DECISION-PRESERVING DIAGNOSTIC COMPRESSION: WHEN KNOWING THE MISMATCH IS NOT ENOUGH

Anonymous authors
Paper under double-blind review

ABSTRACT

A correct diagnosis need not be a sufficient decision representation: it may recover what changed while discarding distinctions required to decide how to intervene. We study this gap between diagnostic recovery and intervention sufficiency by treating diagnosis as a compression $M = h(Z)$ of legal pre-decision information Z. For a fixed action set, we characterize when this compression preserves action-selectable value and decompose the headroom over the best fixed intervention into representation loss, rule defect, and delivered value. In the binary case, representation loss admits an exact mechanism: opposing action advantages cancel within diagnosis cells.

This population distinction also creates an identification problem. We introduce the Counterfactual Adaptation Audit (CAA), which ties each empirical claim to its estimand, information regime, identifying contrast, and status. In a controlled adaptation study, the diagnostic target is recovered near-affinely, intervention preference changes sharply with mismatch source, and a prespecified source-matched rule achieves privileged opportunity value $O(S) = D(r_{\mathrm{gt}}; S) = 0.933$ (95% CI $[0.840, 1.014]$) over the best fixed intervention. Yet the estimator-derived diagnostic becomes available only after one candidate action has been executed, leaving legal pre-decision $O(M)$ and $D(r; M)$ unidentified.

The resulting boundary is structural: information can be recoverable and decision-relevant without being preserved in a form, and available at a time, that can change the decision.

1 INTRODUCTION

After deployment, an adaptive system must turn evidence of mismatch into a choice of intervention. A diagnosis summarizes what changed; a decision rule must determine which available correction will help. These objectives coincide only when the diagnostic representation retains the distinctions needed for that choice. A diagnosis can therefore be semantically correct yet decision-insufficient.

Many adaptive systems address failures by improving the adaptation mechanism itself, for example through richer feedback, stronger correction models, or additional optimization. These approaches can be effective when the representation available to the system preserves the distinctions needed for action selection. This leaves a prior question: whether information sufficient to describe what changed is also sufficient to determine what should change.

This problem requires separating representation quality from decision quality. Decision-focused learning optimizes predictions for the decisions they support Donti et al. (2017); Wilder et al. (2019); Elmachtoub & Grigas (2022); Mandi et al. (2024), while state-abstraction and control-oriented representation work studies which behaviorally relevant distinctions must be retained Givan et al. (2003); Ferns et al. (2004); Abel et al. (2016); Zhang et al. (2021). We ask a complementary question: given an already-defined diagnostic abstraction and a fixed intervention set, which decision-relevant distinctions does that abstraction preserve, and which may already have been lost? The question is not whether a diagnosis is semantically meaningful, but whether its induced partition of pre-decision information is sufficient for action selection.

Let Z denote the information available before intervention, $M = h(Z)$ its diagnostic compression, and $Y(a)$ the utility of intervention a. The natural benchmark is the best action selectable from Z.

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.

Relative to that benchmark, a diagnosis-based rule can lose value in two fundamentally different ways: compression can erase action-relevant distinctions, or the rule can fail to exploit distinctions that the representation still retains. Timing adds a second boundary: useful information may exist but arrive only after the action it would need to influence. Separating these cases prevents a weak rule from being mistaken for a weak representation and explains why diagnostic accuracy alone cannot certify decision value.

Contributions.

1. **Decision-value accounting for intervention-facing representations.** Using classical information-value and abstraction ingredients, we formalize when diagnostic compression preserves the distinctions required for intervention, decompose headroom into representation loss, rule defect, and delivered value, and characterize binary loss through within-diagnosis cancellation. We also formalize *decision-time admissibility*: valuable information may arrive too late to support the decision.
2. **Evidence requirements for justified repair.** CAA links each claim to its estimand, admissible information, and identifying contrast. Its joint audit distinguishes evidence for changing the compression, rule, or acquisition timing from the need to collect a missing contrast. Each proposed repair requires its own supporting evidence.
3. **A controlled instantiation of the representation–decision boundary.** A prespecified privileged source-matched rule yields $O(S) = D(r_{\mathrm{gt}}; S) = 0.933$ over the best fixed action (95% CI $[0.840, 1.014]$), with positive advantage for every nondegenerate source mixture and stability across severity and seed resampling. This establishes that the source distinction carries intervention value. The available estimator-derived diagnostic, however, is observed only after $A1$ execution, so the experiment does not identify legal pre-decision $O(M)$ or $D(r; M)$. The controlled record therefore instantiates the distinction between privileged decision opportunity and deployable diagnostic-rule value.

2 Decision-Preserving Diagnostic Compression

We formalize one intervention decision with a fixed diagnostic compression. Classical information comparison Blackwell (1953) supplies the value ordering; Section 6 maps the preservation and regret criteria to their closest antecedents.

2.1 Decision preservation

Fix a finite intervention set $\mathcal{A}$, integrable potential utilities $Y(a)$, and a measurable diagnostic compression $M = h(Z)$. Expectations use one declared population and horizon. For each intervention $a \in \mathcal{A}$, define

$$q_a(Z) = \mathbb{E}[Y(a) \mid Z]. \tag{1}$$

The best values available from the full legal information and from its diagnostic compression are

$$V_Z = \mathbb{E}\left[\max_{a\in\mathcal{A}} q_a(Z)\right], \qquad V_M = \mathbb{E}\left[\max_{a\in\mathcal{A}} \mathbb{E}[q_a(Z) \mid M]\right]. \tag{2}$$

Conditional Jensen gives

$$\boxed{V_M \le V_Z.} \tag{3}$$

The nonnegative gap $V_Z - V_M$ measures the value of distinctions erased by the diagnostic partition. We call M *decision-preserving* when no intervention value is lost through this compression. Equivalently, an M-measurable rule can select a Z-optimal action almost surely, with ties allowed. Decision preservation is therefore relative to the declared action set, utility, population, and horizon: the representation need only preserve what is required for an optimal intervention, not reconstruct the full state.

2.2 The headroom ledger

Let

$$V_0^\star = \max_{a\in\mathcal{A}} \mathbb{E}[Y(a)] \tag{4}$$

be the value of the best information-independent intervention, and let $r : M \to \mathcal{A}$ be any M-measurable rule. Define

$$O(M) = V_M - V_0^\star, \qquad D(r; M) = \mathbb{E}[Y(r(M))] - V_0^\star, \qquad L_{\text{abs}} = V_Z - V_M. \tag{5}$$

Proposition 1 (Headroom ledger). *Every M-measurable rule satisfies*

$$\boxed{V_Z - V_0^\star = L_{\text{abs}} + \delta(r) + D(r; M),} \tag{6}$$

where $\delta(r) = V_M - \mathbb{E}[Y(r(M))]$. Both $L_{\text{abs}} \geq 0$ and $\delta(r) \geq 0$, while $D(r; M)$ has no sign restriction.

The ledger separates three sources of the observed gap. L_{abs} is *representation loss*: value removed by the compression itself and unrecoverable by any rule using the same M. $\delta(r)$ is *rule defect*: value still available from M but not extracted by the specified mapping. $D(r; M)$ is *delivered value* over the best information-independent intervention and may be positive or negative. The ledger also gives

$$D(r; M) = O(M) - \delta(r) \leq O(M). \tag{7}$$

A rule can therefore fail while the representation remains useful. Observing $D(r; M) \leq 0$ does not imply $O(M) = 0$; a different mapping on the same diagnosis may still improve on the best fixed intervention.

2.3 Compression loss and cancellation

Proposition 2 (Regret form). *Let $a^\star(m)$ be a measurable choice maximizing conditional utility within diagnosis m. Then*

$$L_{\text{abs}} = \mathbb{E}\left[\max_{a \in \mathcal{A}} q_a(Z) - q_{a^\star(M)}(Z)\right]. \tag{8}$$

Thus L_{abs} is the expected regret incurred when action selection is restricted to the diagnostic partition rather than the full legal information. This interpretation does not require a binary action set.

The binary case exposes the mechanism most directly: a diagnosis can be semantically correct while merging states whose optimal interventions disagree.

Proposition 3 (Binary cancellation and zero-loss criterion). *For $\mathcal{A} = \{0, 1\}$, let $\tau(Z) = q_1(Z) - q_0(Z)$. Then*

$$\boxed{L_{\text{abs}} = \frac{1}{2}\left(\mathbb{E}|\tau(Z)| - \mathbb{E}|\mathbb{E}[\tau(Z) \mid M]|\right).} \tag{9}$$

Writing $\tau^+ = \max(\tau, 0)$ and $\tau^- = \max(-\tau, 0)$, the conditional loss is

$$L_{\text{abs}}(M) = \min\{\mathbb{E}[\tau^+ \mid M], \mathbb{E}[\tau^- \mid M]\}. \tag{10}$$

Consequently,

$$L_{\text{abs}} = 0 \iff \Pr(\tau > 0 \mid M) \Pr(\tau < 0 \mid M) = 0 \quad \textit{a.s.} \tag{11}$$

If $\Pr(\tau = 0) = 0$, this is equivalent to $\operatorname{sign} \tau(Z)$ being M-measurable almost surely.

The mechanism is cancellation: fine states that share a diagnosis can prefer different interventions, so their signed advantages shrink or vanish after conditioning on M, even when the underlying effects are large. A representation that perfectly recovers the declared diagnostic target can therefore remain decision-lossy. For illustration, suppose two equally likely fine states share the same correct diagnosis but have $\tau = +1$ and $\tau = -1$. Then $\mathbb{E}[\tau \mid M] = 0$ although each state strongly prefers a different action, giving $L_{\text{abs}} = 1/2$. The example is analytical rather than empirical; it shows how a semantically correct cell can erase the distinction required for action.

Corollary 1 (Abstraction error is not estimation error). *L_{abs} is a functional of the joint law and the fixed partition h. Increasing sample size can improve its estimation but cannot change its population value. A lossy abstraction is not a data shortage.*

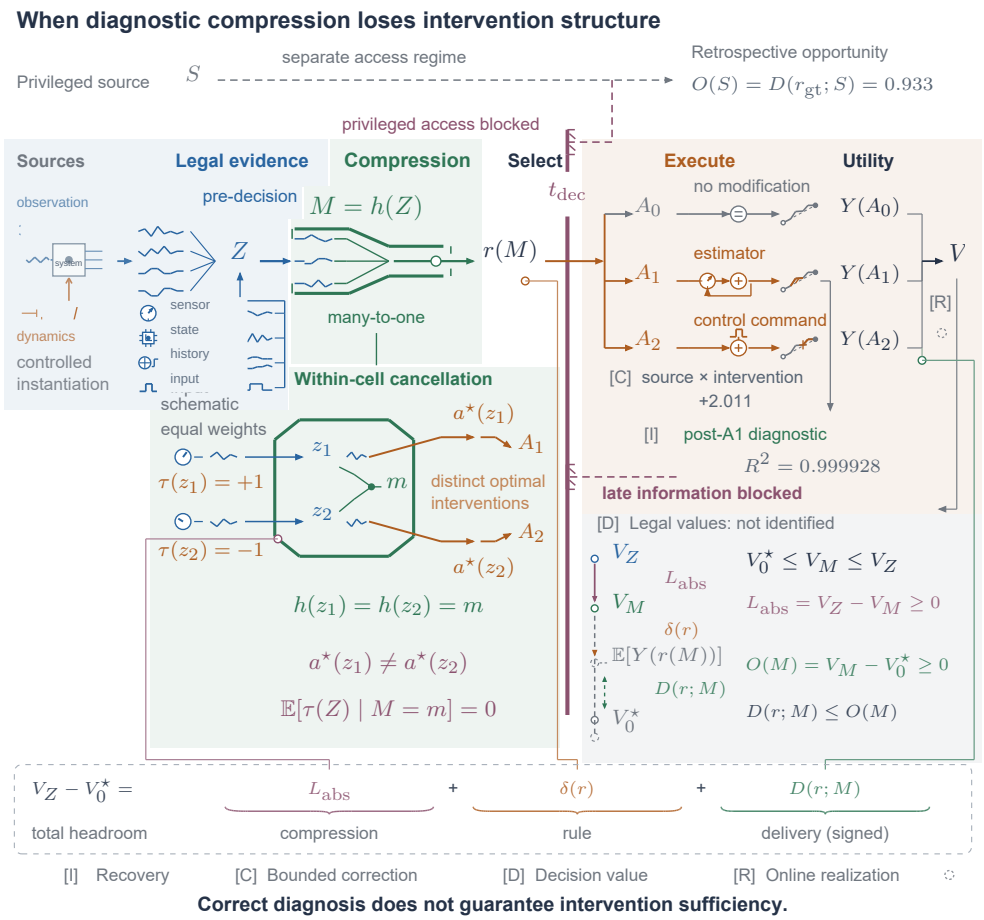

Figure 1: **Representation, decision, and information flow.** The central object is the information flow $Z \to M = h(Z) \to r(M) \to A \to Y(a)$; the controlled system provides one concrete instantiation of this decision problem. A diagnostic compression can remain semantically correct while merging fine states that prefer different interventions, producing within-cell cancellation and compression loss. The schematic illustrates the mechanism of decision insufficiency rather than an empirical estimate of L_{abs}. The displayed $+2.011$ is a PRR interaction; 0.933 is raw-utility opportunity.

2.4 Rule value and decision-time admissibility

Even decision-relevant information cannot support an earlier intervention if it becomes available only after the decision has already been made. For any information object I, define

$$V_I = \mathbb{E}\Big[\max_a \mathbb{E}[Y(a) \mid I]\Big], \qquad O(I) = V_I - V_0^\star, \tag{12}$$

and for a specified rule $r : I \to \mathcal{A}$ define

$$D(r; I) = \mathbb{E}[Y(r(I))] - V_0^\star. \tag{13}$$

Because constant actions and the specified rule are both I-measurable,

$$\boxed{O(I) \geq 0, \qquad D(r; I) \leq O(I).} \tag{14}$$

$O(I)$ is the best decision value that the information could support, whereas $D(r; I)$ is the value extracted by one specified rule.

Definition 1 (Decision-time admissibility)**.** *Let $\mathcal{F}_t$ be the sigma-field generated by information legally available immediately before the action at decision time t. An information object I_t is* decision-time admissible *for that action if it is $\mathcal{F}_t$-measurable.*

A prospective claim about $D(r; I_t)$ requires this admissibility. Information revealed by a candidate action or by its future outcome may identify retrospective opportunity but cannot identify a rule for choosing that same action. Thus *decision-relevant* does not imply *decision-time admissible*.

Corollary 2 (Two-source opportunity value)**.** *Let privileged source $S \in \{o, d\}$ have prevalence $p = \Pr(S = d)$ and $\mu_{s,a} = \mathbb{E}[Y(a) \mid S = s]$. Suppose a_o and a_d are the conditional maximizers for sources o and d, respectively, with*

$$\Delta_o = \mu_{o,a_o} - \mu_{o,a_d} > 0, \qquad \Delta_d = \mu_{d,a_d} - \mu_{d,a_o} > 0,$$

and no other fixed action exceeds the upper envelope of a_o and a_d. For the source-matched rule $r_{\mathrm{gt}}(o) = a_o$, $r_{\mathrm{gt}}(d) = a_d$,

$$\boxed{D(r_{\mathrm{gt}}; S) = O_p(S) = \min\{p\Delta_d, (1-p)\Delta_o\}.} \tag{15}$$

Hence $O_p(S) > 0$ for every nondegenerate mixture $0 < p < 1$, with best-fixed switch point $p^\star = \Delta_o/(\Delta_o + \Delta_d)$.

This characterizes privileged source opportunity, not decision-time admissibility.

3 CAA: The Identification Interface

The population theory distinguishes representation loss, rule defect, and delivered value; an experiment must still determine which of these quantities its design actually identifies. The Counterfactual Adaptation Audit (CAA) provides this empirical interface. For each adaptation claim, it records the target quantity, the identifying contrast, and the information available when the action is chosen.

CAA separates four questions: I asks whether the declared diagnostic target is recoverable under its stated acquisition protocol; C whether the declared intervention class contains a useful intervention; D whether *decision-time admissible* information improves selection beyond the best fixed intervention; and R whether an online mechanism realizes efficacy established by a matched fixed intervention. Each question therefore has its own target and identifying contrast (Table 1 in the appendix). A study may identify some coordinates, partially identify others, and abstain where the required contrast or admissible information is absent. CAA is not a sequential completion checklist. Evidence for one coordinate cannot substitute for the identifying contrast required by another.

The pre-decision firewall fixes what may support a decision claim: selection features must be available before the action, and privileged source labels, future outcomes, and post-action diagnostics are excluded from that earlier choice. A diagnostic can therefore satisfy I under its acquisition protocol while remaining inadmissible for D at an earlier decision point. If the required contrast or admissible information is absent, CAA records the claim as **not identified** rather than substituting a nearby result.

From an explanation to evidence and repair. With the decision problem fixed, an identified $V_Z - V_M > 0$ establishes representation insufficiency; recovering this loss targets h, holding Z fixed. An identified positive advantage of an admissible rule on the same M over r certifies rule defect and supports changing r. Late information requires an acquisition-time record and an admissibility check; repair concerns feasible earlier acquisition or a separately declared later decision. A missing identifying contrast instead calls for action/comparator outcomes with the support and assumptions needed for identification; it warrants neither mechanism diagnosis. These cases can coexist: improving r cannot recover distinctions erased by h, timing evidence alone cannot replace outcome contrasts, and post-action diagnostics cannot legalize an earlier choice. Appendix E shows identical observed logs requiring different repairs. These conditional implications do not identify additional mechanisms in our experiment.

3.1 Reporting normalized effects

CAA's identifying contrasts use raw additive utility. For reporting, let $c = (\mathrm{source}, \mathrm{severity})$ and $b \in \{0, \ldots, 4\}$ index seed. Within each plant, $\bar{J}_{\mathrm{nom}}$ averages severity-zero $A0$ endpoints; the two source-labelled nominal records coincide for each seed. The frozen implementation uses $\epsilon = 10^{-6}$ and

$$D_b(c) = \bar{J}_{\mathrm{nom}} - J_{A0,b}(c), \qquad \Delta J_{a,a',b}(c) = J_{a,b}(c) - J_{a',b}(c),$$

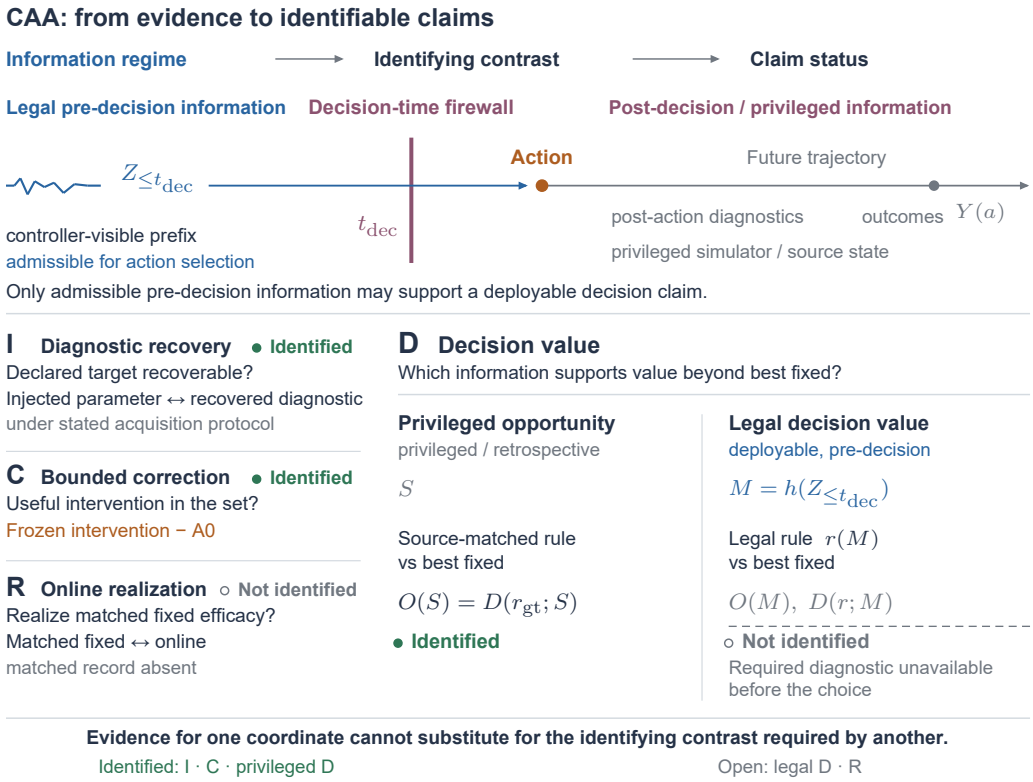

Figure 2: **CAA separates adaptation claims by identifying evidence.** The pre-decision firewall restricts action selection to information available before t_{dec}. The $I/C/D/R$ coordinates separately audit diagnostic recovery, bounded correction, decision value, and online realization. The current record identifies privileged source opportunity but not the value of a legal pre-decision rule; missing contrasts or inadmissible information therefore remain explicitly not identified.

$$\mathrm{PRR}_{a,b}(c) = \begin{cases} 0, & |D_b(c)| < \epsilon, \\ \Delta J_{a,A0,b}(c)/(D_b(c) + \epsilon), & \text{otherwise.} \end{cases}$$

Cell contrasts average within-block PRR differences; they need not equal the ratio of the displayed mean additive contrast to mean denominator. Unresolved denominators trigger a conservative reporting policy. Fieller inference Fieller (1954) concerns a ratio of population contrasts, not this mean of blockwise ratios (Appendix A.1). Neither normalization nor its reporting gate identifies a missing decision contrast.

4 Controlled Instance of the Decision Problem

The controlled dynamical system is an identification instrument with explicit information, intervention, and outcome boundaries. It isolates which information supports which intervention claim.

- Z is controller-visible information available before the initial locus choice; privileged state, future outcomes, and post-action diagnostics are excluded from a legal rule at that point.
- $S \in \{\mathrm{observation}, \mathrm{dynamics}\}$ is privileged injected source ground truth. The estimator-derived $\hat{\omega}_{\mathrm{bias}}$ appears under $A1$ execution and is therefore not a legal $M = h(Z)$ for the preceding $A1$-versus-$A2$ choice.
- $\mathcal{A} = \{A0, A1, A2\}$. $A0$ applies no adaptation, $A1$ a scalar estimator-side heading-bias correction, and $A2$ a bounded low-dimensional residual to commanded angular rate.

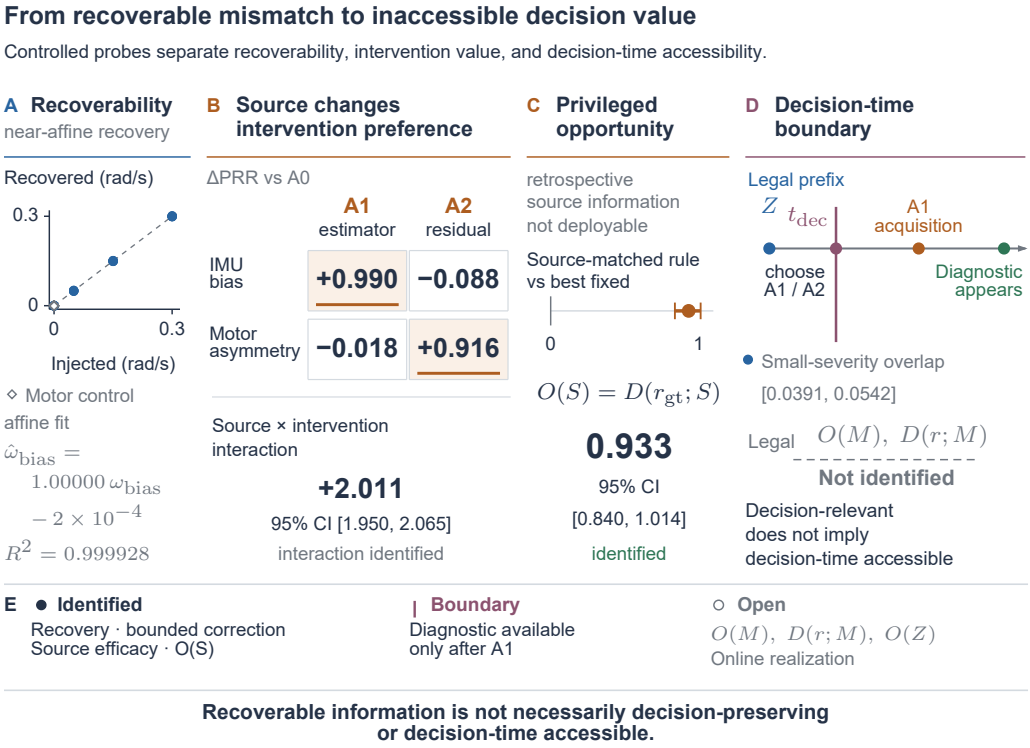

Figure 3: **Controlled evidence for the representation–decision boundary.** The record separates four questions: whether the declared diagnostic target is recoverable, whether mismatch source changes intervention preference, whether privileged source information carries decision value, and whether that value is available at the decision point. Recovery, source-dependent efficacy, and privileged opportunity $O(S)$ are identified; legal pre-decision $O(M)$, $D(r; M)$, and online realization remain unresolved. The recorded scalar overlap is supporting decoder-level evidence, not a population estimate of compression loss. Panel B reports PRR contrasts; panel C reports raw additive utility.

- $Y(a)$ is the negative closed-loop tracking objective, so higher utility means lower tracking error. $V_0^\star$ is the best globally fixed intervention under the same pooling and pairing.

The prespecified privileged map r_{gt} assigns $A1$ to observation mismatch and $A2$ to dynamics mismatch. We first test the source-by-locus interaction and then directly compare r_{gt} with the best fixed action. Because S is privileged, this comparison identifies source *opportunity* value rather than the value of a deployable rule based on legal pre-decision information.

The primary decision record uses a differential-drive kinematic plant derived from the four-wheel system geometry and limits, with two mismatch families, three non-zero severities per family, five paired seed blocks per cell, and 2,000-step horizons. Main contrasts pair interventions within source, severity, and seed. An earlier plant generation is used only for the reporting-scale audit; effect magnitudes are not pooled across generations. Aggregation and artifact provenance are reported in Appendix B.

5 Results

The controlled record separates target recovery, source-dependent intervention preference, privileged opportunity, and decision-time accessibility. These require distinct evidence rather than a single adaptation-performance score.

5.1 DIAGNOSTIC RECOVERY IS NEAR-AFFINE

Under $A1$, the recovered diagnostic is nearly affine in the injected IMU bias:

$$\hat{\omega}_{\text{bias}} = 1.00000\,\omega_{\text{bias}} - 2 \times 10^{-4}\ \text{rad/s}, \qquad R^2 = 0.999928. \tag{16}$$

The motor-asymmetry family provides the negative control and returns values near zero. Thus the declared diagnostic quantity is recoverable under the $A1$ acquisition protocol. Crucially, however, it is exposed only during $A1$ execution and is therefore not decision-time admissible for the preceding $A1$-versus-$A2$ choice. The affine offset changes across plant generations, so diagnostic calibration is not treated as an estimator-intrinsic property. Cross-generation calibration is summarized in Appendix D, Figure 4d.

5.2 INTERVENTION PREFERENCE DEPENDS SHARPLY ON MISMATCH SOURCE

We next ask whether mismatch source changes which intervention is useful. The complete 2×2 source-by-locus response pattern has a matched diagonal that is large and positive, whereas the off-diagonal responses are small or negative, so relative intervention efficacy changes with mismatch source. Directional support is computed from five recurring seeds across three severities; the 15 paired entries are not fifteen independent environments. Exact cell values are listed in Table 3 in the appendix.

On the PRR reporting scale, the source-by-locus interaction is:

$$\overline{\Delta\text{PRR}}_{A1,A2|\text{IMU}} - \overline{\Delta\text{PRR}}_{A1,A2|\text{motor}} = +2.011\ [+1.950, +2.065], \tag{17}$$

where the overbar averages 15 nonzero-severity–seed contrasts, pairing sources before averaging. All 15 are positive, with direction stable across severity. The interval resamples pooled paired tuples, not seed clusters (Appendix B); its scale differs from raw-utility $O(S)$ below.

5.3 PRIVILEGED SOURCE INFORMATION HAS ROBUST OPPORTUNITY VALUE

On the raw additive utility scale, the prespecified rule $r_{\text{gt}}(S)$ is the unique conditional maximizer over $\{A0, A1, A2\}$ for both sources, so $\delta(r_{\text{gt}}) = 0$ and $D(r_{\text{gt}}; S) = O(S)$. At balanced prevalence the best fixed action is $A1$ with $V_0^\star = -1.264$, whereas

$$\boxed{O(S) = D(r_{\text{gt}}; S) = 0.933 \quad [0.840, 1.014].} \tag{18}$$

Conditional optimality remains strict in all 3,125 exhaustive five-seed cluster-bootstrap resamples and all five leave-one-seed-out fits.

Let $p = \Pr(S = \text{dynamics})$. The frozen-population opportunity is

$$O_p(S) = \min\{1.8666p,\ 16.3711(1-p)\}, \qquad p^\star = 0.89765, \tag{19}$$

which is strictly positive for every $0 < p < 1$; the same interior positivity holds for every bootstrap resample, leave-one-seed-out fit, and severity stratum. Appendix C gives the full audit. Because S is privileged and the estimator diagnostic appears only after $A1$, this identifies source opportunity, not a legal pre-decision $O(M)$ or $D(r; M)$.

5.4 THE RESIDUAL-ARM SCALAR PRESERVES MAGNITUDE BUT NOT AN ORDERED SOURCE DISTINCTION

The residual-arm scalar analyzed here is distinct from the estimator diagnostic in Section 5.1 and is used only as a recorded representation exposed by the residual arm. It is the recorded parameter displacement $\|\theta - \theta_{\text{init}}\|_2$. It increases with severity in both families but does not provide an ordered source distinction in this sample. At the smallest severity, IMU values span 0.0320–0.0542 and motor values 0.0391–0.0608, overlapping on $[0.0391, 0.0542]$; no single threshold separates all ten runs. Yet the mean PRR contrast $A2 - A0$ is $+0.939$ for motor asymmetry and -0.055 for IMU bias. This is decoder-level inseparability in the recorded sample, not a population estimate of compression loss.

5.5 CAA SUMMARY AND NORMALIZED-REPORTING CHECK

Recovery under $A1$ (I), bounded correction (C), and privileged source opportunity are identified. Legal pre-decision $O(M)$, $D(r; M)$, prospective $O(Z)$, and matched online realization (R) remain not identified: source is privileged, the estimator diagnostic is late, and the matched online contrast is absent.

Reporting check. Eleven of twelve normalization denominators are resolved ($T_D \geq 15.7$) and agree in sign with the additive contrast. The sole disagreement is the sole unresolved cell ($T_D = 0.14$); its additive contrast is retained and its normalized interpretation is withheld (Appendix D).

6 RELATED WORK

Classical information comparison evaluates information through attainable decision value Blackwell (1953). The feature gap of van Rooyen & Williamson (2015) is the optimal-risk increase induced by a feature map; converting utility to loss gives our L_{abs} and its regret interpretation. Common optimal action within an abstract class has a direct antecedent in π^*-irrelevance Li et al. (2006); this is a single-decision correspondence, not an aggregated-MDP planning guarantee. For nested covariates, Shi et al. (2019) characterize zero incremental treatment value by the absence of conditional sign changes under their regularity conditions. Our binary criterion gives the corresponding almost-sure condition for measurable diagnostic compression. These ingredients are established, not independent novelty claims (Appendix E). Diagnostic-test acquisition, regime value, and causal identification are also connected in Liu et al. (2021).

Decision-focused learning optimizes representations for downstream decisions Donti et al. (2017); Wilder et al. (2019); Elmachtoub & Grigas (2022); Mandi et al. (2024); fault-tolerant control separates diagnosis and accommodation Blanke et al. (2016); Zhang & Jiang (2008); Frei et al. (1999). Relative to these antecedents, our contribution maps each claim to the evidence it requires and the repair it licenses. Value gaps, comparisons of rules on the same representation, acquisition timing, and identifying contrasts must be checked together; none alone diagnoses every failure. CAA makes the next measurement and repair conditional on this audit. The controlled record illustrates the discipline without identifying legal compression loss.

7 SCOPE AND LIMITATIONS

Theoretical scope. Decision preservation is relative to a finite intervention set, utility, population, horizon, and measurable compression; the binary cancellation criterion additionally fixes a two-action contrast. No robotic dynamics assumption enters these value results. **Identification scope.** The design identifies recovery under $A1$, useful bounded correction, source-dependent intervention preference, and privileged $O(S)$, but not legal pre-decision V_Z, V_M, L_{abs}, $O(M)$, $D(r; M)$, matched online realization, or prospective $O(Z)$. The residual-arm scalar result is limited to single-threshold source separation in the recorded sample. **Empirical scope.** Numerical effects are specific to the studied plants, mismatch families, severities, interventions, and five recurring seeds. The PRR intervals retain pooled-tuple resampling with fixed nominal reference; opportunity intervals resample seed clusters. The controlled system is an identification instrument, not a broad-performance benchmark.

8 CONCLUSION

A diagnostic abstraction should be judged by the decisions it preserves, not only by the mismatch it describes. The headroom ledger separates information lost by compression from information left unused by a rule, while decision-time admissibility requires useful information to exist before the decision it is meant to change. In the controlled record, privileged source has robust opportunity value, $O(S) = D(r_{\mathrm{gt}}; S) = 0.933$, yet the available diagnostic appears only after $A1$, so a legal pre-decision rule realizing that opportunity is not identified. The practical consequence is to match each proposed repair to its required evidence before changing the representation, rule, or acquisition protocol.

AI USE STATEMENT

Generative AI tools were used to assist with literature retrieval and synthesis, research ideation and methodological critique, mathematical derivation and proof checking, experimental-design feedback, code and artifact auditing, figure design, and drafting and polishing portions of the manuscript. All AI-assisted scientific claims, derivations, experimental interpretations, and reported numerical results were independently checked against the underlying theory, frozen experimental artifacts, and cited sources. No generative AI system is listed as an author, and the author takes responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

The accompanying anonymous numerical artifact contains frozen primary raw records, verified endpoint-field projections for both plants, analysis code, expected outputs, and SHA-256 manifests. Running `python -B regenerate_all.py` recomputes opportunity and stability checks, PRR tables and interaction, denominator screens, calibration, and residual-arm scalar overlap without simulation or tuning. It regenerates numerical results, not the full trajectories or manuscript figures. Proofs and exact estimands are given below; acquisition logs outside the numerical bundle are not claimed as packaged evidence.

REFERENCES

Abel, D., Hershkowitz, D. E., and Littman, M. L. Near optimal behavior via approximate state abstraction. *International Conference on Machine Learning*, pp. 2915–2923, 2016.

Blackwell, D. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics* 24(2):265–272, 1953.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. *Diagnosis and Fault-Tolerant Control*, 3rd ed. Springer, 2016.

Donti, P. L., Amos, B., and Kolter, J. Z. Task-based end-to-end model learning in stochastic optimization. *Advances in Neural Information Processing Systems* 30:5484–5494, 2017.

Elmachtoub, A. N., and Grigas, P. Smart "predict, then optimize". *Management Science* 68(1):9–26, 2022.

Fieller, E. C. Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B* 16(2):175–185, 1954.

Ferns, N., Panangaden, P., and Precup, D. Metrics for finite Markov decision processes. *Uncertainty in Artificial Intelligence*, pp. 162–169, 2004.

Frei, C. W., Kraus, F. J., and Blanke, M. Recoverability viewed as a system property. *ECC*, pp. 2197–2202, 1999.

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

Mandi, J., Kotary, J., Berden, S., Mulamba, M., Bucarey, V., Guns, T., and Fioretto, F. Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. *Journal of Artificial Intelligence Research* 80:1623–1701, 2024.

Wilder, B., Dilkina, B., and Tambe, M. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. *AAAI* 33(01):1658–1665, 2019.

Zhang, A., McAllister, R., Calandra, R., Gal, Y., and Levine, S. Learning invariant representations for reinforcement learning without reconstruction. *International Conference on Learning Representations*, 2021.

Zhang, Y., and Jiang, J. Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control* 32(2):229–252, 2008.

van Rooyen, B. and Williamson, R. C. A theory of feature learning. *arXiv:1504.00083*, 2015.

Li, L., Walsh, T. J., and Littman, M. L. Towards a unified theory of state abstraction for MDPs. *International Symposium on Artificial Intelligence and Mathematics*, 2006.

Shi, C., Song, R., and Lu, W. On testing conditional qualitative treatment effects. *The Annals of Statistics* 47(4):2348–2377, 2019.

Liu, L., Shahn, Z., Robins, J. M., and Rotnitzky, A. Efficient estimation of optimal regimes under a no direct effect assumption. *arXiv:1908.10448v2*, 2021.

GUIDE TO THE EVIDENCE AUDIT

Navigation. Representation loss and decision preservation: Sections 2.2–2.3 and Propositions 1–3. Decision-time admissibility: Section 2.4. CAA identification map: Section 3 and Appendix B. Exact PRR estimand and aggregation: Appendix D.1. PRR-scale interaction versus raw additive utility: Appendix D.3. Fieller and denominator-resolution gate: Appendices A.1 and D.4. Frozen-number reproduction: the anonymous artifact README and its Paper number map.

A PROOFS AND FULL NORMALIZED-REPORTING CONDITIONS

Purpose. This section records the supporting proofs and the conditions for normalized reporting; it does not introduce a new empirical claim.

Proof of Proposition 1. Add and subtract V_M and $\mathbb{E}[Y(r(M))]$ in $V_Z - V_0^\star$. Nonnegativity of L_{abs} follows from (3); nonnegativity of $\delta(r)$ follows from M-optimality. Equation (7) follows immediately.

Proof of Corollary 2. The source-matched rule is conditionally optimal by assumption, so $\delta(r_{\mathrm{gt}}) = 0$ and $D(r_{\mathrm{gt}}; S) = O_p(S)$. Relative to always choosing a_o, matching the dynamics source gains $p\Delta_d$; relative to always choosing a_d, matching the observation source gains $(1-p)\Delta_o$. Under the stated upper-envelope condition, the best fixed action is one of these two, so the opportunity value is their minimum. Equality gives $p^\star = \Delta_o/(\Delta_o + \Delta_d)$, and both terms are strictly positive for $0 < p < 1$.

Proof of Proposition 2. For an M-measurable maximizer $a^\star(M)$, $V_M = \mathbb{E}[q_{a^\star(M)}(Z)]$. Substitution in $L_{\mathrm{abs}} = V_Z - V_M$ gives (8).

Proof of Proposition 3. For two actions, $\max(q_0, q_1) = q_0 + \frac{1}{2}(\tau + |\tau|)$. Applying the same identity after conditioning on M and subtracting gives (9). For any integrable X, $\frac{1}{2}(\mathbb{E}|X| - |\mathbb{E}X|) = \min\{\mathbb{E}X^+, \mathbb{E}X^-\}$. Applying this identity conditionally yields the cellwise form. It vanishes exactly when strictly positive and strictly negative effects do not both occur in the same conditional law; ties may coexist with either sign.

A.1 NORMALIZED REPORTING CONDITIONS

Purpose. This section states the population ordering and Fieller conditions used by the reporting gate; it does not estimate a new effect.

The main text uses normalization only as a reporting check inside CAA. This appendix gives the population ordering condition, the Fieller resolution result, and the distinction between population rescaling and finite-sample denominator non-resolution. These propositions concern ratios of population contrasts; they are not confidence-set results for means of blockwise ratios. The empirical screen in Table 4 uses $T_D = \sqrt{5}|\widehat{D}|/\sqrt{s_{A0}^2 + s_{\mathrm{nom}}^2}$, with sample SDs from five cell endpoints and five unique nominal seeds. It omits their covariance and is not an exact paired standard error. The archived threshold $T_D > 2$ is a reporting policy, not a calibrated finite-sample Fieller cutoff. Passing it does not supply inference for mean blockwise PRR.

Proposition 4 (Universal ordering preservation)**.** *Fix two cells with positive denominators. Dividing by their respective denominators preserves the additive ordering for every pair of same-sign numerators if and only if the denominators are equal.*

Proposition 5 (Fieller resolution gate)**.** *Let $(\hat{\Delta}, \hat{D})$ estimate an additive contrast and its denominator under a joint-normal working model with positive-definite estimated covariance matrix. Let*

$v_\Delta, v_D, v_{\Delta D}$ denote the estimated variances and covariance, and k_α the declared critical value. The Fieller confidence set for $r = \Delta/D$ is

$$\mathcal{F}_{1-\alpha} = \left\{ r : \ (\hat{\Delta} - r\hat{D})^2 \leq k_\alpha^2 (v_\Delta - 2r v_{\Delta D} + r^2 v_D) \right\}. \tag{20}$$

Its quadratic coefficient is $\hat{D}^2 - k_\alpha^2 v_D$. Thus $T_D = |\hat{D}|/\sqrt{v_D} > k_\alpha$ yields a bounded, nonempty interval under this model. If $T_D \leq k_\alpha$, the Fieller set is unbounded. Equivalently, the denominator interval contains zero exactly when $T_D \leq k_\alpha$. We therefore withhold normalized sign interpretation whenever this gate fails as a conservative policy; this is not a theorem that every unbounded set contains both positive and negative ratios.

Proposition 6 (Two distinct failure modes)**.** *If unequal positive population denominators produce a strict ordering reversal, consistent estimation preserves that reversal asymptotically. If instead the denominator fails the finite-sample gate, the normalized interpretation is unresolved at the attained sample size; the additive contrast remains the appropriate reported target. Further independent blocks may eventually resolve a fixed nonzero denominator.*

Proof of Proposition 4. If the denominators are equal, normalization is a common positive rescaling. If they differ, choose same-sign numerators whose ratio lies strictly between one and the denominator ratio; the additive and normalized orderings reverse.

Proof of Proposition 5. Moving the two sides of the Fieller inequality to one quadratic in r gives leading coefficient $\hat{D}^2 - k_\alpha^2 v_D$. When this is positive, the accepted set is bounded and contains $r_0 = \hat{\Delta}/\hat{D}$ because the left side is zero there and the right side is a nonnegative quadratic form. When the coefficient is nonpositive, the accepted set is unbounded. The denominator interval includes zero exactly when $|\hat{D}|/\sqrt{v_D} \leq k_\alpha$.

Proof of Proposition 6. Consistency sends estimated comparisons to their distinct population targets, so a strict population reversal persists asymptotically. At fixed sample size, failure of the denominator gate leaves the Fieller set unbounded; if the true denominator is nonzero and a valid standard error shrinks with additional independent blocks, T_D diverges in probability and the denominator can eventually be resolved.

B PROTOCOLS AND PROVENANCE

Purpose. This section documents acquisition timing, identifying contrasts, and resampling units for the estimator diagnostic and residual-arm scalar.

Table 1: Operational CAA coordinates and the identifying contrasts used in the main text.

Coordinate	Scientific question	Identifying contrast
I – diagnostic recovery	Is the declared diagnostic target recoverable under its stated acquisition protocol?	Recovery of the injected parameter, with acquisition time and a negative control recorded explicitly
C – bounded correction	Does the declared bounded intervention class contain a useful correction?	Frozen intervention minus $A0$
D – decision value	Does decision-time admissible information support better intervention selection than the best fixed choice?	Legal rule minus fixed comparator; privileged information-value contrasts are reported separately
R – online realization	Does the online mechanism realize efficacy already demonstrated by a matched fixed intervention?	Matched fixed-versus-online and online-versus-$A0$ contrasts

Claim–evidence–status crosswalk. Each row separates the supported claim from a nearby inference that the evidence does not license.

Claim	Evidence	Status	Does not imply
Diagnostic recovery	Injected parameter → estimator diagnostic under $A1$	Identified under the stated acquisition protocol	Decision sufficiency
Privileged source opportunity	Source-matched rule versus best fixed action, using privileged S	Identified	Legal $O(M)$ or legal $D(r; M)$
Legal pre-decision rule value	Required legal rule comparison is absent; the estimator diagnostic is late	Not identified	Zero decision value
Online realization	Matched fixed-versus-online record is absent	Not identified	Online mechanism failure

Invalid inferences.

$D(r; M) \leq 0$	$\not\Rightarrow$	$O(M) = 0$
Accurate recovery of the declared diagnostic target	$\not\Rightarrow$	Decision sufficiency
$O(S) > 0$	$\not\Rightarrow$	Legal $O(M)$ is identified
Post-action information	$\not\Rightarrow$	A valid rule for the earlier action

Stage registry. Each stage retains its own population, seed band, intervention objects and preregistration. The main decision experiments use seeds $\{0, \ldots, 4\}$; live validation probes use three seeds; horizons are 2,000 steps for the decision experiments and 80 steps for the probes.

Statistical procedure. PRR contrasts pair arms within source, severity, and seed before averaging. Table 3 and Eq. (17) pool the 15 nonzero severity–seed tuples equally; the interaction pairs sources before taking their difference. Archived intervals use 10,000 percentile-bootstrap draws of these pooled tuples, with the global nominal reference held fixed. The script retains Python `random.Random(20260914)`, its original shared call order, and sorted-draw indices 250 and 9750 (zero based). The intervals are not seed-cluster intervals and do not re-estimate nominal-reference uncertainty.

The frozen source-opportunity interval used 100,000 random bootstrap draws of the five recurring seed clusters (seed 20260916), retaining both sources, three nonzero severities, and all paired actions in each sampled cluster. The best fixed action is reselected in each draw. Exhaustive enumeration of all $5^5 = 3{,}125$ ordered resamples gives the same percentile endpoints; this is a cross-check, not the historical generating procedure. Conditional optimality and source-mixture analyses use exhaustive cluster resampling. All five leave-one-seed-out fits reselect the best fixed action. Source, severity, and seed weights are respectively $1/2$, $1/3$, and $1/5$ in the balanced opportunity population. These checks use the frozen empirical population and do not establish a legal pre-decision decoder.

Provenance and coverage. The numerical bundle includes 90 unchanged primary nonzero-severity raw records and verified field projections of 240 endpoints covering both plants, all severities, and all arms. The two sets overlap; they are not independent experiments. Projections retain original-record hashes and are not inferred from summary statistics. The utility is $J = -(\mathrm{mean_heading_abs} + 0.1\,\mathrm{mean_lateral_abs})$, using degrees and centimetres in these records; larger is better. Each endpoint averages four evaluation episodes. Full trajectories and live-validation world-definition logs are outside this compact bundle and are not used as additional evidence for the main claims. Expected outputs verify results after calculation; they are not inputs to the estimators.

C PRIVILEGED SOURCE-OPPORTUNITY AUDIT

Purpose. This section reproduces privileged source-opportunity value on the raw additive utility scale; it does not identify a legal pre-decision rule.

The source-opportunity analysis uses raw additive utility from the frozen primary decision record. Higher is better. Conditional means over severity and five recurring seed labels are: At balanced source prevalence, the best globally fixed action is $A1$ with $V_0^\star = -1.264011$, and the matched rule yields $O(S) = D(r_{\mathrm{gt}}; S) = 0.933313$ with seed-cluster bootstrap 95% interval $[0.840099, 1.014482]$. Conditional optimality holds simultaneously for both sources in all $3{,}125$ exhaustive seed-cluster bootstrap resamples; the minimum margins are 14.842865 for observation→ $A1$ and 1.562636 for

Table 2: Conditional action values for privileged mismatch source S. The prespecified source-matched action is the unique conditional optimum in both rows.

Source	$A0$	$A1$	$A2$	Optimum	Margin to runner-up
Observation	-15.235	-0.306	-16.678	$A1$	14.928
Dynamics	-2.187	-2.222	-0.355	$A2$	1.832

Table 3: **Source-by-locus intervention responses on the primary plant.** Entries are mean PRR contrasts against $A0$, equally pooled over 15 nonzero-severity–seed tuples within source, with archived percentile intervals and paired directional support. Five seed labels recur across severities; intervals resample pooled tuples. All six underlying cells clear the denominator screen, and additive contrasts agree in sign.

Injected mismatch	$A1$ (estimator)	$A2$ (residual)
IMU bias	$+0.990$ $[+0.986, +0.993]$, $15/15$	-0.088 $[-0.101, -0.074]$, $0/15$
Motor asymmetry	-0.018 $[-0.082, +0.049]$, $6/15$	$+0.916$ $[+0.892, +0.939]$, $15/15$

dynamics$\rightarrow A2$. It also holds in all five leave-one-seed-out fits, where $O(S)$ ranges from 0.908946 to 0.971312.

For $p = \Pr(S = \text{dynamics})$, the exact frozen-population curve is

$$O_p(S) = D(r_{\text{gt}}; S) = \begin{cases} 1.8666267\, p, & 0 \le p \le 0.8976501, \\ 16.3710796(1-p), & 0.8976501 \le p \le 1. \end{cases}$$

All $3{,}125$ bootstrap resamples, all five leave-one-seed-out fits, and each severity stratum retain strict positivity for every $0 < p < 1$. Because the curve approaches zero as $p \to 0$ or $p \to 1$, this is strict interior positivity, not a uniform margin bounded away from zero. These results concern privileged source S only and do not identify a legal pre-decision decoder.

D Supporting statistics and reporting audit

Purpose. This audit locates the PRR estimand, scale distinctions, and reporting-stability checks; it does not provide evidence that $L_{\text{abs}} > 0$.

The earlier plant generation, full diagnostic calibration details, complete residual-arm scalar values, and the twelve-cell denominator record are retained in the numerical artifact, with summaries below. They are not pooled with the primary effect magnitudes.

D.1 Exact PRR estimand and aggregation

Purpose. This section clarifies the estimand underlying the reported PRR values; it does not introduce a new empirical claim.

The exact within-block PRR definition is given in Section 3.1. Table 3 averages paired PRR differences over the 15 declared source–severity–seed tuples; Appendix B records the pairing and resampling unit.

D.2 Mean of blockwise ratios vs. ratio of cell means

Purpose. This section distinguishes the reported mean of within-block PRR contrasts from a ratio formed from displayed cell means.

D.3 PRR-scale interaction and raw-utility opportunity value

Purpose. This scale crosswalk prevents comparison of quantities with different units; it does not introduce a new estimand.

Table 4: Complete reporting audit over twelve cells. $\widehat{D} = \frac{1}{5}\sum_b D_b$ and $\Delta J = \frac{1}{5}\sum_b \Delta J_{A1,A2,b}$ are separate cell means. The PRR column averages within-block normalized paired contrasts and need not equal $\Delta J/\widehat{D}$ from the displayed means. T_D is the two-component denominator screen in Appendix A.1, with nominal records deduplicated by seed; interpretation is withheld when $T_D \leq 2$.

Plant	Cell	$\hat{D}$	T_D	ΔJ	PRR contrast
1A	IMU / small	9.240	425.0	+9.61	+1.040
1A	IMU / medium	15.365	918.4	+16.59	+1.080
1A	IMU / large	20.597	1057.3	+22.92	+1.113
1A	motor / small	0.942	26.5	−0.92	−0.967
1A	motor / medium	2.018	61.8	−1.87	−0.927
1A	motor / large	3.097	94.9	−2.81	−0.907
0	IMU / small	0.0019	**0.14**	+0.04	**−3.762**
0	IMU / medium	0.189	15.7	+0.76	+4.059
0	IMU / large	0.730	58.1	+1.82	+2.489
0	motor / small	3.108	235.5	−1.22	−0.392
0	motor / medium	7.161	562.9	−1.56	−0.218
0	motor / large	9.768	491.4	−0.81	−0.083

Equation (17) reports $+2.011$ with interval $[1.950, 2.065]$ on the *PRR reporting scale*. Equation (18) reports 0.933 with interval $[0.840, 1.014]$ on the *raw additive utility scale*. The latter is reproduced in Appendix C; neither quantity is a rescaling of the other.

D.4 Denominator-resolution / Fieller reporting gate

Purpose. This audit checks reporting stability and does not provide evidence that $L_{\text{abs}} > 0$.

Eleven cells have resolved denominators with $T_D \geq 15.7$ and agree in sign between additive and normalized contrasts. The sole disagreement is the sole unresolved cell, with $T_D = 0.14$. Full blockwise denominators, including the sign changes that produce unstable blockwise ratios in that cell, are retained in the machine-readable artifact. Appendix A.1 gives the population ordering result and the Fieller resolution condition used to interpret this gate.

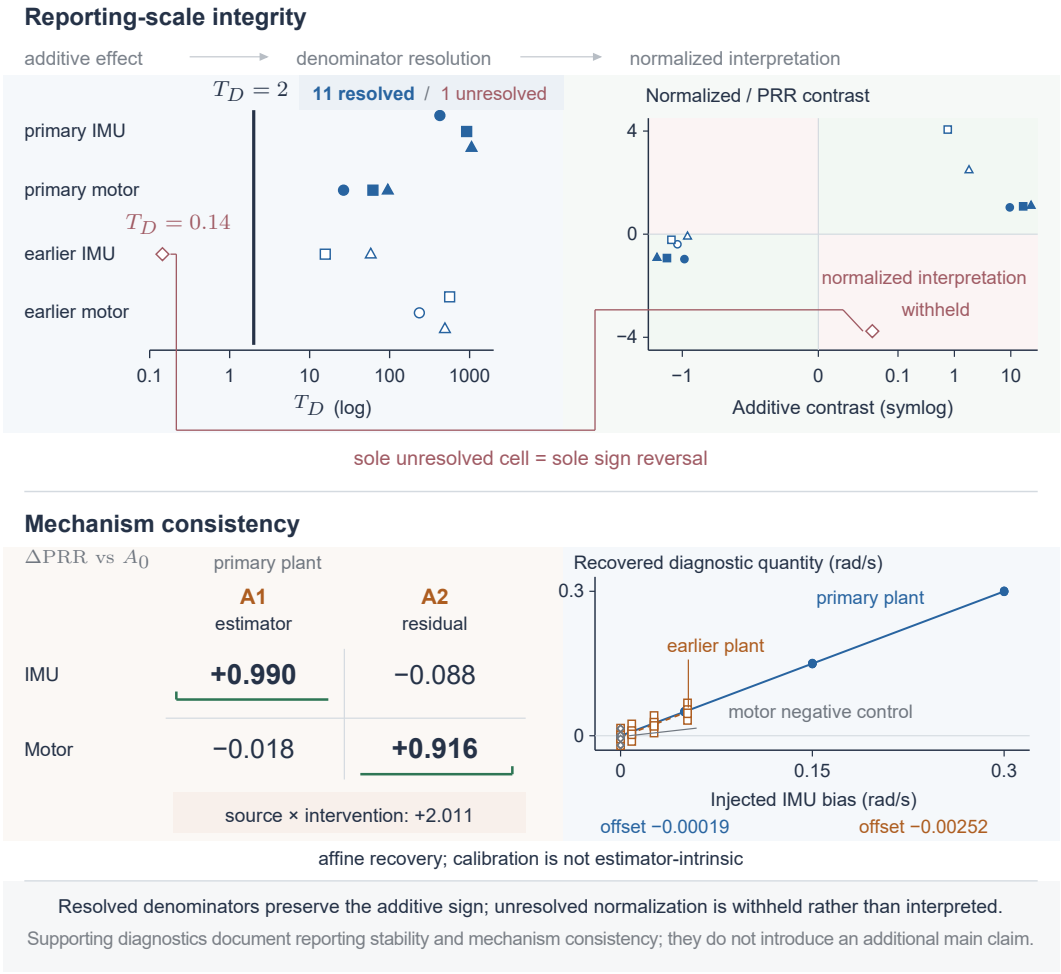

Figure 4: **Supporting reporting and mechanism diagnostics.** Eleven of twelve normalization denominators clear the reporting-resolution gate; the sole unresolved cell is also the sole cross-scale sign reversal. The source-by-intervention matrix and calibration analysis document mechanism consistency and reporting stability. The normalized interpretation of the unresolved cell is withheld, while its additive contrast is retained. These diagnostics do not introduce an additional main claim. The sign agreement is observed in these cells, not a universal implication of passing the gate; the interaction in panel c is on the PRR scale.

E CLOSEST-PRIOR CORRESPONDENCE AND AN INFERENCE ILLUSTRATION

Purpose. This section locates the closest formal correspondences and illustrates why an observed record may not identify the required repair.

The feature-gap correspondence is exact at the level of the value functional. Let $U = (Y(a) : a \in \mathcal{A})$ and use the nonnegative loss $\ell(U, a) = \max_b U_b - U_a$. Finite $\mathcal{A}$ and integrable utilities make this loss integrable. Writing $C = \mathbb{E}[\max_b U_b]$, the optimal risks with Z and M are $C - V_Z$ and $C - V_M$, so their difference is L_{abs}; the feature-gap regret identity of van Rooyen & Williamson (2015, Section 3, Theorem 1) is the direct antecedent. This reduction imports neither uniform loss bounds nor identification of the joint potential-utility vector.

For state abstraction, common optimal action corresponds to π^*-irrelevance in Li et al. (2006, Definition 3(5)). The stronger a^*-irrelevance also retains optimal action values. The single-decision correspondence does not imply the planning guarantee for an abstract MDP; their Theorem 3 explicitly separates these cases. For nested covariates $Z = (X_B, X_C)$ and $M = X_B$, Shi et al. (2019, Eq. (2.4), Theorem 2.2) connect incremental treatment value to conditional sign changes under continuity, support, and moment conditions. Our measurable-compression criterion uses an almost-sure statement and permits ties; the prior's additional regularity-dependent results are not imported. These correspondences motivate the inference discipline below rather than separate claims of new mathematical primitives.

Analytical illustration: the same record, different repairs. Let M, X be independent $\mathrm{Bernoulli}(1/2)$ variables, $Z = (M, X)$, and $h(Z) = M$, all available before the decision. The actions are $\{0, 1\}$, the specified rule is $r(M) = 0$, and $Y(0) = 1/2$. Consider the following two fully specified potential-utility models.

Model	$Y(1)$	V_Z	V_M	L_{abs}	$\delta(r)$	$D(r; M)$
P: omitted decision variable	X	$3/4$	$1/2$	$1/4$	0	0
Q: unused retained variable	M	$3/4$	$3/4$	0	$1/4$	0

In both models, $V_0^\star = 1/2$ and the specified rule has value $1/2$. In P, conditioning on Z chooses between $1/2$ and X, whereas conditioning on M leaves both action means at $1/2$. In Q, M itself permits that choice. Thus changing the rule on the same M can recover the missing $1/4$ in Q, but cannot do so in P.

If the record contains only action $A = 0$, the entire observed law of (Z, M, A, Y) is identical in P and Q. Even recording Z then does not distinguish the required repair: the alternative-action comparison is missing. The displayed $D = 0$ is calculated from the two stipulated models; it is not identified by that record over an unrestricted model class, because the best fixed-action value also depends on the unobserved action. This is an analytical illustration of the framework's inference constraints, not evidence about the controlled system or a new empirical compression-loss claim.