OBSERVATION-REGIME-AWARE BAYESIAN UPDATES FOR CLOSED-LOOP SCIENTIFIC AGENTS

Manish Sai Kota
Vanderbilt University
manish.s.kota@vanderbilt.edu

ABSTRACT

Closed-loop scientific agents that iteratively select experiments and update beliefs must handle outcomes that are not always fully observed. Measurements may fall below instrument detection limits (censored), experiments may be infeasible, or execution may fail. We formalize three observation regimes (full, censored, and absent), each requiring a distinct Bayesian likelihood, connecting this taxonomy to Rubin (1976)'s missing-data classification. In a controlled simulator with discrete hypotheses and Gaussian outcomes, we demonstrate that the most common real-world practice of substituting constants for censored non-detects and applying a standard density likelihood causes iterative Bayesian agents to converge to *wrong* hypotheses, not merely slower, regardless of experiment selection strategy. When the true hypothesis produces 96% censored outcomes, a regime-aware agent using the cumulative-distribution-function likelihood identifies it correctly in 96% of trials, while substitution-based agents identify it in 0%, always converging to the hypothesis whose mean is closest to the substitution constant.

1 INTRODUCTION

Closed-loop "AI scientist" systems can generate hypotheses, invoke tools, and iterate on results (Lu et al., 2024; Bran et al., 2024; Boiko et al., 2023). However, current systems implicitly assume that every experiment returns a clean, fully observed outcome. In practice, experimental pipelines produce outcomes of varying observability: measurements may be censored below a detection limit, experiments may violate feasibility constraints, or apparatus may malfunction.

These observation regimes demand different Bayesian update rules. Applying a probability density function (PDF) likelihood to a censored observation, as occurs when practitioners substitute non-detects with constants such as zero or LOD/2 (Helsel, 2006), is a well-documented source of bias in static settings (Tobin, 1958; Greene, 1981). We show that inside an iterative agent loop, this bias causes not just inaccurate estimates but convergence to a *wrong hypothesis entirely* when a competing hypothesis has mean closer to the substitution constant than the true hypothesis. This condition is common in practice and independent of experiment selection strategy.

Our observation regimes correspond directly to Rubin (1976)'s missing-data taxonomy: full observations are complete data, censored observations satisfy Missing At Random (MAR) when the detection limit is a known instrument property, and execution failures satisfy Missing Completely At Random (MCAR) when failure probability is hypothesis-independent. Lee et al. (2018) mapped Rubin's framework onto classical optimal experimental design; we extend this connection to information-theoretic Bayesian experimental design within autonomous agents.

Contributions. (1) We formalize three observation regimes with distinct Bayesian update rules for closed-loop scientific agents, grounding them in established missing-data theory. (2) We demonstrate empirically that substitution-based conflation of censored observations, the dominant practice in mass spectrometry and automated labs, causes convergence to wrong hypotheses under asymmetric censoring, consistent with KL-minimization under misspecification (Kleijn & van der Vaart, 2012). (3) We provide a minimal, fully reproducible testbed with complete simulator specification (§3; Appendix C).

2 Observation Regimes and Bayesian Update Rules

Let $h \in \mathcal{H}$ be a discrete hypothesis, $x \in \mathcal{X}$ an experiment, and y an outcome drawn from $p(y \mid x, h) = \mathcal{N}(\mu_{h,x}, \sigma^2)$. The agent maintains $p(h \mid \mathcal{D})$ and updates after each observation.

An experimental outcome passes through an *observation channel* that produces one of three regimes, each requiring a distinct likelihood:

Full observation. The outcome y is observed directly:

$$\ell(h) = \frac{1}{\sigma}\,\phi\left(\frac{y - \mu_{h,x}}{\sigma}\right), \tag{1}$$

where ϕ denotes the standard normal density.

Censored observation. The outcome exists but falls below a known detection limit L. The agent updates via the cumulative distribution function (Tobin, 1958):

$$\ell(h) = \Phi\left(\frac{L - \mu_{h,x}}{\sigma}\right). \tag{2}$$

Substituting a constant c and applying Eq. 1 instead is the dominant practice in environmental chemistry (Helsel, 2006; Shoari & Dubé, 2018), despite evidence of resulting bias (Greene, 1981).

Absent observation. No valid outcome is produced. This subsumes two distinct causes: the experiment may be *infeasible* (constraint violation) or may suffer an *execution failure* (operational error). In both cases, the likelihood is constant across hypotheses, producing no posterior update.

Assumptions (detailed in Appendix A): (A1) The agent knows the observation regime. (A2) Execution failures are hypothesis-independent (MCAR). (A3) The detection limit L is fixed and hypothesis-independent (MAR).

3 Experimental Setup

Problem instance. Six hypotheses, twenty experiments, noise $\sigma = 1$, detection limit $L = 1$, execution failure rate $p_{\text{fail}} = 0.05$. Hypotheses are organized as three pairs (Appendix C): a well-separated control pair (h_0: $\mu \approx +3$, 2% censored; h_1: $\mu \approx -3$, $\sim$100% censored), an "Achilles" pair (h_2, h_3: $\mu \approx +1.5$, diverging on 4 experiments), and the critical censoring pair (h_4: $\mu \approx -0.8$, $\sim$96% censored; h_5: $\mu \approx +0.2$, $\sim$78% censored). Two of twenty experiments are infeasible. The mean matrix is fixed across seeds; observation seeds control only outcome noise.

Agent comparison. Both agents select experiments by Expected Information Gain (EIG; Lindley 1956) via numerical quadrature. They differ *only* in censored-observation handling: the **regime-aware** agent applies Eq. 2; the **substitution** agent replaces the censored value with c and applies Eq. 1. We test $c = 0$ and $c = L/2 = 0.5$, both documented as standard practice. Infeasible and failed experiments are skipped by both agents.

Trial protocol. Prior: uniform, $p(h) = 1/K$ for all h. $T = 50$ steps per trial, 50 seeds per condition. Trials terminate early when $\max_h p(h \mid \mathcal{D}) > 0.999$; remaining steps record the converged posterior. Both agents are evaluated on shared pre-generated outcome tables: if both query the same experiment at the same query count, they receive the same outcome (though they are not constrained to select the same experiments). All six hypotheses serve as the true hypothesis (900 total runs).

4 Results

Substitution causes wrong convergence under asymmetric censoring. When h_4 ($\mu \approx -0.8$, 96% censored) is true, the regime-aware agent identifies it correctly in 96% of trials (48/50; 95% Wilson CI: [0.87, 0.99]). Both substitution agents achieve 0% (0/50 for $c = 0$; 0/50 for $c = L/2$), always converging to h_5 ($\mu \approx +0.2$), the hypothesis with mean closest to the substitution constant. When h_1 ($\mu \approx -3$, $\sim$100% censored) is true, the regime-aware agent achieves 100% purely through CDF ratios, while substitution again converges to h_5 on all seeds. When censoring is low (h_0, h_2, h_3, h_5), all agents achieve 100%. The effect requires high, asymmetric censoring (Appendix E).

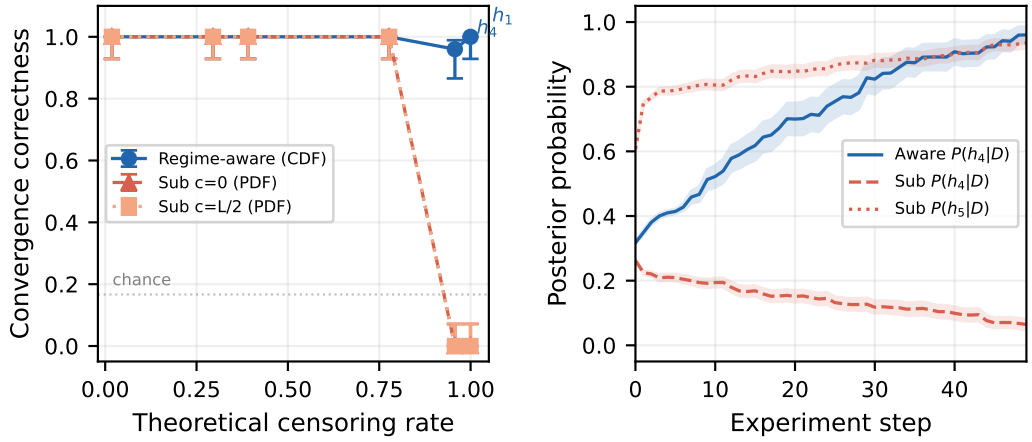

Figure 1: Left: Convergence correctness vs. theoretical censoring rate. Right: Posterior trajectory when h_4 ($\mu \approx -0.8$, 96% censored) is true. The regime-aware agent's posterior concentrates on the true hypothesis (blue, rising). The substitution agent's posterior shifts away from truth (red dashed, declining) and toward the trap h_5 ($\mu \approx 0.2$, red dotted, rising).

The mechanism: PDF evaluation bias creates an inescapable attractor. Substituting $c = 0$ for a censored observation evaluates $\phi(0; \mu_h, \sigma)$, which is maximized for whichever hypothesis has μ_h closest to 0. This is h_5 ($\mu \approx +0.2$), not h_4 ($\mu \approx -0.8$). Per censored observation: -0.30 nats against h_4 relative to h_5. The CDF likelihood instead accumulates $+0.20$ nats for h_4 (Appendix D). Critically, $\phi(y; \mu_{h_4}, \sigma) > \phi(y; \mu_{h_5}, \sigma)$ requires $y < -0.3$, but all observations are $y = c \geq 0$ or $y \geq L = 1$: no observable outcome can ever favor h_4 under substitution. This attractor is inescapable from the uniform prior (Appendix H for boundary conditions). All 100 substitution failures converge to h_5, confirming the KL-minimization prediction under misspecification (Kleijn & van der Vaart, 2012). The $c = L/2$ convention yields $2.7\times$ stronger per-observation bias than $c = 0$.

The failure is selection-independent. Under random (open-loop) experiment selection with substitution updates, the agent also achieves 0/50 correctness, converging to h_5 *more* decisively (final $P(h_5) = 0.999$) than under EIG ($P(h_5) = 0.936$). The failure is driven entirely by the update rule, not by feedback through experiment selection. Under parameters consistent with LC-MS/MS metabolomics (Appendix I), substitution misidentifies a rare low-signal metabolic phenotype ($0/50$) as a subclinical carrier, the condition whose metabolite levels naturally lie closest to LOD$/2$.

5 Discussion

The individual update rules we compare are not new: the CDF likelihood for censored data dates to Tobin (1958), and no-update for failures is standard in crash-constrained optimization (Marco et al., 2021). Our contribution is showing that correct integration of these rules into iterative agents is critical: under asymmetric censoring, the dominant substitution practice produces convergence to a wrong hypothesis, not merely biased estimates.

This failure is persistent across variations. It is selection-independent (random experiment selection also fails, §4), holds across multiple problem instances (Appendix G), and the conditions that produce it arise naturally in mass spectrometry metabolomics (Appendix I). Pipeline-aware EIG that accounts for censoring probability during experiment selection (Hickman et al., 2025) is a natural extension, though it cannot rescue a misspecified update rule.

Limitations and boundary conditions (Appendix H). The substitution failure requires a competing hypothesis with mean closer to c than the true hypothesis. The CDF advantage requires sufficient separation in censoring probabilities (σ not too small relative to $|L - \mu|$); at very low noise, both agents fail and the ordering is uninformative. A strong prior favoring the true hypothesis can resist

the bias. Our three assumptions (oracle regime labels, hypothesis-independent failures, fixed LOD) are idealizations (Appendix A); the regime-aware advantage is stable under $\sim 10\%$ label noise but degrades sharply above 20%. Scaling to continuous hypothesis spaces and real-world benchmarks (Gandhi et al., 2025; Jansen et al., 2024) is future work. We conjecture that the failure mode generalizes: in continuous parameter spaces, substitution-based likelihoods will bias posterior mass toward parameter values whose predictions best match the substitution constant, producing systematic estimation bias that compounds across iterations of an agent's decision loop.

REFERENCES

Boiko, D. A., MacKnight, R., Kline, B., and Gomes, G. Autonomous chemical research with large language models. *Nature*, 624(7992):570–578, 2023.

Bran, A.-M., Cox, S., Schilter, O., Baldassari, C., White, A. D., and Schwaller, P. ChemCrow: Augmenting large-language models with chemistry tools. *Nature Machine Intelligence*, 6:525–535, 2024.

Gandhi, K., Li, M. Y., et al. BoxingGym: Benchmarking progress in automated experimental design and model discovery. *arXiv:2501.01540*, 2025.

Greene, W. H. Sample selection bias as a specification error: Comment. *Econometrica*, 49(3):795–798, 1981.

Helsel, D. R. Fabricating data: How substituting values for nondetects can ruin results. *Chemosphere*, 65(11):2434–2439, 2006.

Hickman, R. J., Tom, G., Zou, Y., Aldeghi, M., and Aspuru-Guzik, A. Anubis: Bayesian optimization with unknown feasibility constraints for scientific experimentation. *Digital Discovery*, 4:2104–2122, 2025.

Jansen, P. A., Côté, M.-A., Khot, T., Bransom, E., Dalvi Mishra, B., Majumder, B. P., Tafjord, O., and Clark, P. DiscoveryWorld: A virtual environment for developing and evaluating automated scientific discovery agents. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024), Datasets and Benchmarks Track*, 2024.

Kleijn, B. J. K. and van der Vaart, A. W. The Bernstein–von-Mises theorem under misspecification. *Electron. J. Statist.*, 6:354–381, 2012.

Lee, K. M., Biedermann, S., and Mitra, R. Optimal design when outcome data are missing. *Statistica Sinica*, 28:1821–1838, 2018.

Lindley, D. V. On a measure of the information provided by an experiment. *Ann. Math. Stat.*, 27(4):986–1005, 1956.

Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., and Ha, D. The AI Scientist: Towards fully automated open-ended scientific discovery. *arXiv:2408.06292*, 2024.

Marco, A., Baumann, D., Khadiv, M., Hennig, P., Righetti, L., and Trimpe, S. Robot learning with crash constraints. *IEEE Robot. Autom. Lett.*, 6(2):1439–1446, 2021.

Rainforth, T., Foster, A., Ivanova, D. R., and Bickford Smith, F. Modern Bayesian experimental design. *Statistical Science*, 39(1):100–127, 2024.

Rubin, D. B. Inference and missing data. *Biometrika*, 63(3):581–592, 1976.

Shoari, N. and Dubé, J.-S. Toward improved analysis of concentration data: Embracing nondetects. *Environ. Toxicol. Chem.*, 37(3):643–656, 2018.

Tobin, J. Estimation of relationships for limited dependent variables. *Econometrica*, 26(1):24–36, 1958.

Gusev, F., Kline, B. C., Quinn, R., Xu, A., Smith, B., Frezza, B., and Isayev, O. Machine learning anomaly detection of automated HPLC experiments in the cloud laboratory. *Digital Discovery*, 4:3445–3454, 2025.

Heitjan, D. F. and Rubin, D. B. Ignorability and coarse data. *Ann. Statist.*, 19(4):2244–2253, 1991.

Lucas, D. D., Klein, R., Tannahill, J., Ivanova, D., Brandon, S., Domyancic, D., and Zhang, Y. Failure analysis of parameter-induced simulation crashes in climate models. *Geoscientific Model Development*, 6(4):1157–1171, 2013.

Wei, R., Wang, J., Su, M., Jia, E., Chen, S., Chen, T., and Ni, Y. Missing value imputation approach for mass spectrometry-based metabolomics data. *Scientific Reports*, 8:663, 2018.

Zhang, Z. and Heitjan, D. F. A simple local sensitivity analysis tool for nonignorable coarsening: Application to dependent censoring. *Biometrics*, 62(4):1260–1268, 2006.

A Assumptions and When They Break

A.1 A1: Oracle Regime Labels

Our simulator provides unambiguous regime labels. In real self-driving laboratories, the boundary between a genuine null result and an instrument failure is often ambiguous. Gusev et al. (2025) documented that air bubbles in HPLC systems cause complete peak loss, built a classifier achieving 0.96 accuracy. Wei et al. (2018) classified missing values in metabolomics into MNAR/MAR/MCAR using random forests.

We treat known-regime labels as a proof-of-concept idealization. Real deployment requires QC modules that classify observation regimes, or joint inference over both hypothesis and regime.

A.2 A2: Hypothesis-Independent Failures (MCAR)

We assume $P(\text{fail} \mid x, h)$ is constant. This fails when failure is mechanism-coupled: S-gene target failure (SGTF) in SARS-CoV-2 PCR identified variants with $> 99\%$ accuracy—the "failure" was diagnostic. Lucas et al. (2013) found 8.5% of climate simulations crashed at specific parameter combinations.

In crash-constrained BO, failure probability is modeled as input-dependent (Marco et al., 2021). When failures are hypothesis-dependent, the no-update rule is incorrect; a separate failure likelihood $P(\text{fail} \mid x, h)$ should be learned.

A.3 A3: Fixed LOD / Ignorable Censoring (MAR)

Our CDF likelihood assumes the detection limit L is an instrument property, independent of the hypothesis. This satisfies Heitjan & Rubin (1991)'s Coarsening At Random condition, making the censoring mechanism ignorable.

This breaks under matrix effects (sample composition changes effective LOD), lab-to-lab variation, or experimenter selection. Informative censoring requires joint modeling via selection models or pattern-mixture models, with sensitivity analysis via the ISNI framework (Zhang & Heitjan, 2006).

B Connection to Rubin's Missing-Data Framework

Regime	**Rubin Class**	**Update**	**Status**
Full observation	Complete	PDF	Fully informative
Censored ($Y < L$)	MAR / CAR	CDF	Partially informative
Infeasible	MCAR	None	Uninformative for H
Exec. failure	MCAR	None	Uninformative for H
Informative failure	MNAR	Failure ℓ	Extension
Informative censoring	MNAR	Joint model	Extension

Table 1: Observation regimes mapped to Rubin (1976)'s taxonomy. The top four are implemented; the bottom two are extensions for non-ignorable settings (Appendix A).

Lee et al. (2018) mapped Rubin's framework onto classical optimal design. We extend this connection to information-theoretic BED (Lindley, 1956; Rainforth et al., 2024) within an autonomous agent's decision loop, demonstrating failure modes from mishandling regimes.

C Full Simulator Specification

C.1 Mean Matrix

$\mu \in \mathbb{R}^{6\times 20}$, constructed from a dedicated RNG (seed $= 0$), fixed across observation seeds.

- h_0: $\mu_{0,x} \sim \mathcal{N}(+3.0, 0.25)$. Censoring: 1.8%.

- h_1: $\mu_{1,x} \sim \mathcal{N}(-3.0, 0.25)$. Censoring: $\sim 100\%$.
- h_2: $\mu_{2,x} \sim \mathcal{N}(+1.5, 0.25)$. Censoring: 39%.
- h_3: $\mu_{3,x} = \mu_{2,x} + \delta_x$, $\delta_x \sim \text{Unif}(2, 3.5)$ on 4 experiments, 0 elsewhere.
- h_4: $\mu_{4,x} \sim \mathcal{N}(-0.8, 0.09)$. Censoring: 95.7%. **True hypothesis in the critical test.**
- h_5: $\mu_{5,x} \sim \mathcal{N}(+0.2, 0.09)$. Censoring: 77.7%. **Trap: mean closest to substitution constant $c = 0$.**

Two of twenty experiments are permanently infeasible.

C.2 Parameters

Parameter	Value	Rationale
K	6	Exact Bayes; confusable pairs
N	20	Sufficient for EIG
σ	1.0	Moderate SNR
L	1.0	Asymmetric censoring across hyps
p_{fail}	0.05	Low; secondary to censoring
T	50	Convergence under heavy censoring
Seeds	50	Workshop-level power

C.3 Observation Channel

For experiment x under true hypothesis h^*: (1) If infeasible: INFEASIBLE. (2) With probability p_{fail}: EXEC_FAILURE. (3) Draw $y \sim \mathcal{N}(\mu_{h^*,x}, \sigma^2)$. (4) If $y < L$: CENSORED. (5) Otherwise: FULL.

Outcomes are pre-generated for all (experiment, query-index) pairs.

C.4 EIG Computation

$\text{EIG}(x) = I(H; Y \mid x, \mathcal{D})$ via numerical quadrature (Riemann sum, 200 points). The quadrature grid spans the prediction range of hypotheses with posterior > 0.001; all K hypotheses contribute to the integrand, weighted by their posterior mass. Both agents use the same EIG formula applied to their respective posteriors; neither accounts for the possibility of censoring during selection.

D Why Substitution Causes Wrong Convergence

When h_4 ($\mu \approx -0.8$) is true and $L = 1$, $\sigma = 1$:

CDF likelihood (correct): $\Phi(1; -0.8, 1) = 0.964$ vs. $\Phi(1; 0.2, 1) = 0.788$. Per censored observation: $\ln(0.964/0.788) = +0.20$ nats for h_4. After 45 censored observations: $+9.1$ nats. Decisive.

PDF at $c = 0$ (wrong): $\phi(0; -0.8, 1) = 0.290$ vs. $\phi(0; 0.2, 1) = 0.391$. Per censored observation: $\ln(0.290/0.391) = -0.30$ nats against h_4. After 45 observations: -13.5 nats. Decisively wrong.

PDF at $c = 0.5$ (also wrong, worse): $\phi(0.5; -0.8, 1) = 0.171$ vs. $\phi(0.5; 0.2, 1) = 0.381$. Per observation: $\ln(0.171/0.381) = -0.80$ nats against h_4. After 45 observations: -36.0 nats.

Impossibility: $\phi(y; -0.8, 1) > \phi(y; 0.2, 1)$ requires $y < -0.3$. All observable outcomes are either $y = c \geq 0$ or $y \geq L = 1$. No outcome can ever favor h_4 under substitution.

E Full Results

Table 2 reports convergence correctness for all six hypotheses across all three agents.

h_{true}	Cens. %	Regime-aware	Sub(0)	Sub($L/2$)	Wrong h
h_0 ($\mu \approx +3$)	1.8%	100%	100%	100%	—
h_2 ($\mu \approx +1.5$)	39.1%	100%	100%	100%	—
h_3 ($\mu \approx +1.5+$)	29.5%	100%	100%	100%	—
h_5 ($\mu \approx +0.2$)	77.7%	100%	100%	100%	—
h_4 ($\mu \approx -0.8$)	95.7%	96%	0%	0%	h_5
h_1 ($\mu \approx -3$)	100%	100%	0%	0%	h_5

Table 2: Convergence correctness (fraction of 50 seeds with correct MAP at final step) sorted by theoretical censoring rate. All 100 substitution failures converge to h_5 (mean closest to $c = 0$).

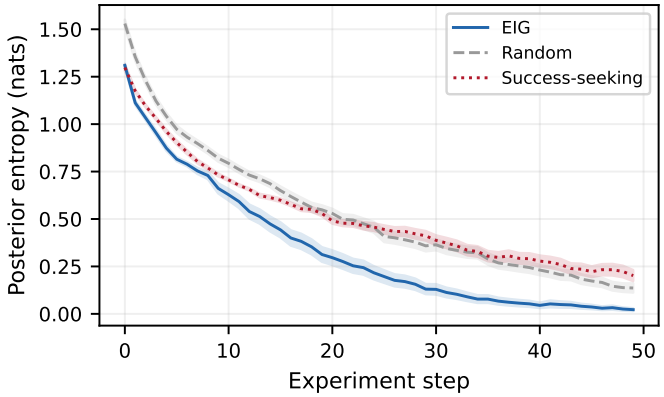

Figure 2: Posterior entropy vs. experiment step for three selection methods (all using regime-aware updates, h_4 true, 50 seeds). Shading: ± 1 SE.

F EIG vs. Alternative Selection Methods

With regime-aware updates held constant, EIG achieves 48/50 correct identifications (final entropy 0.022 ± 0.007 nats), compared to 41/50 for random selection (0.136 ± 0.024) and 45/50 for success-seeking (0.202 ± 0.029). Consistent with prior work (Lindley, 1956; Rainforth et al., 2024), EIG-based selection reduces posterior entropy faster than alternatives. Note that success-seeking achieves higher correctness than random (45/50 vs. 41/50) despite higher final entropy (0.20 vs. 0.14): by selecting experiments with high predicted outcomes, it obtains uncensored observations that happen to favor the correct hypothesis, but does not efficiently eliminate competing hypotheses, producing a less concentrated posterior that nevertheless has the right mode. We focus the main text on the orthogonal question of update-rule correctness.

G Generalization Across Problem Instances

We repeat the critical experiment (h_4 true, 50 seeds) across five matrix seeds (0–4). Table 3 reports correctness, the substitution agent's most common wrong hypothesis, and the hypothesis with mean closest to $c = 0$.

The substitution agent achieves non-zero correctness on seeds 3 and 4. This occurs because different random matrix seeds produce different mean vectors for h_4: on these seeds, h_4's realized mean is closer to $c = 0$, reducing the per-observation PDF bias between h_4 and h_5. Critically, even on these seeds, when the substitution agent converges to a wrong hypothesis, it is *always* h_5, confirming the KL-minimization prediction holds in all five instances without exception. The regime-aware agent remains correct in 82–100% of trials across all seeds.

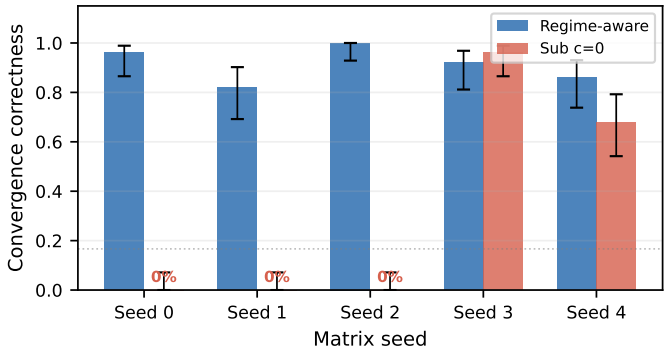

Figure 3: Convergence correctness for h_4 true across five matrix seeds. Error bars: 95% Wilson CI.

Seed	Aware (95% CI)	Sub(0) (95% CI)	Wrong h	Predicted	Match
0	96% [.87,.99]	0% [.00,.07]	h_5	h_5	✓
1	82% [.69,.90]	0% [.00,.07]	h_5	h_5	✓
2	100% [.93,1.0]	0% [.00,.07]	h_5	h_5	✓
3	92% [.81,.97]	96% [.87,.99]	h_5	h_5	✓
4	86% [.74,.93]	68% [.54,.79]	h_5	h_5	✓

Table 3: Generalization across matrix seeds. "Predicted" is $\arg\min_h |\bar{\mu}_h - c|$. When the substitution agent is wrong, it always converges to the predicted trap, confirming the KL-minimization mechanism across all five problem instances.

H Robustness and Boundary Conditions

We probe the sensitivity of our results to parameter changes and assumption violations. All tests use 10–50 seeds; directional findings are consistent across repetitions.

H.1 Noise Level (σ)

The CDF-likelihood advantage depends on separation in censoring probabilities across hypotheses. At $\sigma = 1$ (as reported), $\Phi(1; -0.8, 1) = 0.964$ vs. $\Phi(1; 0.2, 1) = 0.788$: well-separated. At $\sigma = 0.5$, both values approach 1.0 (0.99984 vs. 0.945), collapsing the CDF log-ratio to 0.06 nats/observation. In 50 steps this provides only $\sim$3 nats of evidence, insufficient for reliable identification. Empirically, at $\sigma = 0.5$ the regime-aware agent achieves $1/50$ correctness and the substitution agent achieves $14/50$, *reversing the ordering*. This occurs because both agents effectively fail (neither can distinguish h_4 from h_5 via CDF), and the substitution agent's occasional correct MAP assignment is due to chance among near-uniform posteriors (typical $P(h_4) \approx 0.50$). At $\sigma \geq 1$, the CDF-likelihood advantage is decisive and robust (tested at $\sigma \in \{1, 2, 3\}$).

H.2 Detection Limit (L)

At $L = -1$ (h_4 censoring $\approx 42\%$), substitution achieves $8/10$ correctness; the bias is too weak to overcome the uncensored signal. At $L \geq 2$ (censoring $\geq 99.7\%$), both agents fail: even the CDF likelihood provides insufficient discrimination. Our operating point ($L = 1$, h_4 censoring $= 96\%$) falls in the regime where censoring is heavy enough to cause substitution failure but not so extreme that CDF ratios collapse.

H.3 Prior Sensitivity

From a uniform prior ($P(h_4) = 1/6$), substitution achieves $0/50$. With an informative prior $P(h_4) = 0.9$, substitution achieves $6/10$: the substitution bias (-0.30 nats/observation) requires $\sim$8 censored observations to overcome $\ln(9) = 2.2$ nats of prior advantage, and some seeds exhaust

the prior before converging. At $P(h_4) = 0.999$, substitution achieves $10/10$: the prior is strong enough to resist $\sim$15 nats of accumulated bias within 50 steps.

H.4 Trap-Hypothesis Dependence

Removing h_5 (the trap) from the hypothesis set yields $K = 5$. With h_5 absent, h_4 itself has the mean closest to $c = 0$ among remaining hypotheses, so substitution correctly favors h_4: $10/10$. This confirms that the failure requires a competing hypothesis with mean closer to c than the true hypothesis. Conversely, swapping the means of h_4 and h_5 causes the trap to flip: substitution converges to h_4 ($\mu \approx +0.2$) when h_5 ($\mu \approx -0.8$) is true, confirming the KL-minimization prediction.

H.5 Substitution Value (c)

Sweeping c from -1.5 to $+0.5$ reveals a sharp phase transition at $c \approx -0.3$, the midpoint of the h_4 and h_5 means. For $c < -0.3$ (closer to h_4's mean), substitution favors h_4 and succeeds; for $c \geq 0$ (closer to h_5), it fails completely. The standard conventions ($c = 0$, $c = L/2 = 0.5$) both fall on the failing side.

H.6 Regime-Label Noise (Assumption A1)

Randomly mislabeling $5-10\%$ of CENSORED observations as FULL (passing $y = L$ as the observed value) does not degrade the regime-aware agent ($10/10$). At 20% mislabel rate, correctness drops to $3/10$; at 50%, the agent fails completely ($0/10$). The substitution agent is unaffected by mislabeling (it already treats censored observations as if fully observed). This places a practical bound on Assumption A1: the regime-aware advantage is robust to $\sim 10-15\%$ label noise.

H.7 Non-Gaussian Noise

Replacing Gaussian noise with t-distributed noise (heavy tails) while keeping Gaussian update rules: at df $= 3$, the regime-aware agent achieves $7/10$ (vs. $10/10$ with Gaussian noise) and substitution achieves $0/10$. At df $= 1$ (Cauchy), regime-aware drops to $2/10$. The Gaussian CDF likelihood is itself misspecified, though it still outperforms substitution.

I Illustrative Scenario: LC-MS/MS Metabolomics

To illustrate that the conditions for substitution failure arise in genomics-adjacent domains, we instantiate the simulator with parameters consistent with LC-MS/MS metabolomics, where non-detect rates of $20-50\%$ are routine (Wei et al., 2018) and LOD/2 substitution remains the most common imputation practice. We work in $\log_2$ relative abundance, where the Gaussian assumption is standard for mass spectrometry data. Detection limit: 1% relative abundance ($\log_2 = -6.6$), noise $\sigma = 0.8$ ($\sim 1.7\times$ multiplicative variability, within the typical LC-MS/MS range), substitution $c = \text{LOD}/2$ ($\log_2 = -7.6$).

The six hypotheses represent metabolic conditions producing different analyte abundance profiles:

h	Condition	Rel. abd.	Cens.	Correct
h_0	Healthy control	1.2	0%	50 / 50
h_2	Mild upregulation	0.14	0%	50 / 50
h_3	Drug-altered	0.18	0%	50 / 50
h_5	Subclinical carrier	0.006	82%	50 / 50
h_4	**Rare low-signal phenotype**	**0.003**	**98.5%**	**33 / 0**
h_1	Complete pathway knockout	0.0003	100%	50 / 0

Table 4: Metabolomics scenario (50 seeds). Relative abundance is fraction of reference; values below 0.01 (LOD) are censored. Correct column: regime-aware / substitution. All substitution failures converge to h_5.

The substitution failure is clean: when the true condition is a rare low-signal phenotype (h_4, 98.5% censored), the substitution agent achieves $0/50$ on all seeds tested, always converging to the sub-clinical carrier (h_5), the condition whose metabolite levels ($\sim 0.6\%$ of reference) naturally lie closest to the LOD/2 substitution constant (0.5%). The regime-aware agent achieves $33/50$; its 17 failures converge to the complete pathway knockout (h_1, 100% censored), which produces nearly identical CDF ratios (-0.016 nats/observation relative to h_4). This residual confusion reflects the intrinsic difficulty of distinguishing two near-completely-censored conditions via CDF ratios alone, not a limitation of regime-aware updating. The key comparison ($33/50$ vs. $0/50$) shows that substitution bias is the dominant failure mode, even when the CDF fix does not fully resolve identification in the primary experiment.