EVIDENCE PROVENANCE ROUTING FOR ROBUST DECISION MAKING UNDER PREDICTOR DISAGREEMENT

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern decision pipelines increasingly rely on heterogeneous predictors trained over distinct evidence sources. While these models often achieve comparable aggregate accuracy, they frequently produce conflicting predictions on individual instances—leading to brittle downstream decisions under non-stationary conditions. We show that a fundamental driver of this disagreement is evidence provenance: different predictors observe distinct evidence scopes whose relevance shifts dynamically across environments. To address this, we introduce Evidence Provenance Routing (EPR), a framework that maintains source-specific predictors as explicit decision anchors and dynamically allocates source authority during prediction conflicts. Unlike prior routing methods relying on static confidence metrics or historical accuracy, EPR evaluates authority using context-level evidence available at decision time. Specifically, EPR decouples routing into two stages: (1) Observable State–Decision Alignment (SDA), which identifies candidate source switches based on observable context, and (2) Disagreement Representation & Assessment, where a frozen language model encodes conflict semantics and lightweight probes validate stable route updates. Crucially, the final decision is executed by the selected source anchor, preserving domain-specific calibration rather than delegating task generation to the LLM. Evaluated across three non-stationary financial decision benchmarks and a legal outcome-prediction dataset, EPR consistently outperforms prediction aggregation, direct LLM reasoning, decision-focused learning, and modern router baselines. Finally, we provide a decision-theoretic analysis characterizing the value of context-dependent authority, bounding the decision cost of provenance compression, and decomposing routing risk across information, proposal and assessment stages.

1 INTRODUCTION

Predictive models increasingly serve as foundational components of decision-making systems operating over complex, multi-modal environments. Modern deployments frequently combine predictors trained over distinct, source-specific evidence scopes—such as textual filings, temporal histories, relational graphs, or contextual environment signals. This heterogeneity is ubiquitous: financial systems combine text sentiment with high-frequency market dynamics (Xu & Cohen, 2018; Xie et al., 2024); legal platforms integrate case briefs, precedent networks, and judicial metadata (Semo et al., 2022); and clinical systems reconcile electronic health records with environmental parameters (Farahani et al., 2026).

While individual models trained on distinct scopes may achieve comparable aggregate accuracy, they frequently produce conflicting predictions on individual instances. Resolving these disagreements is not simply a task of selecting the historically strongest predictor: choosing a prediction implicitly determines which evidence scope governs the downstream decision. Predictor disagreement is therefore fundamentally a problem of **evidence provenance**—the origin, scope, and context-dependent relevance of the information available to a predictor.

Because the utility of evidence scopes shifts dynamically under non-stationary environments, traditional routing mechanisms that rely on historical accuracy or static confidence metrics become brittle (Ovadia et al., 2019; Capitaine et al., 2026). Indeed, our diagnostic analysis reveals a striking failure of historical source selection: on the ACL18 financial benchmark, standard source-selection

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.

Table 1: **Positioning Matrix.** Conceptual comparison between EPR and closely related decision paradigms.

Paradigm	Decision Object	Explicit Provenance?	Decision-Time Allocation?	Primary Decision Signal
Prediction Aggregation	Merged Prediction	No	No	Predictor Outputs
LLM-Based Decision	Generated Decision	No	No	Prompt / Retrieved Context
Performance-Oriented Model Routing	Candidate Model	No	Yes	Historical Reward / Quality
EPR (Ours)	**Source Authority**	**Yes**	**Yes**	**Decision-Time Evidence**

AUC collapses from 0.750 on the development split to near-chance (0.512) under chronological test evaluation. Under distribution shift, resolving predictor disagreement requires evaluating source authority directly from evidence observable at decision time.

Existing decision paradigms address predictor disagreement through fundamentally different lenses. Prediction Aggregation & Calibration: Existing reconciliation, aggregation, and decision-calibration approaches (Du et al., 2025; Huang et al., 2026; Tang et al., 2026) generally formulate decision making under a stationary environment, where the underlying predictor–outcome or signal–decision relationship is fixed and disagreement is reconciled, aggregated, or calibrated within that environment. We start from a different premise: the decision relevance of evidence sources is itself non-stationary, so the source that should govern a decision can change across contexts and regimes. LLM-Based Direct Decisioning: End-to-end LLM frameworks use large models for reasoning and direct decision synthesis (Wang et al., 2024a; Handler et al., 2024). However, this delegates final task execution to an uncalibrated generative model, forfeiting the domain-specific reliability of specialized anchors. Performance-Oriented Model Routing: Model routers (Li et al., 2026) select candidate models based on estimated output quality or historical reward signals. They evaluate general model competency rather than determining which evidence provenance should govern the current decision context.

In contrast, we establish evidence provenance itself as the central routing structure. We introduce **Evidence Provenance Routing (EPR)**, a decision framework that maintains source-specific predictors as explicit decision anchors and dynamically assigns context-level authority during prediction conflicts. After route initialization, EPR decouples authority allocation into two distinct stages: Observable State–Decision Alignment (SDA): Constructs bounded source-transition proposals derived strictly from evidence visible at decision time. Disagreement Representation & Assessment: A frozen, task-agnostic language model encodes the contextual disagreement structure between the active and candidate anchors. Lightweight probes then rank proposed source changes to recover useful candidates and retain stable updates through consensus. Crucially, the selected source anchor—rather than the language model—produces the final task prediction, preserving domain-specific calibration throughout.

The primary contributions of this work are threefold. (1) **Formulation & Theoretical Framework:** We establish a decision-theoretic foundation for routing under heterogeneous evidence by viewing predictor disagreement as context-dependent authority allocation over source-indexed anchors. We quantify the decision cost of provenance compression and decompose routing risk into information, proposal, and assessment terms. (2) **Decoupled Routing Architecture:** We propose the EPR architecture, which explicitly decouples state-driven proposal generation from representation-based assessment using lightweight probes over frozen LLM features, preserving anchor provenance at decision time. (3) **Empirical Validation:** Across three chronologically evaluated financial benchmarks and a legal case-outcome benchmark, EPR consistently outperforms prediction-centric, direct LLM reasoning, decision-focused, and competitive router baselines.

2 METHOD

2.1 PROBLEM FORMULATION AND DECISION ANCHORS

Figure 1 summarizes the overall EPR routing pipeline and its three sequential stages.

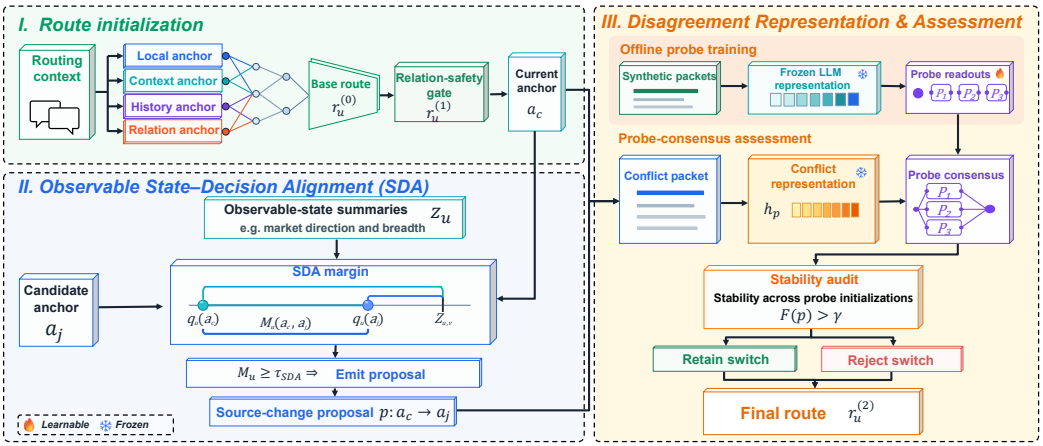


Figure 1: **The EPR routing pipeline. I. Route initialization.** EPR initializes a context-level route over explicit decision anchors and applies the relation-safety gate. **II. Observable State–Decision Alignment (SDA).** SDA proposes bounded source switches by comparing current and candidate anchors against decision-time observable-state summaries. **III. Disagreement Representation & Assessment.** Proposed switches are serialized as conflict packets; offline probes learn conflict readouts from frozen LLM representations of synthetic packets, and Probe-consensus assessment applies these readouts to retain stable switches and produce the final route $r_u^{(2)}$. The selected anchor then produces the final decision.

EPR routes at the **context** level: one source-specific predictor is granted decision authority over all samples in that context. In FinBen, a context is a decision date; in the legal transfer, it is a case together with its visible group evidence.

Let u index a routing context and $i \in \mathcal{I}_u$ a sample within it, with outcome $Y_{u,i} \in \{0, 1\}$. For a fixed routing context u, we suppress the context index when unambiguous. We construct anchors to capture evidence heterogeneity across the distinct levels supported by each task's observable information as comprehensively as possible. In FinBen, these levels comprise signals from individual samples, distributional context on the same day, historical states, and relational structure. Each level is instantiated as a separate predictor, called a **decision anchor**, with $E_{k,i}$ denoting the evidence visible to anchor a_k.

Let $\mathcal{A} = \{a_1, \ldots, a_K\}$ denote the K decision anchors. Given $E_{k,i}$, anchor a_k outputs a binary action and a confidence score:

$$a_k(E_{k,i}) = (\hat{y}_{k,i}, c_{k,i}), \quad \hat{y}_{k,i} \in \{0, 1\}, \quad c_{k,i} \in [0, 1]. \tag{1}$$

Separate anchors preserve evidence provenance, while routing determines which anchor controls the context. Let $r_u \in \mathcal{A}$ denote the anchor granted decision authority for context u. Let $\mathbf{V}_u$ collect the common context summaries provided to the router. Let $\mathbf{O}_{k,u} = \{a_k(E_{k,i}) : i \in \mathcal{I}_u\}$ denote anchor a_k's context-level output set, and let $\mathbf{O}_u = (\mathbf{O}_{1,u}, \ldots, \mathbf{O}_{K,u})$ denote the source-indexed anchor-output state. Let Z_u denote the observable-state summaries used by SDA. Together, $(\mathbf{V}_u, \mathbf{O}_u, Z_u)$ are the routing-visible information.

For task loss ℓ, define the context-level loss when anchor a controls context u as

$$\Lambda_u(a) = \frac{1}{|\mathcal{I}_u|} \sum_{i \in \mathcal{I}_u} \ell(Y_{u,i}, \hat{y}_{a,i}) \,. \tag{2}$$

EPR assigns one anchor to each context to minimize the expected context-level loss $\mathbb{E}_u[\Lambda_u(r_u)]$.

2.2 Route Initialization

The initializer selects a base route $r_u^{(0)} \in \mathcal{A}$ from observable context-level anchor summaries. A relation-safety gate evaluates a paired fallback for relation-anchor assignments using a frozen-representation readout. Let a_{rel} denote the relation anchor and τ_{rel} the relation-safety threshold. When $r_u^{(0)} = a_{\text{rel}}$, the gate selects the fallback if its score is at least τ_{rel} and otherwise retains

$r_u^{(0)}$. The resulting route $r_u^{(1)}$ provides the current anchor for SDA. Appendix B.3 specifies the sequence-refit initializer and the gating criterion.

2.3 Observable State–Decision Alignment (SDA)

SDA is the proposal module of EPR. Given the current anchor, SDA proposes a source switch when a candidate anchor's aggregate action tendency better matches the decision-time observable state. Let $\mathbf{z}_u = (z_{u,1}, \ldots, z_{u,L}) \in [0,1]^L$ denote the realization of Z_u through these L state summaries, and let $q_u(a) \in [0,1]$ denote anchor a's aggregate positive-action tendency in context u. In FinBen, Z_u summarizes contemporaneous market direction and cross-sectional breadth.

For a current anchor a_c and a candidate anchor a_j, SDA computes the normalized alignment margin

$$M_u(a_c, a_j) = \frac{\sum_{\nu=1}^{L} w_\nu \left[|q_u(a_c) - z_{u,\nu}| - |q_u(a_j) - z_{u,\nu}| \right]}{\sum_{\nu=1}^{L} w_\nu}, \tag{3}$$

where $w_\nu \geq 0$ weights state summary ν. A positive margin indicates that the candidate anchor is closer to the visible state than the current anchor under the weighted summaries. SDA emits the context-indexed source-change proposal $p_u : a_c \to a_j$ when $M_u(a_c, a_j) \geq \tau_{\mathrm{SDA}}$, where τ_{SDA} is the normalized proposal margin. The assessment stage determines whether the proposed switch updates the route.

2.4 Disagreement Representation & Assessment

Assessment represents a proposed source change by its source roles and observable conflict state. A frozen LLM encodes the corresponding structured packet as $\mathbf{h}_p$. Probes trained offline on synthetic conflict packets and labels produce candidate-action conflict distributions from this representation.

Across the proposal pool, probe-consensus assessment combines entropy-based proposal ranking with consensus over leave-one-probe rerankings. The ranking score is the ensemble-averaged normalized entropy, while $F(p)$ measures retention frequency under leave-one-probe reranking. Let S_{probe} denote the probe ensemble size, S_{cls} the conflict class size, and $P_s(m \mid p)$ probe s's probability for class m under proposal p. Specifically,

$$\overline{\mathrm{Ent}}(p) = -\frac{1}{S_{\mathrm{probe}} \log S_{\mathrm{cls}}} \sum_{s=1}^{S_{\mathrm{probe}}} \sum_{m=1}^{S_{\mathrm{cls}}} P_s(m \mid p) \log P_s(m \mid p). \tag{4}$$

Proposals are ranked in descending order of $\overline{\mathrm{Ent}}(p)$.

For proposal p, the leave-one-probe consensus score is

$$F(p) = \frac{|\{s \in \{1, \ldots, S_{\mathrm{probe}}\} : p \text{ is retained after leaving out probe } s\}|}{S_{\mathrm{probe}}}. \tag{5}$$

The resulting score $F(p)$ implements the stability audit of the proposed source change. Each reranking applies the same retention rule. For the selected proposal p_u in context u and stability threshold γ, the final route is

$$r_u^{(2)} = \begin{cases} a_j, & p_u : a_c \to a_j \text{ and } F(p_u) > \gamma, \\ r_u^{(1)}, & \text{otherwise.} \end{cases} \tag{6}$$

The selected route $r_u^{(2)}$ determines the anchor used for task prediction. The legal transfer follows the same proposal–assessment sequence with domain-specific proposal rules.

2.5 Decision-Theoretic Characterization

Theorem 2.1 (Value of Context-Dependent Source Authority)**.** *For the finite anchor set $\mathcal{A}$ and integrable context-level losses $\Lambda_u(a)$,*

$$\min_{a \in \mathcal{A}} \mathbb{E}[\Lambda_u(a)] \geq \mathbb{E}\left[\min_{a \in \mathcal{A}} \mathbb{E}[\Lambda_u(a) \mid Z_u] \right]. \tag{7}$$

The inequality is strict when no fixed anchor is conditionally optimal almost surely.

The full statement and proof are given in Appendix E.2.1.

For an information set $\mathcal{J}$, let $R^\star(\mathcal{J})$ denote the lowest expected routing loss over $\mathcal{J}$-measurable policies. For a nonempty random candidate set $\mathcal{C} \subseteq \mathcal{A}$ with $\mathcal{J}$-measurable membership, let $R^\star(\mathcal{J}, \mathcal{C})$ denote the corresponding minimum with anchor choices restricted to $\mathcal{C}$.

Theorem 2.2 (Provenance-Preserving Information Dominance)**.** *For any measurable deterministic compression $\widetilde{\mathbf{O}}_u$ of the source-indexed output state $\mathbf{O}_u$, let $\mathcal{H}_{\mathrm{full}} = \sigma(\mathbf{V}_u, \mathbf{O}_u, Z_u)$ and $\mathcal{H}_{\mathrm{comp}} = \sigma(\mathbf{V}_u, \widetilde{\mathbf{O}}_u, Z_u) \subseteq \mathcal{H}_{\mathrm{full}}$, where $\sigma(\cdot)$ denotes the sigma-field generated by its arguments. Then*

$$R^\star(\mathcal{H}_{\mathrm{full}}) \leq R^\star(\mathcal{H}_{\mathrm{comp}}). \tag{8}$$

Corollary 2.3 (Exact Pairwise Cost of Provenance Compression)**.** *For any fixed pair $a_1, a_2 \in \mathcal{A}$, define $\mu_a^{\mathcal{J}} = \mathbb{E}[\Lambda_u(a) \mid \mathcal{J}]$ for $a \in \{a_1, a_2\}$ and $\mathcal{J} \in \{\mathcal{H}_{\mathrm{full}}, \mathcal{H}_{\mathrm{comp}}\}$. Then*

$$\begin{aligned} &R^\star(\mathcal{H}_{\mathrm{comp}}, \{a_1, a_2\}) - R^\star(\mathcal{H}_{\mathrm{full}}, \{a_1, a_2\}) \\ &\quad = \frac{1}{2}\,\mathbb{E}\big[\big|\mu_{a_1}^{\mathcal{H}_{\mathrm{full}}} - \mu_{a_2}^{\mathcal{H}_{\mathrm{full}}}\big| - \big|\mu_{a_1}^{\mathcal{H}_{\mathrm{comp}}} - \mu_{a_2}^{\mathcal{H}_{\mathrm{comp}}}\big|\big]\,. \end{aligned} \tag{9}$$

In the two-anchor case, the compression gap is therefore strictly positive when the compressed representation aliases full-information states that favor opposite anchors.

The proof and exact sign condition are given in Appendix E.2.2. For any finite anchor set, the same two-source structure extends recursively through the Bayes lower envelope: the provenance-compression gap admits an exact binary recursion whose terminal value is invariant to anchor ordering.

Theorem 2.4 (Information–Action Decomposition)**.** *Let $\mathcal{G}$ be a provenance-preserving information set and let $\mathcal{H} \subseteq \mathcal{G}$ be the information available to the decision rule. Let $\mathcal{C} \subseteq \mathcal{A}$ be a nonempty random candidate set such that $\{a \in \mathcal{C}\}$ is $\mathcal{H}$-measurable for every $a \in \mathcal{A}$, and let $\widehat{A} \in \mathcal{C}$ be an $\mathcal{H}$-measurable selected anchor. Define $R(\widehat{A}) = \mathbb{E}[\Lambda_u(\widehat{A})]$. Then*

$$R(\widehat{A}) - R^\star(\mathcal{G}) = \underbrace{R^\star(\mathcal{H}) - R^\star(\mathcal{G})}_{\mathcal{R}_{\mathrm{info}}} + \underbrace{R^\star(\mathcal{H}, \mathcal{C}) - R^\star(\mathcal{H})}_{\mathcal{R}_{\mathrm{prop}}} + \underbrace{R(\widehat{A}) - R^\star(\mathcal{H}, \mathcal{C})}_{\mathcal{R}_{\mathrm{assess}}}. \tag{10}$$

The information term $\mathcal{R}_{\mathrm{info}}$, proposal term $\mathcal{R}_{\mathrm{prop}}$, and assessment term $\mathcal{R}_{\mathrm{assess}}$ are nonnegative.

For EPR, let $\mathcal{H}_{\mathrm{EPR}}$ be the sigma-field generated by the context and the indexed routing-visible states and proposals used in proposal-pool reranking. With trained components fixed, assume that each context has at most one eligible proposal, whose identity is fixed across leave-one-probe reranking; eligibility, anchor identities, rankings, and tie-breaking are $\mathcal{H}_{\mathrm{EPR}}$-measurable. For an eligible $p : a_c \to a_j$, set $\mathcal{C} = \{a_c, a_j\}$ and define the conditional switch gain $\theta = \mathbb{E}[\Lambda_u(a_c) - \Lambda_u(a_j) \mid \mathcal{H}_{\mathrm{EPR}}]$. Without an eligible proposal, set $\mathcal{C} = \{a_c\}$ and $\theta = F(p) = 0$. Take $\mathcal{H} = \mathcal{H}_{\mathrm{EPR}}$ and let $\widehat{A}$ follow the consensus rule in Section 2.4. Write $(x)_+ = \max\{x, 0\}$ for the positive part, $\mathbf{1}\{\cdot\}$ for an indicator, and $\lfloor x \rfloor$ for the greatest integer not exceeding x.

Theorem 2.5 (Gain-Weighted Assessment Regret and Finite-Consensus Bound)**.** *For integrable context losses, $S_{\mathrm{probe}} \geq 2$, and $0 \leq \gamma < 1$, under the setting above,*

$$\mathcal{R}_{\mathrm{assess}} = \mathbb{E}[(\theta)_+ \mathbf{1}\{F(p) \leq \gamma\} + (-\theta)_+ \mathbf{1}\{F(p) > \gamma\}]\,. \tag{11}$$

It satisfies

$$\mathcal{R}_{\mathrm{assess}} \leq \mathbb{E}\left[\frac{S_{\mathrm{probe}}}{S_{\mathrm{probe}} - \lfloor S_{\mathrm{probe}}\gamma \rfloor}(\theta)_+\big(1 - F(p)\big) + \frac{S_{\mathrm{probe}}}{\lfloor S_{\mathrm{probe}}\gamma \rfloor + 1}(-\theta)_+ F(p)\right]. \tag{12}$$

The two coefficients are individually optimal over this class of retention rules and are attained by leave-one-probe top-one ranking of two proposals with scores in $[0, 1]$. No independence among contexts or probe readouts is required.

These results map directly onto EPR: explicit provenance controls information loss, the proposal stage controls candidate restriction, and assessment trades off gain-weighted rejection of beneficial switches against acceptance of harmful switches. Theorem 2.5 quantifies the latter two losses under finite probe consensus. Its proof is given in Theorem E.8 of Appendix E.2.3.

3 EXPERIMENTS

We evaluate EPR in heterogeneous-evidence decision settings where predictors with distinct evidence provenance may disagree under chronological shifts. The experiments test four claims: **(i)** EPR improves over prediction-centric, LLM-based, decision-focused, and competitive routing baselines; **(ii)** route initialization, observable-state proposal, and conflict-stability assessment each contribute to performance; **(iii)** probe-consensus assessment is robust to design choices and temporal perturbations; and **(iv)** the proposal–assessment framework transfers beyond financial domain.

3.1 EXPERIMENTAL SETUPS

Table 2: Dataset statistics. All three datasets use chronological training, validation, and test partitions, ordered from earliest to latest.

Dataset	Period	Train	Valid	Test	Content
ACL18	2014–2016	20,781	2,555	3,720	Tweets + price
BigData22	2019–2020	4,897	798	1,472	Tweets + price
CIKM18	2017–2018	3,396	431	1,143	News headlines + price

Datasets. We use three public stock-movement datasets, ACL18, BigData22, and CIKM18, standardized by FinBen (Xie et al., 2024). ACL18 follows the StockNet construction (Xu & Cohen, 2018); periods and chronological partitions follow the FinBen-curated source releases (Table 2). These datasets provide chronological financial decision settings with heterogeneous textual and market evidence, making them suitable for evaluating source-authority routing under non-stationarity.

Baselines. We distinguish reported baselines, reproduced baselines, and adapted routing baselines. Reported rows follow public benchmark or paper results; reproduced rows use FinBen records and chronological splits; adapted routers share EPR's four-source action space and pre-decision information. We compare our method against four categories of approaches:

Prediction-Centric Methods: We use the FinBen-reported LR and LSTM baselines with Adv-ALSTM (Feng et al., 2019), DTML (Yoo et al., 2021), and SLOT (Soun et al., 2022), which represent adversarial, multi-level temporal, and self-supervised prediction architectures, respectively.

Direct Language-Model Predictors: We evaluate FinBERT (Araci, 2019) through sentiment transfer, FinGPT-7B-LoRA (Yang et al., 2023) as a financial forecasting model, and DeepSeek-V4-Pro under the FinBen direct-prompt protocol.

LLM-Based Decision Methods: We evaluate LLMFactor (Wang et al., 2024b), which uses sequential knowledge-guided prompting to extract market factors; SEP (Koa et al., 2024), a self-reflective PPO framework, with direct-decoding and reward-ranked 4-shot variants; FinSeer-style RAG + Stock-

Table 3: Main FinBen results (MCC / Acc). Trainable routers report three-seed means; best values are bold. Method families follow Section 3.1.

Category	Method	ACL	BigData	CIKM	Macro
Prediction	LR	0.044/0.522	0.020/0.531	−0.043/0.525	0.007/0.526
	LSTM	0.064/0.528	0.013/0.507	0.022/0.533	0.033/0.522
	Adv-ALSTM	0.069/0.530	0.012/0.500	0.022/0.540	0.034/0.523
	DTML	0.181/0.581	0.065/0.517	0.005/0.539	0.084/0.545
	SLOT	0.207/0.587	0.095/0.548	0.090/0.559	0.131/0.565
Direct LMs	FinBERT	0.000/0.513	0.000/0.552	0.000/0.581	0.000/0.549
	FinGPT-7B-LoRA	0.051/0.528	0.000/0.448	−0.007/0.575	0.015/0.517
	DeepSeek-V4-Pro	0.018/0.515	−0.022/0.547	0.007/0.579	0.001/0.547
LLM decision	LLMFactor	0.013/0.508	0.014/0.496	0.014/0.545	0.013/0.516
	SEP direct decoding	−0.035/0.481	0.020/0.458	0.037/0.444	0.007/0.461
	SEP reward-ranked 4-shot	−0.027/0.482	0.056/0.468	0.049/0.455	0.026/0.468
	FinSeer-style RAG + StockLLM	−0.017/0.502	0.008/0.528	−0.072/0.535	−0.026/0.522
	TradExpert	0.150/0.600	0.120/**0.590**	0.120/0.590	0.130/0.593
DFL/routing	DFL-PGB	0.094/0.542	0.027/0.498	0.076/0.534	0.066/0.525
	RouterDC (adapted)	0.080/0.540	0.054/0.513	0.055/0.537	0.063/0.530
	ZOOTER (adapted)	−0.011/0.494	0.060/0.509	0.047/0.541	0.032/0.515
EPR	EPR (Ours)	**0.247/0.613**	**0.155**/0.573	**0.186/0.603**	**0.196/0.596**

LLM (Xiao et al., 2025), which augments prediction with retrieved historical sequences; and TradExpert (Ding et al., 2025), which aggregates specialized financial LLM experts for stock prediction.

Decision-Focused and Routing Methods: We include DFL-PGB (Gupta & Huang, 2024) as a decision-focused objective, RouterDC (Chen et al., 2024) as a contrastive query–candidate router, and ZOOTER (Lu et al., 2024) as a reward-distilled router. All three are adapted to the common four-source action space; Appendix C.1 gives implementation details and three-seed statistics.

Metrics and implementation. We use Matthews correlation coefficient (MCC) as the primary metric and report accuracy for completeness. EPR uses frozen Qwen2.5-7B-Instruct representations with compact probe-consensus assessment. Full implementation details are provided in Appendices B and C.

3.2 Overall Performance on FinBen (RQ1)

Table 3 reports the main FinBen results. EPR achieves the highest MCC on all three datasets and reaches a macro MCC of 0.196. The strongest non-EPR baselines are SLOT among prediction-centric methods and TradExpert among LLM-based decision methods, with macro MCCs of 0.131 and 0.130, respectively. This shows that routing over evidence-defined anchors improves over both stronger stock-prediction architectures and LLM-based financial decision systems.

The comparison across baseline families clarifies where the gain comes from. Direct language-model predictors remain near zero MCC under the FinBen protocol, indicating that direct prompting or sentiment transfer is not sufficient for these chronological decision tasks. Decision-focused and learned routing baselines also underperform EPR: DFL-PGB, RouterDC, and ZOOTER reach macro MCCs between 0.032 and 0.066 after adaptation to the four-source action space. These results support the need for routing source authority from decision-time evidence instead of training a historical source-quality selector.

3.3 Routing Component Analysis (RQ2)

Table 4: Routing component analysis on chronological FinBen test splits (MCC). Stage progression and outcome-trained alternatives share the same four-source action space and pre-decision information. Best values are bold.

Category	Setting	ACL	BigData	CIKM	Macro
Stage progression	Base route	0.174	0.088	0.113	0.125
	+ Relation-safety gate	0.205	0.109	0.130	0.148
	+ Observable-state proposals	0.207	**0.155**	0.182	0.181
	+ Probe consensus (Full EPR)	**0.247**	**0.155**	**0.186**	**0.196**
Source selection	Supervised LR / RF	0.188	0.088	0.114	0.130
	Supervised GBDT	0.179	0.088	0.125	0.131
	Supervised LightGBM	0.159	0.066	0.122	0.116
	Online Static Selector	0.122	0.095	0.130	0.115
	Online 1-Day Delayed Feedback	0.181	0.095	0.114	0.130
Assessment	SDA + train-only assessment (valid-selected)	0.205	0.109	0.130	0.148
	SDA + train-only assessment ($n_{\mathrm{ret}} = 8$)	0.214	**0.155**	0.135	0.168

Table 4 separates the contribution of EPR's routing stages and compares probe-consensus assessment with outcome-trained alternatives.

Stage contribution. The route improves monotonically from the base route to the full system. The base route obtains 0.125 macro MCC. The relation-safety gate raises performance to 0.148, SDA proposals further raise it to 0.181, and probe-consensus assessment reaches 0.196. This pattern supports the staged design: relation safety improves the initialized route, observable-state proposal contributes the largest route revision gain, and assessment further filters the emitted source changes.

Outcome-trained source selection. Table 4 compares outcome-trained source selectors. Supervised source selectors reach only 0.116–0.131 macro MCC, below the frozen gated route and far below Full EPR. On ACL, the development-set source-selection AUC reaches 0.750, while chronological test

AUC approaches chance level. This degradation suggests that source winners learned from previous outcomes do not transfer reliably across chronological regimes.

Outcome-trained assessment. The assessment rows compare outcome-trained variants. A matched-budget learned assessment reaches 0.168 macro MCC, improving over the gated route but remaining below probe-consensus assessment. This comparison demonstrates that the frozen-representation stability criterion contributes beyond the proposal-retention budget.

Table 5: Representative routing cases. The last column compares correct predictions out of n samples under the current and candidate anchors ($a_c \to a_j$); outcomes are evaluated retrospectively.

Context Date	**SDA proposal** $q_u(a_c) \to q_u(a_j)$	**Visible state** $(z_{u,1}, z_{u,2})$	**Margin** M_u	**Assessment** $F(p)$	**Final anchor**	**Outcome** Correct / n
ACL18 2015-11-23	Context → Local 1.000 → 0.179	$(0.161, 0.179)$	0.821	Retain $101/101$	Local	$14/56 \to 40/56$ Benefit retained
ACL18 2015-12-11	History → Local 0.000 → 0.684	$(0.786, 0.911)$	0.684	Reject $0/101$	History	$75/76 \to 25/76$ Harm avoided
BigData22 2020-11-24	History → Local 0.000 → 0.636	$(0.706, 0.735)$	0.636	Retain $101/101$	Local	$1/44 \to 27/44$ Benefit retained
CIKM18 2018-01-16	Context → Local 1.000 → 0.500	$(0.149, 0.266)$	0.500	Reject $0/101$	Context	$8/24 \to 14/24$ Unstable switch avoided

Representative routing behavior. Table 5 illustrates how SDA proposals and probe-consensus assessment interact on individual test contexts. The examples include retained beneficial switches, rejected harmful switches, and a rejected switch with insufficient stability support that would have been beneficial, highlighting the conservative behavior of the stability criterion, which favors stable route updates even at the cost of occasionally foregoing gains.

3.4 Robustness and Sensitivity (RQ3)

Table 6: Probe-ensemble and proposal-retention sensitivity, where n_{ret} is the retained-proposal count. Full grids are reported in Appendices C.3–C.4.

Probe ensemble	S_{probe}	1	10	50	75	101
	Mean MCC	0.178	0.182	0.185	0.187	0.188
	$\Pr(\text{MCC} > \text{SDA})$ (%)	33.7	51.6	71.8	91.2	100.0
Proposal retention	n_{ret}	0	5	8	9	14
	Macro MCC	0.148	0.180	**0.196**	0.193	0.181
	Harmful proposals	0	0	0	1	4

Table 7: Temporal block-bootstrap robustness (20,000 paired bootstrap resamples per block length). Test days are sampled in circular blocks within each dataset, using identical blocks across compared routes.

Block length	**Mean Full−Gate**	**95% CI**	$\Pr(\text{Full} > \text{Gate})$	$\Pr(\text{Full} > \text{SDA})$
5 Trading Days	$+0.048$	$[+0.014, +0.090]$	99.9%	85.7%
10 Trading Days	$+0.048$	$[+0.016, +0.087]$	$> 99.9\%$	88.3%
20 Trading Days	$+0.048$	$[+0.017, +0.079]$	99.9%	91.9%

Probe robustness. Table 6 shows that probe-consensus assessment becomes increasingly stable as the ensemble grows. Mean MCC rises from 0.178 at $S_{\text{probe}} = 1$ to 0.188 at $S_{\text{probe}} = 101$, while $\Pr(\text{MCC} > \text{SDA})$ increases from 33.7% to 100.0%. These results show that the assessment gain remains stable across probe initializations and becomes more reliable under probe perturbations. Additional robustness analyses are provided in Appendices C.3–C.5.

Proposal-retention sensitivity. Table 6 shows that performance improves from 0.148 macro MCC with no retained proposals to 0.196 at $n_{\text{ret}} = 8$, with no harmful retained changes, but declines once additional proposals introduce harmful switches. This pattern shows that EPR benefits from selectively retaining stable proposals rather than accepting source changes indiscriminately, consistent with the gain-weighted assessment characterization in Appendix E.2.3. Additional design and hyperparameter sensitivity analyses are provided in Appendices C.2 and C.4.

Temporal robustness. Table 7 evaluates whether routing gains persist under temporal dependence. Across block lengths of 5, 10, and 20 trading days, Full EPR improves over the gated route in at least 99.9% of paired resamples and improves over unfiltered SDA in 85.7%–91.9% of paired resamples. The mean Full–Gate gain remains $+0.048$ MCC across all block lengths, with positive confidence intervals. This supports chronological robustness beyond isolated day-level wins.

3.5 Cross-Domain Transfer (RQ4)

To test whether EPR transfers across different structures of evidence heterogeneity, we evaluate it on ClassActionPrediction (Semo et al., 2022). Following the same construction principle, anchors based on case text, group statistics, and citation relations represent heterogeneity at the levels of individual cases, group context, and relational structure. No temporal anchor is instantiated because the benchmark provides no time-series evidence. This tests the same framework for proposal generation and assessment under a different configuration of heterogeneous evidence levels.

Table 8: Legal transfer and routing stages.

Method	Acc	MCC
Majority	0.446	0.000
TF-IDF + LR	0.687	0.386
TF-IDF + SVM	0.674	0.360
Extraction + RF	0.651	0.345
CustomLegalBERT	0.704	0.400
LegalBERT	0.704	0.406
Relation base	0.710	0.417
+ SDA	<u>0.717</u>	<u>0.430</u>
+ Assessment (EPR)	**0.720**	**0.437**

Table 8 reports legal-domain results under this common protocol. EPR achieves 0.437 MCC and 0.720 accuracy, outperforming the strongest reproduced baseline, LegalBERT, by 0.031 MCC and 0.016 accuracy. The standalone sample-text and relation anchors reach 0.406 and 0.417 test MCC, respectively. Starting from the relation anchor, observable-state proposals raise test MCC to 0.430, and probe-consensus assessment further raises it to 0.437.

The same stage-wise pattern appears on validation and test splits. This shows that the proposal–assessment structure transfers beyond finance. Appendix D details the legal anchors, proposal rules, and split-level routing progression.

4 Conclusion

We introduced Evidence Provenance Routing (EPR), a decision-routing framework for predictor disagreement under heterogeneous evidence. EPR treats source-specific predictors as explicit decision anchors, keeping evidence provenance visible before allocating context-level decision authority through a decoupled proposal–assessment sequence. SDA proposes bounded source changes, and frozen-representation probe consensus assesses the stability of the resulting source conflicts.

Across three heterogeneous FinBen-standardized financial benchmarks and a legal case-outcome benchmark, EPR outperforms prediction-centric, LLM-based, decision-focused, and routing baselines. Stage-wise ablations, robustness analyses, and cross-domain transfer show that both observable-state proposal and conflict-stability assessment contribute to the final route. The theoretical analysis further supports the design by characterizing the value of context-dependent source authority, formalizing provenance preservation and compression cost, and decomposing routing risk across information, proposal, and assessment stages.

Future work will extend EPR toward dynamic source pools, larger evidence spaces, and hierarchical authority allocation across multiple evidence granularities. Another promising direction is to integrate downstream utilities and asymmetric decision costs directly into the routing objective, enabling provenance-aware authority allocation in broader multimodal and sequential decision settings.

AI USE STATEMENT

AI-assisted tools were used for LaTeX editing, figure development, citation checking, language polishing, methodology review and assistance, and limited code editing under author supervision. The authors reviewed all AI-assisted material and remain responsible for the study design, implementation, results, and conclusions.

REPRODUCIBILITY STATEMENT

Section 3 describes the datasets, chronological evaluation protocol, metrics, and baseline families. Appendix B reports route configuration, anchor construction, proposal rules, packet representation, probe training, and the full routing procedure, including settings fixed before test evaluation. Appendix C provides diagnostics, sensitivity analyses, and temporal bootstrap tests; Appendix D documents the legal transfer protocol; Appendix E gives the proofs for Section 2.5.

REFERENCES

Dogu Araci. FinBERT: Financial sentiment analysis with pre-trained language models, 2019. arXiv preprint arXiv:1908.10063.

Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3): 1025–1044, March 2020. ISSN 1526-5501.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2): 265–272, 1953. ISSN 0003-4851.

Aymeric Capitaine, Maxime Haddouche, Eric Moulines, Michael Jordan, Etienne Boursier, and Alain Oliviero Durmus. Online decision-focused learning. In *International Conference on Learning Representations*, volume 2026, pp. 46821–46854, 2026.

Shuhao Chen, Weisen Jiang, Baijiong Lin, James T. Kwok, and Yu Zhang. RouterDC: Query-based router by dual contrastive learning for assembling large language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

D. R. Cox. Some problems connected with statistical inference. *The Annals of Mathematical Statistics*, 29(2): 357–372, 1958. ISSN 0003-4851.

Rafael M. O. Cruz, Robert Sabourin, and George D. C. Cavalcanti. Dynamic classifier selection: Recent advances and perspectives. *Information Fusion*, 41:195–216, May 2018. ISSN 1566-2535.

Jasper Dekoninck, Maximilian Baader, and Martin Vechev. A unified approach to routing and cascading for LLMs. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 12987–13010. PMLR, 13–19 Jul 2025.

Qianggang Ding, Haochen Shi, Jiadong Guo, and Bang Liu. Tradexpert: Revolutionizing trading with mixture of expert llms. *arXiv preprint arXiv:2411.00782*, 2025.

Priya Donti, Brandon Amos, and J. Zico Kolter. Task-based end-to-end model learning in stochastic optimization. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Ally Du, Dung Daniel Ngo, and Steven Wu. Reconciling model multiplicity for downstream decision making. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 59133–59164, 2025.

Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1):9–26, January 2022. ISSN 1526-5501.

Saba A. Farahani, Hung Cao, Ramesh Jain, and Amir M. Rahmani. T2d-bench: Evidence-gated evaluation of llm outputs for type 2 diabetes using a multi-layer clinical-lifestyle knowledge graph. *arXiv preprint arXiv:2606.24145*, 2026.

Fuli Feng, Huimin Chen, Xiangnan He, Ji Ding, Maosong Sun, and Tat-Seng Chua. Enhancing stock movement prediction with adversarial training. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pp. 5843–5849. International Joint Conferences on Artificial Intelligence Organization, August 2019.

Yonatan Geifman and Ran El-Yaniv. Selective classification for deep neural networks. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Vishal Gupta and Michael Huang. Decision-focused learning with directional gradients. In *Advances in Neural Information Processing Systems*, volume 37, pp. 79194–79220. Curran Associates, Inc., 2024.

Abram Handler, Kai R. Larsen, and Richard Hackathorn. Large language models present new questions for decision support. *International Journal of Information Management*, 79:102811, 2024. ISSN 0268-4012.

Zhihuan Huang, Yichong Xia, and Yuqing Kong. Decision aggregation under quantal response. In *The Fourteenth International Conference on Learning Representations*, 2026.

Shayan Kiyani, Hamed Hassani, George J. Pappas, and Aaron Roth. Robust decision-making with partially calibrated forecasters. In *The Fourteenth International Conference on Learning Representations*, 2026.

Kelvin J. L. Koa, Yunshan Ma, Ritchie Ng, and Tat-Seng Chua. Learning to generate explainable stock predictions using self-reflective large language models. In *Proceedings of the ACM Web Conference 2024*, pp. 4304–4315. ACM, May 2024.

Hao Li, Yiqun Zhang, Zhaoyan Guo, Chenxu Wang, Shengji Tang, Qiaosheng Zhang, Yang Chen, Biqing Qi, Peng Ye, Lei Bai, Zhen Wang, and Shuyue Hu. LLMRouterBench: A massive benchmark and unified framework for LLM routing. In Maria Liakata, Viviane P. Moreira, Jiajun Zhang, and David Jurgens (eds.), *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 37733–37754, July 2026. ISBN 979-8-89176-395-1.

Keming Lu, Hongyi Yuan, Runji Lin, Junyang Lin, Zheng Yuan, Chang Zhou, and Jingren Zhou. Routing to the expert: Efficient reward-guided ensemble of large language models. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1964–1974, 2024.

Hussein Mozannar and David Sontag. Consistent estimators for learning to defer to an expert. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7076–7087. PMLR, 13–18 Jul 2020.

Isaac Ong, Amjad Almahairi, Vincent Wu, Wei-Lin Chiang, Tianhao Wu, Joseph E Gonzalez, M Kadous, and Ion Stoica. Routellm: Learning to route llms from preference data. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 34433–34448, 2025.

Yaniv Ovadia, Emily Fertig, Jie Ren, Zachary Nado, D. Sculley, Sebastian Nowozin, Joshua Dillon, Balaji Lakshminarayanan, and Jasper Snoek. Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Gil Semo, Dor Bernsohn, Ben Hagag, Gila Hayat, and Joel Niklaus. ClassActionPrediction: A challenging benchmark for legal judgment prediction of class action cases in the US. In *Proceedings of the Natural Legal Language Processing Workshop 2022*, pp. 31–46, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics.

Yejun Soun, Jaemin Yoo, Minyong Cho, Jihyeong Jeon, and U Kang. Accurate stock movement prediction with self-supervised learning from sparse noisy tweets. In *2022 IEEE International Conference on Big Data (Big Data)*, pp. 1691–1700. IEEE, December 2022.

Jingwu Tang, Jiayun Wu, Steven Wu, and Jiahao Zhang. Dimension-free decision calibration for nonlinear loss functions. In *The Fourteenth International Conference on Learning Representations*, 2026.

Amirhossein Vahidi, Hesam Asadollahzadeh, Navid Akhavan Attar, Marie Moullet, Kevin Ly, Xingyi Yang, and Mohammad Lotfollahi. Dirmoe: Dirichlet-routed mixture of experts. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 155806–155824, 2026.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, Wayne Xin Zhao, Zhewei Wei, and Jirong Wen. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), March 2024. ISSN 2095-2236.

Meiyun Wang, Kiyoshi Izumi, and Hiroki Sakaji. LLMFactor: Extracting profitable factors through prompts for explainable stock movement prediction. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 3120–3131, Bangkok, Thailand, August 2024. Association for Computational Linguistics.

Mengxi Xiao, Zihao Jiang, Lingfei Qian, Zhengyu Chen, Yueru He, Yijing Xu, Yuecheng Jiang, Dong Li, Ruey-Ling Weng, Min Peng, Jimin Huang, Sophia Ananiadou, and Qianqian Xie. Retrieval-augmented large language models for financial time series forecasting. *arXiv preprint arXiv:2502.05878*, 2025.

Qianqian Xie, Weiguang Han, Zhengyu Chen, Ruoyu Xiang, Xiao Zhang, Yueru He, Mengxi Xiao, Dong Li, Yongfu Dai, Duanyu Feng, Yijing Xu, Haoqiang Kang, Ziyan Kuang, Chenhan Yuan, Kailai Yang, Zheheng Luo, Tianlin Zhang, Zhiwei Liu, Guojun Xiong, Zhiyang Deng, Yuechen Jiang, Zhiyuan Yao, Haohang Li, Yangyang Yu, Gang Hu, Jiajia Huang, Xiao-Yang Liu, Alejandro Lopez-Lira, Benyou Wang, Yanzhao Lai, Hao Wang, Min Peng, Sophia Ananiadou, and Jimin Huang. FinBen: A holistic financial benchmark for large language models. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 95716–95743. Curran Associates, Inc., 2024.

Yumo Xu and Shay B. Cohen. Stock movement prediction from tweets and historical prices. In Iryna Gurevych and Yusuke Miyao (eds.), *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1970–1979, Melbourne, Australia, July 2018. Association for Computational Linguistics.

Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang. Fingpt: Open-source financial large language models. 2023. FinLLM Workshop at IJCAI 2023.

Jaemin Yoo, Yejun Soun, Yong-chan Park, and U Kang. Accurate multivariate stock movement prediction via data-axis transformer with multi-level contexts. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2037–2045. ACM, August 2021.

Yiyao Zhang, Diksha Goel, Hussain Ahmad, Shixun Huang, and Jun Shen. When the judge should not decide: Evidence-locked, non-compensatory selection bounds llm-judge failure in reasoning pipelines. *arXiv preprint arXiv:2608.07813*, 2026.

Chang Zhou, Xingtong Yu, Minbin Huang, Zhennan Wu, Yuan Fang, Hong Cheng, and Xinming Zhang. Tradingmoe: Routing the right experts in evolving markets. *arXiv preprint arXiv:2608.11785*, 2026.

A Related Work

A.1 prediction-centric and Decision-Aware Learning

Prediction-centric methods primarily optimize predictive performance before their outputs enter downstream decision rules. In stock-movement prediction, recurrent and transformer-based models enrich predictive representations with adversarial training, multi-level temporal contexts, and self-supervision over sparse or noisy text (Xu & Cohen, 2018; Feng et al., 2019; Yoo et al., 2021; Soun et al., 2022). These methods strengthen the forecast supplied to a decision rule, while the evidence scope governing that forecast is already embedded in the model output.

Prescriptive analytics and decision-focused learning connect prediction models more directly to downstream objectives (Bertsimas & Kallus, 2020; Donti et al., 2017). Predict-then-optimize methods estimate problem parameters before solving a downstream optimization problem, and Smart Predict-then-Optimize trains predictors with an optimization-aware loss (Elmachtoub & Grigas, 2022). Recent decision-focused formulations further study directional-gradient surrogates and online decision-making settings (Gupta & Huang, 2024; Capitaine et al., 2026). Recent decision-calibration work studies when forecasts can reliably support downstream actions, including nonlinear losses and partial calibration guarantees (Tang et al., 2026; Kiyani et al., 2026). EPR addresses a different decision problem: when predictors grounded in distinct evidence scopes disagree, it allocates decision authority among the corresponding sources.

A.2 LLM-Based Prediction and Decision Support

Large language models enter decision pipelines as predictors, explainers, retrievers, and expert aggregators. In financial settings, domain-specific encoders and generative models adapt pretrained representations to sentiment transfer and market prediction (Araci, 2019; Yang et al., 2023). Recent LLM-based methods further use prompting or self-reflection to extract market factors, generate explanations, and provide the final prediction for downstream decisions (Wang et al., 2024b; Koa et al., 2024). These methods place the language model close to the prediction or explanation step.

Other LLM-based decision systems ground predictions through retrieval, evidence checks, or expert aggregation. Retrieval-augmented forecasting conditions language models on historical records (Xiao et al., 2025), and TradExpert orchestrates specialized LLM experts for trading decisions (Ding et al., 2025). Evidence-gated evaluation frameworks require explicit evidence support before an output is revised or overridden (Farahani et al., 2026; Zhang et al., 2026). EPR assigns LLM a different role: a frozen LLM represents disagreements between source-specific anchors, lightweight probes assess the stability of proposed source changes, while the selected evidence anchor produces the final task decision.

A.3 Routing, Aggregation, and Evidence-Source Allocation

Model routing assigns an input or context to a suitable candidate from a model pool. Classical conditional-selection paradigms include dynamic classifier selection, selective prediction, and learning to defer (Cruz et al., 2018; Geifman & El-Yaniv, 2017; Mozannar & Sontag, 2020). Recent LLM routing extends this idea to heterogeneous language models, where routing signals are often derived from quality–cost trade-offs, preference data, rewards, or query–candidate representations (Dekoninck et al., 2025; Ong et al., 2025; Lu et al., 2024; Chen et al., 2024). Mixture-of-experts architectures perform a related form of dynamic allocation within a shared model, routing tokens among internal subnetworks rather than selecting among separate evidence-defined predictors (Vahidi et al., 2026).

A complementary decision-theoretic line studies how multiple expert decisions should be aggregated. (Huang et al., 2026) analyze experts receiving private signals under quantal response and derive robust aggregation rules for collective decisions. Performance-oriented routers instead learn query-to-model or query-to-expert assignments from historical outcomes, rewards, preferences, or estimated quality. ZOOTER distills reward signals into a routing function for expert LLM selection (Lu et al., 2024), while RouterDC learns query and candidate representations through dual contrastive objectives (Chen et al., 2024). Concurrent finance-specific routing work further adapts expert selection to evolving market contexts (Zhou et al., 2026). EPR differs from both: rather than aggregating expert decisions

into a fused output or routing by learned source quality, it preserves evidence-defined predictors as distinct decision objects and assigns context-level decision authority from decision-time evidence.

B Implementation Details & Experimental Protocol

Appendix B begins with frequently used notation, then follows the EPR execution flow, covering route configuration, anchor construction and visibility, conflict-packet and probe implementation, routing-time assessment, and computational complexity.

B.1 Notation and Symbols

Table 9: Frequently used notation.

Notation	Description
Decision setting and anchors	
$u,\ i,\ \mathcal{I}_u$	Routing context, sample index, and sample-index set in context u.
$Y_{u,i}$	Binary task outcome for sample i in context u.
$\mathcal{A} = \{a_1, \ldots, a_K\},\ a_k$	Finite set of K source-specific decision anchors and its kth anchor.
$E_{k,i}$	Evidence visible to anchor a_k for sample i.
$(\hat{y}_{k,i}, c_{k,i})$	Binary action and confidence emitted by anchor a_k.
r_u	Anchor granted decision authority for context u.
$\Lambda_u(a)$	Context-level loss when anchor a controls context u.
Routing-visible state and route stages	
$\mathbf{V}_u$	Common context summaries visible to the router.
$\mathbf{O}_{k,u},\ \mathbf{O}_u$	Context-level outputs of anchor a_k and the full source-indexed output state.
Z_u	Observable-state summaries used by SDA.
$r_u^{(0)},\ r_u^{(1)},\ r_u^{(2)}$	Base, relation-gated, and final routes for context u.
$a_{\mathrm{rel}},\ \tau_{\mathrm{rel}}$	Relation anchor and relation-safety threshold.
Observable State–Decision Alignment	
$\mathbf{z}_u = (z_{u,1}, \ldots, z_{u,L}),\ w_\nu$	L observable-state summaries and their weights.
$q_u(a)$	Aggregate positive-action tendency of anchor a in context u.
$M_u(a_c, a_j)$	Normalized alignment margin from current anchor a_c to candidate anchor a_j.
$a_c,\ a_j$	Current and candidate anchors in a proposed source transition.
$p_u : a_c \to a_j,\ \tau_{\mathrm{SDA}}$	Source-change proposal for context u and the SDA proposal threshold.
Disagreement assessment	
$p,\ \mathbf{h}_p$	Generic proposal and its frozen-LLM conflict representation.
$S_{\mathrm{probe}},\ S_{\mathrm{cls}}$	Probe ensemble size and conflict class size.
$P_s(m \mid p)$	Probability assigned by probe s to conflict class m for proposal p.
$\overline{\mathrm{Ent}}(p)$	Aggregated normalized-entropy score used for proposal ranking.
$F(p),\ \gamma$	Leave-one-probe consensus score and its stability threshold.
$\rho,\ \eta$	proposal-pool retention fraction and minimum per-group retention fraction.
Decision-theoretic analysis	
$\mathcal{H}_{\mathrm{full}},\ \mathcal{H}_{\mathrm{comp}}$	Full and compressed routing-information sets in the provenance result.
$\mathcal{G},\ \mathcal{H}$	Provenance-preserving information and information available to the decision rule.
$\mu_a^{\mathcal{H}}$	Conditional risk of anchor a given information $\mathcal{H}$.
$R^\star(\mathcal{H}),\ R^\star(\mathcal{H}, \mathcal{C})$	Optimal routing risk, without and with restriction to candidate set $\mathcal{C}$.
$\mathcal{C},\ \widehat{A}$	Candidate-anchor set and selected anchor.
$\mathcal{R}_{\mathrm{info}},\ \mathcal{R}_{\mathrm{prop}},\ \mathcal{R}_{\mathrm{assess}}$	Information, proposal, and assessment terms in routing excess risk.

Table 9 summarizes frequently used notation in the method and theoretical analysis.

B.2 Routing Configuration and Hyperparameters

Appendix B specifies the operational protocol used to instantiate the abstract routing framework in Section 2. Unless otherwise stated, route configurations are selected on the official development split or synthetic packet validation and then fixed before test evaluation. Test labels are used only for final metric reporting. Tables 10 and 11 collect the thresholds and stage-level route settings used in the reported experiments.

Table 10: Reported hyperparameter settings.

Symbol	Role	Reported setting	Where used
τ_{rel}	Relation-safety threshold	0.4 in reported evaluation	Relation-safety gate
d_h	Hidden-state dimension	Backbone-specific; 3584 for Qwen2.5-7B	LLM representation
τ_{SDA}	Normalized proposal margin	0.5 (equivalent raw-score threshold: 1.0)	Observable-state proposal
S_{probe}	Probe ensemble size	101	Consensus assessment
ρ	proposal-pool retention fraction	0.5	Base probe ranking
η	Minimum per-group retention fraction	0.25	Predefined-group coverage floor
γ	Stability-selection threshold	Stability rule: $F(p) > \gamma$, $\gamma = 0.90$	Leave-one-probe assessment
S_{cls}	Conflict class size	Number of candidate-action conflict classes; defined by the synthetic conflict vocabulary in Appendix B.5	Normalized entropy

The retention parameters ρ and η set the proposal-ranking budget and group-coverage floor used before leave-one-probe stability filtering.

Table 11: Route-configuration summary. Candidate generation and assessment are separate stages.

Component	How it is obtained	Reported setting
Base-route initializer	Sequence refit over development day-level anchor summaries	Candidate order and switch thresholds learned per official development split
Step selector	Scale-matched selector within the same refit family	ACL uses a shallow random forest; BigData and CIKM use standardized logistic regression
Relation-safety gate	Frozen-state factorized readout over relation-vs.-fallback packets	Relation-safety threshold $\tau_{\mathrm{rel}} = 0.4$
SDA proposal	Absolute alignment to decision-time observable state summaries	Equal weights; $\tau_{\mathrm{SDA}} = 0.5$ (equivalent raw-score threshold: 1.0)
Probe-consensus assessment	Candidate-action-family entropy with leave-one-probe stability selection	Qwen2.5-7B MLP readout; $S_{\mathrm{probe}} = 101$; $\rho = 0.5$; $\eta = 0.25$; $\gamma = 0.90$

The step selector is an internal component of the base-route initializer, not an additional routing stage.

B.3 Decision Anchors and Base-Route Construction

Table 12: Anchor inventory. Four routable weak decision sources instantiate the abstract evidence granularities in FinBen.

Anchor	Evidence Provenance	Observable support	Implementation and visibility
Sample	Local	Instance-local text, price, and benchmark fields	Low-capacity local refit or calibrated raw score; current record only.
Same-day	Contextual	Unlabeled candidate set for the target date	Logistic regression on same-day distributional summaries; same-day batch visible, labels unread.
Temporal	Historical-state	Prior-day rolling windows and group histories	Rolling historical statistical source; dates before the target day only.
Relation	Relational-structural	Neighborhood, market-context, and cross-sample relation support	Regularized relation refit; target excluded, no later dates; relation-safety score read from a separate frozen-state relation packet.

Table 12 summarizes the four routable anchors. The key distinction is visibility: sample anchors read only the target record, same-day anchors read the unlabeled candidate set for the target date, temporal anchors read only earlier dates, and relation anchors use relation support without target labels or later records.

Algorithm 1 Anchor Construction and Base-Route Initialization

Require: Development records D_{dev}, Context set $\mathcal{U}$
Ensure: Decision anchor set $\mathcal{A} = \{a_{\mathrm{smp}}, a_{\mathrm{sd}}, a_{\mathrm{tmp}}, a_{\mathrm{rel}}\}$, Base route π_{base}
1: Fit and calibrate a_{smp} on instance-local features in D_{dev}.
2: Construct a_{sd} from contemporaneous unlabeled batch summaries.
3: Form a_{tmp} using rolling historical feature buffers strictly preceding u.
4: Fit regularized relation anchor a_{rel} on neighborhood structural graphs.
5: Set $\mathcal{A} \leftarrow \{a_{\mathrm{smp}}, a_{\mathrm{sd}}, a_{\mathrm{tmp}}, a_{\mathrm{rel}}\}$.
6: Compute the context-level anchor output sets $\mathbf{O}_{k,u}$ for each $u \in \mathcal{U}$.
7: Train stepwise sequence-refit selector π_{base} on D_{dev} anchor summary metrics.
8: **return** $\mathcal{A}$ and π_{base}.

Relation-Safety Gate. Let $\mathcal{A}_{\mathrm{raw}} \subset \mathcal{A}$ denote the ungated anchors eligible for later proposals, and let $b_{\mathrm{rel}}(\mathbf{V}_u) \in \mathcal{A}_{\mathrm{raw}}$ be the fallback anchor paired with the relation route. A frozen-representation readout gives a fallback score $s_{\mathrm{rel}}(\mathbf{V}_u) \in [0, 1]$. The exact gate is

$$r_u^{(1)} = \begin{cases} b_{\mathrm{rel}}(\mathbf{V}_u) & \text{if } r_u^{(0)} = a_{\mathrm{rel}} \text{ and } s_{\mathrm{rel}}(\mathbf{V}_u) \geq \tau_{\mathrm{rel}}, \\ r_u^{(0)} & \text{otherwise.} \end{cases} \tag{13}$$

Subsequent proposals operate over $\mathcal{A}_{\mathrm{raw}}$.

B.4 Input Visibility and Evaluation Protocol

This section records the input visibility of each decision anchor. Each anchor emits a binary action and confidence score, but no anchor has access to future records or evaluation labels. The routing layers operate only on anchor outputs and observable context summaries.

We use $D_{\mathrm{dev}} = D_{\mathrm{train}} \cup D_{\mathrm{valid}}$ as the development pool for route construction and parameter selection. The held-out test labels and outcomes are never used for anchor fitting, route construction, proposal generation, probe training, proposal ranking, or stability assessment. Test outcomes are revealed only for final metric computation after the complete route is fixed.

Table 13: Input visibility for evaluated FinBen methods. A visible field is available under the benchmark files and chronological protocol; architectural use varies by method.

Method	Current record	Past records	Same-day unlabeled set	Future / test labels
LR	Yes	Yes	Yes	No
LSTM	Yes	Yes	Yes	No
Adv-ALSTM	Yes	Yes	Yes	No
DTML	Yes	Yes	Yes	No
SLOT	Yes	Yes	Yes	No
FinBERT sentiment transfer	Yes	Yes	Yes	No
FinGPT-7B-LoRA	Yes	Yes	Yes	No
DeepSeek-V4-Pro direct prompt	Yes	Yes	Yes	No
LLMFactor prompt factors	Yes	Yes	Yes	No
SEP direct decoding	Yes	Yes	Yes	No
SEP reward-ranked 4-shot	Yes	Yes	Yes	No
FinSeer-style RAG + StockLLM	Yes	Yes	Yes	No
TradExpert	Yes	Yes	Yes	No
EPR base route	Yes	Yes	Yes	No
EPR relation-safety gate	Yes	Yes	Yes	No
EPR SDA + probe-consensus assessment	Yes	Yes	Yes	No

B.5 Conflict Packet and Probe Training

Synthetic conflict packets are used to span representative source-conflict states. Each packet describes how two anchors disagree in role, vote, confidence, numeric context, and day size. The resulting synthetic labels define conflict profiles for probe training; Probe supervision is therefore independent of task outcomes.

Packet Serialization. Each candidate transition p is serialized into a deterministic key-value packet:

$$\begin{aligned}\texttt{current anchor: } a_c \,|\, &\texttt{candidate anchor: } a_j\\ \texttt{current vote: } v_c \,|\, &\texttt{candidate vote: } v_j\\ \texttt{confidence gap} \,|\, &\texttt{day-size band}\\ \texttt{action scope: } &[a_c, a_j].\end{aligned}$$

Here, v_c and v_j denote the current- and candidate-anchor votes, respectively. The same schema is used for synthetic packets and real routing-time proposals.

The synthetic teacher assigns a multi-axis conflict profile. The axes are grouped into source-role, vote-relation, confidence-pattern, numeric-context, and candidate-action-family categories. The main assessment uses the candidate-action-family readout, while the full axis inventory supports probe training and diagnostic ablations.

Probe Architecture and Training Protocol. All probes read the final-layer hidden state at the last non-padding packet token from a frozen LLM. The backbone is not updated. Probe heads are trained only on synthetic teacher labels that describe conflict-profile axes. Real routing-time packets are encoded with the same schema and projected through the trained probes, but their task outcomes are never used for probe-label supervision.

The frozen backbone is Qwen2.5-7B-Instruct. A two-layer MLP with GELU activations projects $\mathbf{h}_p$ to $\mathbb{R}^{256}$ and feeds 16 classification heads corresponding to the grouped synthetic conflict attributes. Probes are trained on $5,425$ synthetic conflict packets ($1,088$ validation, $1,137$ test) for 120 epochs using AdamW (learning rate 8×10^{-4}, weight decay 10^{-4}, batch size $4,096$, dropout 0.08). The objective optimizes weighted cross-entropy regularized by $-0.02\ \mathrm{mean}_j\ \mathrm{Std}_{p\in\mathcal{B}}([\mathrm{MLP}(\mathbf{h}_p)]_j)$, where $\mathcal{B}$ denotes a minibatch.

B.6 Routing-Time Assessment Protocol

At evaluation, emitted proposals are ranked jointly within the fixed unlabeled evaluation proposal pool; the ranking uses only proposal representations and probe readouts.

Proposal Ranking. For an ensemble of S_{probe} probes, let $P_s(m \mid p)$ denote probe s's probability for conflict class $m \in \{1, \ldots, S_{\text{cls}}\}$, where S_{cls} is the number of candidate-action conflict classes. The normalized entropy for probe s is

$$\text{Ent}_s(p) = -\frac{1}{\log S_{\text{cls}}} \sum_{m=1}^{S_{\text{cls}}} P_s(m \mid p) \log P_s(m \mid p), \tag{14}$$

and the ensemble mean is

$$\overline{\text{Ent}}(p) = \frac{1}{S_{\text{probe}}} \sum_{s=1}^{S_{\text{probe}}} \text{Ent}_s(p). \tag{15}$$

Higher $\overline{\text{Ent}}(p)$ indicates a stronger distributed conflict signature, so proposals are ranked in descending order of $\overline{\text{Ent}}(p)$.

Proposal Auditing and Selection Protocol. For each proposed transition p_d, we evaluate consensus across the $S_{\text{probe}} = 101$ probe initializations. Leave-one-probe reranking uses the consensus score $F(p)$ from Section 2.4 with proposal-pool retention fraction $\rho = 0.50$ and minimum per-group retention fraction $\eta = 0.25$. Proposals satisfying $F(p_d) > \gamma$ ($\gamma = 0.90$) are committed to update the controlling anchor ($r_d^{(2)} \leftarrow a_j$); otherwise, the current gated route is retained ($r_d^{(2)} \leftarrow r_d^{(1)}$).

B.7 Computational Complexity

Complexity Analysis. Using K, L, and S_{probe} as defined above, let $\mathcal{D}$ denote the evaluation dates, N_d the sample size on date d, $N = \sum_{d \in \mathcal{D}} N_d$ the total sample size, and d_{enc} the per-sample anchor-encoding cost. Anchor inference costs $\mathcal{O}(KNd_{\text{enc}})$. SDA proposal alignment costs $\mathcal{O}(KL)$ per date and $\mathcal{O}(|\mathcal{D}|KL)$ overall. Probe-ensemble readout over cached hidden states costs $\mathcal{O}(S_{\text{probe}}d_h)$ per emitted proposal. Across all proposals, consensus auditing is linear in S_{probe} and quasilinear in the emitted-proposal count.

The operational routing overhead is linear in N for fixed K, L, S_{probe}, and proposal rate.

C Additional Empirical Analyses

Appendix C complements Sections 3.3–3.4 with analyses of alternative routing strategies, proposal-state design, assessment robustness, hyperparameter sensitivity, and temporal stability.

C.1 Outcome-Trained Routing Alternatives

This section details the outcome-trained routing alternatives compared in Table 4. We consider two protocol-aligned alternatives. The first directly learns source selectors from development outcomes. The second applies outcome-trained assessment under the same proposal budget as probe consensus. Both alternatives use the same chronological splits, source-action space, and pre-decision information as EPR.

Table 14: Three-seed MCC robustness for protocol-aligned routing baselines. Values are mean ± sample standard deviation over fixed seeds $\{41, 42, 43\}$.

Method	ACL18	BigData22	CIKM18	Macro Mean
DFL-PGB	0.094 ± 0.016	0.027 ± 0.012	0.076 ± 0.027	0.066 ± 0.010
RouterDC (adapted)	0.080 ± 0.009	0.054 ± 0.022	0.055 ± 0.014	0.063 ± 0.009
ZOOTER (adapted)	-0.011 ± 0.073	0.060 ± 0.057	0.047 ± 0.039	0.032 ± 0.042

Directly supervised source selectors transfer weakly across chronological regimes: LR/RF, GBDT, and LightGBM reach 0.130, 0.131, and 0.116 macro MCC, respectively, below the frozen gated route

at 0.148. Online delayed-feedback variants reach 0.115–0.130. Matched-budget outcome-trained assessment improves over the gated route, reaching 0.168 macro MCC, but remains below probe-consensus assessment at 0.196. These results suggest that development outcomes reveal source-choice signal, but do not by themselves provide a stable future routing rule.

C.2 Proposal-State Design Analysis

This section evaluates whether richer decision-time observable features substantially improve source-improvement separability across chronological periods. The two-axis SDA geometry provides a compact observable state for stable source-change proposals.

Table 15: Cross-period separability of observable proposal states. AUC measures whether a candidate anchor improves the current route; the two temporal directions evaluate representation separability across adjacent periods.

Observable representation	LR forward	LR reverse	RF forward	RF reverse
Two-axis SDA geometry	0.530	0.505	0.596	0.590
Compact conflict-packet fields	0.556	0.542	0.600	0.575
Expanded decision-time state	0.587	0.525	0.548	0.599
All observable state and packet fields	0.589	0.557	0.563	**0.610**
All observable fields plus probe scores	**0.602**	0.550	0.571	0.598

Table 15 shows that the compact two-axis SDA geometry already provides comparable cross-period AUC to larger observable feature sets. Expanded feature sets slightly improve some within-direction scores, but the gains are inconsistent across temporal directions and model families. This supports the use of a compact observable-state proposal rule before frozen-representation assessment.

C.3 Assessment Robustness and Representation Diagnostics

Representation geometry. Principal Component Analysis (PCA) indicates that the leading principal component explains 15.5% of variance in synthetic packets versus 47.8% in real empirical packets, yielding an effective rank reduction from 23.6 to 7.1. This structural compression reflects the discrete template nature of empirical conflict packets, supporting entropy-based stability auditing over raw latent geometry.

Table 16: Probe signal ablation at seven retained proposals. Candidate-action-family entropy, low top probability, and low margin are monotone-equivalent on this proposal set. Harm is the fraction of retained switch days with negative net-correct change.

Assessment score	Macro MCC	Harmful rate	Net correct
Candidate-action entropy (high)	**0.188**	**0.0%**	+118
Candidate-action top probability (low)	**0.188**	**0.0%**	+118
Candidate-action margin (low)	**0.188**	**0.0%**	+118
Candidate-line entropy (high)	0.185	14.3%	+95
Agreement-regime top probability (high)	0.182	14.3%	+64
Mean entropy over all heads (high)	0.153	42.9%	−39
Random seven-proposal subset (mean)	0.164	—	—

Readout Specificity. The assessment score should rank proposal conflicts, not summarize all synthetic axes uniformly. Candidate-action entropy, low top probability, and low margin induce the same retained proposal set on the reported proposals and retain no harmful changes. Broader readouts are less selective: candidate-line entropy and agreement-regime top probability retain harmful proposals, and mean entropy over all heads admits substantially more harmful changes. This supports using the candidate-action-family readout as the default ranking signal.

Table 17: Raw probe-consensus stability with seven proposal-pool retained proposals. Macro MCC and the fraction of probe subsets that outperform the unfiltered SDA proposal route (0.181 macro MCC). The final method subsequently applies the proposal-pool fraction, group floor, and stability-selection rule.

S_{probe}	Mean MCC	5th percentile	Minimum	$\Pr(\text{MCC} > \text{SDA})$
1	0.178	0.163	0.145	33.7%
10	0.182	0.168	0.150	51.6%
50	0.185	0.174	0.158	71.8%
75	0.187	0.179	0.167	91.2%
101	0.188	0.188	0.188	100.0%

Probe-ensemble stability. Increasing the probe ensemble size improves the probability of outperforming unfiltered SDA, reaching 100% at $S_{\text{probe}} = 101$.

Table 18: Frozen-backbone ablation under one probe protocol. Only the cached final-layer packet representation changes; the synthetic split, 16-head MLP, seeds 0–100, SDA proposals, and $(\rho, \eta, \gamma) = (0.5, 0.25, 0.90)$ retention rule are unchanged. Retained reports proposals as total (beneficial/harmful), and Net is their net change in correct predictions. Gap to Gate denotes the macro-MCC difference from the frozen gated route.

Frozen backbone	ACL	BigData	CIKM	Macro
Qwen2.5-7B-Instruct	**0.247**	**0.155**	**0.186**	**0.196**
FinGPT-Llama2-7B	0.191	0.124	**0.186**	0.167
Qwen2.5-14B-Instruct	0.199	**0.155**	0.135	0.163
DeepSeek-R1-Distill-Qwen-7B	0.216	0.140	0.121	0.159
Frozen gated route (no final assessment)	0.205	0.109	0.130	0.148

Frozen backbone	Retained (B/H)	Net	Gap to Gate
Qwen2.5-7B-Instruct	8 (8/0)	+134	+0.048
FinGPT-Llama2-7B	9 (7/2)	+22	+0.019
Qwen2.5-14B-Instruct	7 (5/2)	+20	+0.015
DeepSeek-R1-Distill-Qwen-7B	7 (5/2)	+42	+0.011
Frozen gated route (no final assessment)	0	0	—

Backbone Sensitivity. Representation choice affects route quality, while all tested backbones remain above the gated route under the same probe protocol. The performance ordering is not monotonic in model size, indicating that parameter count alone does not explain representation suitability for conflict assessment.

C.4 Hyperparameter Sensitivity

Table 19: Parameter sensitivity of proposal and stability thresholds.

Parameter sweep	**Setting**	**Switches**	**Beneficial / harmful**	**CIKM MCC**	**Macro MCC**
Stability threshold γ	> 0.50	8	8 / 0	0.186	0.196
	> 0.75	8	8 / 0	0.186	0.196
	> 0.90	8	8 / 0	0.186	**0.196**
	> 0.95	7	7 / 0	0.161	0.188
Proposal margin τ_{SDA}	0.000 (raw 0.00)	112	—	—	0.098
	0.150 (raw 0.30)	54	—	—	0.128
	0.250 (raw 0.50)	46	28 / 18	—	0.133
	0.375 (raw 0.75)	29	20 / 9	—	0.162
	0.500 (raw 1.00)	14	10 / 4	—	**0.181**

Operating Stability. The reported setting lies on a broad stability plateau. The stability threshold is also insensitive over $\gamma \in [0.50, 0.90]$, where the retained proposal set, harmful-proposal count, and macro MCC remain unchanged. Only the stricter $\gamma > 0.95$ setting becomes more conservative and drops to 0.188 macro MCC. Corollary E.9 gives the corresponding monotonicity of missed beneficial and retained harmful mass.

Table 19 reports sensitivity to the stability threshold γ and the SDA proposal margin τ_{SDA}.

C.5 Temporal Block-Bootstrap Robustness

To test whether routing improvements are concentrated in a small number of neighboring test days, Table 7 reports paired circular block-bootstrap evaluations over 20,000 resamples.

D Cross-Domain Transfer to Legal Judgment

Appendix D documents the legal-domain instantiation of EPR. The transfer preserves the context-level authority-allocation abstraction and the proposal–assessment sequence, while replacing finance-specific anchors and observable-state rules with legal-domain counterparts. No finance-specific market summaries are used in the legal experiment.

D.1 Legal Task and Evaluation Protocol

The ClassActionPrediction benchmark evaluates binary case outcomes. We use the public 3,000-case release with a deterministic $2,398/295/307$ train/validation/test split. Each routing context consists of one target case and its visible citation-group neighborhood. The route selects one legal-domain anchor to govern the case-level prediction.

D.2 Legal Decision Anchors

Table 20: Legal-domain anchors.

Anchor	Evidence scope	Implementation / visibility
Sample-text	Target complaint text	Fine-tuned LegalBERT over raw complaint text
Group-statistical	Structured case metadata	Categorical/statistical prior over visible group fields
Relation	Citation neighborhood	Text-neighborhood graph encoder over citation structure

These anchors instantiate the same Evidence Provenance roles as the FinBen anchors: local text, group context, and relational support.

D.3 Legal EPR Instantiation

Legal Observable-State Proposal. The relation anchor serves as the base route. A candidate switch to the sample-text anchor is emitted when three pre-evaluation conditions hold: (i) the sample-text and group-statistical anchors agree against the relation anchor, (ii) the sample-text confidence exceeds the relation-anchor confidence, and (iii) citation-neighborhood support exceeds the development-set median support score. These conditions use only anchor outputs and visible citation-group support; case outcomes are joined only after routes are fixed.

The probe ensemble evaluates proposed switches via consensus stability. The decision-routing sequence is stable under hyperparameter variation: varying the relation confidence floor within $[0.10, 0.25]$ and agreement threshold within $[0.85, 0.95]$ consistently achieves the best results on both validation (0.341 MCC) and evaluation splits (0.437 MCC).

D.4 Transfer Results and Routing Progression

Table 21: Legal baselines under the common 2,398/295/307 split. Text models consume complaint text; extraction-field models consume only the frozen structured fields available to the legal transfer.

Model Class	Method	Input Modality	Accuracy	MCC
Heuristic	Training-Set Majority Predictor	Label Prior	0.446	0.000
Linear / Kernel	TF-IDF + Logistic Regression	Full Text	0.687	0.386
Linear / Kernel	TF-IDF + Linear SVM	Full Text	0.674	0.360
Linear / Kernel	TF-IDF + SGD logistic	Full Text	0.678	0.360
Feature Ensemble	Extraction Fields + Logistic Regression	Structured Metadata	0.632	0.260
Feature Ensemble	Extraction Fields + Random Forest	Structured Metadata	0.651	0.345
Domain Transformer	CustomLegalBERT	Full Text	0.704	0.400
Domain Transformer	LegalBERT	Full Text	0.704	0.406
EPR (Ours)	Evidence Provenance Routing	Multi-Source Anchors	0.720	0.437

Table 22: Legal routing progression by split. Test intervention counts are measured against the text-neighborhood relation base. B/H denotes beneficial/harmful switches measured by test outcomes after the route is fixed; these labels are not used by the proposal or assessment rule.

Routing stage	Train	Validation	Test	Test switches (B/H)
Sample-Text Anchor Standalone	0.319	0.324	0.406	38 (18/20)
Group-Statistical Anchor Standalone	0.353	0.252	0.331	42 (14/28)
Relation Base Anchor $r^{(0)}$	0.300	0.331	0.417	—
+ Observable-State Proposals (SDA)	0.314	0.334	0.430	6 (4/2)
+ Probe-Consensus Assessment (Full EPR)	0.307	0.341	0.437	5 (4/1)

The legal transfer shows that the EPR abstraction can be re-instantiated when a domain provides separable evidence scopes and observable support conditions. In this setting, the relation anchor provides a strong base, observable proposals recover beneficial sample-text switches, and probe consensus removes one harmful switch while retaining the beneficial ones.

E Theoretical Analysis

Appendix E is organized in two layers. Section E.1 establishes the common decision-theoretic setup. Section E.2 gives the three core results corresponding directly to Section 2.5: context-dependent source authority, provenance-preserving information dominance with quantitative compression costs, and the information–action decomposition of routing risk with gain-weighted consensus assessment. Section E.3 collects framework-level theoretical implications for the representation, proposal, and assessment components of EPR.

E.1 Foundations and Bayes Routing

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let U be a random routing context with $|\mathcal{I}_{\mathsf{U}}| \geq 1$. We use U for a random routing context and u for its realization. The finite anchor set is $\mathcal{A} = \{a_1, \ldots, a_K\}$. For $a \in \mathcal{A}$, let $\hat{y}_{a,\mathsf{U},i}$ denote anchor a's prediction for sample i in random context U. For anchor $a \in \mathcal{A}$, define the realized context-level loss

$$\Lambda(a) = \frac{1}{|\mathcal{I}_{\mathsf{U}}|} \sum_{i \in \mathcal{I}_{\mathsf{U}}} \ell(Y_{\mathsf{U},i}, \hat{y}_{a,\mathsf{U},i}) \,. \tag{16}$$

We assume that each $\Lambda(a)$ is integrable. Bounds below additionally assume $0 \le \Lambda(a) \le L_{\max} < \infty$, where $L_{\max}$ is a uniform loss bound. All oracle choices use a fixed tie-breaking rule. Expectations include the random context, outcomes, and any internal randomization.

Let $\mathbf{V} = \mathbf{V}_{\mathsf{U}}$ denote the context summaries visible to the router, let $\mathbf{O} = (\mathbf{O}_1, \dots, \mathbf{O}_K)$ denote the source-indexed anchor-output state, and let Z denote the SDA-specific observable state. Define the routing information sigma-field as $\mathcal{H}_{\mathrm{rt}} = \sigma(\mathbf{V}, \mathbf{O}, Z)$. Let $\mathcal{H} \subseteq \mathcal{F}$ be the information available to a routing policy. Define anchor a's conditional risk under $\mathcal{H}$ as

$$\mu_a^{\mathcal{H}} = \mathbb{E}[\Lambda(a) \mid \mathcal{H}]. \tag{17}$$

Let $\mathbf{1}\{\cdot\}$ denote an event indicator. Define $\Pi(\mathcal{H}; \mathcal{A}) = \{\pi : \Omega \to \mathcal{A} \mid \pi \text{ is } \mathcal{H}\text{-measurable}\}$. For $\pi \in \Pi(\mathcal{H}; \mathcal{A})$, define $R(\pi) = \mathbb{E}[\Lambda(\pi)]$ and

$$R^\star(\mathcal{H}) = \inf_{\pi \in \Pi(\mathcal{H}; \mathcal{A})} R(\pi). \tag{18}$$

For an information set $\mathcal{H}$ and a nonempty random candidate set $\mathcal{C} \subseteq \mathcal{A}$ whose membership events $\{a \in \mathcal{C}\}$ are $\mathcal{H}$-measurable for every $a \in \mathcal{A}$, define the candidate-restricted Bayes risk as

$$R^\star(\mathcal{H}, \mathcal{C}) = \mathbb{E}\left[\min_{a \in \mathcal{C}} \mu_a^{\mathcal{H}}\right]. \tag{19}$$

Lemma E.1 (Bayes Routing under an Information Set)**.** *For any information set $\mathcal{H}$,*

$$R^\star(\mathcal{H}) = \mathbb{E}\left[\min_{a \in \mathcal{A}} \mu_a^{\mathcal{H}}\right]. \tag{20}$$

Proof. For any $\mathcal{H}$-measurable policy π,

$$R(\pi) = \mathbb{E}[\mathbb{E}[\Lambda(\pi) \mid \mathcal{H}]] = \mathbb{E}\left[\sum_{a \in \mathcal{A}} \mathbf{1}\{\pi = a\} \mu_a^{\mathcal{H}}\right] \ge \mathbb{E}\left[\min_{a \in \mathcal{A}} \mu_a^{\mathcal{H}}\right].$$

Because $\mathcal{A}$ is finite, the fixed tie-breaking rule gives an $\mathcal{H}$-measurable minimizer in $\arg\min_{a \in \mathcal{A}} \mu_a^{\mathcal{H}}$ that attains the lower bound. ■

E.2 Core Theoretical Results

E.2.1 Value of Context-Dependent Source Authority

Using the SDA-specific decision-time observable state Z defined above, define $\mu_a(Z) = \mathbb{E}[\Lambda(a) \mid Z]$. The best fixed-anchor risk and the best context-dependent routing risk are

$$R^\star_{\mathrm{static}} = \min_{a \in \mathcal{A}} \mathbb{E}[\Lambda(a)], \qquad R^\star_{\mathrm{ctx}} = \mathbb{E}\left[\min_{a \in \mathcal{A}} \mu_a(Z)\right]. \tag{21}$$

Define the authority heterogeneity gap as $\mathcal{G}_{\mathrm{auth}} = R^\star_{\mathrm{static}} - R^\star_{\mathrm{ctx}}$.

Theorem E.2 (Value of Context-Dependent Source Authority; Theorem 2.1 in the main text)**.** *For any finite anchor set,*

$$\mathcal{G}_{\mathrm{auth}} = \min_{a \in \mathcal{A}} \mathbb{E}\left[\mu_a(Z) - \min_{b \in \mathcal{A}} \mu_b(Z)\right] \ge 0. \tag{22}$$

Moreover, $\mathcal{G}_{\mathrm{auth}} = 0$ if and only if there exists a fixed anchor $a_0 \in \mathcal{A}$ such that $a_0 \in \arg\min_{a \in \mathcal{A}} \mu_a(Z)$ almost surely.

Proof. We have

$$\mathcal{G}_{\mathrm{auth}} = \min_{a \in \mathcal{A}} \mathbb{E}[\mu_a(Z)] - \mathbb{E}\left[\min_{b \in \mathcal{A}} \mu_b(Z)\right] = \min_{a \in \mathcal{A}} \mathbb{E}\left[\mu_a(Z) - \min_{b \in \mathcal{A}} \mu_b(Z)\right].$$

Every integrand is nonnegative, so $\mathcal{G}_{\mathrm{auth}} \ge 0$. If a fixed a_0 minimizes the conditional risk almost surely, its expected gap is zero. Conversely, because $\mathcal{A}$ is finite, $\mathcal{G}_{\mathrm{auth}} = 0$ implies that some a_0 satisfies

$$\mathbb{E}\left[\mu_{a_0}(Z) - \min_{b \in \mathcal{A}} \mu_b(Z)\right] = 0.$$

The integrand is nonnegative, so it must vanish almost surely. ■

Corollary E.3 (Pairwise Authority-Reversal Gap)**.** *For two anchors a_1 and a_2, define $\Xi(Z) = \mu_{a_1}(Z) - \mu_{a_2}(Z)$. Then*

$$\mathcal{G}_{\text{auth}} = \frac{1}{2}\left(\mathbb{E}|\Xi(Z)| - |\mathbb{E}\Xi(Z)|\right). \tag{23}$$

Consequently, $\mathcal{G}_{\text{auth}} > 0$ if and only if $\mathbb{P}(\Xi(Z) > 0) > 0$ and $\mathbb{P}(\Xi(Z) < 0) > 0$.

Proof. The identity $\min(x, y) = (x + y - |x - y|)/2$ gives

$$R^\star_{\text{ctx}} = \tfrac{1}{2}\left(\mathbb{E}[\mu_{a_1}(Z)] + \mathbb{E}[\mu_{a_2}(Z)] - \mathbb{E}|\Xi(Z)|\right),$$
$$R^\star_{\text{static}} = \tfrac{1}{2}\left(\mathbb{E}[\mu_{a_1}(Z)] + \mathbb{E}[\mu_{a_2}(Z)] - |\mathbb{E}\Xi(Z)|\right).$$

Subtracting yields Equation 23. Equality in $|\mathbb{E}\Xi(Z)| \le \mathbb{E}|\Xi(Z)|$ holds exactly when $\Xi(Z)$ has one sign almost surely. ■

E.2.2 Provenance-Preserving Information Dominance

Let $\widetilde{\mathbf{O}} = f(\mathbf{O})$ be a deterministic compression. Define the full and compressed information sets as

$$\mathcal{H}_{\text{full}} := \mathcal{H}_{\text{rt}} = \sigma(\mathbf{V}, \mathbf{O}, Z), \qquad \mathcal{H}_{\text{comp}} := \sigma(\mathbf{V}, \widetilde{\mathbf{O}}, Z). \tag{24}$$

Because $\widetilde{\mathbf{O}}$ is a function of $\mathbf{O}$, $\mathcal{H}_{\text{comp}} \subseteq \mathcal{H}_{\text{full}}$.

Theorem E.4 (Provenance-Preserving Information Dominance; Theorem 2.2 in the main text)**.** *For any measurable deterministic compression f,*

$$R^\star(\mathcal{H}_{\text{full}}) \le R^\star(\mathcal{H}_{\text{comp}}). \tag{25}$$

Define $\Gamma_{\text{full}} = \arg\min_{a\in\mathcal{A}} \mu_a^{\mathcal{H}_{\text{full}}}$ as the set of full-information Bayes anchors. Equality holds if and only if there exists an $\mathcal{H}_{\text{comp}}$-measurable policy π satisfying $\pi \in \Gamma_{\text{full}}$ almost surely. If Γ_{full} is a singleton almost surely, denote its unique element by $A^\star_{\text{full}}$. In this case, equality holds if and only if $A^\star_{\text{full}}$ is $\mathcal{H}_{\text{comp}}$-measurable.

Proof. Since $\mathcal{H}_{\text{comp}} \subseteq \mathcal{H}_{\text{full}}$, every policy in $\Pi(\mathcal{H}_{\text{comp}}; \mathcal{A})$ also belongs to $\Pi(\mathcal{H}_{\text{full}}; \mathcal{A})$. Taking the infimum over the nested policy classes gives Equation 25. If an $\mathcal{H}_{\text{comp}}$-measurable policy selects from Γ_{full} almost surely, it attains $R^\star(\mathcal{H}_{\text{full}})$, so equality follows. Conversely, suppose equality holds and let $\pi^\star_{\mathcal{H}_{\text{comp}}}$ be an $\mathcal{H}_{\text{comp}}$-Bayes policy. Then

$$0 = R(\pi^\star_{\mathcal{H}_{\text{comp}}}) - R^\star(\mathcal{H}_{\text{full}}) = \mathbb{E}\left[\mu^{\mathcal{H}_{\text{full}}}_{\pi^\star_{\mathcal{H}_{\text{comp}}}} - \min_{a\in\mathcal{A}} \mu_a^{\mathcal{H}_{\text{full}}}\right].$$

The integrand is nonnegative, so $\pi^\star_{\mathcal{H}_{\text{comp}}} \in \Gamma_{\text{full}}$ almost surely. Under uniqueness, this condition is equivalent to $\pi^\star_{\mathcal{H}_{\text{comp}}} = A^\star_{\text{full}}$ almost surely, which is possible exactly when $A^\star_{\text{full}}$ is $\mathcal{H}_{\text{comp}}$-measurable. ■

This result connects evidence identity to the routing state. Decision-equivalent compression preserves the information required to recover the optimal authority decision.

Corollary E.5 (Exact Pairwise Cost of Provenance Compression; Corollary 2.3 in the main text)**.** *Fix two anchors a_1 and a_2. Then*

$$R^\star(\mathcal{H}_{\text{comp}}, \{a_1, a_2\}) - R^\star(\mathcal{H}_{\text{full}}, \{a_1, a_2\})$$
$$= \frac{1}{2}\mathbb{E}\left[\left|\mu^{\mathcal{H}_{\text{full}}}_{a_1} - \mu^{\mathcal{H}_{\text{full}}}_{a_2}\right| - \left|\mu^{\mathcal{H}_{\text{comp}}}_{a_1} - \mu^{\mathcal{H}_{\text{comp}}}_{a_2}\right|\right]. \tag{26}$$

The gap is strictly positive if and only if the events $\mu^{\mathcal{H}_{\text{full}}}_{a_1} - \mu^{\mathcal{H}_{\text{full}}}_{a_2} > 0$ and $\mu^{\mathcal{H}_{\text{full}}}_{a_1} - \mu^{\mathcal{H}_{\text{full}}}_{a_2} < 0$ both have positive conditional probability given $\mathcal{H}_{\text{comp}}$ on a set of positive probability.

Proof. The definition of candidate-restricted Bayes risk and $\min(x, y) = \frac{1}{2}(x + y - |x - y|)$ give

$$R^\star(\mathcal{H}_{\text{full}}, \{a_1, a_2\}) = \frac{1}{2}\mathbb{E}\left[\mu^{\mathcal{H}_{\text{full}}}_{a_1} + \mu^{\mathcal{H}_{\text{full}}}_{a_2} - \left|\mu^{\mathcal{H}_{\text{full}}}_{a_1} - \mu^{\mathcal{H}_{\text{full}}}_{a_2}\right|\right],$$

and the same expression holds with $\mathcal{H}_{\text{comp}}$. The tower property gives $\mu_a^{\mathcal{H}_{\text{comp}}} = \mathbb{E}[\mu_a^{\mathcal{H}_{\text{full}}} \mid \mathcal{H}_{\text{comp}}]$, so the unconditional expectations of the sum terms agree. Subtraction gives Equation 26. Conditioning on $\mathcal{H}_{\text{comp}}$ yields

$$\mathbb{E}\left[\left|\mu^{\mathcal{H}_{\text{full}}}_{a_1} - \mu^{\mathcal{H}_{\text{full}}}_{a_2}\right| \middle| \mathcal{H}_{\text{comp}}\right] \ge \left|\mu^{\mathcal{H}_{\text{comp}}}_{a_1} - \mu^{\mathcal{H}_{\text{comp}}}_{a_2}\right|.$$

Equality holds exactly when the full-information conditional loss difference has one sign conditionally almost surely. The conditional inequality is strict exactly when both sign events have positive conditional probability. Integrating proves the stated condition. ■

Compression is costly exactly when it aliases full-information states whose optimal authority differs.

Proposition E.6 (Finite-Anchor Binary Lower-Envelope Recursion)**.** *Fix any ordering $a_1, \dots, a_K$ of the finite anchor set. For each $k \in \{1, \dots, K\}$ and $\mathcal{J} \in \{\mathcal{H}_{\mathrm{full}}, \mathcal{H}_{\mathrm{comp}}\}$, define the conditional-risk lower envelope*

$$\psi_k^{\mathcal{J}} = \min_{1 \le j \le k} \mu_{a_j}^{\mathcal{J}},$$

and define the provenance-compression gap after k anchors as

$$\mathcal{R}_{\mathrm{info},k} = \mathbb{E}\left[\psi_k^{\mathcal{H}_{\mathrm{comp}}} - \psi_k^{\mathcal{H}_{\mathrm{full}}}\right].$$

For a real number x, let $(x)_+ = \max\{x, 0\}$ denote its positive part. Then $\mathcal{R}_{\mathrm{info},1} = 0$, and for $k = 2, \dots, K$,

$$\begin{aligned}\mathcal{R}_{\mathrm{info},k} = \mathcal{R}_{\mathrm{info},k-1} &+ \mathbb{E}\left[\left(\psi_{k-1}^{\mathcal{H}_{\mathrm{full}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{full}}}\right)_+\right] \\ &- \mathbb{E}\left[\left(\psi_{k-1}^{\mathcal{H}_{\mathrm{comp}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{comp}}}\right)_+\right]. \end{aligned} \tag{27}$$

Equivalently,

$$\begin{aligned}\mathcal{R}_{\mathrm{info},k} = \frac{1}{2}\mathcal{R}_{\mathrm{info},k-1} + \frac{1}{2}\mathbb{E}\Big[&\left|\psi_{k-1}^{\mathcal{H}_{\mathrm{full}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{full}}}\right| \\ &- \left|\psi_{k-1}^{\mathcal{H}_{\mathrm{comp}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{comp}}}\right|\Big]. \end{aligned} \tag{28}$$

For $k = 2$, $\psi_1^{\mathcal{J}} = \mu_{a_1}^{\mathcal{J}}$ and $\mathcal{R}_{\mathrm{info},1} = 0$, so Equation 28 equals the pairwise identity in Equation 26. At $k = K$,

$$\mathcal{R}_{\mathrm{info},K} = R^\star(\mathcal{H}_{\mathrm{comp}}) - R^\star(\mathcal{H}_{\mathrm{full}}), \tag{29}$$

which is invariant to the anchor ordering. Intermediate values $\mathcal{R}_{\mathrm{info},k}$ may depend on that ordering.

Proof. Because $\mathcal{H}_{\mathrm{comp}} \subseteq \mathcal{H}_{\mathrm{full}}$, the tower property gives $\mu_a^{\mathcal{H}_{\mathrm{comp}}} = \mathbb{E}[\mu_a^{\mathcal{H}_{\mathrm{full}}} \mid \mathcal{H}_{\mathrm{comp}}]$ for every anchor a. Hence $\mathcal{R}_{\mathrm{info},1} = 0$. For $k \ge 2$, applying $\min(x, y) = x - (x - y)_+$ to each lower envelope gives Equation 27. Applying $\min(x, y) = \frac{1}{2}(x + y - |x - y|)$ instead, together with $\mathbb{E}[\mu_{a_k}^{\mathcal{H}_{\mathrm{comp}}}] = \mathbb{E}[\mu_{a_k}^{\mathcal{H}_{\mathrm{full}}}]$, gives Equation 28. The case $k = 2$ follows by substitution. Finally, $\psi_K^{\mathcal{J}} = \min_{a \in \mathcal{A}} \mu_a^{\mathcal{J}}$, so Lemma E.1 gives Equation 29; the terminal lower envelope is unchanged by permuting the anchors. ■

Incremental value. Equation 27 gives

$$\begin{aligned}\mathcal{R}_{\mathrm{info},k} - \mathcal{R}_{\mathrm{info},k-1} = &\mathbb{E}\left[\left(\psi_{k-1}^{\mathcal{H}_{\mathrm{full}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{full}}}\right)_+\right] \\ &- \mathbb{E}\left[\left(\psi_{k-1}^{\mathcal{H}_{\mathrm{comp}}} - \mu_{a_k}^{\mathcal{H}_{\mathrm{comp}}}\right)_+\right]. \end{aligned}$$

An additional anchor can either increase provenance-sensitive decision opportunities or provide a stable alternative that reduces the decision value of provenance.

Further consequences.

Corollary (Provenance-Aliasing Strictness)**.** *Let $W = (\mathbf{V}, Z, \widetilde{\mathbf{O}})$, so that $\mathcal{H}_{\mathrm{full}} = \sigma(\mathbf{V}, \mathbf{O}, Z)$ and $\mathcal{H}_{\mathrm{comp}} = \sigma(W)$. Assume that the $\mathcal{H}_{\mathrm{full}}$-Bayes authority selector $A_{\mathrm{full}}^\star$ is unique almost surely. Suppose there exists a measurable, nonsingular transformation $\tau : \Omega \to \Omega$ such that $W \circ \tau = W$ almost surely but*

$$\mathbb{P}(A_{\mathrm{full}}^\star \circ \tau \neq A_{\mathrm{full}}^\star) > 0, \tag{30}$$

Then $A_{\mathrm{full}}^\star$ is not $\mathcal{H}_{\mathrm{comp}}$-measurable, and therefore

$$R^\star(\mathcal{H}_{\mathrm{full}}) < R^\star(\mathcal{H}_{\mathrm{comp}}).$$

Proof. Because W is invariant under the nonsingular transformation τ, every $\mathcal{H}_{\mathrm{comp}}$-measurable authority selector is invariant under τ almost surely. Equation 30 therefore implies that $A^\star_{\mathrm{full}}$ is not $\mathcal{H}_{\mathrm{comp}}$-measurable. The equality condition in Theorem E.4 then gives the strict inequality. ■

Thus, when compression aliases source configurations that require different authority decisions, provenance loss induces a strict decision-risk penalty.

Corollary (Decision-Time Information Value)**.** *Let $\mathcal{H}_{\mathrm{hist}}$ denote the information available from historical source preferences and outcomes, and define $\mathcal{H}_{\mathrm{dt}} = \mathcal{H}_{\mathrm{hist}} \vee \sigma(Z)$, where $\vee$ denotes the smallest sigma-field containing both information sets. Then*

$$R^\star(\mathcal{H}_{\mathrm{dt}}) \le R^\star(\mathcal{H}_{\mathrm{hist}}). \tag{31}$$

Equality holds if and only if the historical information still admits an $\mathcal{H}_{\mathrm{dt}}$-optimal authority selector. If the $\mathcal{H}_{\mathrm{dt}}$-optimal anchor is unique almost surely and is not $\mathcal{H}_{\mathrm{hist}}$-measurable, the inequality is strict.

Proof. Since $\mathcal{H}_{\mathrm{hist}} \subseteq \mathcal{H}_{\mathrm{dt}}$, the policy class measurable with respect to $\mathcal{H}_{\mathrm{hist}}$ is contained in the class measurable with respect to $\mathcal{H}_{\mathrm{dt}}$. The inequality and equality condition follow from the same nested-policy argument as Theorem E.4. ■

E.2.3 Information–Action Decomposition of Routing Risk

Let $\mathcal{G}$ be a provenance-preserving information set and let $\mathcal{H} \subseteq \mathcal{G}$ be the information retained by a router. Let $\mathcal{C} \subseteq \mathcal{A}$ be a nonempty random candidate set whose membership events $\{a \in \mathcal{C}\}$ are $\mathcal{H}$-measurable for every $a \in \mathcal{A}$. Let $\widehat{A}$ be an $\mathcal{H}$-measurable assessed anchor satisfying $\widehat{A} \in \mathcal{C}$ almost surely.

Theorem E.7 (Information–Action Decomposition of Routing Risk; Theorem 2.4 in the main text)**.** *The routing excess risk admits the exact decomposition*

$$\begin{aligned} R(\widehat{A}) - R^\star(\mathcal{G}) = {} & \underbrace{R^\star(\mathcal{H}) - R^\star(\mathcal{G})}_{\mathcal{R}_{\mathrm{info}}} + \underbrace{R^\star(\mathcal{H},\mathcal{C}) - R^\star(\mathcal{H})}_{\mathcal{R}_{\mathrm{prop}}} \\ & + \underbrace{R(\widehat{A}) - R^\star(\mathcal{H},\mathcal{C})}_{\mathcal{R}_{\mathrm{assess}}}, \end{aligned} \tag{32}$$

where $\mathcal{R}_{\mathrm{info}}$, $\mathcal{R}_{\mathrm{prop}}$, and $\mathcal{R}_{\mathrm{assess}}$ are the information, proposal, and assessment terms, respectively, and all three are nonnegative.

Proof. Add and subtract $R^\star(\mathcal{H})$ and $R^\star(\mathcal{H},\mathcal{C})$. The information term is nonnegative because $\mathcal{H} \subseteq \mathcal{G}$. The proposal term is nonnegative because $\mathcal{C} \subseteq \mathcal{A}$ implies $\min_{a\in\mathcal{C}} \mu_a^{\mathcal{H}} \ge \min_{a\in\mathcal{A}} \mu_a^{\mathcal{H}}$ pointwise. The assessment term is nonnegative because $\widehat{A} \in \mathcal{C}$ and

$$R(\widehat{A}) = \mathbb{E}[\mu_{\widehat{A}}^{\mathcal{H}}] \ge \mathbb{E}\left[\min_{a\in\mathcal{C}} \mu_a^{\mathcal{H}}\right].$$

This proves the decomposition and nonnegativity. ■

EPR instantiation. For the sampled context U, let $\mathcal{H}_{\mathrm{EPR}}$ denote the sigma-field generated by U and the indexed collection of routing-visible states and proposals used by proposal-pool reranking. With trained components fixed, current routes, candidate sets, conflict packets, frozen representations, probe readouts, proposal rankings, and final selected anchors are $\mathcal{H}_{\mathrm{EPR}}$-measurable. Taking $\mathcal{G} = \mathcal{H} = \mathcal{H}_{\mathrm{EPR}}$ gives $\mathcal{R}_{\mathrm{info}} = 0$. When SDA emits no proposal, $\mathcal{C}$ contains only the current anchor; when it emits $a_c \to a_j$, $\mathcal{C} = \{a_c, a_j\}$. The term $\mathcal{R}_{\mathrm{prop}}$ measures this candidate restriction, and $\mathcal{R}_{\mathrm{assess}}$ measures probe-consensus selection error relative to the candidate-set Bayes rule.

Gain-weighted consensus assessment. After proposal ranking, assume that each routing context has at most one eligible proposal $p : a_c \to a_j$ and that its identity remains fixed throughout leave-one-probe reranking. Eligibility, anchor identities, rankings, and tie-breaking are $\mathcal{H}_{\mathrm{EPR}}$-measurable. The context losses are integrable, and contexts and probe readouts may be arbitrarily dependent. On

an eligible context, let $\mathcal{C} = \{a_c, a_j\}$ and let $g(p) = \Lambda(a_c) - \Lambda(a_j)$ denote the realized switch gain. Define the conditional switch gain as

$$\theta = \mathbb{E}[g(p) \mid \mathcal{H}_{\mathrm{EPR}}] = \mu_{a_c}^{\mathcal{H}_{\mathrm{EPR}}} - \mu_{a_j}^{\mathcal{H}_{\mathrm{EPR}}}. \tag{33}$$

On a context without an eligible proposal, let $\mathcal{C} = \{a_c\}$, set $\theta = 0$ and $F(p) = 0$, and retain a_c. Thus, all expectations below use the full random-context risk from Section E.1. For $S_{\mathrm{probe}} \geq 2$ and $0 \leq \gamma < 1$, let $\widehat{A}_\gamma$ denote the final anchor that selects a_j when $F(p) > \gamma$ and otherwise selects a_c. In the information–action decomposition, set $\widehat{A} = \widehat{A}_\gamma$ and $\mathcal{H} = \mathcal{H}_{\mathrm{EPR}}$.

Theorem E.8 (Gain-Weighted Assessment Regret and Finite-Consensus Bound)**.** *Under the setting above, the assessment term has the exact form*

$$\mathcal{R}_{\mathrm{assess}} = \mathbb{E}[(\theta)_+ \mathbf{1}\{F(p) \leq \gamma\} + (-\theta)_+ \mathbf{1}\{F(p) > \gamma\}]. \tag{34}$$

It satisfies

$$\mathcal{R}_{\mathrm{assess}} \leq \mathbb{E}\left[\frac{S_{\mathrm{probe}}}{S_{\mathrm{probe}} - \lfloor S_{\mathrm{probe}}\gamma \rfloor}(\theta)_+\big(1 - F(p)\big) + \frac{S_{\mathrm{probe}}}{\lfloor S_{\mathrm{probe}}\gamma \rfloor + 1}(-\theta)_+ F(p)\right]. \tag{35}$$

The two coefficients in Equation 35 are individually optimal over this class of retention rules. They are also attained by leave-one-probe top-one ranking of two proposals with scores in $[0, 1]$.

Proof. Conditional on $\mathcal{H}_{\mathrm{EPR}}$, the selected-anchor risk is

$$\mathbb{E}[\Lambda(\widehat{A}_\gamma) \mid \mathcal{H}_{\mathrm{EPR}}] = \mu_{a_c}^{\mathcal{H}_{\mathrm{EPR}}} - \mathbf{1}\{F(p) > \gamma\}\theta,$$

whereas the candidate-set optimal risk is

$$\min_{a \in \mathcal{C}} \mu_a^{\mathcal{H}_{\mathrm{EPR}}} = \mu_{a_c}^{\mathcal{H}_{\mathrm{EPR}}} - (\theta)_+.$$

Their difference is

$$(\theta)_+ - \mathbf{1}\{F(p) > \gamma\}\theta = (\theta)_+ \mathbf{1}\{F(p) \leq \gamma\} + (-\theta)_+ \mathbf{1}\{F(p) > \gamma\}.$$

The zero extension gives the same identity on contexts without an eligible proposal. Taking expectations proves Equation 34.

Because $S_{\mathrm{probe}} F(p)$ is an integer, the strict rule $F(p) > \gamma$ gives

$$F(p) \leq \gamma \Longrightarrow S_{\mathrm{probe}} F(p) \leq \lfloor S_{\mathrm{probe}}\gamma \rfloor,$$
$$F(p) > \gamma \Longrightarrow S_{\mathrm{probe}} F(p) \geq \lfloor S_{\mathrm{probe}}\gamma \rfloor + 1.$$

Hence, pointwise,

$$\mathbf{1}\{F(p) \leq \gamma\} \leq \frac{S_{\mathrm{probe}}}{S_{\mathrm{probe}} - \lfloor S_{\mathrm{probe}}\gamma \rfloor}\big(1 - F(p)\big),$$
$$\mathbf{1}\{F(p) > \gamma\} \leq \frac{S_{\mathrm{probe}}}{\lfloor S_{\mathrm{probe}}\gamma \rfloor + 1} F(p).$$

Multiplying these inequalities by $(\theta)_+$ and $(-\theta)_+$, respectively, summing, and taking expectations proves Equation 35.

For optimality of the first coefficient, consider one proposal with $\theta = 1$ and exactly $\lfloor S_{\mathrm{probe}}\gamma \rfloor$ retention votes, with all other conditional gains equal to zero. The proposal is rejected, its assessment regret is one, and $(\theta)_+(1 - F(p)) = (S_{\mathrm{probe}} - \lfloor S_{\mathrm{probe}}\gamma \rfloor)/S_{\mathrm{probe}}$. For optimality of the second coefficient, take one proposal with $\theta = -1$ and exactly $\lfloor S_{\mathrm{probe}}\gamma \rfloor + 1$ retention votes. The proposal is accepted, its assessment regret is one, and $(-\theta)_+ F(p) = (\lfloor S_{\mathrm{probe}}\gamma \rfloor + 1)/S_{\mathrm{probe}}$. Both gains are realized by losses in $[0, 1]$. Embedding either construction in any collection of contexts multiplies both quantities by the same context probability and leaves the coefficient ratio unchanged.

It remains to realize the required vote counts through leave-one-probe top-one ranking. For a target count $t \in \{1, \ldots, S_{\mathrm{probe}} - 1\}$, use two proposals. In exactly t probe rows, assign the first and second proposals the scores

$$\frac{1}{2} - \frac{S_{\mathrm{probe}} - t}{4S_{\mathrm{probe}}}, \qquad \frac{1}{2} + \frac{S_{\mathrm{probe}} - t}{4S_{\mathrm{probe}}},$$

and in the remaining rows assign them

$$\frac{1}{2} + \frac{t}{4S_{\text{probe}}}, \qquad \frac{1}{2} - \frac{t}{4S_{\text{probe}}}.$$

The row-wise score advantages sum to zero. Leaving out a row reverses the sign of that row's advantage in the remaining mean, so the first proposal ranks first in exactly t leave-one-probe rounds. Constant negative and positive score advantages realize $t = 0$ and $t = S_{\text{probe}}$, respectively. Substituting $t = \lfloor S_{\text{probe}}\gamma \rfloor$ and $t = \lfloor S_{\text{probe}}\gamma \rfloor + 1$ gives the two equality constructions. ■

Reported finite coefficients. For $S_{\text{probe}} = 101$ and $\gamma = 0.90$, the two coefficients are $101/11$ and $101/91$.

Relative to accepting every eligible proposal, threshold assessment changes expected loss by $\mathbb{E}[\theta \mathbf{1}\{F(p) \le \gamma\}]$ and has lower expected loss exactly when this quantity is negative.

Corollary E.9 (Consensus-Threshold Risk Increment)**.** *Fix the candidate sets, $\mathcal{H}_{\text{EPR}}$, and leave-one-probe retention frequencies. For $0 \le \gamma_1 < \gamma_2 < 1$,*

$$R(\widehat{A}_{\gamma_2}) - R(\widehat{A}_{\gamma_1}) = \mathbb{E}[\theta \mathbf{1}\{\gamma_1 < F(p) \le \gamma_2\}]\,. \tag{36}$$

Consequently, $\mathbb{E}[(\theta)_+ \mathbf{1}\{F(p) \le \gamma\}]$ is nondecreasing in γ, while $\mathbb{E}[(-\theta)_+ \mathbf{1}\{F(p) > \gamma\}]$ is nonincreasing in γ.

Proof. Only proposals with $\gamma_1 < F(p) \le \gamma_2$ change from accepted to rejected. Each such change increases conditional loss by $\mu_{a_c}^{\mathcal{H}_{\text{EPR}}} - \mu_{a_j}^{\mathcal{H}_{\text{EPR}}} = \theta$. Taking expectations proves Equation 36; the two monotonicity statements follow from their nested indicator events. ■

E.3 Framework-Level Theoretical Implications

The following results connect the core characterization to additional parts of EPR, including stochastic information transformations, conflict-packet representations, and ex-post routing behavior.

E.3.1 Information-Theoretic Extensions of Provenance Preservation

Proposition E.10 (Conditional Information Companion)**.** *Let $\mathbf{Y}_{\mathsf{U}} := (Y_{\mathsf{U},i})_{i \in \mathcal{I}_{\mathsf{U}}}$ denote the outcome vector for context U, and assume the conditional mutual informations below are finite. If a deterministic or stochastic compression satisfies the Markov chain $\mathbf{Y}_{\mathsf{U}} \to (\mathbf{V}, Z, \mathbf{O}) \to \widetilde{\mathbf{O}}$, then*

$$I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O} \mid \mathbf{V}, Z) - I(\mathbf{Y}_{\mathsf{U}}; \widetilde{\mathbf{O}} \mid \mathbf{V}, Z) = I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O} \mid \widetilde{\mathbf{O}}, \mathbf{V}, Z) \ge 0. \tag{37}$$

Proof. The Markov condition gives $I(\mathbf{Y}_{\mathsf{U}}; \widetilde{\mathbf{O}} \mid \mathbf{V}, Z, \mathbf{O}) = 0$. The two chain-rule expansions are

$$I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O}, \widetilde{\mathbf{O}} \mid \mathbf{V}, Z) = I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O} \mid \mathbf{V}, Z) + I(\mathbf{Y}_{\mathsf{U}}; \widetilde{\mathbf{O}} \mid \mathbf{V}, Z, \mathbf{O}) = I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O} \mid \mathbf{V}, Z),$$
$$I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O}, \widetilde{\mathbf{O}} \mid \mathbf{V}, Z) = I(\mathbf{Y}_{\mathsf{U}}; \widetilde{\mathbf{O}} \mid \mathbf{V}, Z) + I(\mathbf{Y}_{\mathsf{U}}; \mathbf{O} \mid \widetilde{\mathbf{O}}, \mathbf{V}, Z).$$

Equating them gives Equation 37. Nonnegativity follows from conditional mutual information. ■

Proposition E.11 (Stochastic-Garbling Extension)**.** *Let $\widetilde{\mathbf{O}}$ be generated by a stochastic kernel $Q(\cdot \mid \mathbf{V}, Z, \mathbf{O})$. Let $R^{\star}_{\text{rand}}(\mathcal{H})$ denote the infimum risk over randomized policies based on information $\mathcal{H}$. If the anchor-loss vector $(\Lambda(a))_{a \in \mathcal{A}}$ is conditionally independent of $\widetilde{\mathbf{O}}$ given $(\mathbf{V}, Z, \mathbf{O})$, then*

$$R^{\star}_{\text{rand}}(\sigma(\mathbf{V}, \mathbf{O}, Z)) \le R^{\star}_{\text{rand}}\Big(\sigma(\mathbf{V}, \widetilde{\mathbf{O}}, Z)\Big)\,, \tag{38}$$

Proof. Let $\kappa(a \mid \mathbf{V}, Z, \widetilde{\mathbf{O}})$ be any randomized policy based on the compressed observation. A policy observing $(\mathbf{V}, Z, \mathbf{O})$ can first sample $\widetilde{\mathbf{O}}$ from $Q(\cdot \mid \mathbf{V}, Z, \mathbf{O})$ and then sample an anchor from κ. Conditional independence ensures that this simulation produces the same joint distribution of selected anchors and losses. Full information can therefore reproduce every compressed-information policy, so its optimal risk cannot be larger. ■

E.3.2 Decision Value of Mechanism Information

Let D collect packet fields that directly carry prediction or decision signals, such as current and candidate votes. Let M collect fields describing the conflict mechanism or reliability state. Let P be the deterministic structured serialization of (D, M).

Proposition E.12 (Outcome-Independent Mechanism Information)**.** *Suppose* $\mathsf{M} \perp \mathbf{Y}_{\mathsf{U}}$. *Then*

$$R^\star(\sigma(\mathsf{D}, \mathsf{M})) \leq R^\star(\sigma(\mathsf{D}))\,. \tag{39}$$

The inequality is strict whenever the unique Bayes-optimal anchor given (D, M) *cannot be represented as a function of* D *alone on a set of positive probability.*

Proof. Since $\sigma(\mathsf{D}) \subseteq \sigma(\mathsf{D}, \mathsf{M})$, every decision rule based on D is also feasible when (D, M) is observed. Hence enlarging the information set cannot increase optimal risk. Under uniqueness, strict inequality follows whenever the (D, M)-optimal anchor is not $\sigma(\mathsf{D})$-measurable on a set of positive probability. ■

As a scalar illustration, let $Y_{\mathrm{ex}}, M_{\mathrm{ex}} \sim \mathrm{Bernoulli}(1/2)$ independently and define $D_{\mathrm{ex}} = Y_{\mathrm{ex}} \oplus M_{\mathrm{ex}}$, where $\oplus$ denotes exclusive OR (XOR). Consider an instance with two anchors that output 0 and 1, respectively. Then $M_{\mathrm{ex}} \perp Y_{\mathrm{ex}}$ and $D_{\mathrm{ex}} \perp Y_{\mathrm{ex}}$, so neither variable alone predicts the outcome. Yet $(D_{\mathrm{ex}}, M_{\mathrm{ex}})$ determines Y_{ex} exactly. Under 0–1 loss,

$$R^\star(\sigma(D_{\mathrm{ex}})) = \tfrac{1}{2}, \qquad R^\star(\sigma(D_{\mathrm{ex}}, M_{\mathrm{ex}})) = 0.$$

Corollary (Mechanism-Preserving Serialization)**.** *If the structured conflict packet satisfies* $\sigma(\mathsf{P}) = \sigma(\mathsf{D}, \mathsf{M})$*, then*

$$R^\star(\sigma(\mathsf{P})) = R^\star(\sigma(\mathsf{D}, \mathsf{M}))\,.$$

Hence a mechanism-preserving serialization introduces no additional decision-information loss relative to the mechanism variables it encodes.

This result separates outcome prediction from mechanism information. A mechanism variable need not directly predict the task outcome to be useful for authority assessment; it may instead identify how observed decision signals should be interpreted. In EPR, the conflict packet provides a structured interface for preserving such decision-relevant mechanism information before frozen representation and probe-based assessment. This distinction is consistent with classical decision-theoretic comparisons of information structures and with the role of ancillary information in conditional inference (Blackwell, 1953; Cox, 1958).

Corollary (Assessment Representation Sufficiency)**.** *Let* $\mathbf{h}_{\mathsf{P}}$ *be a fixed deterministic representation of packet* P. *Then*

$$R^\star(\sigma(\mathbf{V}, \mathsf{P})) \leq R^\star(\sigma(\mathbf{V}, \mathbf{h}_{\mathsf{P}}))\,.$$

Equality holds if and only if a $\sigma(\mathbf{V}, \mathbf{h}_{\mathsf{P}})$*-measurable assessment policy attains* $R^\star(\sigma(\mathbf{V}, \mathsf{P}))$.

Proof. Since $\mathbf{h}_{\mathsf{P}}$ is a deterministic function of P, $\sigma(\mathbf{V}, \mathbf{h}_{\mathsf{P}}) \subseteq \sigma(\mathbf{V}, \mathsf{P})$. The result follows from Theorem E.4. ■

This characterizes the frozen LLM as a representation interface rather than a source of additional decision evidence.

E.3.3 Exact Ex-Post Proposal–Assessment Decomposition

Unlike Theorem E.7, which compares attainable Bayes risks under information and action restrictions, the following decomposition uses the realized-loss oracle and therefore serves as an ex-post characterization of proposal and assessment quality.

Let $A^\star \in \arg\min_{a \in \mathcal{A}} \Lambda(a)$ be the realized-loss oracle anchor. For the nonempty candidate set $\mathcal{C}$, let $A^\star_{\mathcal{C}} \in \arg\min_{a \in \mathcal{C}} \Lambda(a)$ be the candidate-set oracle, and let $\widehat{A} \in \mathcal{C}$ be the assessed anchor. Fixed tie-breaking makes both oracle choices unique.

Corollary E.13 (Exact Ex-Post Proposal–Assessment Decomposition)**.** *Let* $R_{\mathrm{final}} = \mathbb{E}[\Lambda(\widehat{A})]$ *and* $R_{\mathrm{oracle}} = \mathbb{E}[\Lambda(A^\star)]$. *Then*

$$R_{\mathrm{final}} - R_{\mathrm{oracle}} = \underbrace{\mathbb{E}[\Lambda(A^\star_{\mathcal{C}}) - \Lambda(A^\star)]}_{\textit{proposal term}} + \underbrace{\mathbb{E}[\Lambda(\widehat{A}) - \Lambda(A^\star_{\mathcal{C}})]}_{\textit{assessment term}}, \tag{40}$$

and both terms are nonnegative.

Proof. Pointwise, add and subtract $\Lambda(A^\star_{\mathcal{C}})$:

$$\Lambda(\widehat{A}) - \Lambda(A^\star) = [\Lambda(A^\star_{\mathcal{C}}) - \Lambda(A^\star)] + \left[\Lambda(\widehat{A}) - \Lambda(A^\star_{\mathcal{C}})\right].$$

The first term is nonnegative because $A^\star$ minimizes loss over $\mathcal{A}$; the second is nonnegative because $A^\star_{\mathcal{C}}$ minimizes loss over $\mathcal{C}$ and $\widehat{A} \in \mathcal{C}$. Taking expectations proves the result. ■

This exact decomposition distinguishes the severity of proposal errors: omitting the oracle anchor incurs only the loss gap between the candidate-set oracle and the full oracle, rather than a binary miss penalty.

Corollary E.14 (Probability Bound)**.** *Define* $\Delta(a) = \Lambda(a) - \Lambda(A^\star)$ *and the proposal-miss event* $\mathcal{B}_{\text{miss}} = \{A^\star \notin \mathcal{C}\}$. *Assume* $0 \le \Delta(a) \le \Delta_{\max}$ *for every* $a \in \mathcal{A}$*, where* $\Delta_{\max}$ *is a uniform regret bound. Let* $\epsilon_{\text{prop}} = \mathbb{P}(\mathcal{B}_{\text{miss}})$ *and*

$$\epsilon_{\text{assess}} = \mathbb{P}\Big(\widehat{A} \neq A^\star \mid \mathcal{B}^c_{\text{miss}}\Big),$$

with $\epsilon_{\text{assess}} = 0$ *when* $\mathbb{P}(\mathcal{B}^c_{\text{miss}}) = 0$. *Then*

$$R_{\text{final}} - R_{\text{oracle}} \le \Delta_{\max}[\epsilon_{\text{prop}} + (1 - \epsilon_{\text{prop}})\epsilon_{\text{assess}}]. \tag{41}$$

Proof. The two disjoint error events give

$$\begin{aligned} R_{\text{final}} - R_{\text{oracle}} = \mathbb{E}[\Delta(\widehat{A})\mathbf{1}\{\mathcal{B}_{\text{miss}}\}] \\ + \mathbb{E}[\Delta(\widehat{A})\mathbf{1}\{\mathcal{B}^c_{\text{miss}}, \widehat{A} \neq A^\star\}]. \end{aligned}$$

The first term is at most $\Delta_{\max}\epsilon_{\text{prop}}$. The second is at most $\Delta_{\max}(1 - \epsilon_{\text{prop}})\epsilon_{\text{assess}}$. Summing the bounds proves Equation 41. ■

The exact decomposition formalizes proposal and assessment as distinct loss sources. In EPR, SDA targets proposal quality, while probe-consensus assessment targets stable selection within the emitted candidate set.