ *almost surely. Let* $\hat{\tau}$ *be a stopping time that minimizes* $\mathbb{E}[c_{\mathrm{step}}\tau + g(p_\tau)]$ *over all stopping times* $\tau \le T$ *adapted to the filtration* $(\mathcal{F}_t)$*. At* $\hat{\tau}$*, let* $\hat{a}$ *be any Bayes action under the reported score, allowing arbitrary randomization when* $p_{\hat{\tau}} = p^\dagger$*, and define* $\hat{\pi} := (\hat{\tau}, \hat{a})$*. Let* $\pi^* = (\tau^*, a^*)$ *minimize the true Bayes risk. Then*

$$\mathbb{E}[\mathcal{C}(\hat{\pi})] - \mathbb{E}[\mathcal{C}(\pi^*)] \le 2c_{\mathrm{err}}\varepsilon. \tag{6}$$

Proof sketch. The actual incurred terminal cost $\tilde{g}(p_{\hat{\tau}}, p^*_{\hat{\tau}})$ satisfies $\tilde{g}(p, p^*) \le g(p) + c_{\mathrm{err}}|p - p^*|$ in every case (answer, abstain, or no-abstention). This contributes one $c_{\mathrm{err}}\varepsilon$ term and reduces the comparison to the surrogate objective $\hat{J}$. A second $c_{\mathrm{err}}\varepsilon$ term arises from the c_{err}-Lipschitz transfer $\hat{J}(\tau^*) \le J(\tau^*) + c_{\mathrm{err}}\varepsilon$. The full chain is in Section A.12.

Proposition 7 (Tie-free linear lower bound)**.** *Assume $c_{\mathrm{abs}} < \infty$ and fix any $\varepsilon > 0$ with $p^\dagger \in (\varepsilon, 1-\varepsilon)$. For every $\zeta \in (0,1)$ there exists a one-step problem satisfying $|p_1 - p^*_1| = \varepsilon$ in which the reported and true terminal decisions are both strict, yet the surrogate policy incurs regret at least $(1-\zeta)c_{\mathrm{err}}\varepsilon$. Hence the linear dependence in Theorem 8 is unavoidable under every tie convention.*

Proposition 8 (Tight lower bound)**.** *For every $c_{\mathrm{err}} > 0$ and every $\eta \in (0,2)$, there exist $\varepsilon > 0$, $c_{\mathrm{abs}} \in (0, c_{\mathrm{err}})$, and a 2-step problem with $\sup_{t \le 2} |p_t - p^*_t| \le \varepsilon$ such that every relevant terminal and stopping comparison is strict and the surrogate-optimal policy has regret greater than $(2-\eta)c_{\mathrm{err}}\varepsilon$. Combined with Theorem 8, the constant 2 is tight over the admissible cost class without relying on a tie. The family varies $c_{\mathrm{abs}}/c_{\mathrm{err}}$ with ε; the claim is not that the constant is attained at every fixed reject-cost ratio.*

Remark 1 (Tie-breaking does not affect the bounds)**.** *Theorem 8 allows any randomized Bayes action at $p = p^\dagger$, while both lower bounds use strict comparisons. Their rates and sharp constant are therefore independent of the tie convention.*

4 Related Work and Broader Impact

Sequential decisions and optimal stopping. The full-history problem is classical Bayesian sequential decision with a sampling cost and adapted terminal loss (Wald, 1945; 1947; Wald & Wolfowitz, 1948; Arrow et al., 1949; Chow et al., 1971; Peskir & Shiryaev, 2006; Bertsekas, 2017); its Snell recursion is standard (Snell, 1952). Partially observed control uses a belief state as a sufficient statistic of history (Åström, 1965; Smallwood & Sondik, 1973). Here the controller instead receives $(\hat{Y}_t, p_t)$ from an opaque predictor, and we ask when this report determines the stopping loss and the continuation value.

Calibration, ideal forecasts, and abstention. Confidence calibration requires the conditional correctness probability given the score to equal the score (Dawid, 1982; Guo et al., 2017). Forecasting calls a forecast *ideal relative to an information set* when it equals the conditional law given that information (Gneiting & Ranjan, 2013); Definition 1(iii) is its Bernoulli case. The terminal loss g is Chow's fixed-time reject option (Chow, 1957; 1970), developed in selective classification (Bartlett & Wegkamp, 2008; El-Yaniv & Wiener, 2010; Geifman & El-Yaniv, 2017; 2019). Set-valued and conformal methods control coverage or ambiguity (Sadinle et al., 2019; Angelopoulos & Bates, 2023; Angelopoulos et al., 2025; Hegazy et al., 2025), while decision calibration, U-calibration, multicalibration, and related threshold-score guarantees target chosen losses or a chosen time (Zhao et al., 2021; Sahoo et al., 2021; Kleinberg et al., 2023; Hébert-Johnson et al., 2018; Rothblum & Yona, 2023; Franc et al., 2023; Corvelo Benz & Gomez Rodriguez, 2023). Theorem 2 instead considers a history-dependent stopping time. Online calibration concerns long-run frequency alignment (Oakes, 1985; Foster & Vohra, 1998; Qiao & Zheng, 2024).

State abstraction and confidence-based stopping. Under Proposition 4(iii), confidence is a Markov function of the posterior state (Rogers & Pitman, 1981); the pushforward condition is a form of lumpability (Kemeny & Snell, 1960; Givan et al., 2003; Li et al., 2006). Early-exit and adaptive-computation systems often threshold softmax or learned scores (Graves, 2016; Teerapittayanon et al., 2016; Kaya et al., 2019; Schwartz et al., 2020; Xin et al., 2020; Elbayad et al., 2020); CALM controls the marginal risk of a chosen stopping-rule family (Schuster et al., 2022; Angelopoulos et al., 2025); language-model self-evaluation and uncertainty scores are likewise assessed mainly through marginal summaries (Kadavath et al., 2022; Kuhn et al., 2023; Band et al., 2024; Mao et al., 2025; Kubaty et al., 2025).

Limitations and broader impact. Condition (iii) is a correct-specification target, not an estimator or a distribution-free certificate. Logged histories permit regression and cross-fitted tests, but no finite test on $\mathcal{G}_t$ can certify an identity relative to a richer $\mathcal{F}_t$; held-out tests can only challenge the Markov-kernel condition. Experiments verify synthetic gaps but no real-data cost gap. Estimation remains open, and marginal calibration alone is inadequate for safety-critical stopping.

AI USE STATEMENT

Generative AI tools assisted with drafting and copyediting, preparation of numerical verification code, and checks of cross-references and bibliographic metadata. The authors formulated the research questions, model, and results; independently checked every theorem statement and proof; reproduced every numerical constant reported in the paper; and verified each bibliographic entry against the publisher record. The authors take full responsibility for the manuscript and its claims.

ETHICS STATEMENT

The only non-synthetic data are the public, de-identified Wisconsin Breast Cancer Diagnostic records distributed with scikit-learn. No new human-subjects data were collected. The study is a diagnostic proxy on tabular features, not a clinical decision system.

REPRODUCIBILITY STATEMENT

All theorem statements list their assumptions explicitly. Complete proofs are in Section A, including the tie-breaking and correctness-signal arguments. Synthetic constructions, seeds, and sample sizes are specified in Section A.16; the Wisconsin proxy protocol is specified in Section A.17. Accompanying code reproduces Figures 1 to 4 and the summary table.

REFERENCES

Anastasios N. Angelopoulos and Stephen Bates. Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4):494–591, 2023.

Anastasios N. Angelopoulos, Stephen Bates, Emmanuel J. Candès, Michael I. Jordan, and Lihua Lei. Learn then test: Calibrating predictive algorithms to achieve risk control. *The Annals of Applied Statistics*, 19(2):1641–1662, 2025.

Kenneth J. Arrow, David Blackwell, and M. A. Girshick. Bayes and minimax solutions of sequential decision problems. *Econometrica*, 17(3/4):213–244, 1949.

Karl Johan Åström. Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205, 1965.

Neil Band, Xuechen Li, Tengyu Ma, and Tatsunori Hashimoto. Linguistic calibration of long-form generations. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 2732–2778, 2024.

Peter L. Bartlett and Marten H. Wegkamp. Classification with a reject option using a hinge loss. *Journal of Machine Learning Research*, 9:1823–1840, 2008.

Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, Volume I*. Athena Scientific, 4 edition, 2017.

David Blackwell. Comparison of experiments. In Jerzy Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102. University of California Press, 1951.

C. K. Chow. An optimum character recognition system using decision functions. *IRE Transactions on Electronic Computers*, EC-6(4):247–254, 1957.

C. K. Chow. On optimum recognition error and reject tradeoff. *IEEE Transactions on Information Theory*, 16(1):41–46, 1970.

Yuan Shih Chow, Herbert Robbins, and David Siegmund. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, 1971.

Nina L. Corvelo Benz and Manuel Gomez Rodriguez. Human-aligned calibration for ai-assisted decision making. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

A. Philip Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77 (379):605–610, 1982.

A. Philip Dawid. Calibration-based empirical probability. *The Annals of Statistics*, 13(4):1251–1274, 1985.

Morris H. DeGroot. *Optimal Statistical Decisions*. Wiley Classics Library. John Wiley & Sons, 2004. Originally published 1970; reprinted 2004.

Ran El-Yaniv and Yair Wiener. On the foundations of noise-free selective classification. *Journal of Machine Learning Research*, 11:1605–1641, 2010.

Maha Elbayad, Jiatao Gu, Edouard Grave, and Michael Auli. Depth-adaptive transformer. In *International Conference on Learning Representations*, 2020.

Dean P. Foster and Rakesh V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998.

Vojtech Franc, Daniel Prusa, and Vaclav Voracek. Optimal strategies for reject option classifiers. *Journal of Machine Learning Research*, 24(11):1–49, 2023.

Yonatan Geifman and Ran El-Yaniv. Selective classification for deep neural networks. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Yonatan Geifman and Ran El-Yaniv. SelectiveNet: A deep neural network with an integrated reject option. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2151–2159, 2019.

Robert Givan, Thomas Dean, and Matthew Greig. Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1–2):163–223, 2003.

Tilmann Gneiting and Roopesh Ranjan. Combining predictive distributions. *Electronic Journal of Statistics*, 7:1747–1782, 2013. doi: 10.1214/13-EJS823.

Alex Graves. Adaptive computation time for recurrent neural networks. *arXiv preprint arXiv:1603.08983*, 2016.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1321–1330, 2017.

Úrsula Hébert-Johnson, Michael P. Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Calibration for the (Computationally-Identifiable) masses. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1939–1948, 2018.

Mahmoud Hegazy, Liviu Aolaritei, Michael I. Jordan, and Aymeric Dieuleveut. Valid selection among conformal sets. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Saurav Kadavath, Tom Conerly, Amanda Askell, Tom Henighan, Dawn Drain, Ethan Perez, Nicholas Schiefer, Zac Hatfield-Dodds, Nova DasSarma, Eli Tran-Johnson, Scott Johnston, Sheer El-Showk, Andy Jones, Nelson Elhage, Tristan Hume, Anna Chen, Yuntao Bai, Sam Bowman, Stanislav Fort, Deep Ganguli, Danny Hernandez, Josh Jacobson, Jackson Kernion, Shauna Kravec, Liane Lovitt, Kamal Ndousse, Catherine Olsson, Sam Ringer, Dario Amodei, Tom Brown, Jack Clark, Nicholas Joseph, Ben Mann, Sam McCandlish, Chris Olah, and Jared Kaplan. Language models (mostly) know what they know. *arXiv preprint arXiv:2207.05221*, 2022.

Yigitcan Kaya, Sanghyun Hong, and Tudor Dumitras. Shallow-Deep networks: Understanding and mitigating network overthinking. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 3301–3310, 2019.

John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Van Nostrand, 1960.

Bobby Kleinberg, Renato Paes Leme, Jon Schneider, and Yifeng Teng. U-calibration: Forecasting for an unknown agent. In *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 5143–5145, 2023.

Piotr Kubaty, Filip Szatkowski, Metod Jazbec, and Bartosz Wójcik. Failure prediction is a better performance proxy for early-exit networks than calibration. *arXiv preprint arXiv:2508.21495*, 2025.

Lorenz Kuhn, Yarin Gal, and Sebastian Farquhar. Semantic uncertainty: Linguistic invariances for uncertainty estimation in natural language generation. In *International Conference on Learning Representations*, 2023.

Lihong Li, Thomas J. Walsh, and Michael L. Littman. Towards a unified theory of state abstraction for MDPs. In *Proceedings of the Ninth International Symposium on Artificial Intelligence and Mathematics*, 2006.

Yuzhen Mao, Thibaut Durand, Nazanin Mehrasa, Jiawei He, and Martin Ester. Calibrating LLMs for selective prediction: Balancing coverage and risk. In *NeurIPS 2025 Workshop on Socially Responsible and Trustworthy Foundation Models*, 2025.

David Oakes. Self-calibrating priors do not exist. *Journal of the American Statistical Association*, 80 (390):339, 1985.

Goran Peskir and Albert N. Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Birkhäuser, Basel, 2006.

Mingda Qiao and Letian Zheng. On the distance from calibration in sequential prediction. In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 4307–4357, 2024.

L. C. G. Rogers and J. W. Pitman. Markov functions. *The Annals of Probability*, 9(4):573–582, 1981.

Guy N. Rothblum and Gal Yona. Decision-making under miscalibration. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 92:1–92:20, 2023.

Mauricio Sadinle, Jing Lei, and Larry Wasserman. Least ambiguous set-valued classifiers with bounded error levels. *Journal of the American Statistical Association*, 114(525):223–234, 2019.

Roshni Sahoo, Shengjia Zhao, Alyssa Chen, and Stefano Ermon. Reliable decisions with threshold calibration. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Tal Schuster, Adam Fisch, Jai Gupta, Mostafa Dehghani, Dara Bahri, Vinh Q. Tran, Yi Tay, and Donald Metzler. Confident adaptive language modeling. In *Advances in Neural Information Processing Systems*, volume 35, pp. 17456–17472, 2022. doi: 10.52202/068431-1269.

Roy Schwartz, Gabriel Stanovsky, Swabha Swayamdipta, Jesse Dodge, and Noah A. Smith. The right tool for the job: Matching model and instance complexities. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 6640–6651, 2020.

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.

J. Laurie Snell. Applications of martingale system theorems. *Transactions of the American Mathematical Society*, 73(2):293–312, 1952.

Surat Teerapittayanon, Bradley McDanel, and H. T. Kung. BranchyNet: Fast inference via early exiting from deep neural networks. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 2464–2469, 2016.

Abraham Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16 (2):117–186, 1945.

Abraham Wald. *Sequential Analysis*. John Wiley & Sons, 1947.

Abraham Wald and Jacob Wolfowitz. Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, 19(3):326–339, 1948.

Ji Xin, Raphael Tang, Jaejun Lee, Yaoliang Yu, and Jimmy Lin. DeeBERT: Dynamic early exiting for accelerating BERT inference. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2246–2251, 2020.

Shengjia Zhao, Michael P. Kim, Roshni Sahoo, Tengyu Ma, and Stefano Ermon. Calibrating predictions to decisions: A novel approach to multi-class calibration. In *Advances in Neural Information Processing Systems*, volume 34, pp. 22313–22324, 2021.

A Proofs and Supplementary Material

A.1 Proof of the characterization theorem

Proof of Theorem 1. The parenthetical in (a) is the definition of conditional expectation.

Assume (a) and let $\tau \leq T$ be a bounded stopping time. We show (b). For any $A \in \mathcal{F}_\tau$,

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{\{Y=\hat{Y}_\tau\}}\mathbf{1}_A] &= \sum_{t=1}^{T} \mathbb{E}[\mathbf{1}_{\{Y=\hat{Y}_t\}}\mathbf{1}_{A\cap\{\tau=t\}}] \\ &= \sum_{t=1}^{T} \mathbb{E}[p_t\mathbf{1}_{A\cap\{\tau=t\}}] = \sum_{t=1}^{T} \mathbb{E}[p_\tau\mathbf{1}_{A\cap\{\tau=t\}}] \\ &= \mathbb{E}[p_\tau\mathbf{1}_A].\end{aligned}$$

The second equality uses $A \cap \{\tau = t\} \in \mathcal{F}_t$ and (a), and the third uses $p_\tau = p_t$ on $\{\tau = t\}$. Hence $p_\tau = \mathbb{P}(Y = \hat{Y}_\tau \mid \mathcal{F}_\tau)$ almost surely.

Assume (b). For a fixed t, apply (b) to the deterministic stopping time $\tau \equiv t$ to recover (a).

Finally, if condition (iii) of Definition 1 holds, then $\sigma(p_t) \subseteq \mathcal{F}_t$ and $\sigma(p_\tau) \subseteq \mathcal{F}_\tau$, so

$$\mathbb{P}(Y = \hat{Y}_t \mid p_t) = \mathbb{E}[\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t) \mid p_t] = \mathbb{E}[p_t \mid p_t] = p_t,$$

and analogously $\mathbb{P}(Y = \hat{Y}_\tau \mid p_\tau) = p_\tau$ for any bounded stopping time τ. □

Proposition 9 (When fixed-time calibration implies ideal confidence)**.** *Condition (iii) of Definition 1 is equivalent to confidence calibration at each fixed time together with $\sigma(p_t)$-measurability of $\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t)$ for every t. In particular, if $\sigma(p_t) = \mathcal{F}_t$ a.s. for all t, then fixed-time calibration also implies calibration at every bounded stopping time and ideal confidence.*

Proof of Proposition 9. Let $E_t := \{Y = \hat{Y}_t\}$ and $Q := \mathbb{P}(E_t \mid \mathcal{F}_t)$.

($\Rightarrow$) If condition (iii) holds, then $Q = p_t$, which is $\sigma(p_t)$-measurable. Fixed-time calibration follows from Theorem 1.

($\Leftarrow$) If Q is $\sigma(p_t)$-measurable, then $Q = \mathbb{E}[Q \mid p_t]$. By the tower property and fixed-time calibration,

$$\mathbb{E}[Q \mid p_t] = \mathbb{E}[\mathbb{P}(E_t \mid \mathcal{F}_t) \mid p_t] = \mathbb{P}(E_t \mid p_t) = p_t.$$

Hence $Q = p_t$ a.s., so condition (iii) holds. □

A.2 Proof that a score suffices for every reject cost

Proof of Theorem 2. Let $E_\tau := \{Y = \hat{Y}_\tau\}$, $\mathcal{G}_\tau := \sigma(s_\tau)$, and $r_\tau := \mathbb{P}(E_\tau \mid \mathcal{F}_\tau)$. Write an answer/abstain rule as an answer set $A \in \mathcal{F}_\tau$: answer on A and abstain on A^c. For a fixed threshold $\lambda \in [0, 1]$, the Bayes risk of A under the normalized loss ℓ_λ is

$$\begin{aligned}R_\lambda(A) &:= \mathbb{E}[\mathbf{1}_A\mathbf{1}_{E_\tau^c} + (1-\lambda)\mathbf{1}_{A^c}] \\ &= \mathbb{E}[(1-r_\tau)\mathbf{1}_A + (1-\lambda)\mathbf{1}_{A^c}] \\ &= 1 - \lambda + \mathbb{E}[(\lambda - r_\tau)\mathbf{1}_A].\end{aligned}$$

To minimize this expectation, we must set $\mathbf{1}_A = 1$ when the integrand coefficient $(\lambda - r_\tau)$ is strictly negative, and $\mathbf{1}_A = 0$ when it is strictly positive. Hence A is Bayes-optimal if and only if it answers exactly where $r_\tau \geq \lambda$, up to ties on the set $\{r_\tau = \lambda\}$:

$$\{r_\tau > \lambda\} \subseteq A \subseteq \{r_\tau \geq \lambda\} \qquad \text{a.s.} \tag{7}$$

(b) $\Rightarrow$ (a). If r_τ admits a $\mathcal{G}_\tau$-measurable version, then $A_\lambda := \{r_\tau \geq \lambda\} \in \mathcal{G}_\tau$ is Bayes-optimal by (7). This proves (a) for every λ, hence also for every rational λ.

(a) ⇒ (b). Assume (a). For each rational $q \in [0,1]$, let $A_q \in \mathcal{G}_\tau$ be a Bayes-optimal answer set for threshold q. By (7),

$$\{r_\tau > q\} \subseteq A_q \subseteq \{r_\tau \geq q\} \qquad \text{a.s.}$$

For each rational q, choose a null set N_q outside of which these inclusions hold, and let

$$N := \bigcup_{q \in \mathbb{Q} \cap [0,1]} N_q.$$

Then $\mathbb{P}(N) = 0$, and all inclusions above hold simultaneously on N^c for every rational q. Fix rational $t \in [0,1]$. We claim that

$$\{r_\tau > t\} = \bigcup_{\substack{q \in \mathbb{Q} \cap [0,1] \\ q > t}} A_q \qquad \text{a.s.}$$

Indeed, on N^c, if $r_\tau(\omega) > t$, choose a rational q with $t < q < r_\tau(\omega)$; then $\omega \in \{r_\tau > q\} \subseteq A_q$. Conversely, if $\omega \in A_q$ for some rational $q > t$, then $\omega \in \{r_\tau \geq q\}$, so $r_\tau(\omega) > t$. Therefore each strict superlevel set $\{r_\tau > t\}$ admits a $\mathcal{G}_\tau$-measurable version.

For each rational $q \in [0,1]$, define

$$B_q := \bigcup_{\substack{r \in \mathbb{Q} \cap [0,1] \\ r > q}} A_r \in \mathcal{G}_\tau.$$

By the claim with $t = q$, each B_q is a $\mathcal{G}_\tau$-measurable version of $\{r_\tau > q\}$. Now define

$$\tilde{r}_\tau(\omega) := \sup\big(\{0\} \cup \{q \in \mathbb{Q} \cap [0,1] : \omega \in B_q\}\big).$$

Then $\tilde{r}_\tau$ is $\mathcal{G}_\tau$-measurable. On N^c, for every rational q we have

$$\omega \in B_q \iff r_\tau(\omega) > q \iff \tilde{r}_\tau(\omega) > q.$$

Hence $\tilde{r}_\tau$ and r_τ have the same strict rational superlevel sets on N^c, so $\tilde{r}_\tau = r_\tau$ on N^c and therefore almost surely. Thus r_τ admits a $\mathcal{G}_\tau$-measurable version.

Final clause. If additionally $\mathbb{P}(E_\tau \mid s_\tau) = s_\tau$ almost surely, then

$$s_\tau = \mathbb{P}(E_\tau \mid \mathcal{G}_\tau) = \mathbb{E}[\mathbb{P}(E_\tau \mid \mathcal{F}_\tau) \mid \mathcal{G}_\tau] = \mathbb{E}[r_\tau \mid \mathcal{G}_\tau].$$

Under (a), part (b) implies that r_τ admits a $\mathcal{G}_\tau$-measurable version, so $\mathbb{E}[r_\tau \mid \mathcal{G}_\tau] = r_\tau$ almost surely. Therefore $s_\tau = r_\tau$ almost surely. □

A.3 Selective prediction and martingale facts

Proof of Proposition 1. Let $\alpha_\tau : \Omega \to [0,1]$ be any feasible acceptance probability with $\mathbb{E}[\alpha_\tau] = \gamma$. Implement it by introducing an auxiliary $U \sim \mathrm{Unif}[0,1]$ independent of $\mathcal{F}_\tau$ and accepting on the event $\{U \leq \alpha_\tau\}$. By condition (iii) and iterated expectation, the conditional error of this randomized selective rule equals

$$\frac{\mathbb{E}[(1-p_\tau)\alpha_\tau]}{\mathbb{E}[\alpha_\tau]} = \frac{\mathbb{E}[(1-p_\tau)\alpha_\tau]}{\gamma}.$$

Thus it suffices to minimize $\mathbb{E}[(1-p_\tau)\alpha_\tau]$ over all $\mathcal{F}_\tau$-measurable $\alpha_\tau \in [0,1]$ with expectation γ.

Choose $\theta \in [0,1]$ so that

$$\mathbb{P}(p_\tau > \theta) \leq \gamma \leq \mathbb{P}(p_\tau \geq \theta).$$

Such a threshold exists by monotonicity of the tail probabilities. If $\mathbb{P}(p_\tau = \theta) > 0$, define

$$\rho := \frac{\gamma - \mathbb{P}(p_\tau > \theta)}{\mathbb{P}(p_\tau = \theta)} \in [0,1];$$

otherwise set $\rho := 0$. Then

$$\alpha_\tau^\star := \mathbf{1}_{\{p_\tau > \theta\}} + \rho\, \mathbf{1}_{\{p_\tau = \theta\}}$$

satisfies $\mathbb{E}[\alpha_\tau^\star] = \gamma$.

Now let α_τ be any other feasible acceptance probability. Since $\mathbb{E}[\alpha_\tau - \alpha_\tau^\star] = 0$,

$$\mathbb{E}[(1 - p_\tau)(\alpha_\tau - \alpha_\tau^\star)] = \mathbb{E}[(\theta - p_\tau)(\alpha_\tau - \alpha_\tau^\star)].$$

On $\{p_\tau > \theta\}$, we have $\alpha_\tau^\star = 1$, so $\alpha_\tau - \alpha_\tau^\star \le 0$ and $\theta - p_\tau < 0$. On $\{p_\tau < \theta\}$, we have $\alpha_\tau^\star = 0$, so $\alpha_\tau - \alpha_\tau^\star \ge 0$ and $\theta - p_\tau > 0$. On $\{p_\tau = \theta\}$, the integrand vanishes. Therefore

$$\mathbb{E}[(1 - p_\tau)(\alpha_\tau - \alpha_\tau^\star)] \ge 0,$$

which proves optimality of $\alpha_\tau^\star$. □

Proposition 10 (MAP confidence: submartingale and convergence)**.** *Assume condition (iii) of Definition 1 and let $\pi_t(y) := \mathbb{P}(Y = y \mid \mathcal{F}_t)$. If the predictor chooses a MAP answer, $\hat{Y}_t \in \arg\max_{y \in \mathcal{Y}} \pi_t(y)$, then $p_t = \max_{y \in \mathcal{Y}} \pi_t(y)$ is a bounded submartingale and converges almost surely to an $\mathcal{F}_\infty$-measurable limit $p_\infty \in [1/|\mathcal{Y}|, 1]$, where $\mathcal{F}_\infty := \sigma(\bigcup_{t \ge 0} \mathcal{F}_t)$.*

Proof. For each $y \in \mathcal{Y}$, the process $\pi_t(y)$ is a martingale because

$$\mathbb{E}[\pi_{t+1}(y) \mid \mathcal{F}_t] = \mathbb{E}[\mathbb{P}(Y = y \mid \mathcal{F}_{t+1}) \mid \mathcal{F}_t] = \mathbb{P}(Y = y \mid \mathcal{F}_t) = \pi_t(y).$$

The map $(x_y)_{y \in \mathcal{Y}} \mapsto \max_y x_y$ is convex, so conditional Jensen implies

$$\mathbb{E}[p_{t+1} \mid \mathcal{F}_t] = \mathbb{E}\left[\max_y \pi_{t+1}(y) \mid \mathcal{F}_t\right] \ge \max_y \mathbb{E}[\pi_{t+1}(y) \mid \mathcal{F}_t] = \max_y \pi_t(y) = p_t.$$

Because $p_t \in [1/|\mathcal{Y}|, 1]$, the submartingale is uniformly bounded, so Doob's submartingale convergence theorem yields $p_t \to p_\infty$ almost surely with $p_\infty \in [1/|\mathcal{Y}|, 1]$. □

A.4 Proof of full-information optimal stopping

Proof of Theorem 3. **Part (a).** Fix a stopping time τ with $1 \le \tau \le T$. By Theorem 1,

$$\mathbb{P}(Y = \hat{Y}_\tau \mid \mathcal{F}_\tau) = p_\tau \qquad \text{a.s.}$$

Hence, for any $\mathcal{F}_\tau$-measurable terminal action a_τ,

$$\begin{aligned}\mathbb{E}\Big[c_{\mathrm{err}} \mathbf{1}_{\{a_\tau = \text{ANSWER}\}} \mathbf{1}_{\{Y \ne \hat{Y}_\tau\}} + c_{\mathrm{abs}} \mathbf{1}_{\{a_\tau = \text{ABSTAIN}\}} \;\Big|\; \mathcal{F}_\tau\Big] \\ = c_{\mathrm{err}}(1 - p_\tau) \mathbf{1}_{\{a_\tau = \text{ANSWER}\}} + c_{\mathrm{abs}} \mathbf{1}_{\{a_\tau = \text{ABSTAIN}\}} \ge g(p_\tau).\end{aligned}$$

Equality holds when the action minimizes the right-hand side. The canonical selector $a^\dagger(p_\tau)$ does exactly this. When $c_{\mathrm{abs}} < \infty$, it answers iff $p_\tau \ge p^\dagger$, with abstention also optimal on the boundary $p_\tau = p^\dagger$. If $c_{\mathrm{abs}} = +\infty$, then $g(p) = c_{\mathrm{err}}(1 - p)$ and $a^\dagger(p) = \text{ANSWER}$ always.

Parts (b) and (c). Let $L_t := g(p_t)$. Since $0 \le L_t \le c_{\mathrm{err}}$, the process (L_t) is bounded and integrable. Define

$$J_T := L_T, \qquad J_t := \min\{L_t, c_{\mathrm{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\} \quad \text{for } t = T - 1, \dots, 1.$$

We prove by backward induction that

$$J_t = \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_t} \mathbb{E}[c_{\mathrm{step}}(\tau - t) + L_\tau \mid \mathcal{F}_t]. \tag{8}$$

The claim is immediate at time T. Assume it holds at time $t + 1$ and fix $\tau \in \mathcal{T}_t$. Let $\bar{\tau} := \tau \vee (t + 1)$, so $\bar{\tau} \in \mathcal{T}_{t+1}$ and $\bar{\tau} = \tau$ on $\{\tau > t\}$. Then

$$\begin{aligned}\mathbb{E}[c_{\mathrm{step}}(\tau - t) + L_\tau \mid \mathcal{F}_t] &= \mathbb{E}\big[\mathbf{1}_{\{\tau = t\}} L_t + \mathbf{1}_{\{\tau > t\}}(c_{\mathrm{step}} + c_{\mathrm{step}}(\bar{\tau} - t - 1) + L_{\bar{\tau}}) \;\big|\; \mathcal{F}_t\big] \\ &= \mathbf{1}_{\{\tau = t\}} L_t + \mathbf{1}_{\{\tau > t\}}\big(c_{\mathrm{step}} + \mathbb{E}[c_{\mathrm{step}}(\bar{\tau} - t - 1) + L_{\bar{\tau}} \mid \mathcal{F}_t]\big) \\ &= \mathbf{1}_{\{\tau = t\}} L_t + \mathbf{1}_{\{\tau > t\}} c_{\mathrm{step}} \\ &\quad + \mathbf{1}_{\{\tau > t\}} \mathbb{E}\big[\mathbb{E}[c_{\mathrm{step}}(\bar{\tau} - t - 1) + L_{\bar{\tau}} \mid \mathcal{F}_{t+1}] \;\big|\; \mathcal{F}_t\big] \\ &\ge \mathbf{1}_{\{\tau = t\}} L_t + \mathbf{1}_{\{\tau > t\}}\big(c_{\mathrm{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\big) \\ &\ge \min\{L_t, c_{\mathrm{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\} = J_t,\end{aligned}$$

where the inequality uses the induction hypothesis at time $t+1$. This proves the lower bound in (8).

To attain the bound, define recursively

$$\tau_T^* := T, \qquad \tau_t^* := t\,\mathbf{1}_{A_t} + \tau_{t+1}^*\,\mathbf{1}_{A_t^c}, \qquad A_t := \{L_t \le c_{\text{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\}.$$

Each τ_t^* is a stopping time, and it is equivalent to the first hitting time

$$\tau_t^* = \inf\{s \in \{t, \dots, T\} : J_s = L_s\}.$$

By construction, on A_t we stop immediately and incur conditional cost $L_t = J_t$. On A_t^c, we continue one step and then use τ_{t+1}^*, so

$$\begin{aligned}\mathbb{E}[c_{\text{step}}(\tau_t^* - t) + L_{\tau_t^*} \mid \mathcal{F}_t] &= \mathbf{1}_{A_t} L_t + \mathbf{1}_{A_t^c}\left(c_{\text{step}} + \mathbb{E}[c_{\text{step}}(\tau_{t+1}^* - (t+1)) + L_{\tau_{t+1}^*} \mid \mathcal{F}_t]\right) \\ &= \mathbf{1}_{A_t} L_t + \mathbf{1}_{A_t^c}\left(c_{\text{step}} + \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]\right) = J_t,\end{aligned}$$

where the second equality uses the induction hypothesis. This proves optimality of τ_t^* for the stopping problem with terminal cost $L_t = g(p_t)$.

Finally, part (a) shows that at every stopping time τ the original answer/abstain problem has conditional Bayes terminal loss $g(p_\tau) = L_\tau$. Therefore the policy $(\tau_1^*, a_{\tau_1^*}^*)$ obtained by combining τ_1^* with the optimal terminal action from part (a) is Bayes-optimal for the original finite-horizon problem. □

A.5 Proof of the non-Markov counterexample

Proof of Theorem 4. Let $H \in \{a, b\}$ satisfy $\mathbb{P}(H = a) = \mathbb{P}(H = b) = 1/2$, and let $Y \sim \text{Bernoulli}(3/4)$ be independent of H. Define

$$Z_1 := H, \qquad Z_2 := \begin{cases} Y & \text{if } H = a, \\ \star & \text{if } H = b, \end{cases}$$

where $\star$ is a constant symbol. Let $\mathcal{F}_t := \sigma(Z_1, \dots, Z_t)$.

At time 1, the predictor reports

$$\hat{Y}_1 = 1, \qquad p_1 = 3/4.$$

At time 2, it reports

$$(\hat{Y}_2, p_2) = \begin{cases} (Y, 1) & \text{if } H = a, \\ (1, 3/4) & \text{if } H = b. \end{cases}$$

Ideal confidence relative to the history. At time 1,

$$\mathbb{P}(Y = \hat{Y}_1 \mid \mathcal{F}_1) = \mathbb{P}(Y = 1 \mid H) = 3/4 = p_1 \qquad \text{a.s.}$$

because Y is independent of H. At time 2, on $\{H = a\}$ we reveal Y exactly, so $\mathbb{P}(Y = \hat{Y}_2 \mid \mathcal{F}_2) = 1 = p_2$. On $\{H = b\}$, the signal $Z_2 = \star$ carries no new information, so

$$\mathbb{P}(Y = \hat{Y}_2 \mid \mathcal{F}_2) = \mathbb{P}(Y = 1 \mid H = b, Z_2 = \star) = 3/4 = p_2.$$

Hence condition (iii) holds at both times.

Optimal stopping decisions. Because $c_{\text{abs}} = +\infty$ and $c_{\text{err}} = 1$, the terminal loss is $g(p) = 1 - p$. Therefore

$$L_1 = g(p_1) = 1/4, \qquad L_2 = g(p_2) = \begin{cases} 0 & \text{if } H = a, \\ 1/4 & \text{if } H = b. \end{cases}$$

By Theorem 3,

$$J_2 = L_2, \qquad J_1 = \min\{L_1, c_{\text{step}} + \mathbb{E}[J_2 \mid \mathcal{F}_1]\}.$$

On $\{H = a\}$, $\mathbb{E}[J_2 \mid \mathcal{F}_1] = 0$, so

$$J_1 = \min\{1/4, c_{\text{step}}\} = c_{\text{step}} \qquad \text{on } \{H = a\},$$

and since $c_{\text{step}} < 1/4$, continuing is strictly optimal. On $\{H = b\}$, $\mathbb{E}[J_2 \mid \mathcal{F}_1] = 1/4$, so

$$J_1 = \min\{1/4, c_{\text{step}} + 1/4\} = 1/4 \qquad \text{on } \{H = b\},$$

and stopping is strictly optimal.

Confidence is not a sufficient statistic. The random variable p_1 is constant almost surely, so $\sigma(p_1)$ is trivial. But

$$J_1 = \begin{cases} c_{\text{step}} & \text{on } \{H = a\}, \\ 1/4 & \text{on } \{H = b\}, \end{cases}$$

which is nonconstant because $c_{\text{step}} \in (0, 1/4)$. Hence J_1 is not $\sigma(p_1)$-measurable, and no representation $J_1 = V_1(p_1)$ is possible. □

A.6 Proof of the confidence-only policy gap

Proof of Proposition 2. Fix $\theta \in (0,1)$, set $c_{\text{step}} := c_{\text{err}}(1-\theta)/2$ and $c_{\text{abs}} := +\infty$. Let $H \in \{a, b\}$ with $\mathbb{P}(H = a) = \mathbb{P}(H = b) = 1/2$, and let $Y \sim \text{Bernoulli}(\theta)$ be independent of H. Reveal $Z_1 := H$ at $t = 1$; at $t = 2$, reveal $Z_2 := Y$ on $\{H = a\}$ and $Z_2 := \star$ on $\{H = b\}$. Reports: $\hat{Y}_1 = 1$, $p_1 = \theta$; $(\hat{Y}_2, p_2) = (Y, 1)$ on $\{H = a\}$ and $(1, \theta)$ on $\{H = b\}$.

Condition (iii) follows by the same argument as in Section A.5, since $Y \perp H$ and the time-2 signal preserves the $H = b$ posterior $\mathbb{P}(Y = 1 \mid H = b) = \theta$.

The total cost (including the mandatory first step) of each policy on each branch is

$$\text{stop at } t = 1\text{: } c_{\text{step}} + c_{\text{err}}(1-\theta), \quad \text{continue on } a\text{: } 2c_{\text{step}}, \quad \text{continue on } b\text{: } 2c_{\text{step}} + c_{\text{err}}(1-\theta).$$

Because $c_{\text{step}} < c_{\text{err}}(1-\theta)$, the full-information Bayes policy continues on a and stops on b, with

$$R^*_{\text{full}} = \tfrac{1}{2}(2c_{\text{step}}) + \tfrac{1}{2}\big(c_{\text{step}} + c_{\text{err}}(1-\theta)\big) = \tfrac{3}{2}c_{\text{step}} + \tfrac{1}{2}c_{\text{err}}(1-\theta) = \tfrac{5}{4}c_{\text{err}}(1-\theta).$$

Since $p_1 = \theta$ is constant, every $(\sigma(p_1, \ldots, p_t))$-adapted policy must commit at $t = 1$ to "always stop" or "always continue":

$$R_{\text{stop}} = c_{\text{step}} + c_{\text{err}}(1-\theta) = \tfrac{3}{2}c_{\text{err}}(1-\theta), \qquad R_{\text{cont}} = 2c_{\text{step}} + \tfrac{1}{2}c_{\text{err}}(1-\theta) = \tfrac{3}{2}c_{\text{err}}(1-\theta).$$

Both equal $\frac{3}{2}c_{\text{err}}(1-\theta)$, so $R^*_{\text{scalar}} = \frac{3}{2}c_{\text{err}}(1-\theta)$ and the gap is $c_{\text{err}}(1-\theta)/4$. Choosing $\theta < 4\delta/c_{\text{err}}$ gives gap $\geq c_{\text{err}}/4 - \delta$.

□

A.7 Proof of the Markov-state characterization

Proof of Theorem 5. **(b) ⇒ (c).** This is immediate by taking the first coordinate of the future path: for every Borel set $B \subseteq [0,1]$,

$$\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = \mathbb{P}\big((p_{t+1}, \ldots, p_T) \in B \times [0,1]^{T-t-1} \mid \mathcal{F}_t\big),$$

which is $\sigma(p_t)$-measurable under (b). Since $[0,1]$ is a standard Borel space, these conditional probabilities admit a Borel kernel representation $K_t(p_t, B)$.

(c) ⇒ (b). Fix $t < T$. The one-step kernels may be composed recursively to obtain a Borel kernel Q_t from $[0,1]$ to $[0,1]^{T-t}$ such that for every bounded Borel function Φ on $[0,1]^{T-t}$,

$$\mathbb{E}[\Phi(p_{t+1}, \ldots, p_T) \mid \mathcal{F}_t] = \int \Phi(\mathbf{p})\, Q_t(p_t, d\mathbf{p}) \qquad \text{a.s.}$$

This means the conditional law of the future confidence path given $\mathcal{F}_t$ depends on the past only through p_t, which is exactly (b).

(c) ⇒ (a). Fix $u \leq T$, a bounded Borel loss ℓ, and $c > 0$. We prove by backward induction on $t \leq u$ that there exist Borel functions $V_t^{u,\ell,c}$ with $J_t^{u,\ell,c} = V_t^{u,\ell,c}(p_t)$.

At time u, we must stop, so $J_u^{u,\ell,c} = \ell(p_u)$ and we may set $V_u^{u,\ell,c} = \ell$. Suppose $J_{t+1}^{u,\ell,c} = V_{t+1}^{u,\ell,c}(p_{t+1})$ for some Borel $V_{t+1}^{u,\ell,c}$. Since the only choice at time t is to stop immediately or continue one step, the full-information Snell-envelope argument gives

$$J_t^{u,\ell,c} = \min\Big\{\ell(p_t),\ c + \mathbb{E}[J_{t+1}^{u,\ell,c} \mid \mathcal{F}_t]\Big\}.$$

Using the kernel property,

$$\mathbb{E}[J_{t+1}^{u,\ell,c} \mid \mathcal{F}_t] = \int_0^1 V_{t+1}^{u,\ell,c}(p')\, K_t(p_t, dp').$$

Therefore

$$J_t^{u,\ell,c} = \min\left\{\ell(p_t),\ c + \int_0^1 V_{t+1}^{u,\ell,c}(p')\, K_t(p_t, dp')\right\} = V_t^{u,\ell,c}(p_t),$$

for the obvious Borel function $V_t^{u,\ell,c}$. This proves (a).

(a) ⇒ (c). Fix $t < T$ and a bounded Borel function $h : [0,1] \to \mathbb{R}$. Let $e_h := \mathbb{E}[h(p_{t+1}) \mid \mathcal{F}_t]$. Apply (a) to the two-time problem with horizon $u = t+1$, terminal loss $\ell = h$, and step cost $c_n := 1/n$. Then

$$J_t^{t+1,h,c_n} = \min\{h(p_t), c_n + e_h\}$$

is $\sigma(p_t)$-measurable for every $n \in \mathbb{N}$. Because $\min\{h(p_t), c_n + e_h\} \to \min\{h(p_t), e_h\}$ pointwise as $n \to \infty$, the limit remains $\sigma(p_t)$-measurable, so

$$m_h := \min\{h(p_t), e_h\}$$

is $\sigma(p_t)$-measurable. Repeating the same argument with terminal loss $-h$ shows that

$$m_{-h} := \min\{-h(p_t), -e_h\} = -\max\{h(p_t), e_h\}$$

is also $\sigma(p_t)$-measurable. Hence

$$e_h = m_h - m_{-h} - h(p_t)$$

is $\sigma(p_t)$-measurable.

Now take $h = \mathbf{1}_B$ for any Borel set $B \subseteq [0,1]$. Then

$$\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = \mathbb{E}[\mathbf{1}_B(p_{t+1}) \mid \mathcal{F}_t]$$

is $\sigma(p_t)$-measurable. By the regular conditional distribution theorem on standard Borel spaces, these conditional probabilities assemble into a Borel kernel K_t satisfying part (c).

Bellman recursion. Under any of the equivalent conditions, the proof of (c)⇒(a) already yields the scalar Bellman recursion for every choice of u, ℓ, c. The argument used only that $[0,1]$ is a standard Borel space (regular conditional distributions, kernel composition, and bounded Borel test functions). The same equivalence therefore holds with $[0,1]$ replaced by any standard Borel space S, in particular $S = \Delta^{|\mathcal{Y}|-1}$ with state π_t. □

A.8 Proof of the i.i.d. sufficient condition

Proof of Proposition 3. Ideal confidence is immediate: $\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t) = \mathbb{P}(Y = 1 \mid \mathcal{F}_t) = p_t$ a.s.

For the kernel condition, choose any common dominating measure ν for μ_0, μ_1 and let $f_y = d\mu_y/d\nu$. By Bayes' theorem,

$$p_{t+1} = \frac{p_t f_1(Z_{t+1})}{p_t f_1(Z_{t+1}) + (1-p_t) f_0(Z_{t+1})} = \phi(p_t, Z_{t+1}),$$

which is a Borel function of (p_t, Z_{t+1}). The assumed conditional independence $Z_{t+1} \perp \mathcal{F}_t \mid Y$ implies

$$\begin{aligned}\mathbb{P}(Z_{t+1} \in dz \mid \mathcal{F}_t) &= \pi_t(1)\,\mathbb{P}(Z_{t+1} \in dz \mid Y = 1) + \pi_t(0)\,\mathbb{P}(Z_{t+1} \in dz \mid Y = 0) \\ &= p_t\,\mu_1(dz) + (1-p_t)\,\mu_0(dz),\end{aligned}$$

which depends on $\mathcal{F}_t$ only through p_t. Hence for every Borel $B \subseteq [0,1]$,

$$\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = \int \mathbf{1}_{\{\phi(p_t,z) \in B\}} \left[p_t\,\mu_1(dz) + (1-p_t)\,\mu_0(dz)\right] = K(p_t, B),$$

which is the stated kernel and is Borel in p_t.

Concavity preservation. We show that $(\mathcal{K}f)(p) := \mathbb{E}_p[f(\phi(p, Z))]$ is concave in p when f is concave (an instance of the classical Blackwell information ordering (Blackwell, 1951; DeGroot, 2004)). For $p_1, p_2 \in [0,1]$, $\lambda \in [0,1]$, and $\bar{p} := \lambda p_1 + (1-\lambda)p_2$, introduce an auxiliary mixture index $J \in \{1,2\}$ with $\mathbb{P}(J=1) = \lambda$; sampling Y from p_J and then $Z \mid Y$ from μ_Y produces the same joint (Y,Z)-distribution as sampling under prior $\bar{p}$. The agent who observes the finer signal (J,Z) has posterior $\phi(p_J, Z)$, and the agent who observes only Z has posterior $\phi(\bar{p}, Z)$. Jensen's inequality applied to the conditional law of $\phi(p_J, Z)$ given Z (whose mean is $\phi(\bar{p}, Z)$ by the tower property) gives, for concave f,

$$\mathbb{E}\big[f(\phi(p_J, Z))\big] \;\le\; \mathbb{E}\big[f(\mathbb{E}[\phi(p_J, Z) \mid Z])\big] \;=\; \mathbb{E}\big[f(\phi(\bar{p}, Z))\big] \;=\; (\mathcal{K}f)(\bar{p}),$$

and the left-hand side equals $\lambda(\mathcal{K}f)(p_1) + (1-\lambda)(\mathcal{K}f)(p_2)$. Hence $(\mathcal{K}f)$ is concave, and Corollary 1 applies to the Bellman recursion induced by K. □

A.9 Proof for conditionally independent correctness signals

Proof of Proposition 4. Part (i) is Theorem 3(a): once $\hat{Y}_\tau$ is fixed, the conditional expected cost of answering is $c_{\mathrm{err}}(1-p_\tau)$ and the cost of abstaining is c_{abs}, for any finite $\mathcal{Y}$.

Part (ii). Ideal confidence: $\hat{Y}_t = \hat{Y}_1$ gives $\{Y = \hat{Y}_t\} = \{C = 1\}$, so $\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t) = \mathbb{P}(C = 1 \mid \mathcal{F}_t) = p_t$ by definition. For the kernel, the assumption gives $Z_{t+1} \perp \mathcal{F}_t \mid C$ with fixed conditional laws ν_c. Note that $\mathcal{F}_t = \sigma(\mathcal{F}_1, Z_2, \ldots, Z_t)$ and the signals are i.i.d. given C and independent of $\mathcal{F}_1$ given C. Hence

$$\mathbb{P}(Z_{t+1} \in dz \mid \mathcal{F}_t) = p_t \nu_1(dz) + (1-p_t)\nu_0(dz).$$

Let ν dominate ν_0 and ν_1, and write $f_c = d\nu_c/d\nu$. Bayes' rule yields

$$p_{t+1} = \frac{p_t f_1(Z_{t+1})}{p_t f_1(Z_{t+1}) + (1-p_t) f_0(Z_{t+1})} = \phi(p_t, Z_{t+1})$$

almost surely on the predictive support. These two identities are exactly those used in the proof of Proposition 3 with (Y, μ_y) replaced by (C, ν_c). The remainder of that proof, including concavity preservation via the Blackwell ordering, applies verbatim. Thus Assumption 1(iii) holds with the stationary kernel

$$K(p, B) = \int \mathbf{1}_{\{\phi(p,z) \in B\}} \, [p\, \nu_1(dz) + (1-p)\, \nu_0(dz)],$$

and $K(1, \{1\}) = 1$ because $p = 1$ forces $C = 1$ almost surely. Thus Theorem 6 applies with Markov state p_t for any $|\mathcal{Y}|$, and Corollary 1 applies whenever $c_{\mathrm{abs}} = +\infty$.

Part (iii). For every Borel $B \subseteq [0,1]$, the Markov property of π_t and the stated pushforward condition give

$$\begin{aligned}\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) &= Q_t\big(\pi_t, m_{t+1}^{-1}(B)\big) \\ &= K_t\big(m_t(\pi_t), B\big) = K_t(p_t, B).\end{aligned}$$

Thus Assumption 1(iii) holds. If the candidate is allowed to switch, the correctness event $\{Y = \hat{Y}_t\}$ changes with t, so the binary reduction in part (ii) need not apply. □

A.10 Proofs for the Markov confidence model

Proof of Theorem 6. By Theorem 3, there is a value process (J_t) satisfying (8). We show by backward induction that J_t factors through p_t.

At time T, define $V_T := g$, so $J_T = V_T(p_T)$. Suppose $J_{t+1} = V_{t+1}(p_{t+1})$ for some Borel function V_{t+1}. Then the kernel assumption gives

$$\mathbb{E}[J_{t+1} \mid \mathcal{F}_t] = \mathbb{E}[V_{t+1}(p_{t+1}) \mid \mathcal{F}_t] = \int_0^1 V_{t+1}(p')\, K_t(p_t, dp').$$

Hence

$$J_t = \min\left\{ g(p_t),\; c_{\mathrm{step}} + \int_0^1 V_{t+1}(p')\, K_t(p_t, dp') \right\} = V_t(p_t),$$

where V_t is exactly the Bellman recursion stated in the theorem. This proves (5).

The optimal stopping time from Theorem 3 therefore becomes

$$\tau^* = \inf\{t \le T : V_t(p_t) = g(p_t)\}.$$

The optimal terminal action at τ^* is the canonical selector $a^\dagger(p_{\tau^*})$ from Theorem 3(a). Because the first observation must be made before any stopping decision is possible, the total Bayes risk equals $c_{\text{step}} + \mathbb{E}[J_1] = c_{\text{step}} + \mathbb{E}[V_1(p_1)]$. □

Proof of Corollary 1. We prove by backward induction that each V_t is concave. The base case $V_T(p) = c_{\text{err}}(1-p)$ is linear and therefore concave. Assume V_{t+1} is concave and define

$$W_t(p) := c_{\text{step}} + \int_0^1 V_{t+1}(p')\, K_t(p, dp').$$

By assumption, W_t is concave. Since $g(p) = c_{\text{err}}(1-p)$ is also concave and

$$V_t(p) = \min\{g(p), W_t(p)\},$$

the function V_t is concave because the hypograph of $\min\{g, W_t\}$ is the intersection of the two convex hypographs.

Now define $\phi_t(p) := g(p) - W_t(p)$. Since g is linear and W_t is concave, ϕ_t is convex. The stopping set is $S_t := \{p \in [0,1] : g(p) \le W_t(p)\} = \{p : \phi_t(p) \le 0\}$, hence S_t is a convex subset of $[0,1]$.

It remains to show that $1 \in S_t$. First, $V_T(1) = 0$. If $V_{t+1}(1) = 0$, then certainty preservation gives

$$W_t(1) = c_{\text{step}} + \int_0^1 V_{t+1}(p')\, K_t(1, dp') = c_{\text{step}} + V_{t+1}(1) = c_{\text{step}}.$$

Since $g(1) = 0$, we have $\phi_t(1) = -c_{\text{step}} < 0$, so $1 \in S_t$. Thus S_t is a convex subset of $[0,1]$ containing 1, so there exists $\beta_t \in [0,1]$ such that either $S_t = [\beta_t, 1]$ or $S_t = (\beta_t, 1]$. Therefore stopping is optimal for $p > \beta_t$, continuation is optimal for $p < \beta_t$, and only a boundary tie can remain at $p = \beta_t$. □

Proof of Proposition 5. We first observe $\|g\|_\infty \le c_{\text{err}}$: if $c_{\text{abs}} < \infty$ then $g(p) \le c_{\text{abs}} < c_{\text{err}}$, and if $c_{\text{abs}} = +\infty$ then $g(p) = c_{\text{err}}(1-p) \le c_{\text{err}}$. Hence both J_t and $\widetilde{V}_t$ take values in $[0, c_{\text{err}}]$.

We prove the bound by backward induction on t. At $t = T$, $J_T = g(p_T) = \widetilde{V}_T(p_T)$, so the bound holds with constant 0.

Assume $|J_{t+1} - \widetilde{V}_{t+1}(p_{t+1})| \le 2(T-t-1)\eta\, c_{\text{err}}$ a.s. Writing $A := \mathbb{E}[J_{t+1} \mid \mathcal{F}_t]$, $B := \mathbb{E}[\widetilde{V}_{t+1}(p_{t+1}) \mid \mathcal{F}_t]$, and $C := \int \widetilde{V}_{t+1}(p')\, \widetilde{K}_t(p_t, dp')$,

$$\left|J_t - \widetilde{V}_t(p_t)\right| \le |A - C| \le |A - B| + |B - C|.$$

The first term is at most $\mathbb{E}[|J_{t+1} - \widetilde{V}_{t+1}(p_{t+1})| \mid \mathcal{F}_t] \le 2(T-t-1)\eta\, c_{\text{err}}$ by the inductive hypothesis. The second term is bounded by $2c_{\text{err}}\,\eta$ via the standard total-variation duality $|\int f\, d\mu - \int f\, d\nu| \le 2\|f\|_\infty \|\mu - \nu\|_{\text{TV}}$, with $\|\mu - \nu\|_{\text{TV}} := \sup_A |\mu(A) - \nu(A)|$, with $\|\widetilde{V}_{t+1}\|_\infty \le c_{\text{err}}$. Summing gives $|J_t - \widetilde{V}_t(p_t)| \le 2(T-t)\eta\, c_{\text{err}}$, completing the induction. The Bayes-risk gap follows by taking expectations at $t = 1$. □

A.11 Proof of the reported-output separation theorem

Proof of Theorem 7. Let $H \in \{a, b\}$ with $\mathbb{P}(H = a) = \mathbb{P}(H = b) = 1/2$, and define

$$\mathbb{P}(Y = 1 \mid H = a) = 0.6, \qquad \mathbb{P}(Y = 1 \mid H = b) = 0.9.$$

Set $Z_1 := H$, so $\mathcal{F}_1 = \sigma(H)$, and let $Z_2 := Y$, so time 2 reveals perfect information.

At time 1, the predictor always reports

$$\hat{Y}_1 = 1, \qquad p_1 = 0.75.$$

At time 2, it reports

$$\hat{Y}_2 = Y, \qquad p_2 = 1.$$

Take $c_{\text{step}} = 0.25$, $c_{\text{err}} = 1$, and $c_{\text{abs}} = +\infty$.

Confidence calibration at each fixed time. At time 1,

$$\mathbb{P}(Y = \hat{Y}_1 \mid p_1 = 0.75) = \mathbb{P}(Y = 1) = \tfrac{1}{2} \cdot 0.6 + \tfrac{1}{2} \cdot 0.9 = 0.75.$$

At time 2, $\mathbb{P}(Y = \hat{Y}_2 \mid p_2 = 1) = 1$. Hence confidence is calibrated at $t = 1, 2$.

Failure of ideality relative to the history. Because $\mathcal{F}_1 = \sigma(H)$,

$$\mathbb{P}(Y = \hat{Y}_1 \mid \mathcal{F}_1) = \begin{cases} 0.6 & \text{on } \{H = a\}, \\ 0.9 & \text{on } \{H = b\}, \end{cases}$$

which is not equal to the constant $p_1 = 0.75$. Thus condition (iii) fails.

Bayes-optimal policy under the full filtration. If $H = a$, stopping at time 1 costs $0.25 + (1 - 0.6) = 0.65$, while continuing to time 2 costs 0.5. Thus it is optimal to continue. If $H = b$, stopping at time 1 costs $0.25 + (1 - 0.9) = 0.35$, while continuing costs 0.5. Thus it is optimal to stop. The full-information Bayes risk is therefore

$$R^* = \tfrac{1}{2} \cdot 0.5 + \tfrac{1}{2} \cdot 0.35 = 0.425.$$

Any policy based only on reported outputs is suboptimal. The reported pair $(\hat{Y}_1, p_1)$ is constant, so $\mathcal{G}_1 = \sigma((\hat{Y}_1, p_1))$ is trivial. Any $\mathcal{G}_t$-adapted policy must therefore choose one of two deterministic actions at time 1:

(i) stop immediately, which costs $0.25 + \mathbb{P}(Y \neq 1) = 0.25 + 0.25 = 0.5$;
(ii) continue to time 2, which costs 0.5 because the answer is then always correct.

Any exogenous randomization only mixes these same two costs and therefore also achieves expected cost 0.5. Hence every $\mathcal{G}_t$-adapted policy has expected cost at least 0.5. The suboptimality gap is $0.5 - 0.425 = 0.075$. □

Proof of Proposition 6. Parametrize the construction of Theorem 7 by $\delta \in (0, 1/4)$: let $H \in \{a, b\}$ be equiprobable with $\mathbb{P}(Y = 1 \mid H = a) = q_a := 3/4 - \delta$ and $\mathbb{P}(Y = 1 \mid H = b) = q_b := 3/4 + \delta$. Take $\mathcal{F}_1 := \sigma(H)$, $\mathcal{F}_2 := \sigma(H, Y)$, and the same reports as in Theorem 7: $\hat{Y}_1 = 1$, $p_1 = 3/4$, $\hat{Y}_2 = Y$, $p_2 = 1$. Take $c_{\text{step}} := c_{\text{err}}/4$ and $c_{\text{abs}} = +\infty$.

Confidence calibration at each fixed time. At $t = 1$, $\mathbb{P}(Y = \hat{Y}_1) = (q_a + q_b)/2 = 3/4 = p_1$; at $t = 2$, $\mathbb{P}(Y = \hat{Y}_2 \mid p_2 = 1) = 1$.

Full-information Bayes risk. On $\{H = a\}$, stopping costs $c_{\text{step}} + c_{\text{err}}(1 - q_a) = c_{\text{err}}/2 + c_{\text{err}}\delta$, while continuing costs $2c_{\text{step}} = c_{\text{err}}/2$, so continuation is strictly optimal. On $\{H = b\}$, stopping costs $c_{\text{step}} + c_{\text{err}}(1 - q_b) = c_{\text{err}}/2 - c_{\text{err}}\delta$, while continuing costs $2c_{\text{step}} = c_{\text{err}}/2$, so stopping is strictly optimal. Hence

$$R^*_{\text{full}} = \tfrac{1}{2}(2c_{\text{step}}) + \tfrac{1}{2}\big(c_{\text{step}} + c_{\text{err}}(1 - q_b)\big) = \tfrac{c_{\text{err}}}{2} - \tfrac{c_{\text{err}}\delta}{2}.$$

Reported-output Bayes risk. Because $(\hat{Y}_1, p_1)$ is constant, every $\mathcal{G}_t$-adapted policy must commit at $t = 1$ to "always stop" or "always continue":

$$R_{\text{stop}} = c_{\text{step}} + c_{\text{err}}\,\mathbb{P}(Y \neq 1) = \tfrac{c_{\text{err}}}{4} + \tfrac{c_{\text{err}}}{4} = \tfrac{c_{\text{err}}}{2}, \qquad R_{\text{cont}} = 2c_{\text{step}} = \tfrac{c_{\text{err}}}{2},$$

so $R^*_{\text{scalar}} = c_{\text{err}}/2$. The gap is $R^*_{\text{scalar}} - R^*_{\text{full}} = c_{\text{err}}\delta/2$. □

A.12 Proof of the approximate-probability regret bound

Proof of Theorem 8. We record three elementary bounds.

Step 1: g is c_{err}-Lipschitz. For any $p, q \in [0, 1]$, $|g(p) - g(q)| \leq c_{\text{err}}|p - q|$, because $p \mapsto c_{\text{err}}(1 - p)$ is c_{err}-Lipschitz and the pointwise minimum of a c_{err}-Lipschitz function and a constant is c_{err}-Lipschitz.

Step 2: a sharp comparison of $\tilde{g}$ to g. Let $\tilde{g}(p, p^*)$ denote the expected terminal cost when any Bayes action under the reported score p is evaluated under the true posterior p^*. We claim

$$\tilde{g}(p, p^*) \le g(p) + c_{\mathrm{err}}|p - p^*| \qquad \text{for all } p, p^* \in [0, 1]. \tag{9}$$

If $c_{\mathrm{abs}} = +\infty$, then $\tilde{g}(p, p^*) = c_{\mathrm{err}}(1 - p^*)$ and $g(p) = c_{\mathrm{err}}(1 - p)$, so $\tilde{g}(p, p^*) - g(p) = c_{\mathrm{err}}(p - p^*) \le c_{\mathrm{err}}|p - p^*|$. If $c_{\mathrm{abs}} < \infty$ and $p > p^\dagger$, the unique Bayes action is to answer, so $\tilde{g}(p, p^*) = c_{\mathrm{err}}(1 - p^*)$ and $g(p) = c_{\mathrm{err}}(1 - p)$, giving the same bound. This is the action-mismatch branch that makes the constant 2 tight when $p^* < p^\dagger$. If $c_{\mathrm{abs}} < \infty$ and $p < p^\dagger$, the unique Bayes action is to abstain, and $\tilde{g}(p, p^*) = c_{\mathrm{abs}} = g(p)$, so the difference is zero. At $p = p^\dagger$, the answer and abstain branches both satisfy (9); any randomized Bayes action is their convex combination and therefore satisfies it as well.

Step 3: Lipschitz transfer between true and surrogate objectives. For a stopping time $\tau \le T$, define

$$J(\tau) := \mathbb{E}[c_{\mathrm{step}}\tau + g(p^*_\tau)], \qquad \hat{J}(\tau) := \mathbb{E}[c_{\mathrm{step}}\tau + g(p_\tau)].$$

By Step 1 and the uniform approximation assumption,

$$|J(\tau) - \hat{J}(\tau)| \le c_{\mathrm{err}}\varepsilon \qquad \text{for every stopping time } \tau \le T. \tag{10}$$

Putting the pieces together. Because $\hat{\tau}$ minimizes $\hat{J}$,

$$\hat{J}(\hat{\tau}) \le \hat{J}(\tau^*). \tag{11}$$

By Theorem 3(a) we may assume that $\pi^* = (\tau^*, a^*)$ uses a Bayes-optimal terminal action at τ^*, so $\mathbb{E}[\mathcal{C}(\pi^*)] = J(\tau^*)$. The actual cost of $\hat{\pi}$ is

$$\begin{aligned}
\mathbb{E}[\mathcal{C}(\hat{\pi})] &= \mathbb{E}[c_{\mathrm{step}}\hat{\tau} + \tilde{g}(p_{\hat{\tau}}, p^*_{\hat{\tau}})] \\
&\le \mathbb{E}[c_{\mathrm{step}}\hat{\tau} + g(p_{\hat{\tau}})] + c_{\mathrm{err}}\varepsilon && \text{by (9)} \\
&= \hat{J}(\hat{\tau}) + c_{\mathrm{err}}\varepsilon \\
&\le \hat{J}(\tau^*) + c_{\mathrm{err}}\varepsilon && \text{by (11)} \\
&\le J(\tau^*) + 2c_{\mathrm{err}}\varepsilon && \text{by (10)} \\
&= \mathbb{E}[\mathcal{C}(\pi^*)] + 2c_{\mathrm{err}}\varepsilon.
\end{aligned}$$

This proves (6). The improvement over the naive three-term bound comes from (9), which compares $\tilde{g}$ directly to $g(p)$ (the surrogate-evaluated cost) rather than routing through $g(p^*)$ and incurring a separate Lipschitz term. □

Proof of Proposition 7. Fix $\zeta \in (0, 1)$ and choose $d \in (0, \zeta\varepsilon]$. Take $T = 1$, trivial $\mathcal{F}_1$, $\hat{Y}_1 \equiv 1$, and

$$p_1 := p^\dagger + d, \qquad Y \sim \mathrm{Bernoulli}(q), \qquad q := p^*_1 = p^\dagger + d - \varepsilon.$$

The assumption $p^\dagger \in (\varepsilon, 1 - \varepsilon)$ and the choice $0 < d < \varepsilon$ ensure that $p_1, q \in (0, 1)$. Moreover, $|p_1 - p^*_1| = \varepsilon$, $p_1 > p^\dagger$, and $p^*_1 < p^\dagger$. Thus every Bayes selector under the reported score strictly answers, while the true Bayes action strictly abstains. Using $c_{\mathrm{abs}} = c_{\mathrm{err}}(1 - p^\dagger)$, the regret is

$$c_{\mathrm{err}}(1 - p^*_1) - c_{\mathrm{abs}} = c_{\mathrm{err}}(\varepsilon - d) \ge (1 - \zeta)c_{\mathrm{err}}\varepsilon.$$

□

Proof of Proposition 8. Fix $c_{\mathrm{err}} > 0$, $\eta \in (0, 2)$, and $\kappa > 1$. Choose $\varepsilon > 0$ so that $\varepsilon < 1/(2(\kappa + 1))$ and $(\kappa + 1)\varepsilon < \eta/2$, then choose d with $0 < d < \min\{\varepsilon, (\kappa - 1)\varepsilon/2\}$. Set

$$c_{\mathrm{abs}} := \kappa c_{\mathrm{err}}\varepsilon, \qquad p^\dagger := 1 - \kappa\varepsilon, \qquad p_1 := p^\dagger + d, \qquad q := p^\dagger + d - \varepsilon.$$

Let $\mathcal{F}_1$ be trivial, $\mathcal{F}_2 := \sigma(Y)$, $\hat{Y}_1 = \hat{Y}_2 \equiv 1$, and $Y \sim \mathrm{Bernoulli}(q)$. At time 2, report $p_2 = 1 - \varepsilon$ when $Y = 1$ and $p_2 = \varepsilon$ when $Y = 0$. Then $p^*_1 = q$, $p^*_2 = Y$, and $\sup_{t \le 2} |p_t - p^*_t| = \varepsilon$. The chosen range gives $q > 1/2$ and $q < p^\dagger$, so the unique true terminal action at time 1 is to abstain, whereas every Bayes selector under p_1 answers.

Define $L := q(c_{\mathrm{abs}} - c_{\mathrm{err}}\varepsilon) - c_{\mathrm{err}}d$. Since $q\varepsilon(\kappa - 1) > d$, we have $L > 0$, and $qc_{\mathrm{abs}} - L = qc_{\mathrm{err}}\varepsilon + c_{\mathrm{err}}d > 0$. The surrogate stop-now cost is $c_{\mathrm{step}} + c_{\mathrm{abs}} - c_{\mathrm{err}}d$, while its continuation cost is $2c_{\mathrm{step}} + qc_{\mathrm{err}}\varepsilon + (1-q)c_{\mathrm{abs}}$; hence it strictly stops when $c_{\mathrm{step}} > L$. The true stop-now and continuation costs are $c_{\mathrm{step}} + c_{\mathrm{abs}}$ and $2c_{\mathrm{step}} + (1-q)c_{\mathrm{abs}}$, respectively, so the true policy strictly continues when $c_{\mathrm{step}} < qc_{\mathrm{abs}}$. Thus the interval (L, qc_{abs}) is nonempty and all decisions are strict.

Choose c_{step} in that interval with $0 < c_{\mathrm{step}} - L < \eta c_{\mathrm{err}}\varepsilon/2$. The regret is

$$[c_{\mathrm{step}} + c_{\mathrm{err}}(1-q)] - [2c_{\mathrm{step}} + (1-q)c_{\mathrm{abs}}] = (1-q)(c_{\mathrm{err}} - c_{\mathrm{abs}}) - c_{\mathrm{step}}.$$

At $c_{\mathrm{step}} = L$ this expression equals $c_{\mathrm{err}}\{2\varepsilon - (\kappa+1)\varepsilon^2 + \varepsilon d\}$. Our choices therefore make the actual regret greater than $(2-\eta)c_{\mathrm{err}}\varepsilon$. No equality case is used at either the terminal or stopping decision. □

A.13 Adapted step costs

We outline the extension to non-uniform step costs. The constant c_{step} in (1) may be replaced by any $\mathcal{F}_t$-adapted nonnegative cost c_t with $\mathbb{E}[\sum_{t=1}^{T} c_t] < \infty$. Theorems 1 and 2 and Proposition 1 are unchanged, because they concern only the terminal answer/abstain problem at a fixed stopping time. The Snell recursion of Theorem 3 becomes $J_t = \min\{g(p_t), \mathbb{E}[c_{t+1} + J_{t+1} \mid \mathcal{F}_t]\}$, now representing

$$J_t = \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_t} \mathbb{E}\Big[\sum_{s=t}^{\tau-1} c_{s+1} + L_\tau \;\Big|\; \mathcal{F}_t\Big].$$

The scalar recursion of Theorem 6 extends when the conditional law of (c_{t+1}, p_{t+1}) given $\mathcal{F}_t$ factors through the current confidence; in particular, the earlier form is recovered if c_{t+1} is a Borel function of p_t. The standing main-text results use uniform step cost; a discounted infinite-horizon variant is recorded next.

A.14 Discounted infinite-horizon extension

We outline the extension of Theorems 3 and 6 to a stationary discounted infinite-horizon variant. As in the finite-horizon Bellman recursion, this appendix studies the continuation problem after the current confidence has already been revealed; from the start of the overall sequential process, one adds the mandatory first-step cost separately. Fix a discount factor $\gamma \in (0,1)$ and a stationary Borel kernel K on $[0,1]$. Allow stopping times to take values in $\{1, 2, \ldots\} \cup \{\infty\}$. For $\tau \geq t$, define the discounted terminal-optimized objective

$$D_t(\tau) := \mathbb{E}\left[\sum_{s=t}^{\tau-1} \gamma^{s-t} c_{\mathrm{step}} + \mathbf{1}_{\{\tau<\infty\}} \gamma^{\tau-t} g(p_\tau) \;\middle|\; \mathcal{F}_t\right].$$

On $\{\tau = \infty\}$, the sum is the infinite discounted sum and the terminal term is zero. For finite τ, ideal confidence makes $g(p_\tau)$ the conditional Bayes terminal loss.

Proposition 11 (Discounted Bellman equation)**.** *Assume condition (iii) of Definition 1 and that $\mathbb{P}(p_{t+1} \in B \mid \mathcal{F}_t) = K(p_t, B)$ for every Borel $B \subseteq [0,1]$ and every $t \geq 1$. Then the operator*

$$(\mathcal{T}f)(p) := \min\left\{g(p),\ c_{\mathrm{step}} + \gamma \int_0^1 f(p')\, K(p, dp')\right\}$$

is a γ-contraction on $\mathcal{B}_b([0,1])$, the Banach space of bounded Borel functions equipped with the sup norm. Its unique fixed point V_γ satisfies, for every $t \geq 1$,

$$V_\gamma(p_t) = \operatorname*{ess\,inf}_{\tau \geq t} D_t(\tau) \quad \textit{a.s.},$$

where the infimum includes the value $\tau = \infty$. The rule $\tau_t^* := \inf\{s \geq t : V_\gamma(p_s) = g(p_s)\}$, with the infimum of the empty set defined as ∞, attains it. On $\{\tau_t^* < \infty\}$, the canonical terminal action $a^\dagger(p_{\tau_t^*})$ is Bayes-optimal.

Proof. For $f, h \in \mathcal{B}_b([0,1])$ and $p \in [0,1]$, the elementary identity $|\min(a,x) - \min(a,y)| \leq |x-y|$ gives

$$|\mathcal{T}f(p) - \mathcal{T}h(p)| \leq \gamma \Big|\int (f-h)\, dK(p,\cdot)\Big| \leq \gamma \|f-h\|_\infty,$$

since $K(p,\cdot)$ is a probability measure. Hence $\mathcal{T}$ is a γ-contraction on a Banach space, and Banach's fixed-point theorem yields a unique fixed point $V_\gamma \in \mathcal{B}_b([0,1])$.

Write $V := V_\gamma$. The fixed-point equation gives both $V(p_s) \le g(p_s)$ and

$$V(p_s) \le c_{\text{step}} + \gamma \mathbb{E}[V(p_{s+1}) \mid \mathcal{F}_s].$$

Iterating these inequalities up to $\tau \wedge (n+1)$ shows that, for every extended stopping time $\tau \ge t$,

$$V(p_t) \le \mathbb{E}\Big[\sum_{s=t}^{(\tau\wedge(n+1))-1} \gamma^{s-t} c_{\text{step}} + \mathbf{1}_{\{\tau \le n\}} \gamma^{\tau-t} g(p_\tau) + \mathbf{1}_{\{\tau > n\}} \gamma^{n+1-t} V(p_{n+1}) \;\Big|\; \mathcal{F}_t \Big].$$

Because V is bounded and $\gamma < 1$, the last term vanishes as $n \to \infty$, proving $V(p_t) \le D_t(\tau)$. For $\tau = \tau_t^*$, the continuation inequality is an equality before stopping and $V(p_{\tau_t^*}) = g(p_{\tau_t^*})$ at every finite stop. The same limiting argument therefore gives equality, proving optimality and the value identity. Finally, condition (iii) and the stopped-event decomposition used in Theorem 1 identify $g(p_{\tau_t^*})$ as the Bayes terminal loss on every finite stopping event, attained by $a^\dagger(p_{\tau_t^*})$. □

Corollary 2 (Algorithmic convergence of value iteration)**.** *Let V_γ be the unique fixed point of $\mathcal{T}$ in Proposition 11, and define the value-iteration sequence $V_0 := g$ and $V_{k+1} := \mathcal{T}V_k$. Then for every $k \ge 0$,*

$$\|V_k - V_\gamma\|_\infty \;\le\; \gamma^k \|V_0 - V_\gamma\|_\infty \;\le\; \gamma^k \cdot c_{\text{err}}.$$

In particular, for $\varepsilon \in (0, c_{\text{err}}]$, $\lceil \log(c_{\text{err}}/\varepsilon) / \log(1/\gamma) \rceil$ iterations suffice to compute V_γ to sup-norm accuracy ε.

Proof. The contraction inequality $\|\mathcal{T}f - \mathcal{T}h\|_\infty \le \gamma \|f - h\|_\infty$ established in the proof of Proposition 11 iterates to $\|V_k - V_\gamma\|_\infty \le \gamma^k \|V_0 - V_\gamma\|_\infty$. Since $0 \le V_\gamma(p) \le g(p) \le c_{\text{err}}$ (stopping immediately is always feasible and $\|g\|_\infty \le c_{\text{err}}$ as in the proof of Proposition 5) and $V_0 = g \in [0, c_{\text{err}}]$, both functions take values in $[0, c_{\text{err}}]$, giving $\|V_0 - V_\gamma\|_\infty \le c_{\text{err}}$. Solving $\gamma^k c_{\text{err}} \le \varepsilon$ yields the iteration count. □

A.15 Worked numerical examples

Two finite worked examples help fix the decision-theoretic intuition behind Theorems 2 and 4.

Example: recoverable decisions without score equality. Fix a stopping time τ and suppose the conditional probability that the current answer is correct given $\mathcal{F}_\tau$ is

$$r_\tau = \begin{cases} 0.6, & \text{on } E_1, \\ 0.9, & \text{on } E_2, \end{cases} \qquad \mathbb{P}(E_1) = \mathbb{P}(E_2) = 1/2.$$

Define the reported score $s_\tau := r_\tau^2$, so $s_\tau \in \{0.36, 0.81\}$. For any reject cost with $c_{\text{err}} = 1$, the Bayes rule answers iff $r_\tau \ge 1 - c_{\text{abs}}$. Because $r_\tau = \sqrt{s_\tau}$ on the support, the same rule can be implemented from s_τ alone for every reject cost. For example, if $c_{\text{abs}} = 0.25$, then the answer threshold is $r_\tau \ge 0.75$, so one abstains on E_1 and answers on E_2; equivalently, one thresholds s_τ at $0.75^2 = 0.5625$. Thus s_τ supports Bayes-optimal terminal answer/abstain decisions for every reject cost even though $s_\tau \ne r_\tau$; the additional condition $\mathbb{P}(Y = \hat{Y}_\tau \mid s_\tau) = s_\tau$ in Theorem 2 is what upgrades recoverability to equality.

Same current confidence, opposite continuation values. In the construction of Theorem 4, set $c_{\text{step}} = 0.1$, $c_{\text{err}} = 1$, and $c_{\text{abs}} = +\infty$. At time 1 both histories report $(\hat{Y}_1, p_1) = (1, 0.75)$, so stopping immediately has Bayes cost $g(0.75) = 0.25$. On branch a, continuing reveals Y perfectly at time 2, so the continuation cost is $0.1 + 0 = 0.1$. On branch b, continuing reveals no new information and keeps $p_2 = 0.75$, so the continuation cost is $0.1 + 0.25 = 0.35$. Hence continuing is optimal on branch a while stopping is optimal on branch b, despite the common current confidence $p_1 = 0.75$.

Table 1: Concrete numerical instance of the non-Markov phenomenon from Theorem 4.

History	Current score p_1	Stop now	Continue
a	0.75	0.25	0.10
b	0.75	0.25	0.35

Table 2: Selected theorem instances for the synthetic numerical illustrations.

Setting	Parameters	Bayes risk	Restricted/surrogate risk
Fixed-time calibration gap	$\delta = 0.15$	0.425	0.500
Ideal confidence, non-Markov	$c_{\text{step}} = 0.125$	0.3125	0.3750
Approx. regret	$\varepsilon = 0.10$	0.500	0.590

A.16 Synthetic numerical illustrations

We instantiate the finite constructions underlying Theorems 4, 7 and 8 and Propositions 6 and 7 and compare Monte Carlo estimates with the exact Bayes-risk formulas. These experiments are not benchmarks of learned estimators; they confirm that the predicted phenomena appear numerically with the claimed slopes and magnitudes. Selected theorem instances and the corresponding Bayes-risk numbers are summarized in Table 2.

All plots are produced by the accompanying supplementary code. Each marker averages 10 random seeds with 100,000 trajectories per seed; error bars show 95% confidence intervals across seeds, while solid lines show the exact theory. All runs use standard CPU-based Python 3.12 with NumPy and Matplotlib; the full synthetic script runs in tens of seconds on a laptop-class CPU.

Figure 1 reports the three canonical numerical checks. In all panels, markers show Monte Carlo means, solid lines show the exact formulas, and error bars show 95% confidence intervals across seeds. Panel (A) sweeps the heterogeneity parameter δ in the family from Theorem 7. At time 1, the reported pair is always $(\hat{Y}_1, p_1) = (1, 0.75)$, so confidence is calibrated at that fixed time, but the two histories have conditional probabilities $0.75 - \delta$ and $0.75 + \delta$. The exact Bayes risk of the full-information policy is $0.5 - \delta/2$, whereas the best policy measurable with respect to the reported outputs remains at risk 0.5.

Panel (B) sweeps the step cost in the ideal-confidence construction from Theorem 4. Again $p_1 = 0.75$ on both histories, but the optimal full-information value J_1 is c_{step} on one branch and 0.25 on the other. The plot visualizes the resulting gap between the unrestricted Bayes policy and the best policy measurable with respect to the scalar confidence alone.

Panel (C) uses a tie-free one-step member of the family in Proposition 7: for $\varepsilon \in (0, 0.4]$, let $p^\dagger = 1/2$, $p_1 = 1/2 + \varepsilon^2$, and $p_1^* = 1/2 + \varepsilon^2 - \varepsilon$. The surrogate strictly answers and the true Bayes rule strictly abstains, yielding regret $\varepsilon - \varepsilon^2$. The origin is included for reference. The empirical curve checks the linear lower-bound rate without relying on a tie; the sharper two-step construction of Proposition 8 approaches the general upper bound 2ε from Theorem 8.

Empirical validation of the extension theorems. Figures 2 and 3 validate Propositions 2 and 11 and Corollary 2 on the binary symmetric channel from Proposition 3, and visualize the regime in which the surrogate stops while the true problem continues, which drives Proposition 8.

Panel (D) sweeps $\theta \in [0, 0.9]$ in the parametrized family of Proposition 2 with $c_{\text{step}} = c_{\text{err}}(1-\theta)/2$: the empirical confidence-only gap $R^*_{\text{scalar}} - R^*_{\text{full}}$ matches the theory $(1-\theta)/4$ exactly and approaches the supremum $c_{\text{err}}/4$ as $\theta \downarrow 0$. Panel (E) computes the optimal stopping threshold $\beta(\gamma)$ of the discounted Bellman fixed point of Proposition 11 via Banach iteration (signal accuracy 0.7, $c_{\text{step}} = 0.02$, no abstention); β decreases monotonically with γ. Panel (F) plots the value functions $V_\gamma(p)$ for $\gamma \in \{0.5, 0.7, 0.85, 0.95\}$, showing how the continuation region (where $V_\gamma(p) < g(p)$) widens and the threshold $\beta(\gamma)$ moves leftward as γ grows.

Panel (G) reports the per-iteration sup-norm gap of value iteration $\|V_k - V_\gamma\|_\infty$, with the theoretical envelope γ^k from Corollary 2 (dashed) closely tracking the empirical rate. Panel (H) shows the empirical surrogate-policy regret over a 2-D $(\varepsilon, c_{\text{step}})$ sweep in a 2-step family with $c_{\text{abs}} = 1/2$ (so

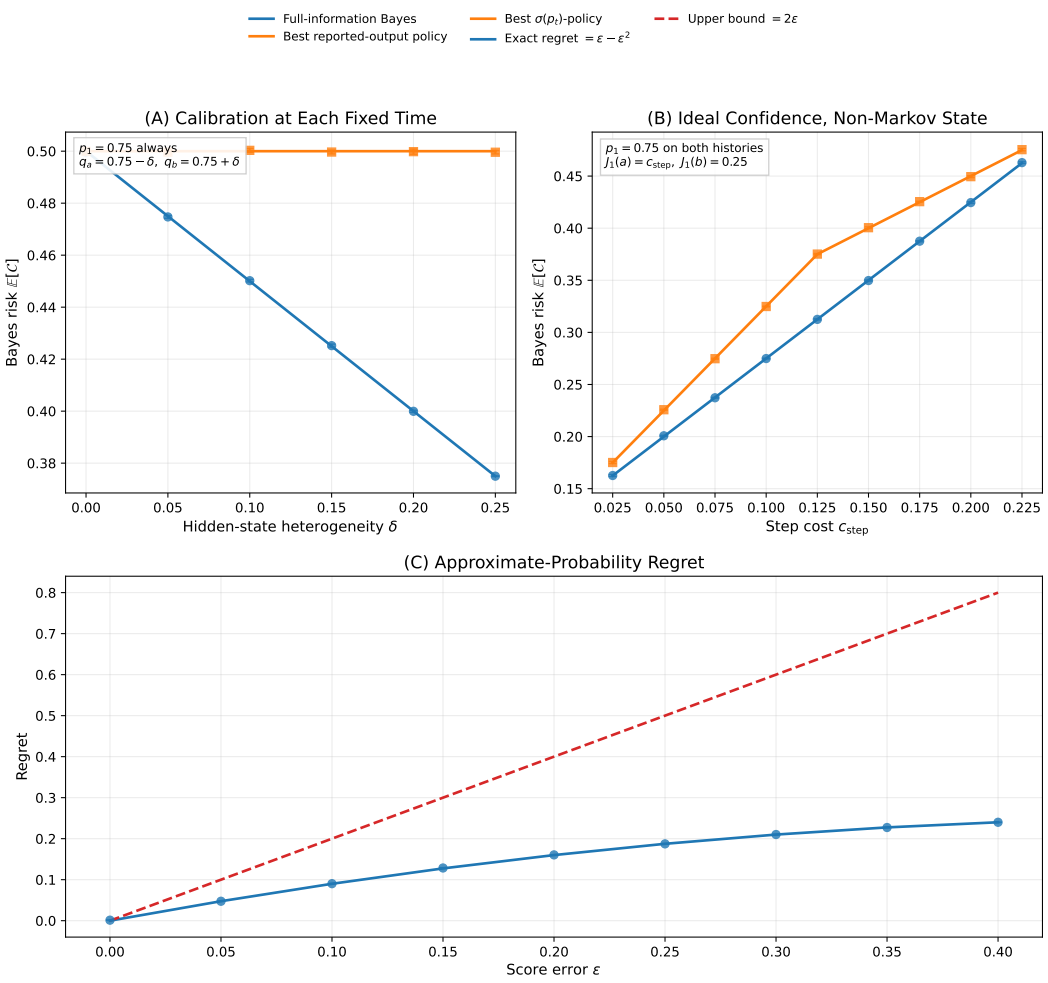

Figure 1: Synthetic numerical checks for three illustrative finite families: reports calibrated at each fixed time but heterogeneous across histories (A), ideal confidence that is not a sufficient statistic for stopping (B), and the approximate-probability regret slope (C).

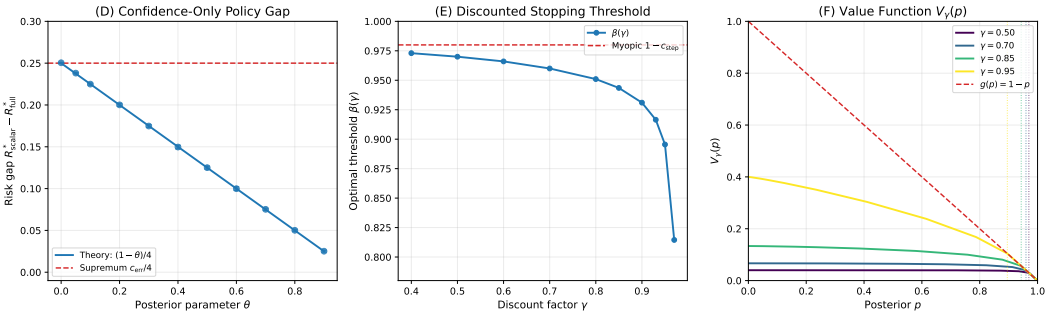

Figure 2: Diagnostics for the Markov confidence model and discounted dynamic program: the confidence-only policy gap in the binary channel (D), the stopping threshold as the discount factor varies (E), and the corresponding value functions (F).

$p^\dagger = 1/2$) in which the surrogate stops while the true problem continues, underlying Proposition 8; the analytic ridge $c_{\text{step}} = 1/4 - \varepsilon$ is overlaid (cyan dashed). The construction of Proposition 8 uses a different scaling ($c_{\text{abs}} = \kappa c_{\text{err}} \varepsilon$ with $\kappa > 1$) that pushes regret toward the upper bound $2c_{\text{err}}\varepsilon$ as $\varepsilon \downarrow 0$. Panel (I) is the bias diagnostic on the same $c_{\text{abs}} = 1/2$ family: empirical surrogate regret

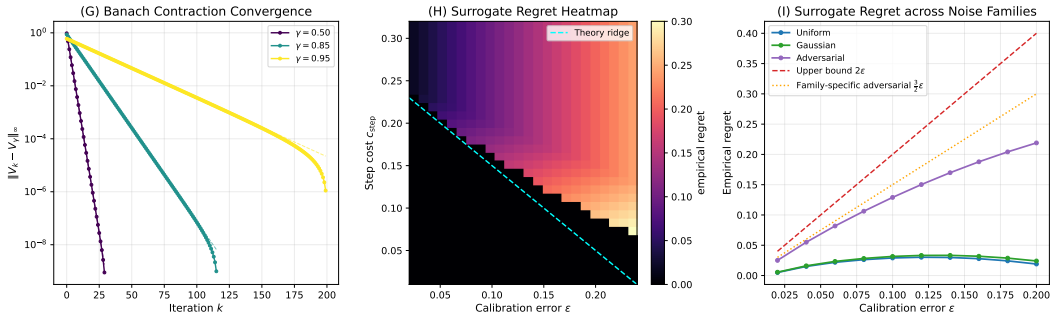

Figure 3: Diagnostics for value iteration and surrogate regret: Banach contraction convergence (G), the surrogate-policy regret heatmap with its analytic ridge (H), and the noise-family bias diagnostic (I).

with the calibration error $p_t - p_t^*$ generated by three noise families (zero-mean uniform, truncated Gaussian, and an adversarial fixed sign). All three remain at or below the $2c_{\text{err}}\varepsilon$ bound of Theorem 8. The orange dotted line at $\frac{3}{2}\varepsilon$ is the analytical worst case for this specific $c_{\text{abs}} = 1/2$ family, not the universal lower bound; the latter approaches $2c_{\text{err}}\varepsilon$ in the family of Proposition 8 as $\kappa \to \infty$.

A.17 Real-data diagnostic for residual history dependence

Theorem 7 and Proposition 6 predict that systems calibrated at each fixed time can hide history-dependent structure relevant to optimal stopping. We illustrate this empirically on a real binary classification benchmark (the Wisconsin breast cancer diagnostic tabular dataset distributed with scikit-learn, with 15% label noise added to keep the task non-degenerate). We train a 4-exit feedforward network with hidden widths 24-16-12-8 and one exit head per layer, using Adam (learning rate 3×10^{-3}, weight decay 10^{-4}) for 120 full-batch epochs on a $50/25/25$ train/calibration/test split, then apply per-exit temperature scaling with LBFGS on the calibration split. On the test split, we compute (i) the standard marginal expected calibration error (ECE) and (ii) a residual-stratified conditional ECE proxy that also bins by the per-sample residual $r_k = p_k - p_{k-1}$ for exit index $k \geq 1$, which is $\mathcal{F}_k$-measurable and generally not $\sigma(p_k)$-measurable; at exit 0, with no previous confidence, the proxy reduces to the marginal ECE. If condition (iii) of Definition 1 held, then within any joint bin of (p_k, r_k) the conditional correctness mean would equal the mean score, so the residual-stratified and marginal ECEs would coincide population-wise for this proxy construction, up to binning approximation and sampling error. We use 8 equal-width confidence bins on $[0, 1]$ and 3 equal-frequency residual strata, implemented by residual quantiles. Each label is independently flipped with probability 0.15 before the train/calibration/test split, and all figure error bars report 95% confidence intervals across 25 random seeds. This experiment is a diagnostic sanity check, not a benchmark claim: we ask whether a standard recalibration pipeline can leave a measurable symptom of hidden history-dependent structure.

Figure 4 shows the result, averaged over 25 random seeds. After recalibration, the marginal ECE is around 0.10 and stable across exits. We do not interpret this as near-perfect calibration; in this noisy small-tabular proxy it is only the baseline against which the residual-stratified gap is measured. The residual-stratified conditional ECE proxy is roughly 25–37% larger at exits 1, 2, and 3 (0.137–0.149), with seed-averaged values clearly above the marginal counterparts at those exits. This is a proxy demonstration. The exact conditional probability $\mathbb{P}(Y = \hat{Y}_t \mid \mathcal{F}_t)$ is not directly measurable in real data because the $\mathcal{F}_t$-conditional law is unknown. The consistent gap between the two ECE estimators is nevertheless a measurable symptom of history-dependent structure that standard calibration metrics do not detect.

From metric gap to Bayes-risk gap. Because this appendix is diagnostic, we also checked whether the ECE proxy gap translates into a Bayes-risk gap under one concrete cost regime. With the same 4-exit network and a stopping problem with $c_{\text{step}} = 0.02$, $c_{\text{err}} = 1$, and no abstention, we compare per-exit temperature scaling against per-(exit, residual stratum) temperature scaling using a backward-DP threshold rule (as detailed in the supplementary code). On this dataset and cost regime, both recalibrations produce the same aggregate mean cost and 95% interval (0.233 ± 0.012

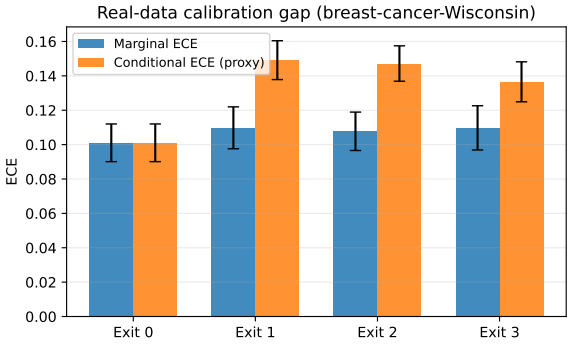

Figure 4: Real-data proxy after per-exit temperature scaling: marginal expected calibration error (ECE; blue) versus residual-stratified conditional ECE (orange). Under the tested costs both recalibrations had the same mean risk, below the single-final-exit baseline.

vs. 0.233 ± 0.012 at 25 seeds), well below the single-final-exit baseline (0.293 ± 0.014). The metric-level proxy gap is present, but its decision-level impact is dataset- and cost-regime-dependent; the constructions of Theorem 7 and Proposition 6 realize the decision-level gap. A direct semi-synthetic extension could match examples on p_t, suppress later-exit information in one group, and compare confidence-only with full-information stopping. We leave this proposed stress test to future work rather than treating the present diagnostic as a positive decision-gap result.