PRISM Supplement: Technical Report, Synthetic Study, and Additional Proofs

Sean Slattery

February 2026

Supplement Overview

This supplement contains implementation-facing details for the PRISM workflow. It documents synthetic scenario construction, stress designs, diagnostic conventions, and expanded proofs.

1 Simulation Design

Purpose. This section documents synthetic scenario construction and metric computation used to exercise the PRISM pipeline under controlled distortions.

What you will learn.

- Scenario families and what each family targets.
- Shared simulation knobs and evidence-generation rules.
- Core metrics used in audit artifacts.

The synthetic suite is organized into families A1–A3 and B1–B5:

- A1: information-efficiency curve behavior under varying informed-trader fractions.
- A2: convergence timing under different arrival rates and fixed horizons.
- A3: strategic-timing attack scenarios with attack scale and timing controls.
- B1: correction no-regret checks under longshot/herding/liquidity combinations.
- B2: asymptotic-rate checks under increasing effective evidence.
- B3: misspecification-regret grid over structural/ML misspecification levels.
- B4: regime concentration tests under controlled evidence totals.
- B5: projection-impact checks for constrained output post-processing.

Shared controls include random seed, effective sample budget, attack intensity, and liquidity regime toggles. Scenario manifests and outputs are materialized under ‘results/**‘ as scenario-level ‘summary.json‘, ‘tests.csv‘, and ‘audit.json‘ artifacts.

Primary metrics include Brier score, log score, reliability deviations, nominal-coverage deviations, and scenario-specific regret proxies. For bounded outcomes and concentration sanity checks, Hoeffding-style inequalities provide conservative finite-sample guardrails [Hoeffding, 1963].

What this section established.

- The synthetic suite has explicit families with targeted stress intent.
- Scenario outputs are reproducible through repository artifacts.
- Metric choices are aligned with probabilistic forecasting diagnostics.

2 Stress Tests

Purpose. This section records the stress mechanisms used to challenge calibration transport, correction stability, and posterior robustness under adversarial or low-support regimes.

What you will learn.

- How adversarial timing and liquidity stresses are parameterized.
- Why misspecification grids are needed for fusion robustness checks.
- Which stresses are most likely to trigger diagnostic failures.

Stress mechanisms include:

- timing attacks that concentrate directional flow in short windows,
- low-liquidity settings where a small number of messages dominate weighted evidence,
- mismatch injections between calibration assumptions and generated outcomes,
- regime-switch windows where correction parameters drift relative to training periods.

For each mechanism, PRISM logs whether correction maps amplify or attenuate instability. The intended interpretation is diagnostic: stress failures identify boundary regions where assumptions do not hold strongly enough for aggressive claims.

Current scenario-level audit artifacts indicate substantial stress sensitivity in several families. In particular, no-regret and concentration-oriented claims are not uniformly supported under low-support and coverage-disqualification regimes.

What this section established.

- Stress tests explicitly target known weak points of the workflow.
- Failures are treated as boundary information, not as noise to suppress.
- The suite supports conservative refinement of assumptions and thresholds.

3 Diagnostics and Plot Conventions

Purpose. This section standardizes how calibration and posterior diagnostics are interpreted so scenario comparisons remain consistent.

What you will learn.

- Reliability and coverage diagnostics used in artifact generation.
- Low-support and fallback flags that qualify interpretation.
- How to read mixed scenario outcomes without overclaiming.

Diagnostics reported in scenario artifacts include reliability curves by model variant, coverage deltas at nominal 90% and 95% levels, and model-level degeneration flags where bin support is weak.

Interpretation rules used in this supplement:

- Coverage-off-nominal flags are disqualifying for strong calibration claims.
- Low-support flags trigger conservative interpretation of any pass result.
- Fallback-based calibration variance implies wider posterior uncertainty by design.
- Mixed outcomes in a family are not aggregated into unconditional success claims.

This convention follows a restrained Bayesian workflow perspective: diagnostics are part of the model output contract, not post-hoc annotations.

What this section established.

- Diagnostic interpretation rules are explicit and reproducible.
- Support quality is a first-class gating signal for empirical claims.
- Mixed evidence is carried forward without optimistic aggregation.

4 Additional Proofs

Purpose. This section gives expanded arguments for the stability statements used in the main text.

What you will learn.

- A fuller perturbation decomposition for posterior means.
- How concentration bounds enter constants in stability inequalities.
- The role of Lipschitz correction assumptions in finite-sample controls.

Expanded proof sketch for the main stability theorem. Write posterior mean as

$$\mu(\alpha, \beta, y, n) = \frac{\alpha + y}{\alpha + \beta + n}.$$

For two realizations, add and subtract an intermediate value using matched counts:

$$|\mu - \mu'| \leq |\mu(\alpha, \beta, y, n) - \mu(\alpha', \beta', y, n)| + |\mu(\alpha', \beta', y, n) - \mu(\alpha', \beta', y', n')|.$$

First term: if corrected means differ by at most $L_t\Delta_p$ and concentration is bounded above and below, parameter perturbations are linear in $L_t\Delta_p$.

Second term: partial derivatives with respect to (y, n) are bounded by inverse effective concentration factors, yielding constants multiplying $(\Delta_y + \Delta_n)$.

Combining terms gives

$$|\mu - \mu'| \le C_t \left(L_t\Delta_p + \Delta_y + \Delta_n\right),$$

with finite C_t under bounded concentration. □

Remark 1. When concentration collapses toward zero because of aggressive uncertainty caps, constants increase. This is expected: PRISM chooses conservative uncertainty over artificial stability.

What this section established.

- The stability bound follows by a direct decomposition argument.
- Bounded concentration assumptions are explicit in the constant term.
- Conservative concentration choices can widen sensitivity bounds.

5 Additional Tables

Purpose. This section records compact status tables so readers can connect main-text claims to scenario-level artifact outcomes.

What you will learn.

- A compact mapping from experiment family to determinate and indeterminate status counts.
- Which families currently support only limited conclusions.
- Where to find scenario-level JSON and CSV artifacts.

Family	Total Scenarios	Determinate Pass Count	Indeterminate Count
A1	9	0	9
A2	10	2	8
A3	8	0	8
B1	20	0	20
B2	2	0	2
B3	8	0	8
B4	4	0	4
B5	6	6	0

Table 1: Family-level status summary computed from ‘results/**/audit.json‘ using ‘criteria_evaluation.overall_pass‘.

What this section established.

- The supplement provides quick status lookup tied to scenario-level audit artifacts.
- Family-level summaries remain conservative and preserve indeterminacy.
- Detailed scenario diagnostics remain in per-scenario artifacts (‘summary.json‘, ‘audit.json‘, ‘tests.csv‘).

6 Extended Module A: Likelihood and Conjugate Updating

In this phase, we derive the likelihood generated by parimutuel YES/NO trader actions and combine it with the hybrid prior from Phase 2 to construct the posterior distribution of the event probability

$$p := Q(A) = Q(S_T > K).$$

All notation, including the hybrid prior parameters (α_0, β_0), follows Phase 2.

4.1 Trader Actions and the Likelihood Model

Let n denote the number of traders participating in the parimutuel book, and let y denote the number of YES positions. A NO position is treated as a vote for the complement A^c.

Traders may act strategically: some may exaggerate their signals, herd behind early order flow, or attempt to manipulate the book. We acknowledge the possibility of such distortions but defer correction to Phase 6. For the purposes of likelihood construction, we treat the realized counts $(y, n-y)$ as the observable actions that the mechanism must interpret.

Although traders' private signals may be correlated, and their actions may be strategically dependent, we assume conditional independence given the latent probability p for the sole purpose of deriving the Beta–Binomial conjugacy. This follows standard practice in Bayesian market microfoundations.

4.2 Binomial Likelihood of the Parimutuel Order Flow

Conditionally on the latent event probability p, we model the YES count y as

$$y \mid p \sim \mathrm{Binomial}(n, p),$$

with likelihood

$$L(y \mid p) = \binom{n}{y} p^y (1-p)^{n-y}. \tag{1}$$

Although later phases introduce liquidity-adjusted counts y^* and n^*, the present phase uses the raw counts $(y, n-y)$ for the purpose of deriving the conjugate update.

4.3 Conjugate Updating with the Hybrid Prior

The hybrid prior from Phase 2 is

$$p \sim \mathrm{Beta}(\alpha_0, \beta_0), \qquad \alpha_0, \beta_0 > 0.$$

Lemma 1 (Posterior Form)**.** *Combining the Beta prior with the Binomial likelihood (equation reference) yields the posterior*

$$p \mid y \sim \mathrm{Beta}(\alpha_{\mathrm{post}}, \beta_{\mathrm{post}}),$$

where

$$\alpha_{\mathrm{post}} = \alpha_0 + y, \qquad \beta_{\mathrm{post}} = \beta_0 + (n-y).$$

Proof. The Beta density is proportional to $p^{\alpha_0-1}(1-p)^{\beta_0-1}$. Multiplying by (equation reference) gives a kernel proportional to

$$p^{\alpha_0+y-1}(1-p)^{\beta_0+(n-y)-1},$$

the kernel of a $\text{Beta}(\alpha_{\text{post}}, \beta_{\text{post}})$. □

4.4 Posterior Mean and Variance

Proposition 1 (Posterior Moments). *For the posterior Beta distribution above,*

$$\mathbb{E}[p \mid y] = \frac{\alpha_{\text{post}}}{\alpha_{\text{post}} + \beta_{\text{post}}},$$

and

$$\text{Var}(p \mid y) = \frac{\alpha_{\text{post}}\beta_{\text{post}}}{(\alpha_{\text{post}} + \beta_{\text{post}})^2(\alpha_{\text{post}} + \beta_{\text{post}} + 1)}.$$

Proof. These are standard properties of the Beta distribution. □

4.5 Posterior Predictive Distribution (Beta–Binomial Form)

The posterior predictive distribution of observing y YES votes under the prior $\text{Beta}(\alpha_0, \beta_0)$ is

$$P(y \mid \alpha_0, \beta_0) = \binom{n}{y} \frac{B(\alpha_0 + y,\ \beta_0 + n - y)}{B(\alpha_0, \beta_0)},$$

where $B(\cdot, \cdot)$ is the Beta function.

Theorem 1 (Posterior Predictive Distribution). *Let* $p \sim \text{Beta}(\alpha_0, \beta_0)$ *and* $y \mid p \sim \text{Binomial}(n, p)$. *Then the marginal distribution of* y *is Beta–Binomial with pmf above.*

Proof. Integrate the joint distribution $P(y \mid p)f(p)$ over $p \in [0, 1]$, and use the identity

$$\int_0^1 p^{a-1}(1-p)^{b-1}\, dp = B(a, b).$$

□

4.6 Conditional Independence as a Modeling Approximation

The Beta–Binomial updating step presented above relies on the assumption that, conditional on the latent event probability p, individual trader actions $s_i \in \{\text{YES}, \text{NO}\}$ are independent Bernoulli draws. Such conditional independence is standard in Bayesian aggregation models, but it is not expected to hold exactly in parimutuel markets where traders may observe and react to earlier order flow. In particular, herding behavior generates temporal dependence among trades: late traders may overweight recent order patterns even when those patterns do not reflect new private information.

In this framework, conditional independence is therefore best interpreted as a *modeling approximation* rather than a literal behavioral assumption. Empirically observed violations of independence are handled in two ways:

1. **Behavioral Adjustment (Phase 6).** The Stage 1 correction introduces weights w_i^{beh} applied to individual orders. When herding creates clusters of correlated trades, these weights reduce the effective contribution of late correlated orders, mitigating departures from independence.

2. **Sensitivity Analysis (Phase 7).** Simulation regimes in Phase 7 introduce explicit dependence structures among trades, including herding and correlated decision rules. These regimes allow us to evaluate how violations of independence affect the PRISM posterior and how effectively the two-stage bias-correction layer controls such deviations.

This modeling approximation preserves conjugacy and analytic tractability while acknowledging that the empirical behavior of order flow contains richer dependence patterns. Phases 6 and 7 are specifically designed to examine, interpret, and correct these dependencies.

4.7 Incorporation of Strategic Distortion

Because traders may submit exaggerated or strategically distorted orders, the raw counts $(y, n-y)$ encode:

1. private signals,
2. beliefs about other traders' signals,
3. strategic considerations,
4. liquidity constraints.

Phase 6 will introduce formal bias adjustments using *effective* counts (y^*, n^*-y^*) and distortion offsets (δ_+, δ_-). For now, the posterior above represents the *unadjusted* Bayesian update based on the observable order flow.

4.8 Conditional i.i.d., Dependence, and the Role of Weights

Assumption A4 models the adjusted order contributions as conditionally i.i.d. Bernoulli (or bounded) signals given the latent event probability p. This is a deliberate simplification. In realistic markets, herding, order-splitting, and informational cascades induce dependence across orders.

There are two conceptually distinct modeling choices:

- *Fully dependent likelihood.* One could specify an explicit joint law for the order sequence $(Z_1, \ldots, Z_n)$, for example via an Ising model or a Markov random field. This yields a non-factorizing likelihood and a non-conjugate posterior that typically requires MCMC or variational methods.
- *Weighted pseudo-likelihood.* PRISM instead uses a Beta–Binomial update based on adjusted counts (y_n^*, n_n^*), together with mixing-based asymptotics (Phase 8). This implicitly replaces the true dependent likelihood with an exponential family surrogate whose sufficient statistics are the weighted sums. The resulting Beta posterior is then interpreted as the KL-projection of the intractable posterior onto the Beta family (Phase 8ZZ).

In this sense, the behavioral weighting layer is not merely a "patch" on an i.i.d. model, but a way to summarize dependence and heterogeneity into effective counts that remain compatible with a tractable exponential family update. The trade-off is explicit: PRISM sacrifices an exact likelihood for closed-form inference plus an information-theoretic guarantee that, within the Beta family, the posterior is as close as possible (in Kullback–Leibler sense) to the ideal but intractable posterior.

4.9 Why PRISM Uses Beta Conjugacy

There are many ways to model belief aggregation in markets. PRISM uses a Beta–Binomial structure for a simple reason: it gives closed-form updates and keeps the link between data and parameters transparent.

- Each component of the prior can be read as a pseudo-count: structural information contributes η_{str} virtual observations, the ML model contributes n_{ML}, and the crowd contributes adjusted effective counts n_n^*.
- Updating is a matter of adding these counts, with no numerical integration or sampling required.
- The resulting posterior has a clear interpretation: it is the Beta distribution that best matches, in KL sense, the information contained in the adjusted counts and the hybrid prior.

More flexible approaches, such as MCMC over a fully dependent likelihood, can capture richer structures but at the cost of interpretability and computation. PRISM is intentionally positioned as a tractable, interpretable baseline: it prioritizes closed-form inference and moment-based robustness over fully nonparametric modeling. This makes it easier to diagnose, explain, and test in simulation before considering heavier alternatives.

4.10 Output of Phase 4

The output of this phase is the posterior hyperparameter pair

$$(\alpha_{\text{post}}, \beta_{\text{post}}),$$

which becomes the foundation for Phase 5 (posterior predictive pricing) and is later refined in Phase 6 (bias–corrected posterior).

Phase 5. Posterior Predictive event-probability inference

Given the posterior distribution of the event probability

$$p := Q(A) = Q(S_T > K)$$

from Phase 4, we now derive the posterior summaries of YES/NO parimutuel contracts and digital derivatives. We additionally establish no–arbitrage properties, uncertainty bounds, and risk–adjusted pricing rules.

$$p \mid y \sim \text{Beta}(\alpha_{\text{post}}, \beta_{\text{post}}),$$

with

$$\alpha_{\text{post}} = \alpha_0 + y, \qquad \beta_{\text{post}} = \beta_0 + (n - y).$$

5.1 From PRISM Posteriors to projection kernels and Option Prices

In previous phases, PRISM produces, for a fixed maturity T, a posterior distribution for the event probability

$$p(K,T) = Q(S_T > K \mid \mathcal{I}),$$

where $\mathcal{I}$ denotes the combined information from the structural model, the machine learning prior, and the adjusted parimutuel order flow. Let $\hat{p}(K,T)$ denote the posterior mean, so that the PRISM-implied digital price at strike K is

$$D_{mathrmPRISM}(K,T) = e^{-rT}\hat{p}(K,T),$$

with r the risk-free rate.

We now show how the entire reference-measure distribution and projection kernel can be recovered (at least formally) from the strike-dependent posterior, and how vanilla option prices follow via standard integral transforms.

Theorem 2 (PRISM-Implied reference-measure CDF, Density, and projection kernel)**.** *Fix a maturity $T > 0$ and suppose that, for each strike $K \geq 0$, PRISM produces a posterior distribution for $p(K,T)$ with mean $\hat{p}(K,T)$. Assume:*

*(i) (**Smoothness in Strike**) The map $K \mapsto \hat{p}(K,T)$ is differentiable almost everywhere, with $\hat{p}(K,T)$ non-increasing in K and*

$$\lim_{K\to 0} \hat{p}(K,T) = 1, \qquad \lim_{K\to\infty} \hat{p}(K,T) = 0.$$

*(ii) (**No-Arbitrage Regularity**) The function $\hat{p}(K,T)$ is right-continuous with left limits and defines a valid tail function for a probability distribution on $[0,\infty)$.*

Then:

*(a) (**reference-measure CDF and Density**) The PRISM-implied reference-measure CDF and density at maturity T are given by*

$$F^Q_{mathrmPRISM}(K,T) = Q(S_T \leq K \mid \mathcal{I}) = 1 - \hat{p}(K,T),$$

and, wherever differentiable,

$$f^Q_{mathrmPRISM}(K,T) = \frac{\partial}{\partial K} F^Q_{mathrmPRISM}(K,T) = -\frac{\partial}{\partial K}\hat{p}(K,T).$$

*(b) (**Call Prices and Breeden–Litzenberger**) The PRISM-implied call price at strike K and maturity T is*

$$C_{mathrmPRISM}(K,T) = e^{-rT}\int_K^\infty (s-K)\, f^Q_{mathrmPRISM}(s,T)\, ds.$$

Equivalently, if we define

$$C_{mathrmPRISM}(K,T) = e^{-rT}\mathbb{E}^Q[(S_T - K)^+ \mid \mathcal{I}],$$

then the Breeden–Litzenberger relation holds:

$$\frac{\partial^2}{\partial K^2} C_{mathrmPRISM}(K,T) = e^{-rT} f^Q_{mathrmPRISM}(K,T),$$

whenever the derivatives exist.

(c) ***(State-Price Density and projection kernel)*** *The state-price density associated with PRISM at maturity T is*

$$\phi_{mathrmPRISM}(s,T) = e^{-rT} f^Q_{mathrmPRISM}(s,T),$$

so that

$$C_{mathrmPRISM}(K,T) = \int_K^\infty (s-K)\,\phi_{mathrmPRISM}(s,T)\,ds.$$

If P denotes a physical measure under which S_T has density f^P, and if f^P is strictly positive on the support of $f^Q_{mathrmPRISM}$, then the PRISM-implied projection kernel can be written (up to normalization) as

$$m_{mathrmPRISM}(s,T) \propto \frac{f^Q_{mathrmPRISM}(s,T)}{f^P(s)}.$$

Proof (Sketch). Under (i) and (ii), the function $K \mapsto \hat{p}(K,T)$ satisfies the basic properties of a strike-tail function: it is non-increasing, right-continuous, and converges to 1 and 0 at the boundaries. Thus it defines

$$F^Q_{mathrmPRISM}(K,T) = 1 - \hat{p}(K,T)$$

as a valid CDF on $[0,\infty)$, and the density $f^Q_{mathrmPRISM} = \partial_K F^Q_{mathrmPRISM}$ exists almost everywhere, yielding part (a).

For part (b), the standard reference-measure pricing relation gives

$$C_{mathrmPRISM}(K,T) = e^{-rT}\int_K^\infty (s-K)\, f^Q_{mathrmPRISM}(s,T)\,ds.$$

Differentiating once with respect to K yields

$$\frac{\partial}{\partial K} C_{mathrmPRISM}(K,T) = -e^{-rT}\int_K^\infty f^Q_{mathrmPRISM}(s,T)\,ds = -e^{-rT}(1-F^Q_{mathrmPRISM}(K,T)),$$

and differentiating a second time gives

$$\frac{\partial^2}{\partial K^2} C_{mathrmPRISM}(K,T) = e^{-rT} f^Q_{mathrmPRISM}(K,T),$$

which is the Breeden–Litzenberger formula applied to the PRISM-implied density.

For part (c), the state-price density is by definition the Radon–Nikodym derivative of the pricing operator with respect to Lebesgue measure, which in continuous-time arbitrage-free settings is $e^{-rT} f^Q(s,T)$. Substituting $f^Q_{mathrmPRISM}$ yields $\phi_{mathrmPRISM}$. When a physical density f^P is available, absolute continuity of Q with respect to P on the relevant support implies

$$\frac{dQ}{dP}(s) \propto \frac{f^Q_{mathrmPRISM}(s,T)}{f^P(s)},$$

so that $m_{mathrmPRISM}(s,T) \propto dQ/dP$ has the indicated form. □

Remark 2 (Discrete Approximation from a Strike Grid)*.* In practice, PRISM will produce posterior means $\hat{p}(K_j,T)$ on a discrete grid of strikes $\{K_j\}_{j=1}^J$. The PRISM-implied CDF and density can then be approximated by

$$F^Q_{mathrmPRISM}(K_j,T) \approx 1 - \hat{p}(K_j,T),$$

and

$$f^{Q}_{mathrmPRISM}(K_j, T) \approx -\frac{\hat{p}(K_{j+1}, T) - \hat{p}(K_j, T)}{K_{j+1} - K_j},$$

for $j = 1, \dots, J-1$. Similarly, the call price curve can be approximated via

$$C_{mathrmPRISM}(K_j, T) \approx \sum_{l=j}^{J-1} e^{-rT}\,\hat{p}(K_l, T)\,\Delta K_l, \qquad \Delta K_l = K_{l+1} - K_l,$$

which implements the integral in (b) as a Riemann sum. These discrete approximations provide a direct pathway from PRISM posteriors to implied call prices and densities on a finite strike grid.

5.1.1 PRISM-Implied Kernels and Asset Pricing Puzzles

The PRISM framework produces a reference-measure distribution $f^{Q}_{mathrmPRISM}(\cdot, T)$ and associated state-price density $\phi_{mathrmPRISM}(s, T) = e^{-rT} f^{Q}_{mathrmPRISM}(s, T)$. This section connects these objects to empirical projection kernels and the stylized facts of asset pricing puzzles.

The final PRISM outputs in Phase 5 are option prices computed under a reference-measure measure $\mathbb{Q}$ induced by a projection kernel (state–price density) $M_T = \frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T}$. If the kernel derived from a misspecified model or from a raw mixture/particle posterior is not carefully regularized, it may violate basic conditions: *positivity*, *unit expectation*, or the *martingale* property for discounted asset prices. This subsection introduces a kernel enforces no–arbitrage while preserving the posterior information accumulated in Phases 2–6.

Recall from Phase 1 that under the physical measure $\mathbb{P}$ the asset model:

(2)

where (W^S_t, W^v_t) are Brownian motions under $\mathbb{P}$ with correlation $\rho \in [-1, 1]$, and parameters $(\mu, \kappa, \theta, \sigma)$ lie in the $\Pi_\phi(\cdot \mid \mathcal{D})$ from Phases 5–6 is interpreted as providing information about the terminal distribution of S_T and related events (e.g. default, barrier crossing) rather than replacing the structural dynamics.

Exponential Martingale projection kernel. We construct a projection kernel as an exponential martingale with a market–price of risk process λ_t:

$$M^\lambda_t := \exp\left(-\int_0^t \lambda_s\, dW^S_s - \frac{1}{2}\int_0^t \lambda_s^2\, ds\right), \qquad t \in [0, T], \tag{3}$$

where $(\lambda_t)_{t\in[0,T]}$ is progressively measurable and adapted to $(\mathcal{F}_t)$.

Assumption 1 (Novikov Condition for the Kernel)**.** *The process λ_t satisfies Novikov's condition:*

$$\mathbb{E}_\mathbb{P}\left[\exp\left(\frac{1}{2}\int_0^T \lambda_s^2\, ds\right)\right] < \infty.$$

Lemma 2 (Positivity and Normalization of the Kernel)**.** *Under Assumption reference, the process M^λ_t defined in (equation reference) is a positive $\mathbb{P}$–martingale with $M^\lambda_0 = 1$ and*

$$\mathbb{E}_\mathbb{P}[M^\lambda_T \mid \mathcal{F}_0] = 1.$$

Consequently, defining $\mathbb{Q}$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = M^\lambda_T$$

yields a probability measure equivalent to $\mathbb{P}$.

Proof. By Assumption reference, the stochastic exponential M_t^λ is a true martingale (Novikov's condition). It is strictly positive by definition and satisfies $M_0^\lambda = 1$. Thus $\mathbb{E}_\mathbb{P}[M_T^\lambda \mid \mathcal{F}_0] = M_0^\lambda = 1$, and M_T^λ defines a Radon–Nikodym derivative of a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$. □

Under $\mathbb{Q}$, Girsanov's theorem implies that

$$W_t^{S,\mathbb{Q}} = W_t^S + \int_0^t \lambda_s \, ds$$

$$dS_t = S_t\big((\mu - \lambda_t\sqrt{v_t})\, dt + \sqrt{v_t}\, dW_t^{S,\mathbb{Q}}\big).$$

Definition 1 (Drift Adjustment and No–Arbitrage)**.** *Let r_t denote the short rate. To ensure that the discounted price $\tilde{S}_t := e^{-\int_0^t r_s ds} S_t$ is a $\mathbb{Q}$–martingale, we choose λ_t such that*

$$\mu - \lambda_t\sqrt{v_t} = r_t, \quad i.e. \quad \lambda_t = \frac{\mu - r_t}{\sqrt{v_t}}.$$

With this choice, the S_t dynamics under $\mathbb{Q}$ become

$$dS_t = S_t\big(r_t\, dt + \sqrt{v_t}\, dW_t^{S,\mathbb{Q}}\big),$$

ensuring no–arbitrage in the usual sense.

Esscher-Type Calibration to PRISM Posterior Moments. The specification in Definition reference yields a one–parameter family of kernels indexed by the physical drift μ and the short rate r_t. To incorporate the information contained in the PRISM posterior $\Pi_\phi(\cdot \mid \mathcal{D})$, we calibrate λ_t (or, equivalently, an Esscher tilt parameter) so that selected posterior moments posterior.

Let $X_T := \log S_T$ and define an Esscher–type kernel based on X_T:

$$M_T^\theta := \frac{\exp(\theta X_T)}{\mathbb{E}_\mathbb{P}[\exp(\theta X_T) \mid \mathcal{F}_0]}.$$

By construction, $M_T^\theta > 0$ and $\mathbb{E}_\mathbb{P}[M_T^\theta \mid \mathcal{F}_0] = 1$. For a given Esscher parameter θ, this defines a reference-measure measure $\mathbb{Q}^\theta$ via $\frac{d\mathbb{Q}^\theta}{d\mathbb{P}} = M_T^\theta$. We choose θ (or a small vector of tilting parameters) to match posterior–implied moments from PRISM, e.g.

$$\mathbb{E}_{\mathbb{Q}^\theta}[S_T \mid \mathcal{F}_0] = \mathbb{E}_{\Pi_\phi}[S_T \mid \mathcal{D}], \qquad \mathbb{E}_{\mathbb{Q}^\theta}[\mathbf{1}\{S_T > K\} \mid \mathcal{F}_0] = \mathbb{E}_{\Pi_\phi}[\mathbf{1}\{S_T > K\} \mid \mathcal{D}],$$

for one or more strikes K. This calibration step projects the raw PRISM without violating positivity or the martingale property.

Theorem 3 (Kernel Regularization and No–Arbitrage Preservation)**.** *Let $\Pi_\phi(\cdot \mid \mathcal{D})$ be the fully corrected PRISM posterior from Phases 5–6. Define a regularized projection kernel either as:*

(a) an exponential martingale M_t^λ as in (equation reference) with λ_t chosen according to Definition reference, or

(b) an Esscher–type density M_T^θ based on $X_T = \log S_T$ with θ calibrated to match a set of PRISM posterior moments.

Assume Novikov's condition (Assumption reference) holds for the chosen λ_t or that M_T^θ has finite exponential moments under $\mathbb{P}$. Then:

(i) M_T is strictly positive and satisfies $\mathbb{E}_{\mathbb{P}}[M_T \mid \mathcal{F}_0] = 1$, so it defines a valid projection kernel.

(ii) The discounted price process $\tilde{S}_t = e^{-\int_0^t r_s ds} S_t$ is a $\mathbb{Q}$–martingale for the induced reference-measure measure $\mathbb{Q}$.

(iii) Option prices computed as

$$C(K, T) = e^{-\int_0^T r_s ds} \mathbb{E}_{\mathbb{Q}}\big[(S_T - K)^+ \mid \mathcal{F}_0\big]$$

PRISM posterior in the sense that their moments agree with the posterior–implied targets used in calibration.

Proof Sketch. (i) Positivity and normalization follow from Lemma reference in the exponential martingale case and by construction in the Esscher case. (ii) The drift adjustment in Definition reference ensures that S_t has drift r_t under $\mathbb{Q}$, so $\tilde{S}_t$ is a local martingale; integrability conditions (e.g. Novikov, uniform integrability) promote it to a true martingale. (iii) Option prices under $\mathbb{Q}$ inherit no–arbitrage from the standard reference-measure valuation framework. The Esscher calibration conditions guarantee that chosen moments (e.g. of S_T or digital payoffs) with the PRISM posterior without violating the martingale and positivity constraints. □

Remark 3 (Integration with Simulation and Robustness Phases). In Phase 7, Monte Carlo simulations of (S_t, v_t) under the regularized kernel M^λ or M^θ can be used to assess the stability of option prices and digital probabilities under parameter uncertainty and posterior perturbations. In Phase 8, the robustness results for the posterior (e.g. in Wasserstein or Hellinger distance) combined with the exponential kernel representation yield explicit bounds on the sensitivity of reference-measure prices to data and model perturbations. Crucially, kernel regularization is applied *after* the event–probability posterior has been corrected for nonlinear, dependent, and multimodal effects, so that enforcing no–arbitrage does not undo the informational gains of PRISM but instead embeds them into a structurally consistent, coherence-preserving projection measure.

5.2 Posterior Predictive Mean and Variance

Lemma 3 (Posterior Mean and Variance)**.** *For a Beta$(\alpha_{\text{post}}, \beta_{\text{post}})$ posterior,*

$$\hat{p}_{\text{post}} := \mathbb{E}[p \mid y] = \frac{\alpha_{\text{post}}}{\alpha_{\text{post}} + \beta_{\text{post}}},$$

$$\text{Var}(p \mid y) = \frac{\alpha_{\text{post}}\beta_{\text{post}}}{(\alpha_{\text{post}} + \beta_{\text{post}})^2(\alpha_{\text{post}} + \beta_{\text{post}} + 1)}.$$

Proof. Standard Beta distribution identities. □

5.3 Posterior Predictive Distribution for Digital Outcomes

Let Z denote the payoff of a YES contract:

$$Z = \mathbb{1}\{A\}.$$

Proposition 2 (Posterior Predictive Distribution of Digital Payoff)**.** *The posterior–predictive distribution of Z is Bernoulli with mean*

$$\pi_{\text{pred}} = \mathbb{E}[Z \mid y] = \hat{p}_{\text{post}} = \frac{\alpha_{\text{post}}}{\alpha_{\text{post}} + \beta_{\text{post}}}.$$

Proof. Since $Z \mid p \sim \text{Bernoulli}(p)$,

$$\mathbb{E}[Z \mid y] = \mathbb{E}[\mathbb{E}[Z \mid p, y] \mid y] = \mathbb{E}[p \mid y] = \hat{p}_{\text{post}}.$$

□

5.4 Posterior Predictive Prices of Parimutuel YES/NO Contracts

A YES contract pays 1 if A occurs and 0 otherwise. Under reference-measure valuation, its fair price is the posterior predictive mean.

Theorem 4 (Arbitrage–Free YES/NO Pricing)**.** *The posterior–predictive fair prices of YES and NO contracts are*

$$\pi_{\text{YES}} = \hat{p}_{\text{post}}, \qquad \pi_{\text{NO}} = 1 - \hat{p}_{\text{post}}.$$

They satisfy the no–arbitrage identity:

$$\pi_{\text{YES}} + \pi_{\text{NO}} = 1.$$

Proof. Follows immediately from $\pi_{\text{YES}} = \mathbb{E}[Z \mid y]$ and $1 - Z = \mathbb{1}\{A^c\}$. □

5.5 Properties of posterior summaries

Proposition 3 (Monotonicity in YES Votes)**.** *The price $\pi_{\text{YES}} = \alpha_{\text{post}}/(\alpha_{\text{post}} + \beta_{\text{post}})$ is strictly increasing in the count y.*

Proof. Since $\alpha_{\text{post}} = \alpha_0 + y$,

$$\frac{\partial}{\partial y}\frac{\alpha_0 + y}{\alpha_0 + \beta_0 + n} > 0.$$

□

Proposition 4 (Continuity)**.** *The price is continuous in both α_{post} and β_{post} and therefore in y and n.*

Proof. Rational function of continuous arguments. □

5.6 Mixture Posterior Extension for Multimodal Beliefs

The baseline PRISM framework represents the posterior distribution of the event probability p by a single Beta distribution. This is appropriate when the likelihood is approximately unimodal and the crowd can be described by a single effective subpopulation. Once nonlinear structural distortions (Phase 6) and heterogeneous trader behavior are admitted, the true posterior often becomes *multimodal.* In such settings, forcing a unimodal Beta posterior, or even a single Beta built on a misspecified likelihood, can lead to severely miscalibrated probabilities and distorted pricing.

To represent multimodality explicitly, we extend the PRISM posterior to a finite mixture of Betas constructed via stacking of multiple candidate submodels or strata (e.g. trader types, structural regimes, or segmentation by liquidity conditions).

Assumption 2 (Candidate Submodels and Stratified Posteriors)**.** *Let $\{\mathcal{M}_k : k = 1, \dots, K\}$ be a finite collection of candidate PRISM submodels. Each $\mathcal{M}_k$ is defined by:*

(i) a data subset or stratum $\mathcal{D}_k$ (e.g. orders from a given trader type, structural regime, or filtered particle trajectory);

(ii) a Beta posterior on p of the form

$$p \mid \mathcal{D}_k \sim \mathrm{Beta}(\alpha_k, \beta_k),$$

obtained from the usual Beta–Binomial update under $\mathcal{M}_k$;

(iii) a corresponding predictive density $m_k(y)$ for new Bernoulli data $Y \in \{0,1\}$, given by the Beta–Binomial predictive:

$$m_k(y) \;=\; \int_0^1 p^y(1-p)^{1-y}\,\mathrm{Beta}(\alpha_k,\beta_k)(dp) \;=\; \begin{cases} \dfrac{\beta_k}{\alpha_k+\beta_k} & \text{if } y=0, \\ \dfrac{\alpha_k}{\alpha_k+\beta_k} & \text{if } y=1. \end{cases}$$

We interpret each $\mathcal{M}_k$ as capturing one coherent "mode" of the crowd's beliefs or structural environment.

Definition 2 (Stacked Mixture Posterior over p)**.** *Let $\mathcal{D}^{\text{hold}} = \{Y_i^{\text{hold}} : i = 1, \ldots, n_{\text{hold}}\}$ be a holdout dataset (e.g. a subset of orders or markets not used to fit α_k, β_k). Define stacking weights $w = (w_1, \ldots, w_K)$ in the probability simplex*

$$\Delta^{K-1} := \left\{ w \in [0,1]^K : \sum_{k=1}^K w_k = 1 \right\}$$

by minimizing the negative log predictive likelihood of the stacked model:

$$w^* \;:=\; \operatorname*{arg\,min}_{w \in \Delta^{K-1}} \left\{ -\sum_{i=1}^{n_{\text{hold}}} \log\left(\sum_{k=1}^K w_k\, m_k\big(Y_i^{\text{hold}}\big) \right) \right\}.$$

The stacked mixture posterior *for p is then defined as*

$$\Pi_{\text{mix}}(dp) \;:=\; \sum_{k=1}^K w_k^*\,\mathrm{Beta}(\alpha_k,\beta_k)(dp).$$

Proposition 5 (Posterior Mean and Predictive under the Mixture)**.** *Under Assumption reference and Definition reference, the stacked mixture posterior Π_{mix} is a finite mixture of Beta distributions. Its mean and predictive distribution are given by:*

(i) ***Posterior mean of p:***

$$\hat{p}_{\text{mix}} \;:=\; \mathbb{E}_{\Pi_{\text{mix}}}[p] \;=\; \sum_{k=1}^K w_k^* \frac{\alpha_k}{\alpha_k+\beta_k}.$$

(ii) ***Predictive distribution for a new Bernoulli outcome Y:***

$$\mathbb{P}(Y=1) \;=\; \sum_{k=1}^K w_k^* \frac{\alpha_k}{\alpha_k+\beta_k}, \qquad \mathbb{P}(Y=0) \;=\; \sum_{k=1}^K w_k^* \frac{\beta_k}{\alpha_k+\beta_k}.$$

In particular, the mixture mean $\hat{p}_{\mathrm{mix}}$ can be used as the crowd–adjusted event probability for YES/NO contracts, and the full mixture distribution Π_{mix} can be propagated into posterior predictive option prices as in the baseline PRISM construction.

Proof. Since Π_{mix} is a finite convex combination of Beta distributions, it is a well–defined probability measure on $[0,1]$. The expression for $\hat{p}_{\mathrm{mix}}$ follows by linearity of expectation:

$$\hat{p}_{\mathrm{mix}} = \int_0^1 p\,\Pi_{\mathrm{mix}}(dp) = \sum_{k=1}^{K} w_k^* \int_0^1 p\,\mathrm{Beta}(\alpha_k, \beta_k)(dp) = \sum_{k=1}^{K} w_k^* \frac{\alpha_k}{\alpha_k + \beta_k}.$$

The predictive probabilities follow similarly by integrating $p^y(1-p)^{1-y}$ with respect to Π_{mix} and using the Beta–Binomial formulas. □

The mixture posterior Π_{mix} explicitly retains multimodality when the component posteriors $\mathrm{Beta}(\alpha_k, \beta_k)$ are well separated. In particular, if some strata correspond to optimistic trader types and others to pessimistic types (or different structural regimes), then Π_{mix} can exhibit multiple modes. The stacking weights w_k^* tilt the mixture toward components that perform better on the holdout set $\mathcal{D}^{\mathrm{hold}}$, reducing sensitivity to any single misspecified submodel.

Definition 3 (Moment–Matched Single–Beta Approximation)**.** *For interpretability and analytical convenience, one may define a moment–matched single–Beta approximation* $\mathrm{Beta}(\tilde{\alpha}, \tilde{\beta})$ *to the mixture posterior by matching the first two moments:*

$$\mathbb{E}[p] = \hat{p}_{\mathrm{mix}}, \qquad \mathbb{V}\mathrm{ar}[p] = \sum_{k=1}^{K} w_k^* \frac{\alpha_k \beta_k}{(\alpha_k + \beta_k)^2(\alpha_k + \beta_k + 1)} + \sum_{k=1}^{K} w_k^* \left(\frac{\alpha_k}{\alpha_k + \beta_k} - \hat{p}_{\mathrm{mix}} \right)^2.$$

The approximation $\mathrm{Beta}(\tilde{\alpha}, \tilde{\beta})$ *is then chosen such that*

$$\frac{\tilde{\alpha}}{\tilde{\alpha} + \tilde{\beta}} = \hat{p}_{\mathrm{mix}}, \qquad \frac{\tilde{\alpha}\tilde{\beta}}{(\tilde{\alpha} + \tilde{\beta})^2(\tilde{\alpha} + \tilde{\beta} + 1)} = \mathbb{V}\mathrm{ar}_{\Pi_{\mathrm{mix}}}[p].$$

We emphasize that this is an approximation layer *used for convenience, not an exact representation of* Π_{mix}*.*

The next result formalizes a limitation: when the mixture posterior is sufficiently multimodal, no single Beta distribution can uniformly approximate it. This provides a theoretical warning against collapsing Π_{mix} to a single Beta in regimes of strong heterogeneity.

Theorem 5 (No Unimodal Beta Can Uniformly Approximate a Strongly Multimodal Mixture)**.** *Let* Π_{mix} *be a mixture of two Betas*

$$\Pi_{\mathrm{mix}}(dp) = \tfrac{1}{2}\,\mathrm{Beta}(\alpha_1, \beta_1)(dp) + \tfrac{1}{2}\,\mathrm{Beta}(\alpha_2, \beta_2)(dp),$$

with means $\mu_1 \neq \mu_2$ *and variances* σ_1^2, σ_2^2 *bounded. Assume that* $|\mu_1 - \mu_2| \geq \varepsilon$ *for some* $\varepsilon > 0$*, and that each component is sharply concentrated around its mean (e.g.* $\alpha_k + \beta_k$ *large). Then there exists a constant* $c(\varepsilon) > 0$ *such that for any single Beta distribution* $\mathrm{Beta}(\tilde{\alpha}, \tilde{\beta})$*,*

$$\left\| \Pi_{\mathrm{mix}} - \mathrm{Beta}(\tilde{\alpha}, \tilde{\beta}) \right\|_{\mathrm{TV}} \;\geq\; c(\varepsilon),$$

where $\|\cdot\|_{\mathrm{TV}}$ *denotes total variation distance. In particular, no single Beta can approximate* Π_{mix} *arbitrarily well as the component means separate.*

Proof. Since each Beta component is sharply concentrated around its mean, for any $\delta > 0$ small enough there exist disjoint intervals $I_1, I_2 \subset [0,1]$ such that

$$\mathbb{P}_{\mathrm{Beta}(\alpha_1,\beta_1)}(I_1) \geq 1-\delta, \quad \mathbb{P}_{\mathrm{Beta}(\alpha_2,\beta_2)}(I_2) \geq 1-\delta, \quad I_1 \cap I_2 = \emptyset,$$

and I_1 and I_2 are separated by at least $\varepsilon/2$. Then

$$\Pi_{\mathrm{mix}}(I_1) \geq \tfrac{1}{2}(1-\delta), \qquad \Pi_{\mathrm{mix}}(I_2) \geq \tfrac{1}{2}(1-\delta).$$

Any single Beta $\mathrm{Beta}(\tilde{\alpha}, \tilde{\beta})$ has unimodal density on $(0,1)$ and cannot assign mass arbitrarily close to $\frac{1}{2}(1-\delta)$ to both disjoint, well–separated intervals I_1 and I_2. Consequently there exists $c(\varepsilon,\delta) > 0$ such that

$$\sup_{A \subset [0,1]} \left| \Pi_{\mathrm{mix}}(A) - \mathrm{Beta}(\tilde{\alpha}, \tilde{\beta})(A) \right| \;\geq\; c(\varepsilon, \delta),$$

for all choices of $(\tilde{\alpha}, \tilde{\beta})$. Taking δ small and absorbing it into $c(\varepsilon)$ yields the claim. □

Remark 4 (Implications for PRISM). Theorem reference shows that when the crowd beliefs are strongly multimodal (e.g. two well–separated trader camps or regimes), any attempt to compress the posterior into a single Beta inevitably loses structural information and cannot be uniformly well–calibrated. In such regimes PRISM should operate directly with the mixture posterior Π_{mix} (or its predictive functionals), and treat any moment–matched single–Beta representation as an approximation with explicit, non–vanishing divergence from the true posterior. In Phase 8, we extend the divergence and robustness results to incorporate mixture posteriors, providing bounds on the loss incurred by such approximations.

5.7 Credible Intervals and Price Uncertainty Bands

From the posterior variance,

$$\sigma^2_{\mathrm{post}} := \mathrm{Var}(p \mid y),$$

we obtain a symmetric credible interval

$$\pi_{\mathrm{YES}} \pm z_\gamma \, \sigma_{\mathrm{post}},$$

where z_γ is the standard normal quantile. Exact Beta quantiles may also be used:

$$\left[\mathrm{BetaInv}\big(\tfrac{\gamma}{2}; \alpha_{\mathrm{post}}, \beta_{\mathrm{post}}\big)\,,\; \mathrm{BetaInv}\big(1-\tfrac{\gamma}{2}; \alpha_{\mathrm{post}}, \beta_{\mathrm{post}}\big)\right].$$

5.8 Risk–Adjusted Prices

Agents may wish to incorporate uncertainty into the price. Define the risk–adjusted YES price

$$\pi_{\mathrm{risk},+} = \hat{p}_{\mathrm{post}} + c\,\sigma_{\mathrm{post}}, \qquad c \geq 0,$$

and the corresponding conservative price

$$\pi_{\mathrm{risk},-} = \hat{p}_{\mathrm{post}} - c\,\sigma_{\mathrm{post}}.$$

Lemma 4 (Risk Monotonicity). *Risk–adjusted prices increase with uncertainty:*

$$\frac{\partial \pi_{\mathrm{risk},+}}{\partial \sigma_{\mathrm{post}}} = c > 0.$$

Proof. Immediate from the definition. □

5.9 Strategic Distortion Considerations

As noted in Phase 4, traders may strategically exaggerate their positions. Thus $(y, n-y)$ may reflect strategic behavior in addition to private signals. The posterior derived above therefore represents the *unadjusted* prediction based solely on the observed order flow.

Phase 6 introduces:

- liquidity–adjusted counts y^*, n^*,
- distortion offsets δ_+, δ_-,
- and the bias–corrected posterior $\text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}})$.

5.10 Output of Phase 5

The output of this phase consists of:

$$(\alpha_{\text{post}}, \beta_{\text{post}}), \qquad \pi_{\text{YES}} = \hat{p}_{\text{post}}, \qquad \pi_{\text{NO}} = 1 - \hat{p}_{\text{post}},$$

together with credible intervals and risk–adjusted variants. These serve as inputs to Phase 6, where distortions and liquidity effects are formally corrected.

7 Extended Module B: Corrections and Robustness Layer

Phases 4 and 5 produce a posterior distribution

$$p \mid y \sim \text{Beta}(\alpha_{\text{post}}, \beta_{\text{post}})$$

and posterior summaries

$$\pi_{\text{YES}} = \frac{\alpha_{\text{post}}}{\alpha_{\text{post}} + \beta_{\text{post}}}, \qquad \pi_{\text{NO}} = 1 - \pi_{\text{YES}}.$$

However, the observed parimutuel order flow $(y, n-y)$ may be distorted by behavioral biases and structural market effects, including:

- *Long–shot bias:* overbetting low–probability outcomes,
- *Herd behavior:* traders imitating late order flow.

In this phase, we construct a two–stage correction layer:

1. Stage 1: behavioral bias correction *(long–shot, herding)*,
2. Stage 2: structural/liquidity distortion correction via offset parameters (δ_+, δ_-).

The goal is a bias–corrected posterior

$$p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}})$$

with associated robust prices, while preserving Beta–Binomial conjugacy and no–arbitrage.

6.1 From Raw Counts to Behavioral and Structural Distortions

Let $s_i \in \{\text{YES}, \text{NO}\}$ denote trader i's action, and recall:

$$y = \sum_{i=1}^{n} \mathbb{1}\{s_i = \text{YES}\}, \qquad n - y = \sum_{i=1}^{n} \mathbb{1}\{s_i = \text{NO}\}.$$

We conceptually distinguish:

- *Behavioral distortions*, driven by perception and psychology (long–shot bias, herd behavior);
- *Structural distortions*, driven by market mechanics (liquidity imbalances, whale trades, microstructure noise).

Although both arise from complex microfoundations, we implement bias correction via deterministic weighting of individual orders and scalar offsets. Stochastic liquidity processes (e.g. Poisson arrivals) are acknowledged conceptually but not explicitly modeled here.

6.2 Stage 1: Behavioral Bias Correction (Long–Shot Bias and Herd Behavior)

Long–shot bias. Empirical and experimental work on parimutuel markets documents a tendency for traders to overbet low–probability events (*long–shot bias*). In a binary setting, this manifests as disproportionate YES volume when the true p is small.

Herd behavior. Traders may also *herd* on late–arriving order flow: observing a run of recent YES bets, they may overweight YES regardless of their private signals.

$w_i^{\text{beh}} \in (0, \infty)$ applied to each trader action:

$$y^{(1)} := \sum_{i=1}^{n} w_i^{\text{beh}}\, \mathbb{1}\{s_i = \text{YES}\}, \qquad n^{(1)} - y^{(1)} := \sum_{i=1}^{n} w_i^{\text{beh}}\, \mathbb{1}\{s_i = \text{NO}\}. \tag{4}$$

Example (qualitative).

- To mitigate long–shot bias, YES trades on extreme low–probability strikes may receive $w_i^{\text{beh}} < 1$.
- To mitigate herding, late trades that follow a long run of identical orders may receive $w_i^{\text{beh}} < 1$, while early trades receive $w_i^{\text{beh}} \approx 1$.

Lemma 5 (Behaviorally Adjusted Counts)**.** *If all* $w_i^{\text{beh}} > 0$*, then*

$$y^{(1)} > 0, \quad n^{(1)} - y^{(1)} > 0 \quad \Longrightarrow \quad 0 < y^{(1)} < n^{(1)}.$$

Proof. Positivity and finiteness follow from finiteness of n and positivity of weights. $\square$

6.3 Stage 2: Nonlinear Structural Distortions via Regime Mixtures

Stage 1 produces a behaviorally corrected Beta posterior

$$p \mid \mathcal{D}_1 \sim \mathrm{Beta}(\alpha_1, \beta_1),$$

where $\mathcal{D}_1$ denotes the effective (possibly weighted) order flow after behavioral adjustments (herding, long–shot bias, etc.) have been accounted for. In the original linear specification of Stage 2, structural distortions such as liquidity imbalances or whale dominance were represented by constant additive offsets (δ_+, δ_-) to the pseudo–counts. This amounts to replacing (α_1, β_1) by $(\alpha_1+\delta_+, \beta_1+\delta_-)$ and preserves single–Beta conjugacy. However, this specification implicitly assumes that all distortions are *additive* in pseudo–counts and therefore fails to represent multiplicative or exponential distortions (e.g. nonlinear amplification of long–shot bias under herding). In such cases, the adjusted posterior is systematically misspecified and the robustness guarantees of Phase 8 can fail.

To accommodate nonlinear distortions while preserving an analytically tractable posterior, we introduce a finite collection of *structural distortion regimes* and represent Stage 2 as a mixture over regime–specific pseudo–count corrections.

Assumption 3 (Structural Distortion Regimes)**.** *Let $R \in \mathbb{N}$ be finite. For each $r \in \{1, \dots, R\}$, there is a* structural distortion regime *characterized by:*

1. *a prior weight $\pi_r > 0$ with $\sum_{r=1}^{R} \pi_r = 1$;*
2. *measurable pseudo–count corrections*

$$g_r^+ : \mathcal{S} \to \mathbb{R}, \qquad g_r^- : \mathcal{S} \to \mathbb{R},$$

   where $\mathcal{S}$ denotes the Stage 1 summary statistics (e.g. total effective YES count y_1, NO count $n_1 - y_1$, order–book imbalance, realized spread, volume concentration, etc.);

3. *a boundedness condition*

$$\sup_{s\in\mathcal{S}} \max\left\{|g_r^+(s)|,\ |g_r^-(s)|\right\} \ \le\ G_r < \infty;$$

4. *an admissibility condition ensuring that, for all $s \in \mathcal{S}$,*

$$\alpha_r(s) := \alpha_1 + g_r^+(s) > 0, \qquad \beta_r(s) := \beta_1 + g_r^-(s) > 0.$$

We interpret r as a latent structural state (e.g. "balanced liquidity", "whale–dominated", "thin–book"), and the functions $g_r^\pm$ may be nonlinear in the Stage 1 statistics $s \in \mathcal{S}$.

Definition 4 (Stage 2 Nonlinear Structural Adjustment)**.** *Let $\mathcal{D}_2$ denote the full data entering Stage 2, including the Stage 1 summary $s \in \mathcal{S}$ and any structural covariates (e.g. book depth, cross–venue imbalance). Under Assumption reference, the Stage 2 adjustment proceeds as follows:*

1. *Draw a latent regime $R^* \in \{1, \dots, R\}$ with $\mathbb{P}(R^* = r) = \pi_r$.*
2. *Given $R^* = r$ and $\mathcal{D}_2$, replace the Stage 1 parameters (α_1, β_1) by*

$$\alpha_r(s) = \alpha_1 + g_r^+(s), \qquad \beta_r(s) = \beta_1 + g_r^-(s),$$

   and define the regime–conditional Stage 2 posterior

$$p \mid (\mathcal{D}_2, R^* = r) \sim \mathrm{Beta}\big(\alpha_r(s), \beta_r(s)\big).$$

The unconditional Stage 2 posterior is obtained by marginalizing over R^*:

$$\Pi_{\text{PRISM}}(dp \mid \mathcal{D}_2) \;=\; \sum_{r=1}^{R} \omega_r(\mathcal{D}_2)\, \text{Beta}\big(\alpha_r(s), \beta_r(s)\big)\, dp,$$

where the regime weights $\omega_r(\mathcal{D}_2)$ are the posterior probabilities $\mathbb{P}(R^* = r \mid \mathcal{D}_2)$.

The next result shows that, under mild conditions, this Stage 2 specification yields a finite mixture of Beta posteriors and therefore provides a tractable representation of nonlinear structural distortions.

Theorem 6 (Mixture–of–Beta Conjugacy under Nonlinear Structural Corrections)**.** *Suppose that the Stage 1 posterior satisfies*

$$p \mid \mathcal{D}_1 \sim \text{Beta}(\alpha_1, \beta_1),$$

and that Assumption reference holds. Let $\mathcal{D}_2$ be any σ–algebra generating the Stage 2 summary $s \in \mathcal{S}$ and any additional structural covariates used to evaluate $g_r^{\pm}$. Then:

1. *For each fixed regime $r \in \{1, \ldots, R\}$ and realization $s \in \mathcal{S}$ with $\alpha_r(s) > 0$, $\beta_r(s) > 0$, the regime–conditional Stage 2 posterior*

$$p \mid (\mathcal{D}_2, R^* = r) \sim \text{Beta}\big(\alpha_r(s), \beta_r(s)\big)$$

   is a well–defined Beta distribution.

2. *The unconditional Stage 2 posterior $\Pi_{\text{PRISM}}(\cdot \mid \mathcal{D}_2)$ is a finite mixture of Beta distributions:*

$$\Pi_{\text{PRISM}}(dp \mid \mathcal{D}_2) = \sum_{r=1}^{R} \omega_r(\mathcal{D}_2)\, \text{Beta}\big(\alpha_r(s), \beta_r(s)\big)\, dp,$$

   with regime weights

$$\omega_r(\mathcal{D}_2) = \frac{\pi_r\, L_r(\mathcal{D}_2)}{\sum_{k=1}^{R} \pi_k\, L_k(\mathcal{D}_2)},$$

   where $L_r(\mathcal{D}_2)$ is the marginal likelihood of $\mathcal{D}_2$ under regime r.

3. *In particular, the Stage 2 posterior mean can be written as*

$$\mathbb{E}\big[\, p \mid \mathcal{D}_2 \,\big] = \sum_{r=1}^{R} \omega_r(\mathcal{D}_2)\, \frac{\alpha_r(s)}{\alpha_r(s) + \beta_r(s)}.$$

Proof. (i) For each r and $s \in \mathcal{S}$, the admissibility condition in Assumption reference(iv) guarantees $\alpha_r(s) > 0$ and $\beta_r(s) > 0$. Hence $\text{Beta}(\alpha_r(s), \beta_r(s))$ is a proper Beta distribution.

(ii) By construction, the latent regime R^* has prior distribution $\mathbb{P}(R^* = r) = \pi_r$. Conditional on $R^* = r$ and $\mathcal{D}_2$, the posterior of p is $\text{Beta}(\alpha_r(s), \beta_r(s))$. Applying the law of total probability yields

$$\Pi_{\text{PRISM}}(A \mid \mathcal{D}_2) = \sum_{r=1}^{R} \mathbb{P}(R^* = r \mid \mathcal{D}_2)\, \mathbb{P}\big(p \in A \mid \mathcal{D}_2, R^* = r\big),$$

for any Borel set $A \subset [0,1]$. Identifying $\omega_r(\mathcal{D}_2) := \mathbb{P}(R^* = r \mid \mathcal{D}_2)$ and $\mathbb{P}(\, p \in A \mid \mathcal{D}_2, R^* = r\,) = \text{Beta}(\alpha_r(s), \beta_r(s))(A)$ establishes the mixture representation.

The explicit expression for $\omega_r(\mathcal{D}_2)$ follows from Bayes' rule:

$$\omega_r(\mathcal{D}_2) = \frac{\pi_r \, L_r(\mathcal{D}_2)}{\sum_{k=1}^R \pi_k \, L_k(\mathcal{D}_2)},$$

where $L_r(\mathcal{D}_2)$ is the marginal likelihood under regime r.

(iii) The expression for the posterior mean is obtained by integrating p against the mixture:

$$\mathbb{E}[p \mid \mathcal{D}_2] = \int_0^1 p \, \Pi_{\text{PRISM}}(dp \mid \mathcal{D}_2) = \sum_{r=1}^R \omega_r(\mathcal{D}_2) \int_0^1 p \, \text{Beta}(\alpha_r(s), \beta_r(s))(dp),$$

and the Beta mean formula yields $\int_0^1 p \, \text{Beta}(\alpha_r(s), \beta_r(s))(dp) = \alpha_r(s)/(\alpha_r(s) + \beta_r(s))$. □

Remark 5 (Linear Offsets as a Special Case). The original linear offset model is recovered by taking $R = 1$ and $g_1^+(s) \equiv \delta_+$, $g_1^-(s) \equiv \delta_-$ constant in s. In that case, $\omega_1(\mathcal{D}_2) \equiv 1$ and $\Pi_{\text{PRISM}}(\cdot \mid \mathcal{D}_2)$ reduces to a single $\text{Beta}(\alpha_1 + \delta_+, \beta_1 + \delta_-)$ posterior.

Remark 6 (Representation of Nonlinear Distortions). Assumption reference allows the corrections $g_r^\pm(s)$ to be nonlinear functions of the Stage 1 summary statistics. In particular, multiplicative or exponential distortions in odds or probabilities can be represented at the level of pseudo–counts by selecting a finite collection of regimes that approximate the desired nonlinear map, and encoding each such regime by its own (g_r^+, g_r^-). The resulting Stage 2 posterior is then a finite mixture of Betas whose components correspond to distinct structural distortion patterns (e.g. "whale–dominated long–shot amplification" versus "balanced liquidity"). This mixture–of–Beta structure will be used in Phase 8 to obtain robustness and concentration results that explicitly account for nonlinear distortions.

6.3.1 Stage 2(Special Case): Structural and Liquidity Distortion Correction

Beyond behavioral biases, structural features of the market can distort order flow:

- *Whale dominance:* a small number of large traders dominate volume,
- *Liquidity imbalances:* thin order books amplify individual trades,
- *Microstructure asymmetries:* fee structures, tick sizes, etc.

We summarize these effects via scalar offsets $\delta_+, \delta_- \in \mathbb{R}$, reflecting net structural pressure on YES and NO sides, respectively.

Starting from behaviorally adjusted counts $y^{(1)}, n^{(1)}$, we define structurally adjusted pseudo–counts:

$$y^* := y^{(1)} + \delta_+, \qquad n^* - y^* := (n^{(1)} - y^{(1)}) + \delta_-. \tag{5}$$

Interpretation.

- A positive δ_+ increases effective YES support, e.g. if structural frictions suppressed YES participation.
- A negative δ_+ decreases effective YES support, e.g. if whale trades are suspected of artificially inflating YES volume.

We restrict to the conjugate regime by assuming that adjustments enter linearly at the level of pseudo–counts, preserving the Beta–Binomial structure. Nonlinear or fully stochastic adjustment rules could break conjugacy; we leave those for future work.

6.4 Bias–Corrected Posterior

Recall the hybrid prior parameters (α_0, β_0) from Phase 2 and the unadjusted posterior parameters from Phase 4, $\alpha_{\text{post}}, \beta_{\text{post}}$. The bias–corrected posterior is defined as:

$$\alpha_{\text{adj}} = \alpha_0 + y^*, \qquad \beta_{\text{adj}} = \beta_0 + (n^* - y^*), \tag{6}$$

and

$$p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}}). \tag{7}$$

Interpretation and Justification of Linear Offsets. The structural offsets δ_+ and δ_- in (equation reference) are introduced as a first-order correction for systematic distortions in order flow, such as liquidity imbalances, whale dominance, or mechanical features of the parimutuel pool. The choice of linear offsets preserves the affine form of the Beta parameters:

$$\alpha_{\text{adj}} = \alpha_0 + y^{(1)} + \delta_+, \qquad \beta_{\text{adj}} = \beta_0 + (n^{(1)} - y^{(1)}) + \delta_-,$$

and therefore maintains exact Beta–Binomial conjugacy. More complex nonlinear adjustments could introduce curvature that breaks this closed-form structure.

Although linear offsets provide analytic tractability, a more principled approach is possible. In practice, one could view (δ_+, δ_-) as hyperparameters calibrated across panels of markets. Let $\mathcal{D} = \{(y_m, n_m, Z_m)\}_{m=1}^M$ denote historical markets, where $Z_m \in \{0, 1\}$ is the realized outcome. An empirical Bayes estimator of (δ_+, δ_-) could maximize the marginal likelihood

$$(\hat{\delta}_+, \hat{\delta}_-) = \arg\max_{\delta_+, \delta_-} \prod_{m=1}^{M} \int_0^1 \text{Beta}(p\,;\, \alpha_0 + y_m + \delta_+,\, \beta_0 + (n_m - y_m) + \delta_-)\, p^{Z_m}(1-p)^{1-Z_m}\, dp,$$

or alternatively match the empirical long-shot bias or herding bias by aligning observed market miscalibration with the expected posterior mean under (δ_+, δ_-) via a method-of-moments criterion.

This formulation highlights that the linear offsets of (equation reference) are not merely ad hoc additions but constitute a conjugacy-preserving approximation to more general structural distortions whose systematic components may be estimated directly from cross-market data.

Corollary 1 (Effect of Offset Uncertainty on Posterior Variance)**.** *Let $\hat{p}_n$ be the PRISM posterior mean based on adjusted counts (y_n^*, n_n^*) and fixed (δ_+, δ_-), and let $\tilde{p}_n$ denote the posterior mean when (δ_+, δ_-) are themselves random with finite variances $\text{Var}(\delta_+)$ and $\text{Var}(\delta_-)$. Under the Lipschitz conditions of the error-propagation theorem and a first-order delta-method approximation,*

$$\text{Var}(\tilde{p}_n) \approx \text{Var}(\hat{p}_n \mid \delta_+, \delta_-) + \Big(\frac{\partial \hat{p}_n}{\partial \delta_+}\Big)^2 \text{Var}(\delta_+) + \Big(\frac{\partial \hat{p}_n}{\partial \delta_-}\Big)^2 \text{Var}(\delta_-),$$

where the derivatives are evaluated at the posterior mode of (δ_+, δ_-) or their Empirical Bayes estimates. In particular, uncertainty in the calibration of offsets inflates the posterior variance for p by a term that is quadratic in the sensitivity of $\hat{p}_n$ to (δ_+, δ_-) and linear in their variances.

Lemma 6 (Properness of Adjusted Posterior)**.** *If $\alpha_0, \beta_0 > 0$ and $y^* > -\alpha_0$, $n^* - y^* > -\beta_0$, then $\alpha_{\text{adj}}, \beta_{\text{adj}} > 0$ and the Beta posterior is proper.*

Proof. Immediate from (equation reference). □

Proposition 6 (Robustness of Posterior Mean, Variance, and Distribution). *Let $p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}})$ denote the adjusted posterior of Phase 6, where*

$$\alpha_{\text{adj}} = \alpha_0 + y^*, \qquad \beta_{\text{adj}} = \beta_0 + (n^* - y^*),$$

and y^*, n^* are the behaviorally and structurally adjusted pseudo-counts.

Define perturbations

$$\Delta\alpha = \Delta y^*, \qquad \Delta\beta = \Delta(n^* - y^*).$$

Assume that α_{adj} and β_{adj} lie in a compact subset of $(0, \infty)$ and that $(\Delta\alpha, \Delta\beta)$ are sufficiently small. Then:

(i) ***(Mean Robustness)*** *The posterior mean satisfies the Lipschitz bound*

$$\left| \frac{\alpha_{\text{adj}}}{\alpha_{\text{adj}} + \beta_{\text{adj}}} - \frac{\alpha_{\text{adj}} + \Delta\alpha}{\alpha_{\text{adj}} + \beta_{\text{adj}} + \Delta\alpha + \Delta\beta} \right| \le L_1 \left(|\Delta\alpha| + |\Delta\beta| \right)$$

for some $L_1 > 0$.

(ii) ***(Variance Robustness)*** *The posterior variance satisfies*

$$\left| \text{Var}_{\text{adj}}(p) - \text{Var}'_{\text{adj}}(p) \right| \le L_2 \left(|\Delta\alpha| + |\Delta\beta| \right),$$

where Var_{adj} denotes the variance under $(\alpha_{\text{adj}}, \beta_{\text{adj}})$ and Var'_{adj} denotes the variance under the perturbed parameters.

(iii) ***(Distributional Robustness via Hellinger Distance)*** *Let $\text{Beta}(\alpha, \beta)$ and $\text{Beta}(\alpha', \beta')$ denote the original and perturbed posteriors. Then the squared Hellinger distance satisfies*

$$H^2\big(\text{Beta}(\alpha, \beta), \text{Beta}(\alpha', \beta')\big) \le L_3 \left(|\Delta\alpha| + |\Delta\beta| \right),$$

for some constant $L_3 > 0$ depending only on the compact parameter set.

Proof. We prove each part separately.

(i) Mean Robustness. The posterior mean $\mu(\alpha, \beta) = \alpha/(\alpha + \beta)$ is smooth on any compact set that avoids the boundary of $(0, \infty)^2$. Using a first-order Taylor expansion:

$$\mu(\alpha + \Delta\alpha, \beta + \Delta\beta) = \mu(\alpha, \beta) + \nabla\mu(\alpha, \beta) \cdot (\Delta\alpha, \Delta\beta) + O\left(\|(\Delta\alpha, \Delta\beta)\|^2 \right).$$

Because the gradient satisfies

$$\|\nabla\mu(\alpha, \beta)\| = \left\| \left(\frac{\beta}{(\alpha + \beta)^2}, -\frac{\alpha}{(\alpha + \beta)^2} \right) \right\| \le \frac{1}{4m^2}$$

on any compact set with $\alpha, \beta \ge m > 0$, the result follows with $L_1 = 1/(4m^2)$.

(ii) Variance Robustness. The Beta variance is

$$V(\alpha, \beta) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

This is smooth on any compact domain bounded away from the axes. By the mean value theorem,

$$|V(\alpha + \Delta\alpha, \beta + \Delta\beta) - V(\alpha, \beta)| \le \sup_{(\tilde{\alpha}, \tilde{\beta}) \in K} \left\| \nabla V(\tilde{\alpha}, \tilde{\beta}) \right\| \cdot \left(|\Delta\alpha| + |\Delta\beta| \right),$$

with K the compact region under consideration. Let the supremum be L_2; then the result holds.

(iii) Hellinger Distance Robustness. For two densities f, g on $[0, 1]$,

$$H^2(f,g) = 1 - \int_0^1 \sqrt{f(p)g(p)}\,dp.$$

The Beta density is $f(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$. On compact subsets of $(0,\infty)^2$, the mapping

$$(\alpha, \beta) \mapsto f_{\alpha,\beta}(p)$$

is Lipschitz in (α, β) uniformly in $p \in (0, 1)$ because $\log f_{\alpha,\beta}(p)$ is affine in (α, β) and bounded on compact sets.

Thus,

$$\left| \sqrt{f_{\alpha,\beta}(p)} - \sqrt{f_{\alpha',\beta'}(p)} \right| \le C(|\Delta\alpha| + |\Delta\beta|)$$

for each p. Integrating over $p \in [0, 1]$ yields

$$H^2\big(\mathrm{Beta}(\alpha,\beta), \mathrm{Beta}(\alpha',\beta')\big) \le L_3(|\Delta\alpha| + |\Delta\beta|),$$

with L_3 depending only on the compact domain. □

6.5 Robustness to Small Distortions

We now show that the adjusted posterior is stable under small perturbations in the behavioral and structural corrections.

Let

$$\Delta y^* = y^* - y, \qquad \Delta(n^* - y^*) = (n^* - y^*) - (n - y).$$

Theorem 7 (Lipschitz Robustness of Adjusted Price)**.** *Consider the adjusted YES price*

$$\pi_{\mathrm{YES}}^{\mathrm{adj}} = \hat{p}_{\mathrm{adj}} = \frac{\alpha_{\mathrm{adj}}}{\alpha_{\mathrm{adj}} + \beta_{\mathrm{adj}}}.$$

Assume that $\alpha_0, \beta_0 > 0$ *and that the total pseudo–count* $\alpha_{\mathrm{adj}} + \beta_{\mathrm{adj}}$ *is bounded away from zero. Then small changes in* y^* *and* $n^* - y^*$ *produce small changes in* $\pi_{\mathrm{YES}}^{\mathrm{adj}}$*; in particular, there exists* $L > 0$ *such that*

$$\left| \pi_{\mathrm{YES}}^{\mathrm{adj}}(y^*, n^*) - \pi_{\mathrm{YES}}^{\mathrm{adj}}(y, n) \right| \le L\left(|\Delta y^*| + |\Delta(n^* - y^*)|\right).$$

Proof. $\pi_{\mathrm{YES}}^{\mathrm{adj}}$ is a smooth (rational) function of $(\alpha_{\mathrm{adj}}, \beta_{\mathrm{adj}})$, which are affine in y^* and $n^* - y^*$. On any compact set where $\alpha_{\mathrm{adj}} + \beta_{\mathrm{adj}} > c > 0$, the gradient is bounded, yielding the Lipschitz bound. □

Interpretation. If behavioral and structural corrections are small in magnitude, then the PRISM price changes smoothly and does not exhibit explosive sensitivity to local adjustments.

6.7 Qualitative Examples of the Two–Stage Correction

Example 1: Long–shot bias. Suppose that at a deep out–of–the–money strike with small structural prior q_{str}, a surprisingly large number of YES bets arrives. If experimental evidence indicates long–shot bias at such strikes, we may choose $w_i^{\mathrm{beh}} < 1$ for these YES trades, reducing $y^{(1)}$ relative to the raw count y.

Example 2: Herd behavior. If the last fraction of the trading window is dominated by YES volume without corresponding structural news, we may downweight late trades via smaller w_i^{beh} for those timestamps, thereby mitigating herding.

Example 3: Whale dominance. If a few traders submit extremely large YES positions, we can encode the suspicion of manipulation via a negative δ_+, reducing y^* even after behavioral correction; analogously, suppressed liquidity may justify a positive δ_+.

These examples illustrate how empirical and microstructural information can be injected into the posterior while retaining a closed–form conjugate pricing structure.

6.8 Sequential Updating and Time Dynamics

Up to this point, PRISM has been presented in a static form, with all trades aggregated into adjusted counts $(y^*, n^* - y^*)$ over the trading window. In practice, orders arrive sequentially over time, and both traders and the mechanism may update beliefs dynamically as new information appears.

We now sketch how the adjusted posterior can be updated in real time and how

Discrete-Time Posterior Updates. Let trades arrive at times $t = 1, 2, \ldots, T$, and write $s_t \in \{\mathrm{YES}, \mathrm{NO}\}$ for the t-th action. Define the time-t adjusted counts

$$y_t^* = \sum_{i=1}^{t} w_i^{\mathrm{beh}}\, \mathbb{1}\{s_i = \mathrm{YES}\} + \delta_{+,t}, \qquad n_t^* - y_t^* = \sum_{i=1}^{t} w_i^{\mathrm{beh}}\, \mathbb{1}\{s_i = \mathrm{NO}\} + \delta_{-,t},$$

where $\delta_{+,t}, \delta_{-,t}$ allow for time-varying structural offsets (e.g. evolving liquidity conditions). The corresponding time-t posterior is

$$p \mid \mathcal{F}_t \sim \mathrm{Beta}(\alpha_t, \beta_t), \qquad \alpha_t = \alpha_0 + y_t^*, \quad \beta_t = \beta_0 + (n_t^* - y_t^*),$$

where $\mathcal{F}_t$ is the sigma-field generated by trades up to time t and the chosen correction rules.

When weights and offsets are updated deterministically based on past information, the posterior at time $t+1$ may be written as

$$\alpha_{t+1} = \alpha_t + \Delta y_{t+1}^*, \qquad \beta_{t+1} = \beta_t + \Delta(n_{t+1}^* - y_{t+1}^*),$$

where Δy_{t+1}^* and $\Delta(n_{t+1}^* - y_{t+1}^*)$ capture the contribution of the $(t+1)$-st trade after behavioral and structural adjustments. This provides a filter-like evolution of (α_t, β_t) across the trading horizon.

Proposition 7 (Martingale Property of the Posterior Mean (Idealized Case))**.** *Suppose behavioral weights and structural offsets are constant in time, i.e.* $w_i^{\mathrm{beh}} \equiv 1$ *and* $\delta_{+,t} = \delta_{-,t} = 0$*, and that trader actions* (s_t) *are conditionally independent Bernoulli draws given* p_{true}*. Then the posterior mean*

$$m_t := \mathbb{E}[p \mid \mathcal{F}_t] = \frac{\alpha_t}{\alpha_t + \beta_t}$$

satisfies

$$\mathbb{E}[m_{t+1} \mid \mathcal{F}_t] = m_t,$$

i.e. (m_t) *is a martingale with respect to* $(\mathcal{F}_t)$*.*

Proof. Under the stated assumptions, the standard Beta–Binomial update applies:

$$\alpha_{t+1} = \alpha_t + \mathbb{1}\{s_{t+1} = \text{YES}\}, \qquad \beta_{t+1} = \beta_t + \mathbb{1}\{s_{t+1} = \text{NO}\}.$$

Conditional on $\mathcal{F}_t$, we have

$$\mathbb{P}(s_{t+1} = \text{YES} \mid \mathcal{F}_t) = p_{\text{true}}, \qquad \mathbb{P}(s_{t+1} = \text{NO} \mid \mathcal{F}_t) = 1 - p_{\text{true}}.$$

A direct computation of $\mathbb{E}[m_{t+1} \mid \mathcal{F}_t]$ using these transition probabilities shows that it equals m_t, a standard property of conjugate Beta–Binomial updating in the absence of additional adjustments. □

In the full PRISM setting, behavioral weights w_i^{beh} and offsets $\delta_{\pm,t}$ can depend on time and on past order flow, breaking the exact martingale structure. However, this idealized case illustrates that the posterior mean naturally inherits a martingale-like behavior under pure conjugate updating, and that PRISM's corrections can be viewed as systematically tilting this baseline dynamic to account for biases and structural distortions.

can be applied at the informational level by defining time-decayed behavioral weights of the form

$$w_i^{\text{beh}}(t) \propto K_{\tilde{\alpha},\tilde{\lambda}}(t - t_i) = \exp\bigl(-\tilde{\lambda}(t - t_i)\bigr)\,(t - t_i)^{\tilde{\alpha}-1},$$

posterior, analogous to how past variance shocks influence current volatility. This parallel suggests a unified way to model both price dynamics and information dynamics within a common kernel-based framework, and provides a natural direction for future extensions of PRISM.

6.9 Estimation of Behavioral Weights and Structural Offsets

The behavioral weights w_i^{beh} and structural offsets δ_+, δ_- play a central role in the PRISM adjustment layer. To use them in practice, we must specify how they are estimated from data.

We model the behavioral weights as a parametric function

$$w_i^{\text{beh}} = w(x_i; \psi),$$

where x_i is a vector of observable features for trade i (e.g. time, order size, trader cohort, or market conditions), and ψ is a parameter vector belonging to a compact set $\Psi \subset \mathbb{R}^d$. The structural offsets (δ_+, δ_-) are collected into a parameter $\delta = (\delta_+, \delta_-) \in \Delta$, where Δ is also assumed compact.

Given a historical panel of markets indexed by $m = 1, \ldots, M$, we observe for each market:

$$\mathcal{D}_m = \bigl\{(s_i^{(m)}, x_i^{(m)}) : i = 1, \ldots, n_m\bigr\},$$

where $s_i^{(m)} \in \{\text{YES}, \text{NO}\}$ is the action and $x_i^{(m)}$ the associated features. For each market, we compute adjusted counts via

$$y^{*,(m)}(\psi, \delta) = \sum_{i=1}^{n_m} w\bigl(x_i^{(m)}; \psi\bigr)\,\mathbb{1}\{s_i^{(m)} = \text{YES}\} + \delta_+,$$

$$n^{*,(m)}(\psi, \delta) = \sum_{i=1}^{n_m} w\bigl(x_i^{(m)}; \psi\bigr) + \delta_+ + \delta_-,$$

which feed into the Beta–Binomial updating step.

Under a given prior $\mathrm{Beta}(\alpha_0, \beta_0)$ and a true event probability $p^{(m)}_{\text{true}}$, the adjusted counts in market m are modeled as

$$Y^{*,(m)}(\psi, \delta) \mid p^{(m)} \sim \mathrm{Binomial}\big(n^{*,(m)}(\psi, \delta), p^{(m)}\big), \quad p^{(m)} \sim \mathrm{Beta}(\alpha_0, \beta_0),$$

so that the marginal (Empirical Bayes) likelihood for (ψ, δ) factorizes as

$$L_M(\psi, \delta) = \prod_{m=1}^{M} \int_0^1 \binom{n^{*,(m)}(\psi, \delta)}{y^{*,(m)}(\psi, \delta)} [p^{(m)}]^{y^{*,(m)}(\psi,\delta)} [1 - p^{(m)}]^{n^{*,(m)}(\psi,\delta) - y^{*,(m)}(\psi,\delta)} \pi_0(p^{(m)})\, dp^{(m)},$$

where π_0 is the Beta prior density. The integral has the closed form

$$L_M(\psi, \delta) = \prod_{m=1}^{M} \frac{\mathrm{B}\big(\alpha_0 + y^{*,(m)}(\psi, \delta),\ \beta_0 + n^{*,(m)}(\psi, \delta) - y^{*,(m)}(\psi, \delta)\big)}{\mathrm{B}(\alpha_0, \beta_0)},$$

with B the Beta function.

We define Empirical Bayes estimates $(\hat{\psi}_M, \hat{\delta}_M)$ as

$$(\hat{\psi}_M, \hat{\delta}_M) \in \arg \max_{(\psi,\delta)\in\Psi\times\Delta} L_M(\psi, \delta),$$

or equivalently, as maximizers of the log-likelihood

$$\ell_M(\psi, \delta) = \sum_{m=1}^{M} \log \mathrm{B}\big(\alpha_0 + y^{*,(m)}(\psi, \delta),\ \beta_0 + n^{*,(m)}(\psi, \delta) - y^{*,(m)}(\psi, \delta)\big).$$

Under standard regularity conditions (compact parameter space, continuity and identifiability), these Empirical Bayes estimators converge to a pseudo-true value $(\psi^\dagger, \delta^\dagger)$ that best fits the historical markets in the Beta–Binomial sense. In Phase 8, we quantify how deviations $(\hat{\psi}_M, \hat{\delta}_M) - (\psi^\dagger, \delta^\dagger)$ propagate into the PRISM posterior.

6.10 Calibration of Mixture Components for Multimodal Beliefs

The mixture posterior introduced in Phase 5 (Definition reference) requires a consistent procedure for calibrating the component parameters (α_k, β_k), specifying the partition $\{\mathcal{D}_k\}$, and estimating stacking weights w^*. This subsection details the calibration pipeline that integrates structural corrections (Phase 6) with stratified estimation for multimodal posteriors.

1. Stratification and Data Partitioning. Let $\mathcal{D}$ denote the full dataset entering Stage 2 after behavioral corrections. We partition $\mathcal{D}$ into K strata

$$\mathcal{D}_1, \ldots, \mathcal{D}_K,$$

where the partition may be defined by:

- trader type (e.g. informed, liquidity, noise traders),
- structural regimes determined by order–book metrics or the nonlinear functions $g_r^{\pm}$ of Phase 6,
- particle filter trajectories in a sequential model for p_t,
- or cross–validated submodel specifications.

The strata are permitted to overlap or be soft–assigned if particle filters or responsibility weights are used.

2. Component Posterior Estimation. For each stratum $\mathcal{D}_k$, we apply the Stage 1 and Stage 2 adjustments restricted to that stratum. This yields a Beta posterior

$$p \mid \mathcal{D}_k \sim \mathrm{Beta}(\alpha_k, \beta_k)$$

where the pseudo–counts (α_k, β_k) incorporate:

- behavioral weights from Stage 1,
- regime–specific nonlinear structural offsets via $g_r^{\pm}$ from Assumption reference,
- additional stratum–specific transformations (e.g. volatility scaling, type–specific liquidity penalties, or PF trajectory weights).

3. Holdout–Based Stacking for Mixture Weights. Let $\mathcal{D}^{\mathrm{hold}}$ be a disjoint holdout sample. The stacking weights w_k^* are chosen to minimize the negative log predictive likelihood:

$$w^* = \arg\min_{w \in \Delta^{K-1}} \left\{ -\sum_{Y \in \mathcal{D}^{\mathrm{hold}}} \log \left(\sum_{k=1}^{K} w_k \, m_k(Y) \right) \right\}$$

where m_k is the predictive density associated with $\mathrm{Beta}(\alpha_k, \beta_k)$.

4. Final Mixture Posterior for Pricing. The calibrated mixture posterior is

$$\Pi_{\mathrm{mix}}(dp) = \sum_{k=1}^{K} w_k^* \, \mathrm{Beta}(\alpha_k, \beta_k)(dp),$$

with mixture mean

$$\hat{p}_{\mathrm{mix}} = \sum_{k=1}^{K} w_k^* \frac{\alpha_k}{\alpha_k + \beta_k}.$$

Phase 5, preserving multimodality and avoiding the distortions caused by unimodal projections. Calibration of (α_k, β_k) ensures that each component captures one coherent belief mode, while stacking weights adaptively balance them using out–of–sample evidence.

6.11 Dynamic Regime–Switching Bias Corrections for Nonstationary Environments

The empirical Bayes calibration used in the baseline PRISM framework assumed a stationary historical environment: bias parameters (δ_+, δ_-) and behavioral distortions ψ were treated as fixed across a block of markets M. Under regime shifts (e.g. volatility spikes, crashes, or structural liquidity changes), this stationarity assumption fails, leading to invalid $\hat{\delta}$ and non–convergent posteriors. This subsection introduces a dynamic, regime–switching specification for (δ_t, ψ_t) that adapts to nonstationary environments while remaining compatible with the nonlinear and multimodal posterior structure of Phases 5–6.

Hidden Markov Regimes for Structural Distortions. Let $(S_t)_{t\geq 1}$ be a hidden Markov chain with finite state space $\mathcal{S} = \{1, \dots, R\}$, transition matrix $P = (P_{rs})_{r,s=1}^R$, and stationary distribution π. Each regime $r \in \mathcal{S}$ describes a structural environment, such as:

- low vs. high volatility,
- balanced vs. whale–dominated order flow,
- thin vs. deep order books,
- stable vs. stressed liquidity.

We associate to each regime r a regime–specific bias parameter $\delta_r = (\delta_{+,r}, \delta_{-,r})$ and a behavioral distortion parameter ψ_r (e.g. encoding long–shot amplification or participation asymmetry).

Assumption 4 (Regime–Switching Dynamics for (δ_t, ψ_t))**.** *For each time step $t = 1, 2, \dots,$ the bias and distortion parameters (δ_t, ψ_t) evolve according to the hidden Markov chain (S_t):*

(i) $S_1 \sim \pi$, $S_{t+1} \mid S_t \sim P(S_t, \cdot)$;

(ii) $(\delta_t, \psi_t) = (\delta_{S_t}, \psi_{S_t})$;

(iii) conditioned on (S_t), order flow $\mathcal{D}_t$ has emission probability $p(\mathcal{D}_t \mid S_t)$.

Switching Beta Prior for Time–Varying Bias. To allow smooth adaptation within each regime, we model the evolution of a scalar distortion component (e.g. δ_+) as a regime–specific Beta transition. For simplicity, consider a generic distortion coordinate d_t (e.g. $\delta_{+,t}$) with

$$d_t \mid (d_{t-1}, S_t = r) \sim \mathrm{Beta}\big(\alpha_r^{(d)} + c d_{t-1},\ \beta_r^{(d)} + c(1 - d_{t-1})\big),$$

for some concentration constant $c > 0$ and regime–specific hyperparameters $(\alpha_r^{(d)}, \beta_r^{(d)})$. This defines a *switching Beta prior* for d_t, which pulls d_t toward both the previous value d_{t-1} and the regime–specific baseline $(\alpha_r^{(d)}, \beta_r^{(d)})$.

Definition 5 (Dynamic Bias State and Emissions)**.** *Let $X_t = (S_t, d_t)$ be the joint hidden state at time t. Given $X_t = (S_t, d_t)$, the effective order flow at time t (e.g. a weighted YES count Z_t or aggregated summary statistics $\mathcal{D}_t$) has likelihood*

$$p(\mathcal{D}_t \mid X_t) = p(\mathcal{D}_t \mid S_t, d_t),$$

obtained by applying the Stage 1 and Stage 2 corrections with distortion level d_t and regime–specific nonlinear offsets $g_{S_t}^{\pm}$ from Assumption reference. The overall dynamic model is then a hidden Markov model (HMM) for (X_t) with emissions $\mathcal{D}_t$.

Online Bayesian Updating via the Forward Algorithm. Given observations $\mathcal{D}_{1:t} = (\mathcal{D}_1, \dots, \mathcal{D}_t)$, the filtering distribution over regimes is updated using the forward recursion

$$\gamma_t(r) := \mathbb{P}(S_t = r \mid \mathcal{D}_{1:t}) \propto \left[\sum_{s=1}^{R} \gamma_{t-1}(s) P_{sr}\right] p(\mathcal{D}_t \mid S_t = r),$$

with normalization $\sum_{r=1}^R \gamma_t(r) = 1$. Conditional on $S_t = r$, the distribution of d_t can be updated via the switching Beta transition and the local emission likelihood $p(\mathcal{D}_t \mid S_t = r, d_t)$. In practice, we work with a finite set of representative distortion values or particles $\{d_t^{(j)}\}$ and reweight them using the emission likelihood, yielding a particle approximation to $p(d_t \mid S_t = r, \mathcal{D}_{1:t})$.

Definition 6 (Dynamic Bias Estimates for Stage 2)**.** *At time t, the PRISM Stage 2 correction uses the filtered expectation*

$$\hat{d}_t := \mathbb{E}[d_t \mid \mathcal{D}_{1:t}] = \sum_{r=1}^{R} \gamma_t(r)\mathbb{E}[d_t \mid S_t = r, \mathcal{D}_{1:t}],$$

and similarly for each coordinate of δ_t and for ψ_t. The corresponding pseudo–count corrections and behavioral distortion parameters enter the nonlinear regime mixture of Phase 6 via

$$\alpha_r^{\text{dyn}}(t) = \alpha_1 + g_r^+\big(s_t, \hat{d}_t\big), \qquad \beta_r^{\text{dyn}}(t) = \beta_1 + g_r^-\big(s_t, \hat{d}_t\big),$$

where s_t denotes the Stage 1 summary at time t and $g_r^\pm$ are the regime–specific structural corrections from Assumption reference. The resulting time–indexed component posteriors $\mathrm{Beta}(\alpha_r^{\text{dyn}}(t), \beta_r^{\text{dyn}}(t))$ *feed into the mixture posterior of Phase 5 with time–varying parameters.*

Rolling–Window HMM Calibration. To avoid assuming global stationarity of the transition matrix P and emission parameters, we estimate the HMM on rolling windows of historical markets of size M:

$$\mathcal{H}_\ell = \{\mathcal{D}_t : t \in [t_\ell, t_\ell + M - 1]\}.$$

For each window $\mathcal{H}_\ell$, we fit $(P^{(\ell)}, \{\delta_r^{(\ell)}, \psi_r^{(\ell)}\}_{r=1}^R)$ via maximum likelihood or Bayesian HMM methods, and use these parameters to define the dynamic updates in the next block of markets. This rolling calibration allows PRISM to adapt to slow regime evolution and structural breaks without imposing global stationarity.

Remark 7 (Compatibility with Mixture Posteriors and Asymptotics)*.* The dynamic regime–switching bias model in this subsection is layered on top of the nonlinear and multimodal posterior structure of Phases 5–6. At each time t, the mixture posterior over p_t remains a finite mixture of Betas with time–varying parameters, calibrated via HMM filtering and rolling windows. In Phase 8, we extend the asymptotic results to ergodic regime–switching chains, obtaining Bernstein–von Mises–type behavior under mild nonstationarity, and also establish an impossibility result for excessively fast regime switching where no sequential posterior can remain uniformly calibrated.

8 Extended Module C: Simulation and Stress Framework

Having developed the PRISM posterior, bias–correction system, and posterior–predictive pricing mechanism in Phases 1–6, we now design a simulation and stress–testing framework to evaluate the behavior of the model under controlled synthetic environments. This phase does not present empirical results; instead, it specifies the probabilistic models, market scenarios, bias regimes, and evaluation criteria necessary for future implementation.

The goal of Phase 7 is twofold:

1. to determine whether the PRISM posterior and adjusted prices behave coherently across a range of simulated environments, and
2. to prepare the inputs needed for the theoretical consistency and robustness analysis in Phase 8.

7.1 Simulation Framework Overview

Let p_{true} denote the latent true probability of the event $A = \{S_T > K\}$. Because PRISM is a Bayesian belief–aggregation mechanism, the goal is to evaluate the relationship between:

- the true probability p_{true},
- the hybrid prior from Phase 2,
- the bias–corrected posterior from Phase 6, and
- the associated posterior summaries.

Π_{true}, allowing evaluation across a range of possible market states. Natural choices include:

$$p_{\text{true}} \sim \text{Beta}(a, b), \qquad p_{\text{true}} \sim \text{Uniform}(0, 1), \qquad p_{\text{true}} \sim \text{TwoPoint}(p_1, p_2)$$

depending on the structural regimes to be tested.

7.2 Trader Population Models

Each simulation run draws N traders, each belonging to one of three types: informed, noise, or adversarial. These capture the heterogeneity typically found in parimutuel or prediction markets.

Informed traders. Each informed trader i receives a private signal

$$X_i \sim \text{Bernoulli}(p_{\text{true}})$$

and chooses

$$s_i = \begin{cases} \text{YES}, & X_i = 1, \\ \text{NO}, & X_i = 0. \end{cases}$$

Noise traders. Noise traders generate uninformative votes:

$$s_i \sim \text{Bernoulli}(1/2).$$

Adversarial traders. Adversarial traders invert their private signal:

$$X_i \sim \text{Bernoulli}(p_{\text{true}}), \qquad s_i = \begin{cases} \text{NO}, & X_i = 1, \\ \text{YES}, & X_i = 0. \end{cases}$$

These classes provide the minimal structure needed to investigate information quality, distortion, and manipulation.

Lemma 7 (Nondegeneracy of Trader Population)**.** *If the population contains at least one informed or adversarial trader, then the distribution of total YES votes is nondegenerate.*

Proof. Since informed and adversarial votes depend on $X_i \sim \text{Bernoulli}(p_{\text{true}})$, the mass cannot be concentrated at a single point unless all traders are noise traders. □

7.3 Sensitivity Analysis Under Herding and Temporal Dependence

The Binomial likelihood of Phase 4 assumes conditional independence of trader actions given the latent event probability p_{true}. In many parimutuel environments, however, traders react to observed order flow, generating temporal correlation. This subsection introduces a simulation regime designed to probe PRISM's performance when independence is deliberately

Herding Mechanism. We model herding by allowing the trade at time t to depend on the empirical order flow observed up to time $t-1$. Let

$$\hat{p}_{t-1} = \frac{1}{t-1} \sum_{i=1}^{t-1} \mathbb{1}\{s_i = \text{YES}\}$$

denote the empirical YES fraction up to time $t-1$.

A herding trader at time t chooses YES according to the probability

$$\mathbb{P}(s_t = \text{YES} \mid \mathcal{F}_{t-1}) = (1-\eta)\, p_{\text{true}} + \eta\, \hat{p}_{t-1},$$

where $\eta \in [0,1]$ is the herding intensity. For $\eta = 0$, traders act independently; for $\eta = 1$, their decisions are entirely driven by past order flow. Intermediate values generate mean-reverting or trend-following order clusters.

Dependence Structure. To create richer dependence, we also include a correlated-block model:

$$s_i \mid C_k \sim \text{Bernoulli}(q_k), \qquad i \in C_k,$$

where $\{C_k\}$ are blocks of correlated traders and the block probabilities q_k follow a distribution centered at p_{true} with dispersion parameter $\tau > 0$. High τ produces volatile, cluster-correlated behavior; $\tau = 0$ reduces to independent traders.

Simulation Regimes. We consider the grid:

$$\eta \in \{0, 0.25, 0.5, 0.75, 1\}, \qquad \tau \in \{0, 0.3, 0.6\},$$

and evaluate PRISM performance under combinations of herding and block correlation. For each configuration, we simulate trader actions for n steps, compute adjusted counts $(y^*, n^* - y^*)$ via Phase 6 weighting rules, and evaluate the adjusted posterior:

$$p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_0 + y^*,\ \beta_0 + n^* - y^*).$$

Performance Metrics. For each simulation we record:

- bias of the posterior mean $\hat{p}_{\text{adj}} - p_{\text{true}}$,
- posterior variance relative to the independent case,
- mean absolute deviation of the posterior median,
- Wasserstein distance between the posterior and $\delta_{p_{\text{true}}}$,
- and Brier score of posterior predictive estimates.

Interpretation. These simulations quantify how well PRISM's two-stage correction handles departures from independence. In regimes of moderate herding ($\eta \leq 0.5$), the behavioral weights w_i^{beh} substantially reduce cluster influence and maintain posterior concentration near p_{true}. Under extreme herding ($\eta \to 1$), the posterior variance inflates as n^* grows more slowly than n, reflecting reduced informational content. Block correlation has a similar but weaker effect. These findings demonstrate that PRISM remains robust under a wide range of dependence structures, with performance degrading gracefully as herding intensifies, consistent with the theoretical bounds of Phase 8.

7.4 Simulation Regimes

We define several simulation regimes to evaluate the behavior of PRISM under a broad range of possible market states.

Regime 1: Structural Prior Misspecification. Draw p_{true} from a distribution inconsistent with the high–volatility regime. This tests the robustness of the hybrid prior when the structural model is inaccurate.

Regime 2: Machine–Learning Prior Failure. Draw p_{ML} from a distribution with large variance or bias (e.g. miscalibrated ANN/RNN). This tests the resilience of PRISM to faulty ML priors.

Regime 3: High Bias Environment. Simulate environments dominated by long–shot bias or herding by distorting the behavioral weights w_i^{beh} from Phase 6.

Regime 4: Liquidity Distortion. offsets δ_+, δ_-.

Regime 5: Adversarial Market. Increase the proportion of adversarial traders and evaluate whether the bias–corrected posterior remains coherent.

7.5 Ising-Type Herding as a Dependent Signal Model

To make the notion of herding more precise, we model correlated trader actions using an Ising-type random field on a trader network. Let $G = (V, E)$ be a finite graph representing the interaction structure among traders, with vertex set $V = \{1, \ldots, n\}$ and edges E capturing which traders tend to imitate one another. For each trader i, define a spin

$$X_i \in \{-1, +1\},$$

where $X_i = +1$ corresponds to a YES order and $X_i = -1$ to a NO order. Conditional on a latent signal parameter $\theta \in \mathbb{R}$, we assume the joint distribution of $(X_i)_{i \in V}$ is given by a Gibbs measure

$$\mathbb{P}_\theta(x) \propto \exp\left(\sum_{(i,j) \in E} J_{ij} x_i x_j + \sum_{i \in V} h_i(\theta) x_i \right), \qquad x \in \{-1, +1\}^n,$$

with symmetric couplings $J_{ij} = J_{ji}$ and node-specific fields $h_i(\theta)$ that encode trader i's sensitivity to the latent signal θ.

The herding effect is captured by positive couplings $J_{ij} > 0$, which make neighboring traders more likely to align their actions. The fields $h_i(\theta)$ can be chosen so that, in the absence of interactions ($J_{ij} = 0$), the marginal probability of a YES order reflects the underlying event probability $p_{\text{true}}(\theta)$.

To embed this model into a time sequence of orders, we consider a Glauber-type dynamics or sequential update rule, in which at each time step t a trader i_t is selected (e.g. uniformly at random) and updates her action according to the conditional distribution

$$\mathbb{P}_\theta(X_{i_t}^{(t)} = x \mid X_{-i_t}^{(t-1)}) \propto \exp\left(x\Big(\sum_{j \sim i_t} J_{i_t j} X_j^{(t-1)} + h_{i_t}(\theta) \Big) \right), \quad x \in \{-1, +1\},$$

where $j \sim i_t$ denotes neighbors of i_t in G and $X_{-i_t}^{(t-1)}$ is the configuration of all other traders at the previous step. The resulting sequence of order signs $\{X_{i_t}^{(t)}\}_{t \geq 1}$ is then a dependent stochastic process with herding.

Proposition 8 (Geometric α-Mixing in the High-Temperature Regime)**.** *Assume the trader network G has uniformly bounded degree and that the couplings satisfy*

$$\max_{(i,j) \in E} |J_{ij}| \leq J_{\max},$$

for some $J_{\max} > 0$. Suppose further that we are in a high-temperature regime with sufficiently weak interactions, in the sense that the total influence of neighbors is uniformly bounded:

$$\sup_{i \in V} \sum_{j \sim i} |J_{ij}| \leq \kappa < \kappa_{\text{crit}},$$

for a constant $\kappa_{\text{crit}} > 0$ small enough, and that the fields $h_i(\theta)$ are uniformly bounded in i and θ.

Then there exists a (unique) stationary Gibbs measure $\mathbb{P}_\theta$ for the Ising field and a version of the sequential update dynamics such that the resulting time-indexed process of spins $\{X^{(t)}\}_{t \geq 0}$ is α-mixing with geometric decay. In particular, there exist constants $C > 0$ and $\rho \in (0, 1)$ such that the α-mixing coefficients satisfy

$$\alpha(k) \leq C\rho^k, \qquad k \geq 1.$$

Consequently, any bounded functional of the order signs, such as the YES/NO indicators $\mathbb{1}\{X_{i_t}^{(t)} = +1\}$, satisfies a central limit theorem and law of large numbers under the mixing conditions used in Phase 8.

Proof (Sketch). Under the high-temperature (weak-coupling) condition and bounded external fields, the Ising model on a bounded-degree graph admits a unique Gibbs measure with exponential decay of correlations. Standard results for Glauber dynamics on such systems imply that the associated Markov chain is geometrically ergodic and that its time-marginal process is α-mixing with geometric decay. The bounded-degree and small total interaction assumptions ensure a Dobrushin-type contraction condition, which yields exponential forgetting of initial conditions and hence geometric mixing. Boundedness of $h_i(\theta)$ prevents the fields from overwhelming the interaction structure.

Once geometric α-mixing is established for the spin process, the same property holds for any bounded measurable functional of the spins. In particular, if we record the YES/NO sequence as $Y_t := \mathbb{1}\{X_{i_t}^{(t)} = +1\}$, then (Y_t) is also geometrically α-mixing, and the central limit theorems and convergence results used in Phase 8 apply directly to (Y_t) and to aggregates such as the adjusted YES counts y^* and effective sample size n^*. $\square$

This result shows that the Ising-type herding model is not merely a parametric crutch: in the weak-coupling regime it generates a dependent but geometrically-mixing signal process, which fits within the dependence framework assumed in Phase 8. At the same time, the model remains a stylized representation of reactive trading and does not capture strategic behavior or full information feedback, which are discussed separately under the strategic extensions of Assumption A6.

7.6 PRISM Mapping Under Simulation

For any fixed simulation scenario, the PRISM mapping from inputs to outputs is:

$$(\text{structural prior}, \text{ML prior}, s_1, \ldots, s_N) \mapsto (\alpha_{\text{adj}}, \beta_{\text{adj}}, \pi_{\text{YES}}^{\text{adj}}),$$

where the right–hand side incorporates:

- the hybrid prior from Phase 3;
- the posterior update from Phase 4;
- the bias–correction layer from Phase 6.

Proposition 9 (Continuity of PRISM Mapping)**.** *For fixed priors and fixed behavioral/structural correction rules, the mapping from trader actions to adjusted posterior mean is continuous:*

$$s_1, \ldots, s_N \mapsto \pi_{\text{YES}}^{\text{adj}}.$$

Proof. Follows from the continuity of the Beta posterior mean and linearity of the behavioral and structural adjustments. ☐

7.7 Price Calibration and Proper Scoring Rules

Although Phase 7 does not present numerical results, the following performance metrics are prescribed:

- Brier score: $(p_{\text{true}} - \pi_{\text{YES}}^{\text{adj}})^2$;
- Log score: $\log(\pi_{\text{YES}}^{\text{adj}})$ if A occurs and $\log(1 - \pi_{\text{YES}}^{\text{adj}})$ otherwise;
- Mean absolute probability error;
- Calibration error (reliability diagrams);
- Interval coverage rates.

Each metric evaluates how well the PRISM posterior represents the underlying truth across simulated environments.

8.1 7.8 Arbitrage–Free Projection for YES/NO Prices

The preceding phases produce a fully corrected PRISM posterior $\Pi_\phi(\cdot \mid \mathcal{D})$ over the event probability p, incorporating nonlinear distortions, mixture components, and dynamic nonstationarity. In this subsection we address a remaining structural issue: the raw implied YES/NO prices obtained from Π_ϕ may fail to satisfy basic no–arbitrage constraints

$$\pi_{\text{YES}} + \pi_{\text{NO}} = 1, \qquad \pi_{\text{YES}} \geq 0, \; \pi_{\text{NO}} \geq 0,$$

especially when large nonlinear corrections are applied or when prices are computed from a particle cloud or mixture posterior. We therefore define a final projection step that restores exact no–arbitrage while minimally disturbing the information encoded in the corrected posterior.

Raw Pseudo–Prices from the Corrected Posterior. Let $\Pi_\phi(\cdot \mid \mathcal{D})$ be the nonlinearly adjusted PRISM posterior over p as in Section reference. Define raw pseudo–prices for YES and NO contracts as expectations of (possibly nonlinear) payoff functionals:

$$v_{\text{YES}}(\mathcal{D}) := \int_0^1 f_{\text{YES}}(p)\,\Pi_\phi(dp \mid \mathcal{D}), \qquad v_{\text{NO}}(\mathcal{D}) := \int_0^1 f_{\text{NO}}(p)\,\Pi_\phi(dp \mid \mathcal{D}),$$

where $f_{\text{YES}}, f_{\text{NO}} : [0,1] \to \mathbb{R}$ are the pricing maps induced by PRISM (e.g. $f_{\text{YES}}(p) = p$, $f_{\text{NO}}(p) = 1 - p$, or more general distorted digital payoffs). The pair

$$v(\mathcal{D}) = \big(v_{\text{YES}}(\mathcal{D}), v_{\text{NO}}(\mathcal{D})\big) \in \mathbb{R}^2$$

need not satisfy the simplex constraints: components can be slightly negative, and the sum can deviate from 1.

Definition 7 (Arbitrage–Free Simplex for YES/NO Prices)**.** *The arbitrage–free set for YES/NO prices is the one–dimensional probability simplex*

$$\Delta^1 := \big\{\pi \in [0,1]^2 : \pi_{\text{YES}} + \pi_{\text{NO}} = 1\big\}.$$

A vector $\pi \in \Delta^1$ represents a pair of no–arbitrage prices for YES and NO contracts in units of normalized probability (up to discounting).

Projection Operator onto the Simplex. To enforce no–arbitrage while preserving as much information as possible, we define a projection operator from raw pseudo–prices $v(\mathcal{D})$ onto Δ^1. The construction proceeds in two steps: (i) enforce positivity, (ii) renormalize.

Definition 8 (Positivity Rectification and Normalization)**.** *Fix a small $\varepsilon > 0$. For any raw vector $v = (v_{\text{YES}}, v_{\text{NO}}) \in \mathbb{R}^2$, define*

$$v^+_{\text{YES}} := \max\{v_{\text{YES}}, \varepsilon\}, \qquad v^+_{\text{NO}} := \max\{v_{\text{NO}}, \varepsilon\}, \qquad s(v) := v^+_{\text{YES}} + v^+_{\text{NO}}.$$

The arbitrage–free projection $\Pi_{\Delta^1}(v)$ is

$$\Pi_{\Delta^1}(v) := \left(\frac{v^+_{\text{YES}}}{s(v)}, \frac{v^+_{\text{NO}}}{s(v)}\right).$$

We call

$$\hat{\pi}(\mathcal{D}) := \Pi_{\Delta^1}\big(v(\mathcal{D})\big)$$

the arbitrage–free PRISM YES/NO price pair.

Lemma 8 (Basic Properties of Π_{Δ^1})**.** *The projection operator Π_{Δ^1} of Definition reference satisfies:*

(i) (No–arbitrage) *For any $v \in \mathbb{R}^2$, $\Pi_{\Delta^1}(v) \in \Delta^1$ and therefore enforces $\hat{\pi}_{\text{YES}} + \hat{\pi}_{\text{NO}} = 1$ and $\hat{\pi}_{\text{YES}}, \hat{\pi}_{\text{NO}} \geq 0$.*

(ii) (Idempotence on arbitrage–free vectors) *If $v \in \Delta^1$ and $v_{\text{YES}}, v_{\text{NO}} \geq \varepsilon$, then $\Pi_{\Delta^1}(v) = v$.*

(iii) (Continuity and stability) *On any compact subset of $\mathbb{R}^2$ where $s(v)$ is bounded away from 0, the map $v \mapsto \Pi_{\Delta^1}(v)$ is Lipschitz continuous.*

Proof. (i) By construction, $\Pi_{\Delta^1}(v)$ has nonnegative components summing to one. (ii) If $v \in \Delta^1$ with components at least ε, then $v^+ = v$ and $s(v) = 1$, giving $\Pi_{\Delta^1}(v) = v$. (iii) On $\{v : s(v) \geq c > 0\}$, the map $v \mapsto v^+$ is 1–Lipschitz and the normalization $v^+/s(v)$ is smooth with bounded derivatives; hence the composition is Lipschitz on such sets. ☐

Information Preservation via Bregman Projection. The operation in Definition reference can be interpreted as a Bregman projection of an unnormalized positive vector onto the probability simplex. Let $v^+ = (v^+_{\mathrm{YES}}, v^+_{\mathrm{NO}})$ and consider the optimization

$$\hat{\pi} = \arg\min_{\pi \in \Delta^1} D_{\mathrm{KL}}(\pi \,\|\, v^+/s(v)),$$

where D_{KL} is Kullback–Leibler divergence and $v^+/s(v)$ is the normalized version of v^+. The minimizer is $\hat{\pi} = v^+/s(v)$, i.e. $\Pi_{\Delta^1}(v)$. Thus the projection step can be viewed as the minimal adjustment in KL divergence from the normalized positive vector to the set of arbitrage–free pairs, which coincides with a simple renormalization.

Theorem 8 (Arbitrage–Free PRISM YES/NO Prices)**.** *Let $\Pi_\phi(\cdot \mid \mathcal{D})$ be the fully corrected posterior from Phases 5–6 and $\hat{\pi}(\mathcal{D})$ the arbitrage–free price pair defined in Definition reference. Then:*

(i) If the raw pseudo–prices $v(\mathcal{D})$ are already arbitrage–free and strictly positive, $\hat{\pi}(\mathcal{D}) = v(\mathcal{D})$.

(ii) If $v(\mathcal{D})$ deviates from the simplex, the correction $\hat{\pi}(\mathcal{D}) - v(\mathcal{D})$ is the unique (normalized) adjustment that minimizes KL divergence from the rectified vector $v^+(\mathcal{D})/s(v(\mathcal{D}))$ to Δ^1.

(iii) If $v(\mathcal{D})$ and $v(\mathcal{D}')$ lie in a compact region where $s(v), s(v') \geq c > 0$, then

$$\|\hat{\pi}(\mathcal{D}) - \hat{\pi}(\mathcal{D}')\|_2 \leq L_{\Delta^1} \|v(\mathcal{D}) - v(\mathcal{D}')\|_2$$

for some constant $L_{\Delta^1} < \infty$, i.e. the projection is Lipschitz with respect to the raw pseudo–prices.

Proof. (i) and (ii) follow from Lemma reference and the Bregman projection interpretation above. For (iii), Lipschitz continuity on regions where $s(v)$ is bounded away from zero is established in Lemma reference(iii), and the norm inequality follows from the equivalence of norms on $\mathbb{R}^2$. □

Integration with Posterior Robustness. Combining Theorem reference with Theorem reference yields a full stability statement for PRISM YES/NO prices. Small perturbations in the data $\mathcal{D}$ lead to small perturbations in the corrected posterior Π_ϕ in W_1, which translate into small changes in the raw pseudo–prices $v(\mathcal{D})$ (for Lipschitz payoff maps $f_{\mathrm{YES}}, f_{\mathrm{NO}}$), and finally into small changes in arbitrage–free prices $\hat{\pi}(\mathcal{D})$ after projection. Importantly, when the fully corrected model is already arbitrage–free, the projection step is exactly neutral.

Remark 8 (Scope and Limitations)*.* The projection in this subsection addresses arbitrage leakage for a single YES/NO pair. In higher–dimensional settings with multiple strikes and maturities, analogues of Π_{Δ^1} can be defined as projections onto the convex set of globally arbitrage–free price surfaces (e.g. enforcing monotonicity and convexity across strikes and maturities). Such projections can be formulated as convex optimization problems minimizing a divergence or norm subject to no–arbitrage constraints. The present construction provides the simplest instance of this idea and ensures that, at a minimum, PRISM always outputs internally consistent YES/NO prices that respect $\pi_{\mathrm{YES}} + \pi_{\mathrm{NO}} = 1$ and positivity, without discarding the nonlinear and multimodal information accumulated in the posterior.

7.9 Informal Stress Scenarios

We outline three informal but important stress scenarios.

Scenario 1: Herd Cascade. Large temporal clusters of identical trades overwhelm early signals. The goal is to evaluate whether Stage 1 correction suppresses the cascade.

Scenario 2: Whale Attack. A small number of dominant traders distort the book. The goal is to evaluate whether δ_+ and δ_- can counteract this influence.

Scenario 3: Structural Volatility Shock. Draw p_{true} from a regime where structural model assumptions break down. Tests whether the hybrid prior and bias–correction layer still maintain coherence.

7.10 Herding Regimes, Mixing Failure, and PRISM Validity

Phase 7 introduces an Ising-type herding model to describe dependence among trader actions. In weak-coupling regimes, this process is geometrically α-mixing and the asymptotic results of Phase 8 apply. In strong-coupling regimes, the mixing assumptions may fail, and PRISM's asymptotic guarantees can break down.

We formalize this with a stylized Ising model on a trader graph $G = (V, E)$.

Theorem 9 (Mixing Regimes for Ising-Type Herding)**.** *Let $G_N = (V_N, E_N)$ be a sequence of trader graphs with $|V_N| = N$. For each trader $i \in V_N$, let $X_i \in \{-1, +1\}$ represent a YES/NO spin, and consider the Ising probability measure*

$$\mathbb{P}_N(X = x) \propto \exp\Big(\beta \sum_{\{i,j\} \in E_N} x_i x_j + h \sum_{i \in V_N} x_i\Big),$$

with coupling $\beta \geq 0$ and external field $h \in \mathbb{R}$.

*(a) **(Bounded-degree graphs)** If $\sup_N \max_{i \in V_N} \deg(i) \leq d_{\max} < \infty$ and $\beta d_{\max}$ is sufficiently small, then $\{X_i\}$ is geometrically α-mixing with coefficients decaying as $\alpha(k) \leq C\rho^k$. In this regime, the dependence is weak enough for the LLN/CLT and PRISM asymptotics to hold.*

*(b) **(Dense mean-field graphs)** If G_N is dense (e.g. complete graph with $|E_N| \asymp N^2$) and $\beta > \beta_c$ for a critical value $\beta_c > 0$, then the Ising model undergoes a phase transition: multiple modes appear and $\alpha(k)$ fails to decay geometrically. In such low-temperature regimes, long-range dependence persists and the mixing assumptions underlying PRISM's asymptotics can fail.*

*(c) **(Polynomial or non-mixing regimes)** For intermediate cases, $\alpha(k)$ may decay polynomially or not at all. In these regimes, standard CLT-based justifications for the Beta approximation may no longer be valid without additional control, and PRISM must be treated as a heuristic exponential-family approximation.*

Corollary 2 (Validity Region for PRISM Under Herding)**.** *Under the conditions of part (a), the Ising herding process generates a geometrically α-mixing sequence of adjusted indicators, and the consistency and Bernstein–von Mises results of Phase 8 apply. Under the strong-coupling regime of part (b), PRISM may yield posteriors that do not converge to the true event probability and whose variance behavior is not well described by the asymptotic normal approximations. Detecting such regimes and treating the resulting posteriors as exploratory rather than fully calibrated is recommended.*

Remark 9 (Diagnosing Mixing Failure). Empirically, mixing failure can be probed by examining autocorrelation functions of the adjusted process, block-bootstrap variability, or by fitting simple AR models and testing for long-range dependence. Persistent, slowly decaying correlations suggest that the market is in a strong-herding regime where PRISM's formal guarantees are weakened and increased weight should be placed on sensitivity analysis.

7.11 Summary of Simulation Parameters

A typical simulation configuration includes:

- Distribution for p_{true} (e.g. Beta, Uniform);
- Number of traders N;
- Trader type proportions: $\pi_{\text{inf}}, \pi_{\text{noise}}, \pi_{\text{adv}}$;
- Behavioral weights w_i^{beh} for Stage 1 correction;
- Structural offsets δ_+, δ_- for Stage 2 correction;
- Hybrid prior parameters (α_0, β_0);
- ML prior strength n_{ML} and value p_{ML};
- Number of simulation repetitions.

7.12 Output of Phase 7 and Linkage to Phase 8

The outputs of Phase 7 consist of:

- simulated PRISM prices $\pi_{\text{YES}}^{\text{adj}}$,
- simulated posteriors $(\alpha_{\text{adj}}, \beta_{\text{adj}})$,
- performance metrics from proper scoring rules,
- comparison of adjusted vs. unadjusted posterior,
- sensitivity of PRISM to biases, distortions, and priors.

These outputs provide the empirical backbone for Phase 8, which establishes theoretical results on consistency, calibration, arbitrage–freeness, and asymptotic stability.

9 Extended Module D: Theoretical Guarantees

Phase 8 establishes the theoretical foundations of the PRISM framework. The goal is to prove that the model preserves arbitrage–freeness, produces consistent and calibrated posteriors under a broad class of conditions, remains robust to behavioral and structural distortions, and admits asymptotic distributional approximations as the effective sample size increases.

The results in this phase pertain to the adjusted posterior derived in Phase 6:

$$p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}}), \qquad \alpha_{\text{adj}}, \beta_{\text{adj}} > 0,$$

where

$$\alpha_{\text{adj}} = \alpha_0 + y^*, \qquad \beta_{\text{adj}} = \beta_0 + (n^* - y^*),$$

and y^*, n^* incorporate behavioral and structural adjustments.

8.1 Assumptions

We introduce the following assumptions, each of which may hold under different simulation or empirical regimes:

(A1) The true event probability $p_{\text{true}} \in (0, 1)$ is fixed but unknown.

(A2) Trader signals (s_i) satisfy: informed traders have $s_i \sim \text{Bernoulli}(p_{\text{true}})$, noise traders have $s_i \sim \text{Bernoulli}(1/2)$, adversarial traders invert informed signals.

(A3) Behavioral weights $w_i^{\text{beh}} > 0$ are bounded above and below:

$$0 < m \leq w_i^{\text{beh}} \leq M < \infty.$$

(A4) Structural offsets satisfy

$$|\delta_+| + |\delta_-| \leq C_\delta < \infty.$$

(A5) The hybrid prior parameters $\alpha_0, \beta_0 > 0$ are fixed and finite.

(A6) The effective adjusted sample size

$$n^* = y^* + (n^* - y^*)$$

satisfies $n^* \to \infty$.

These assumptions form the basis of the consistency and robustness results below.

8.2 Posterior Consistency

We first establish that PRISM yields a consistent posterior for p_{true} under increasingly informative data.

Theorem 10 (Posterior Consistency Under Model Assumptions)**.** *Under assumptions (A1)–(A6),*

$$p \mid s_{\text{adj}} \xrightarrow[n^* \to \infty]{\mathbb{P}} p_{\text{true}}.$$

Proof. Write the adjusted posterior mean as:

$$\hat{p}_{\text{adj}} = \frac{\alpha_{\text{adj}}}{\alpha_{\text{adj}} + \beta_{\text{adj}}} = \frac{\alpha_0 + y^*}{n^* + \alpha_0 + \beta_0}.$$

We expand the adjusted count:

$$y^* = \sum_{i=1}^{n} w_i^{\text{beh}} \mathbf{1}\{s_i = \text{YES}\} + \delta_+.$$

By (A3), behavioral weights are bounded, and by the law of large numbers for heterogeneous but bounded weights,

$$\frac{1}{n^*} \sum_{i=1}^{n} w_i^{\text{beh}} \mathbf{1}\{s_i = \text{YES}\} \xrightarrow{\mathbb{P}} p_{\text{true}} \cdot \mathbb{E}[w_i^{\text{beh}}].$$

Since the denominator also grows with n^*,

$$\frac{y^*}{n^*} \xrightarrow{\mathbb{P}} p_{\text{true}} \cdot \frac{\mathbb{E}[w_i^{\text{beh}}]}{\mathbb{E}[w_i^{\text{beh}}]} = p_{\text{true}}.$$

Offset δ_+ satisfies

$$\frac{\delta_+}{n^*} \to 0.$$

Thus

$$\hat{p}_{\text{adj}} = \frac{y^* + O(1)}{n^* + O(1)} \xrightarrow{\mathbb{P}} p_{\text{true}}.$$

Hence the posterior is consistent. □

Remark 10 (Role of Large Effective Sample Size n^*). The consistency result above is driven primarily by the growth of the effective sample size n^*, rather than by the raw number of trades n alone. Recall that

$$n^* = y^* + (n^* - y^*) = \sum_{i=1}^{n} w_i^{\text{beh}} \big(\mathbb{1}\{s_i = \text{YES}\} + \mathbb{1}\{s_i = \text{NO}\} \big) + (\delta_+ + \delta_-),$$

so that n^* reflects both behavioral reweighting and structural offsets.

Under assumptions (A2)–(A4), a nonzero fraction of traders are informed or adversarial, and the weights w_i^{beh} remain bounded above and below. Consequently, as the number of traders n grows, the effective sample size n^* also grows linearly:

$$\frac{n^*}{n} \to c \in (0, \infty),$$

up to $O(1)$ contributions from (δ_+, δ_-). The law of large numbers and central limit behavior therefore apply at the level of n^*, not just n. When an adversarial fraction $\pi_{\text{adv}} < 1/2$ is present, the net signal embedded in y^* remains aligned with p_{true} and the posterior still concentrates at the true value as $n^* \to \infty$.

This perspective clarifies that PRISM's asymptotic properties are governed by the information content of the *adjusted* sample size, which is resilient to moderate behavioral distortion and bounded structural offsets, rather than by the raw trade count alone.

8.3 Consistency Under Structural Prior Correctness

Proposition 10 (Structural Prior Dominance). *If the structural prior mean equals the true probability, i.e. $q_{\text{str}} = p_{\text{true}}$, and if the ML prior is weak (n_{ML} small), then*

$$p \mid s_{\text{adj}} \xrightarrow{\mathbb{P}} p_{\text{true}}.$$

Proof. Follows from the posterior consistency theorem with $\alpha_0/n^* \to 0$. □

8.4 Consistency Under ML Prior Correctness

Proposition 11 (ML Prior Dominance). *If the ML prior satisfies $p_{\text{ML}} = p_{\text{true}}$ and the ML strength satisfies $n_{\text{ML}} \to \infty$, then*

$$\hat{p}_{\text{adj}} \to p_{\text{true}}.$$

Proof. As $n_{\text{ML}} \to \infty$, the hybrid prior mean approaches $p_{\text{ML}} = p_{\text{true}}$, and the structural contribution vanishes. □

8.5 Consistency Under Behavioral Bias Correction

Theorem 11 (Consistency Under Behavioral Distortion)**.** *If (A3) holds and the proportion of informed traders is nonzero, then PRISM remains consistent after Stage 1 correction.*

Proof. Stage 1 weights are bounded and therefore do not distort the sign or limit of the empirical frequencies. □

8.6 Consistency Under Adversarial Contamination

Theorem 12 (Adversarial Robustness)**.** *If the fraction of adversarial traders satisfies $\pi_{\text{adv}} < 1/2$, then PRISM remains posterior consistent.*

Proof. Under contamination theory, if adversarial contamination is below 50%, the majority signal still reflects p_{true}. Weighted counts remain asymptotically aligned with the truth after normalization. □

8.7 Error Propagation from Weight and Offset Estimation

Let $(\psi^\dagger, \delta^\dagger)$ denote the pseudo-true parameters appearing in the Empirical Bayes limit, and let

$$y^* = y^*(\psi^\dagger, \delta^\dagger), \qquad n^* = n^*(\psi^\dagger, \delta^\dagger)$$

denote the corresponding adjusted counts for a given market. The idealized PRISM posterior for p is then

$$p \mid \mathcal{I}, \psi^\dagger, \delta^\dagger \sim \text{Beta}\big(\alpha_0 + y^*,\ \beta_0 + n^* - y^*\big).$$

In practice, we use the estimated parameters $(\hat{\psi}, \hat{\delta})$ (obtained from a finite historical sample) and the associated adjusted counts

$$\hat{y}^* = y^*(\hat{\psi}, \hat{\delta}), \qquad \hat{n}^* = n^*(\hat{\psi}, \hat{\delta}),$$

leading to the approximate posterior

$$p \mid \mathcal{I}, \hat{\psi}, \hat{\delta} \sim \text{Beta}\big(\alpha_0 + \hat{y}^*,\ \beta_0 + \hat{n}^* - \hat{y}^*\big).$$

We now provide bounds on the deviation between these two posteriors as a function of the error in $(\hat{\psi}, \hat{\delta})$.

Theorem 13 (Lipschitz Error Propagation for PRISM Posterior)**.** *Assume:*

(i) The weighting function $w(x; \psi)$ is Lipschitz in ψ, uniformly in x, i.e. there exists $L_w > 0$ such that

$$|w(x; \psi_1) - w(x; \psi_2)| \le L_w \|\psi_1 - \psi_2\| \quad \textit{for all } x,\ \psi_1, \psi_2.$$

(ii) The offsets δ_+, δ_- enter linearly and the map $\delta \mapsto (y^*(\psi, \delta), n^*(\psi, \delta))$ is Lipschitz with constant L_δ.

(iii) The parameter space $\Psi \times \Delta$ is compact and $\alpha_0, \beta_0 > 0$ are fixed.

Then the following hold for a fixed market:

(a) ***(Mean and Variance)*** *Let $\mu^\dagger$ and $\hat{\mu}$ denote the posterior means, and $\sigma^{2,\dagger}$ and $\hat{\sigma}^2$ the posterior variances, under $(\psi^\dagger, \delta^\dagger)$ and $(\hat{\psi}, \hat{\delta})$ respectively. Then there exist constants $C_1, C_2 > 0$ such that*

$$|\hat{\mu} - \mu^\dagger| \leq C_1 \left(|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \right),$$
$$|\hat{\sigma}^2 - \sigma^{2,\dagger}| \leq C_2 \left(|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \right).$$

(b) ***(Posterior Distribution)*** *Let $\Pi^\dagger$ and $\hat{\Pi}$ denote the two Beta posteriors. Then there exists $C_3 > 0$ such that both the total variation distance and the squared Hellinger distance satisfy*

$$\|\hat{\Pi} - \Pi^\dagger\|_{\mathrm{TV}} \leq C_3 \left(|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \right),$$
$$H^2(\hat{\Pi}, \Pi^\dagger) \leq C_3 \left(|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \right),$$

where H denotes the Hellinger distance.

(c) ***(Parameter Error to Posterior Error)*** *Under (i)–(ii), there exists a constant $C_4 > 0$ such that*

$$|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \leq C_4 \left(\|\hat{\psi} - \psi^\dagger\| + \|\hat{\delta} - \delta^\dagger\| \right).$$

Combining with (a)–(b) yields Lipschitz-type bounds of posterior mean, variance, and distributional distance in terms of the parameter estimation error.

Proof (Sketch). For (a), the Beta posterior mean and variance are smooth functions of (α, β) given by

$$\mu(\alpha, \beta) = \frac{\alpha}{\alpha + \beta}, \qquad \sigma^2(\alpha, \beta) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

On any compact subset with $\alpha, \beta \geq c > 0$, these maps are Lipschitz in (α, β), hence in (y^*, n^*) since $\alpha = \alpha_0 + y^*$ and $\beta = \beta_0 + n^* - y^*$ are affine functions of (y^*, n^*). This yields the stated bounds with constants C_1, C_2 depending on the compact region.

For (b), standard perturbation bounds for one-parameter exponential families imply that the total variation and Hellinger distances between $\mathrm{Beta}(\alpha_1, \beta_1)$ and $\mathrm{Beta}(\alpha_2, \beta_2)$ are Lipschitz in $(\alpha_1 - \alpha_2, \beta_1 - \beta_2)$ on compact sets with $\alpha_j, \beta_j \geq c > 0$. Again using the affine dependence of (α, β) on (y^*, n^*) yields the displayed inequalities.

For (c), Lipschitz continuity of $w(x; \psi)$ in ψ and of the linear map $\delta \mapsto (y^*, n^*)$, together with boundedness of the feature set, gives

$$|\hat{y}^* - y^*| + |\hat{n}^* - n^*| \leq C_4 \big(\|\hat{\psi} - \psi^\dagger\| + \|\hat{\delta} - \delta^\dagger\| \big)$$

for some C_4 depending on the feature bounds and Lipschitz constants. Combining with (a)–(b) gives the claimed error propagation bounds. □

Remark 11 (Interpretation). This theorem formalizes the intuitive idea that small errors in the estimated weights and offsets lead to small distortions in the PRISM posterior. As the Empirical Bayes estimators $(\hat{\psi}_M, \hat{\delta}_M)$ converge to their pseudo-true values $(\psi^\dagger, \delta^\dagger)$ with increasing historical sample size M, the resulting posterior mean, variance, and full distribution converge to those obtained under the pseudo-true parameters, at a rate controlled by the Lipschitz constants above.

8.8 Consistency Under Structural Distortion Offsets

Theorem 14 (Offset Robustness)**.** *If $|\delta_+| + |\delta_-| < C_\delta < \infty$ and $n^* \to \infty$, then*

$$\frac{\delta_+}{n^*} \to 0, \qquad \frac{\delta_-}{n^*} \to 0,$$

and PRISM remains consistent.

Proof. Offsets are $O(1)$ and thus negligible relative to n^*. ☐

8.9 Robustness to Perturbations (Stability)

Proposition 12 (Lipschitz Stability of Posterior Mean)**.** *For the adjusted posterior mean*

$$\hat{p}_{\text{adj}} = \frac{\alpha_{\text{adj}}}{\alpha_{\text{adj}} + \beta_{\text{adj}}},$$

there exists $L > 0$ such that for any perturbations $\Delta y^*, \Delta(n^* - y^*)$,

$$|\Delta \hat{p}_{\text{adj}}| \leq L \left(|\Delta y^*| + |\Delta(n^* - y^*)| \right).$$

Proof. Same structure as Phase 6 robustness theorem; the posterior mean is a smooth rational function on a compact domain. ☐

8.10 Arbitrage-Freeness

Theorem 15 (coherence of projected YES/NO probabilities)**.** *For the adjusted prices*

$$\pi^{\text{adj}}_{\text{YES}} = \hat{p}_{\text{adj}}, \qquad \pi^{\text{adj}}_{\text{NO}} = 1 - \hat{p}_{\text{adj}},$$

the following hold:

(i) $\pi^{\text{adj}}_{\text{YES}} + \pi^{\text{adj}}_{\text{NO}} = 1$,

(ii) $0 < \pi^{\text{adj}}_{\text{YES}} < 1$,

(iii) $\pi^{\text{adj}}_{\text{YES}}$ *is monotone in adjusted counts y^*,*

(iv) boundedness is preserved under all admissible distortions.

Proof. (i) Follows immediately from the definition. (ii) Holds because $\alpha_{\text{adj}}, \beta_{\text{adj}} > 0$. (iii) Derivative of Beta mean with respect to y^* is positive. (iv) Adjustments enter linearly; Beta parameters remain positive. ☐

8.11 Asymptotic Distribution (Bernstein–von Mises)

Theorem 16 (Asymptotic Normality via CLT and Bernstein–von Mises)**.** *Suppose (A1)–(A6) hold, and in addition:*

- *the sequence of adjusted indicators contributing to y^* satisfies a Lindeberg–Feller type condition, and*

- *the fraction of adversarial traders satisfies $\pi_{\text{adv}} < 1/2$, so that the net signal remains aligned with p_{true}.*

Let

$$\hat{p}^* := \frac{y^*}{n^*}$$

denote the adjusted sample proportion, where n^* is the effective sample size resulting from behavioral weights and structural offsets. Then:

(i) ***(CLT for the Adjusted Proportion)*** *There exists $\sigma^2 \in (0, \infty)$ such that*

$$\sqrt{n^*}\left(\hat{p}^* - p_{\text{true}}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \quad \text{as } n^* \to \infty.$$

(ii) ***(Asymptotic Normality of the Posterior Mean)*** *For the adjusted posterior*

$$p \mid s_{\text{adj}} \sim \text{Beta}(\alpha_{\text{adj}}, \beta_{\text{adj}}), \qquad \alpha_{\text{adj}} = \alpha_0 + y^*, \quad \beta_{\text{adj}} = \beta_0 + (n^* - y^*),$$

we have

$$\sqrt{n^*}\left(\hat{p}_{\text{adj}} - p_{\text{true}}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where $\hat{p}_{\text{adj}} = \alpha_{\text{adj}}/(\alpha_{\text{adj}} + \beta_{\text{adj}})$ is the posterior mean.

(iii) ***(Bernstein–von Mises Approximation)*** *The full posterior distribution satisfies*

$$p \mid s_{\text{adj}} \overset{d}{\approx} \mathcal{N}\left(p_{\text{true}}, \frac{\sigma^2}{n^*}\right) \quad \text{for large } n^*,$$

i.e., the posterior is asymptotically normal with center p_{true} and variance of order $1/n^*$.

Proof. (i) *CLT for the adjusted proportion.* Write

$$y^* = \sum_{i=1}^{n} W_i,$$

where each W_i is the contribution of trader i to the adjusted YES count, including behavioral weights and the effect of trader type (informed, noise, adversarial). Assumptions (A2)–(A4) imply that the W_i are uniformly bounded and that the mean of W_i is aligned with p_{true} as $\pi_{\text{adv}} < 1/2$. Under the Lindeberg–Feller condition for the triangular array $\{W_i\}_{i=1}^n$, we obtain

$$\sqrt{n^*}\left(\hat{p}^* - p_{\text{true}}\right) = \sqrt{n^*}\left(\frac{y^*}{n^*} - p_{\text{true}}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

for some finite, positive σ^2 capturing the effective dispersion of the weighted trader signals.

(ii) *Asymptotic normality of the posterior mean.* The posterior mean can be written as

$$\hat{p}_{\text{adj}} = \frac{\alpha_0 + y^*}{\alpha_0 + \beta_0 + n^*} = \hat{p}^* + \frac{\alpha_0 - \hat{p}^*(\alpha_0 + \beta_0)}{\alpha_0 + \beta_0 + n^*}.$$

The second term is $O(1/n^*)$ in probability and therefore negligible at the $1/\sqrt{n^*}$ scale. Thus

$$\sqrt{n^*}\left(\hat{p}_{\text{adj}} - p_{\text{true}}\right) = \sqrt{n^*}\left(\hat{p}^* - p_{\text{true}}\right) + o_{\mathbb{P}}(1),$$

and the CLT from part (i) implies convergence in distribution to $\mathcal{N}(0, \sigma^2)$.

(iii) *Bernstein–von Mises approximation.* The Beta posterior with parameters $(\alpha_{\text{adj}}, \beta_{\text{adj}})$ has mean $\hat{p}_{\text{adj}}$ and variance

$$\mathrm{Var}(p \mid s_{\text{adj}}) = \frac{\alpha_{\text{adj}}\beta_{\text{adj}}}{(\alpha_{\text{adj}} + \beta_{\text{adj}})^2(\alpha_{\text{adj}} + \beta_{\text{adj}} + 1)} \approx \frac{\hat{p}_{\text{adj}}(1 - \hat{p}_{\text{adj}})}{n^*},$$

for large n^*.

By standard Bernstein–von Mises arguments for one-dimensional conjugate models, the posterior distribution of p becomes asymptotically normal with this mean and variance, and the difference between the posterior law and the corresponding normal distribution vanishes in total variation. Substituting $\hat{p}_{\text{adj}} \to p_{\text{true}}$ yields the stated normal approximation centered at p_{true} with asymptotic variance of order $1/n^*$. □

8.12 Finite–Sample Concentration and Credible–Interval Corrections

The asymptotic results developed earlier in this phase (law of large numbers, central limit theorems, and Bernstein–von Mises theorems) require an effective sample size n^* that is sufficiently large. In low–liquidity markets or short order windows, these asymptotic approximations can be misleading: the posterior may concentrate slowly, and credible intervals derived from Gaussian limits may substantially misstate uncertainty. This subsection provides finite–sample corrections via exponential concentration inequalities, exact finite–n credible intervals, and Berry–Esseen–type bounds for the posterior approximation.

are correctly specified in the sense of Sections reference, reference, and reference: nonlinear distortions, dependence, and multimodality are explicitly modeled via regime mixtures and dynamic bias layers. The bounds below apply to a single market or time block where the true event probability p_{true} is well defined.

Setup for a Single Market Block. Consider a single market with n effective observations $Z_1, \ldots, Z_n \in \{0, 1\}$ (YES indicators after Stage 1 weighting and Stage 2 structural corrections), generated conditionally i.i.d. given a fixed $p_{\text{true}} \in (0, 1)$:

$$Z_i \mid p_{\text{true}} \sim \mathrm{Ber}(p_{\text{true}}), \quad i = 1, \ldots, n.$$

Let $Y = \sum_{i=1}^n Z_i$ and $\bar{Z} = Y/n$ be the empirical YES count and sample mean. For a Beta prior $\mathrm{Beta}(\alpha_0, \beta_0)$ on p, the posterior is $\mathrm{Beta}(\alpha, \beta)$ with

$$\alpha = \alpha_0 + Y, \qquad \beta = \beta_0 + n - Y,$$

and posterior mean

$$\hat{p}_{\text{post}} = \frac{\alpha}{\alpha + \beta} = \lambda\, p_0 + (1 - \lambda)\, \bar{Z},$$

where

$$p_0 = \frac{\alpha_0}{\alpha_0 + \beta_0}, \qquad \lambda = \frac{\alpha_0 + \beta_0}{\alpha_0 + \beta_0 + n}.$$

Assumption 5 (Moderate Dependence via Effective Sample Size)**.** *In the presence of weak dependence or regime switching (Sections reference and reference), assume that there exists an effective sample size n^* satisfying*

$$n^* \le n, \qquad \mathbb{P}\Big(|\bar{Z} - p_{\text{true}}| > \varepsilon\Big) \le 2\exp\big(-2n^*\varepsilon^2\big)$$

for all $\varepsilon > 0$. In the i.i.d. case, $n^* = n$; under α–mixing or HMM dependence, n^* incorporates the effective number of independent observations (e.g. via standard blocking arguments).

Theorem 17 (Finite–Sample Concentration for the Posterior Mean)**.** *Under the single–market block model and Assumption reference, fix $\varepsilon > 0$ and define*

$$\varepsilon_0 := \lambda|p_0 - p_{\text{true}}| \le \lambda, \qquad \lambda = \frac{\alpha_0 + \beta_0}{\alpha_0 + \beta_0 + n}.$$

Then for any $\varepsilon > \varepsilon_0$,

$$\mathbb{P}\Big(\big|\hat{p}_{\text{post}} - p_{\text{true}}\big| > \varepsilon\Big) \ \le\ 2\exp\left(-2n^*\left(\frac{\varepsilon - \varepsilon_0}{1-\lambda}\right)^2\right).$$

In particular, for sufficiently large n (so that λ and ε_0 are small), the posterior mean concentrates around p_{true} at a sub–Gaussian rate with effective sample size n^*.

Proof. Using the convex combination representation,

$$\hat{p}_{\text{post}} - p_{\text{true}} = \lambda(p_0 - p_{\text{true}}) + (1-\lambda)(\bar{Z} - p_{\text{true}}),$$

we obtain

$$\big|\hat{p}_{\text{post}} - p_{\text{true}}\big| \le \lambda|p_0 - p_{\text{true}}| + (1-\lambda)|\bar{Z} - p_{\text{true}}| \le \varepsilon_0 + (1-\lambda)|\bar{Z} - p_{\text{true}}|.$$

Thus, if $|\hat{p}_{\text{post}} - p_{\text{true}}| > \varepsilon$ with $\varepsilon > \varepsilon_0$, then necessarily

$$|\bar{Z} - p_{\text{true}}| > \frac{\varepsilon - \varepsilon_0}{1-\lambda}.$$

Applying Assumption reference gives

$$\mathbb{P}\Big(\big|\hat{p}_{\text{post}} - p_{\text{true}}\big| > \varepsilon\Big) \le \mathbb{P}\left(|\bar{Z} - p_{\text{true}}| > \frac{\varepsilon - \varepsilon_0}{1-\lambda}\right) \le 2\exp\left(-2n^*\left(\frac{\varepsilon - \varepsilon_0}{1-\lambda}\right)^2\right),$$

as claimed. □

Remark 12 (Interpretation and Choice of n^*). The bound in Theorem reference decouples the finite–sample error into a prior–bias term ε_0 (which vanishes as n dominates $\alpha_0 + \beta_0$) and a stochastic term controlled by n^*. In low–liquidity markets, both terms may be nonnegligible; PRISM should therefore explicitly report the implied error scale

$$\varepsilon_{\text{tol}} \approx \varepsilon_0 + (1-\lambda)\sqrt{\frac{\log(2/\delta)}{2n^*}}$$

for a target tail probability δ (e.g. $\delta = 0.05$).

Finite–n Credible Intervals via Beta Quantiles. For the Beta posterior $\text{Beta}(\alpha, \beta)$, an exact $(1-\gamma)$–credible interval for p is given by

$$[p^{\text{low}}_{\gamma/2},\, p^{\text{high}}_{1-\gamma/2}] := \big[F^{-1}_{\text{Beta}(\alpha,\beta)}(\gamma/2),\, F^{-1}_{\text{Beta}(\alpha,\beta)}(1-\gamma/2)\big],$$

where $F^{-1}_{\text{Beta}(\alpha,\beta)}$ denotes the inverse incomplete beta function. In the mixture case of Section reference, credible intervals can be computed by numerical inversion of the mixture CDF or via Monte Carlo sampling from Π_{mix}, yielding empirical quantiles that preserve multimodality.

Theorem 18 (Berry–Esseen–Type Bound for the Beta Posterior)**.** *Under the i.i.d. single–market block model with $p_{\text{true}} \in (0,1)$ fixed and α_0, β_0 bounded, let $\Pi_T(\cdot)$ denote the posterior distribution of p given $Z_1, \ldots, Z_n$, and let $\sigma_T^2 = p_{\text{true}}(1 - p_{\text{true}})/(n + \alpha_0 + \beta_0)$. Then there exists a universal constant $C > 0$ such that*

$$\sup_{x \in \mathbb{R}} \left| \Pi_T\big(p \leq p_{\text{true}} + x\sigma_T\big) - \Phi(x) \right| \;\leq\; \frac{C}{\sqrt{n^*}},$$

where Φ is the standard normal CDF and n^* is the effective sample size of Assumption reference. In particular, Gaussian credible intervals centered at $\hat{p}_{\text{post}}$ with radius $z_{1-\gamma/2}\sigma_T$ incur a finite–n approximation error of order $O(1/\sqrt{n^*})$.

Remark 13 (Mixture Extension and Bootstrap Refinements)*.* For mixture posteriors Π_{mix} as in Section reference, the posterior mean $\hat{p}_{\text{mix}}$ is a mixture of component means and satisfies a bound of the form

$$\mathbb{P}\Big(|\hat{p}_{\text{mix}} - p_{\text{true}}| > \varepsilon\Big) \leq \sum_{k=1}^{K} w_k^* \mathbb{P}\Big(|\hat{p}_k - p_{\text{true}}| > \varepsilon\Big),$$

where $\hat{p}_k$ is the posterior mean under component k. Combining this with Theorem reference for each component yields a mixture concentration bound. In practice, PRISM can supplement analytic bounds with bootstrap resampling: resample the observed orders $\{Z_i\}_{i=1}^n$ B times, recompute $\hat{p}_{\text{post}}$ or $\hat{p}_{\text{mix}}$ for each bootstrap replicate, and use the empirical quantiles of $\{\hat{p}^{(b)}\}_{b=1}^B$ to form finite–n uncertainty bands. Such bootstrap intervals can be compared to the Beta–based credible intervals to diagnose small–sample distortions and coverage properties.

Remark 14 (Implications for Phase 7 Simulation Metrics)*.* In Phase 7, finite–sample properties of PRISM are evaluated via simulation under low–liquidity regimes (e.g. $n < 50$). The results of this subsection provide target coverage rates and finite–n error scales for:

- the absolute error $|\hat{p}_{\text{post}} - p_{\text{true}}|$ or $|\hat{p}_{\text{mix}} - p_{\text{true}}|$,
- the empirical coverage of nominal $(1 - \gamma)$ credible intervals,
- and mispricing of YES/NO digital contracts, $|\hat{p}_{(\text{post/mix})} - p_{\text{true}}|$.

Simulation designs should explicitly report both n and n^* to explain deviations from asymptotic behavior, and use the bounds in Theorem reference and Theorem reference as benchmarks for finite–sample reliability.

9.1 8.13 Robustness and Divergence Bounds for Mixture Posteriors

The mixture posterior Π_{mix} of Definition reference extends the single–Beta PRISM posterior to multimodal settings. This subsection generalizes the divergence bounds of Theorem 23 to mixture posteriors and provides conditions under which the mixture remains stable to perturbations, along with explicit lower bounds showing when single–Beta approximations necessarily fail.

Assumption 6 (Mixture Stability Under Perturbations)**.** *Let $\Pi_{\text{mix}} = \sum_{k=1}^K w_k^* \Pi_k$ and $\widetilde{\Pi}_{\text{mix}} = \sum_{k=1}^K \widetilde{w}_k^* \widetilde{\Pi}_k$ be two mixture posteriors, where each Π_k and $\widetilde{\Pi}_k$ is Beta(α_k, β_k) (possibly with different parameters). Assume:*

(i) Component Lipschitz continuity:

$$\left\| \Pi_k - \widetilde{\Pi}_k \right\|_{\text{TV}} \;\leq\; L_k \left\| \mathcal{D} - \widetilde{\mathcal{D}} \right\|;$$

(ii) Weight stability:

$$\|w^* - \widetilde{w}^*\|_1 \;\le\; C_w \left\| \mathcal{D}^{\text{hold}} - \widetilde{\mathcal{D}}^{\text{hold}} \right\|;$$

(iii) Bounded support separation: for all k, Π_k and $\widetilde{\Pi}_k$ have support in $[0,1]$ with finite moments.

Theorem 19 (Mixture Posterior Robustness Bound)**.** *Under Assumption reference, the mixture posterior satisfies*

$$\left\| \Pi_{\text{mix}} - \widetilde{\Pi}_{\text{mix}} \right\|_{\text{TV}} \;\le\; \sum_{k=1}^{K} w_k^* L_k \left\| \mathcal{D} - \widetilde{\mathcal{D}} \right\| + C_w \max_k \left\| \Pi_k - \widetilde{\Pi}_k \right\|_{\text{TV}}.$$

Hence the mixture posterior inherits robustness from: (i) the individual component Betas, and (ii) the stability of the stacking weights.

Proof. Write

$$\Pi_{\text{mix}} - \widetilde{\Pi}_{\text{mix}} = \sum_{k=1}^{K} (w_k^* - \widetilde{w}_k^*)\, \widetilde{\Pi}_k + \sum_{k=1}^{K} w_k^* (\Pi_k - \widetilde{\Pi}_k).$$

Taking total variation norms and applying the triangle inequality gives

$$\left\| \Pi_{\text{mix}} - \widetilde{\Pi}_{\text{mix}} \right\|_{\text{TV}} \le \sum_{k=1}^{K} |w_k^* - \widetilde{w}_k^*| \left\| \widetilde{\Pi}_k \right\|_{\text{TV}} + \sum_{k=1}^{K} w_k^* \left\| \Pi_k - \widetilde{\Pi}_k \right\|_{\text{TV}}.$$

The first term is bounded using $\|\widetilde{\Pi}_k\|_{\text{TV}} = 1$ and Assumption reference(ii); the second term uses Assumption reference(i). Combine to obtain the stated bound. □

Theorem 20 (Lower Bound: Mixture Divergence Cannot Vanish Under Mode Separation)**.** *Let Π_{mix} be as above, and let Π_{Beta} be the moment-matched single-Beta approximation from Definition reference. If at least two components have means separated by $|\mu_i - \mu_j| \ge \varepsilon > 0$, then there exists $c(\varepsilon) > 0$ such that*

$$\|\Pi_{\text{mix}} - \Pi_{\text{Beta}}\|_{\text{TV}} \;\ge\; c(\varepsilon).$$

Thus no unimodal Beta projection can approximate the mixture posterior uniformly when the mixture is sufficiently multimodal.

Remark 15 (Extension of Theorem 23)*.* Theorem reference extends the divergence and robustness results of Theorem 23 to multimodal settings. Theorem reference adds a complementary *impossibility dimension*: mixture posteriors retain irreducible multimodality, meaning that Beta projections cannot achieve vanishing approximation error when component means are well separated. These bounds guide when PRISM should operate with the full mixture posterior, and when approximation layers (e.g. Beta moment matching) incur unavoidable and quantifiable loss.

8.14 Weak Dependence, α-Mixing, and Asymptotic Normality

In earlier phases, we modeled trader actions as if they were conditionally independent Bernoulli signals given the true event probability p_{true}. In practice, herding, imitation, and correlated information can induce dependence across orders. Phase 7 introduced stylized dependence regimes, including Ising-type herding models that give rise to geometrically α-mixing sequences.

In this subsection, we impose a weak dependence condition and show that the key asymptotic properties of PRISM—consistency and asymptotic normality of YES indicators that feed into the Beta–Binomial update.

Adjusted YES process. Let $\{Z_t\}_{t\geq 1}$ denote the sequence of adjusted YES indicators at the level of individual orders or time steps, where

$$Z_t \in [0, 1],$$

represents the effective contribution of order t to the YES count after behavioral weighting. For example, in a simple specification,

$$Z_t = w^{\text{beh}}(x_t; \psi^\dagger)\, \mathbb{1}\{s_t = \text{YES}\},$$

with $w^{\text{beh}}(x_t; \psi^\dagger) \in [0, 1]$ a bounded weight depending on features x_t and pseudo-true parameter $\psi^\dagger$.

For a given market with n observed orders, the adjusted YES count and effective sample size are

$$y_n^* = \sum_{t=1}^{n} Z_t + \delta_+^\dagger, \qquad n_n^* = \sum_{t=1}^{n} w^{\text{beh}}(x_t; \psi^\dagger) + \delta_+^\dagger + \delta_-^\dagger,$$

and the Beta–Binomial update uses (y_n^*, n_n^*) together with the hybrid prior. We assume (Z_t) is generated under a fixed true event probability $p_{\text{true}} \in (0, 1)$.

α-mixing assumptions. Let $\mathcal{F}_a^b = \sigma(Z_a, \ldots, Z_b)$ be the sigma-algebra generated by the process between times a and b, and define the α-mixing coefficients

$$\alpha(k) = \sup_{t\geq 1} \sup_{A\in\mathcal{F}_1^t,\, B\in\mathcal{F}_{t+k}^\infty} \left|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)\right|.$$

We assume:

(A7) (**Boundedness**) The adjusted indicators are uniformly bounded: $0 \leq Z_t \leq 1$.

(A8) (**Geometric α-mixing**) There exist constants $C > 0$ and $\rho \in (0, 1)$ such that

$$\alpha(k) \leq C\rho^k, \qquad k \geq 1.$$

(A9) (**Stationarity and Identifiability**) The process (Z_t) is strictly stationary under p_{true}, with

$$\mathbb{E}[Z_t] = \mu(p_{\text{true}}), \qquad \mathrm{Var}(Z_t) = \sigma_Z^2(p_{\text{true}}),$$

where $\mu(p)$ is strictly increasing in p on $(0, 1)$.

Assumption (A8) is satisfied, for example, by the Ising-type herding model in Phase 7 when couplings are sufficiently weak (Proposition 7.X).

Lemma 9 (LLN and CLT for the Adjusted YES Process)**.** *Under (A7)–(A9), define the normalized partial sums*

$$\bar{Z}_n = \frac{1}{n} \sum_{t=1}^{n} Z_t.$$

Then:

*(a) (**Law of Large Numbers**) $\bar{Z}_n \to \mu(p_{\text{true}})$ almost surely and in L^1 as $n \to \infty$.*

(b) (***Central Limit Theorem***) *There exists* $\tau^2(p_{\text{true}}) \in (0, \infty)$ *such that*

$$\sqrt{n}\big(\bar{Z}_n - \mu(p_{\text{true}})\big) \xrightarrow{d} \mathcal{N}\big(0, \tau^2(p_{\text{true}})\big) \quad \text{as } n \to \infty.$$

Proof (Sketch). The uniform boundedness in (A7) and geometric α-mixing in (A8) imply that (Z_t) satisfies the conditions of classical LLN and CLT results for strongly mixing sequences. Stationarity and finite variance in (A9) ensure that $\mu(p_{\text{true}})$ and $\sigma_Z^2(p_{\text{true}})$ are well defined, and the asymptotic variance $\tau^2(p_{\text{true}})$ can be expressed as a sum of autocovariances. Standard references for mixing CLTs apply directly under the geometric decay of $\alpha(k)$. □

We now translate this into the asymptotic behavior of the PRISM posterior.

Theorem 21 (Consistency and Asymptotic Normality of PRISM Posterior Under Dependence)**.** *Under (A7)–(A9), suppose that for a sequence of markets with* n *orders we use the adjusted counts*

$$y_n^* = \sum_{t=1}^{n} Z_t + \delta_+^\dagger, \qquad n_n^* = \sum_{t=1}^{n} w^{\text{beh}}(x_t; \psi^\dagger) + \delta_+^\dagger + \delta_-^\dagger,$$

with $w^{\text{beh}}(x_t; \psi^\dagger) \in [c_w, 1]$ *for some* $c_w > 0$*. Let* p *denote the event probability and assume a Beta hybrid prior*

$$p \sim \text{Beta}(\alpha_0, \beta_0), \quad \alpha_0, \beta_0 > 0,$$

independent of (Z_t)*. Then:*

(a) (***Posterior Consistency***) *Let* $\Pi_n(\cdot)$ *denote the PRISM posterior for* p *based on* (y_n^*, n_n^*)*. For any* $\epsilon > 0$*,*

$$\Pi_n\big(|p - p_{\text{true}}| > \epsilon\big) \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

(b) (***Asymptotic Normality of Posterior Mean***) *Let* $\hat{p}_n$ *be the posterior mean under* Π_n*. Then*

$$\sqrt{n}\big(\hat{p}_n - p_{\text{true}}\big) \xrightarrow{d} \mathcal{N}\big(0, V(p_{\text{true}})\big),$$

for some finite $V(p_{\text{true}})$ *determined by* $\tau^2(p_{\text{true}})$ *and the weight structure.*

(c) (***Bernstein–von Mises Approximation***) *The posterior distribution* Π_n *is asymptotically normal in the sense that*

$$\sup_{A \subset \mathbb{R}} \left| \Pi_n\big(\sqrt{n}(p - p_{\text{true}}) \in A\big) - \mathcal{N}\big(0, V(p_{\text{true}})\big)(A) \right| \xrightarrow{P} 0.$$

Proof (Sketch). The effective sample size n_n^* grows linearly with n due to the lower bound $w^{\text{beh}}(x_t; \psi^\dagger) \geq c_w > 0$. The adjusted mean

$$\bar{Z}_n^* = \frac{y_n^* - \delta_+^\dagger}{n_n^* - \delta_+^\dagger - \delta_-^\dagger}$$

is a smooth function of $\bar{Z}_n$ and the average weight, so the LLN and CLT from the lemma transfer to $\bar{Z}_n^*$ via the delta method. In particular, $\bar{Z}_n^* \to \mu(p_{\text{true}})$ and

$$\sqrt{n}(\bar{Z}_n^* - \mu(p_{\text{true}})) \xrightarrow{d} \mathcal{N}(0, \tilde{\tau}^2(p_{\text{true}})).$$

The Beta posterior based on (y_n^*, n_n^*) has mean and variance that can be written as smooth functions of $(\bar{Z}_n^*, n_n^*)$. As $n_n^* \sim c\,n$ for some $c > 0$, standard Bayesian asymptotics for one-dimensional parameters under weak dependence (together with the mixing LLN/CLT) yield posterior consistency and a Bernstein–von Mises type result. The asymptotic variance $V(p_{\text{true}})$ incorporates both the intrinsic $\tilde{\tau}^2(p_{\text{true}})$. □

Remark 16 (Small Samples and Long-Shot Regimes). The results above are asymptotic in nature and require the effective sample size n_n^* to grow without bound. In low-liquidity markets (small n) or extreme-probability regimes (long-shot events with p_{true} near 0 or 1), the Beta posterior can be highly skewed and heavy-tailed, and the normal approximations may be poor. In such regimes, PRISM should be used with caution, relying on full posterior credible intervals rather than Gaussian approximations, and sensitivity analysis with respect to the prior and adjustment parameters is particularly important.

Proposition 13 (Finite-Sample Concentration under Mixing)**.** *Let (Z_t) be the adjusted YES process satisfying the boundedness and geometric α-mixing conditions of Phase 8. Let*

$$\bar{Z}_n = \frac{1}{n}\sum_{t=1}^{n} Z_t, \qquad \mu = \mathbb{E}[Z_t].$$

Then, for all $x > 0$ and $n \geq 1$, there exist constants $C_1, C_2 > 0$ (depending on the mixing coefficients and bounds on Z_t) such that

$$\mathbb{P}\big(|\bar{Z}_n - \mu| > x\big) \leq C_1 \exp\Big(-C_2 n x^2\Big),$$

i.e. a Bernstein-type exponential inequality holds for the empirical average. In particular, for any confidence level $\delta \in (0,1)$, with probability at least $1-\delta$,

$$|\bar{Z}_n - \mu| \lesssim \sqrt{\frac{\log(1/\delta)}{n}},$$

up to constants depending on the mixing structure.

Remark 17 (Practical Implications in Low-n and Long-Shot Regimes). The concentration bound above is asymptotic in spirit: it guarantees that, for moderate n and weak dependence, $\bar{Z}_n$ will concentrate around μ at a rate comparable to the i.i.d. case. However, in low-liquidity markets (small n), or when p is very close to 0 or 1, three issues arise:

- The constants C_1, C_2 may be unfavorable, leading to loose finite-sample bounds.
- The Beta posterior can be highly skewed, so Gaussian approximations to credible intervals may be misleading.
- Long-shot events make the empirical process more volatile relative to the natural scale of p, further weakening normal approximations.

In these situations, PRISM posteriors should be interpreted via full credible intervals and sensitivity checks, rather than relying solely on asymptotic normality or point estimates.

8.15 Metric–Based Robustness for Nonlinear Posterior Updates

The original Lipschitz robustness results in this phase implicitly assumed a linear or affine adjustment of pseudo–counts, so that the posterior mapping from data to distribution was essentially linear. Once nonlinear distortions, mixture posteriors, and dynamic bias layers are introduced (Sections reference, reference, reference), the adjustment of the posterior can no longer be modeled as a simple additive transformation. This subsection replaces the linear Lipschitz arguments with metric–based robustness results formulated in Wasserstein and Hellinger distances for general nonlinear updates.

Setup. Let $\Theta = [0, 1]$ denote the parameter space for the event probability p. For a given dataset $\mathcal{D}$, let $\Pi_{\text{base}}(\cdot \mid \mathcal{D})$ be the *base* PRISM posterior over p, constructed as in Phases 5–6 (e.g. a mixture of Beta components calibrated via stacking and dynamic bias layers). We model nonlinear bias correction at the posterior level via a measurable map

$$\phi : \Theta \to \Theta,$$

which acts on p to produce an adjusted parameter $\phi(p)$.[1]

Definition 9 (Nonlinear Posterior Update via Pushforward)**.** *For a given dataset $\mathcal{D}$ and base posterior $\Pi_{\text{base}}(\cdot \mid \mathcal{D})$, the* nonlinearly adjusted *posterior is defined as the pushforward measure*

$$\Pi_\phi(\cdot \mid \mathcal{D}) := \phi_\# \Pi_{\text{base}}(\cdot \mid \mathcal{D}),$$

i.e. for any Borel set $A \subseteq \Theta$,

$$\Pi_\phi(A \mid \mathcal{D}) = \Pi_{\text{base}}\big(\phi^{-1}(A) \mid \mathcal{D}\big).$$

We are interested in how sensitive $\Pi_\phi(\cdot \mid \mathcal{D})$ is to small changes in $\mathcal{D}$, measured in appropriate probability metrics (e.g. Wasserstein W_1 or Hellinger distance d_H).

Assumption 7 (Base Posterior Robustness in Wasserstein Distance)**.** *There exists a data metric $d_{\mathcal{D}}$ on the space of datasets such that for any two datasets $\mathcal{D}, \mathcal{D}'$,*

$$W_1\big(\Pi_{\text{base}}(\cdot \mid \mathcal{D}), \Pi_{\text{base}}(\cdot \mid \mathcal{D}')\big) \;\leq\; C_{\text{base}}\, d_{\mathcal{D}}(\mathcal{D}, \mathcal{D}'),$$

for some constant $C_{\text{base}} < \infty$. Here W_1 is the 1–Wasserstein distance on Θ with respect to the Euclidean metric.

Assumption reference is a metric formulation of the robustness results established earlier (e.g. those analogous to Theorem 13 and Theorem reference), expressed at the level of the base posterior Π_{base}.

Assumption 8 (Lipschitz Nonlinear Adjustment)**.** *The nonlinear adjustment map $\phi : \Theta \to \Theta$ is globally Lipschitz with constant $L_\phi < \infty$:*

$$|\phi(p_1) - \phi(p_2)| \leq L_\phi |p_1 - p_2| \quad \textit{for all } p_1, p_2 \in \Theta.$$

Theorem 22 (Wasserstein Robustness of Nonlinear Posterior Updates)**.** *Under Assumptions reference and reference, the adjusted posterior mapping $\mathcal{D} \mapsto \Pi_\phi(\cdot \mid \mathcal{D})$ is Lipschitz in W_1:*

$$W_1\big(\Pi_\phi(\cdot \mid \mathcal{D}), \Pi_\phi(\cdot \mid \mathcal{D}')\big) \;\leq\; L_\phi\, C_{\text{base}}\, d_{\mathcal{D}}(\mathcal{D}, \mathcal{D}') \quad \textit{for all } \mathcal{D}, \mathcal{D}'.$$

Proof. By Definition reference, $\Pi_\phi(\cdot \mid \mathcal{D}) = \phi_\# \Pi_{\text{base}}(\cdot \mid \mathcal{D})$. The 1–Wasserstein distance is contracting under Lipschitz maps: for any two measures μ, ν on Θ and any L–Lipschitz map ϕ,

$$W_1(\phi_\# \mu, \phi_\# \nu) \leq L\, W_1(\mu, \nu).$$

Applying this with $\mu = \Pi_{\text{base}}(\cdot \mid \mathcal{D})$, $\nu = \Pi_{\text{base}}(\cdot \mid \mathcal{D}')$, and $L = L_\phi$, we get

$$W_1\big(\Pi_\phi(\cdot \mid \mathcal{D}), \Pi_\phi(\cdot \mid \mathcal{D}')\big) \leq L_\phi\, W_1\big(\Pi_{\text{base}}(\cdot \mid \mathcal{D}), \Pi_{\text{base}}(\cdot \mid \mathcal{D}')\big)\,.$$

Combining this with Assumption reference yields the claimed bound. □

[1] In applications, ϕ may depend on additional covariates or summaries of $\mathcal{D}$; here we treat these as fixed when conditioning on $\mathcal{D}$.

Remark 18 (Lipschitz Regularization in Calibration). Theorem reference shows that the robustness constant for the adjusted posterior is the product $L_\phi C_{\text{base}}$. Thus, in calibrating ϕ (e.g. via empirical risk minimization on historical markets), it is natural to penalize large Lipschitz constants. A practical approach is to include a regularization term of the form

$$\lambda \operatorname{Lip}(\phi) \quad \text{or} \quad \lambda \|\nabla \phi\|_{L^2},$$

in the calibration objective, trading off fit against robustness. This yields an explicit statistical justification for Lipschitz regularization of nonlinear posterior updates.

Extension to Hellinger Distance. In some arguments, it is convenient to work with Hellinger distance d_H between posterior densities. For one–dimensional models with sufficiently smooth densities and strictly monotone ϕ, we can relate Hellinger distances before and after the nonlinear transformation.

Assumption 9 (Smooth Monotone Nonlinear Adjustment)**.** *Assume $\phi : (0,1) \to (0,1)$ is a C^1 diffeomorphism with derivative bounded away from 0 and ∞:*

$$0 < m \le \phi'(p) \le M < \infty \quad \textit{for all } p \in (0,1).$$

Let $\pi_{\text{base}}(p \mid \mathcal{D})$ and $\pi_{\text{base}}(p \mid \mathcal{D}')$ denote posterior densities, and $\pi_\phi(q \mid \mathcal{D})$ the density of $\Pi_\phi(\cdot \mid \mathcal{D})$ under the change of variables $q = \phi(p)$.

Theorem 23 (Hellinger Stability Under Smooth Monotone Transformations)**.** *Under Assumption reference, there exists a constant $C_H = C_H(m, M)$ such that for any two datasets $\mathcal{D}, \mathcal{D}'$,*

$$d_H\big(\Pi_\phi(\cdot \mid \mathcal{D}), \Pi_\phi(\cdot \mid \mathcal{D}')\big) \;\le\; C_H\, d_H\big(\Pi_{\text{base}}(\cdot \mid \mathcal{D}), \Pi_{\text{base}}(\cdot \mid \mathcal{D}')\big)\,.$$

In particular, if the base posterior mapping is Hellinger–Lipschitz in $\mathcal{D}$, then so is the adjusted posterior mapping.

Proof Sketch. Under the change of variables $q = \phi(p)$, densities transform via

$$\pi_\phi(q \mid \mathcal{D}) = \pi_{\text{base}}(\phi^{-1}(q) \mid \mathcal{D})\, \big|\phi^{-1})'(q)\big|.$$

The Hellinger distance between the transformed densities is controlled by the Hellinger distance between the original densities and the Jacobian factors, with constants depending only on bounds for ϕ' and $(\phi^{-1})'$. Using $m \le \phi'(p) \le M$ and standard change-of-variable bounds, one obtains the stated inequality with C_H depending only on (m, M). □

Implications for Pricing Functionals. Let $f : \Theta \to \mathbb{R}$ be a 1–Lipschitz payoff functional of p (e.g. $f(p) = \mathbf{1}\{p > K\}$ smoothed, or a bounded Lipschitz proxy for digital pricing). Then for any two datasets $\mathcal{D}, \mathcal{D}'$,

$$\left| \int f(p)\, \Pi_\phi(dp \mid \mathcal{D}) - \int f(p)\, \Pi_\phi(dp \mid \mathcal{D}') \right| \le W_1\big(\Pi_\phi(\cdot \mid \mathcal{D}), \Pi_\phi(\cdot \mid \mathcal{D}')\big)\,.$$

Combining with Theorem reference yields a bound on the sensitivity of PRISM prices to data perturbations under nonlinear adjustments:

$$\left| \hat{C}_\phi(\mathcal{D}) - \hat{C}_\phi(\mathcal{D}') \right| \le L_\phi C_{\text{base}}\, d_{\mathcal{D}}(\mathcal{D}, \mathcal{D}'),$$

for any Lipschitz payoff functional f used in price computation.

Remark 19 (Failure Modes and Relation to F2–F3)*.* The robustness results in this subsection rely critically on two properties: (i) the base posterior must be stable in Wasserstein or Hellinger distance, and (ii) the nonlinear update ϕ must be Lipschitz (or smooth monotone with bounded derivative). In the strong herding and fast regime–switching regimes (Sections reference and reference), condition (i) fails: small changes in data can produce large changes in the base posterior. In such regimes, no choice of ϕ (Lipschitz or otherwise) can restore uniform robustness. Thus the metric–based Lipschitz results here apply to the "regular" region of the model space where dependence and nonstationarity remain within the bounds of the PRISM assumptions.

8.16 Failure of Robustness Under Strong Herding: A Threshold Auto–Regressive Counterexample

This section establishes that the robustness guarantees of PRISM fail under *strong herding*, understood as threshold–based majority–following behavior in which order flow becomes nearly deterministic once the fraction of YES orders crosses a critical threshold. In this regime, the dependence in the adjusted order process Z_t violates the mixing assumptions (A7)–(A9), the likelihood becomes multimodal, and neither a single–Beta posterior nor any finite mixture of Betas can recover the true event probability uniformly over the strong–herding parameter region.

Definition 10 (Threshold Auto–Regressive Herding Model)**.** *Let $(Z_t)_{t\geq 1}$ denote the effective YES–indicator process entering Stage 1 after behavioral weighting. For parameters (θ, ρ) with $\theta \in (0,1)$ and $\rho \in [0,1]$, define*

$$Z_{t+1} = \begin{cases} 1 & \textit{with probability } \rho \quad \textit{if } \dfrac{1}{t}\displaystyle\sum_{i=1}^{t} Z_i \geq \theta, \\ 0 & \textit{with probability } \rho \quad \textit{if } \dfrac{1}{t}\displaystyle\sum_{i=1}^{t} Z_i < \theta, \\ Ber(p_{\text{true}}) & \textit{with probability } 1-\rho. \end{cases}$$

When ρ is close to 1, the process follows a majority rule as soon as the empirical fraction of YES exceeds the threshold θ, and otherwise follows a pure "NO–cascade". We call this the strong–herding *regime.*

Lemma 10 (Loss of Mixing Under Strong Herding)**.** *Fix $\theta \in (0,1)$ and let $\rho \to 1$. Under the TAR model of Definition reference, the process (Z_t) fails to satisfy α–mixing with any summable mixing rate. In particular,*

$$\alpha_Z(k) \not\to 0 \qquad \textit{as } k \to \infty$$

whenever the event $\{\frac{1}{t}\sum_{i=1}^{t} Z_i \geq \theta\}$ occurs with positive probability.

Proof. When $\rho \to 1$, the transition becomes

$$Z_{t+1} = \mathbf{1}\left\{\frac{1}{t}\sum_{i=1}^{t} Z_i \geq \theta\right\} \quad \text{a.s.},$$

so once the empirical mean crosses θ, the process becomes identically 1 thereafter. Similarly, if the empirical mean remains below θ, the process becomes identically 0. Hence (Z_t) becomes asymptotically constant and perfectly predictable from events arbitrarily far in the past, implying $\alpha_Z(k) = 1$ for all k. Summability fails, completing the proof. □

Lemma 11 (Bimodal Likelihood Under Strong Herding). *Let $L_n(p)$ denote the likelihood of p_{true} based on $Z_1, \dots, Z_n$ under the TAR model. If ρ is sufficiently close to 1, then with positive probability*

$$L_n(p) \;\; \textit{is asymptotically bimodal on } [0,1],$$

with modes concentrated near 0 and 1 corresponding to NO–cascades and YES–cascades respectively.

Proof. From Lemma reference, the process enters a deterministic regime once the empirical mean crosses θ. If it crosses upward, $Z_t = 1$ for all large t, which for the Bernoulli likelihood behaves as if $p_{\text{true}} = 1$. If it crosses downward, $Z_t = 0$ eventually, behaving as if $p_{\text{true}} = 0$. Since both cascades occur with positive probability whenever p_{true} is not exactly equal to θ, the likelihood assigns mass to neighborhoods of both 0 and 1. Bimodality follows. □

Theorem 24 (Impossibility of Uniform Consistency Under Strong Herding). *Let $\Pi_{\text{PRISM}}(\cdot \mid \mathcal{D}_1, \mathcal{D}_2)$ denote the Stage 2 posterior of PRISM, modeled as a finite mixture of Beta distributions. Under the TAR herding model (Definition reference), the following statements hold.*

*(i) (*No uniform consistency*) For any finite mixture of Beta distributions with $R < \infty$ components,*

$$\sup_{p_{\text{true}} \in (0,1)} \mathbb{E}_{p_{\text{true}}} \left[\left| \mathbb{E}[p \mid Z_1, \dots, Z_n] - p_{\text{true}} \right| \right] \not\to 0 \qquad \textit{as } \rho \to 1.$$

*(ii) (*Failure of finite–mixture representation*) The bimodality of the likelihood (Lemma reference) implies that any finite mixture of Betas fails to approximate the true posterior uniformly over the strong–herding region $\{\rho > \rho^*\}$ for any fixed $\rho^* < 1$.*

*(iii) (*Breakdown of robustness theorems*) The Lipschitz–type continuity results of Phase 8 cannot hold under strong herding: no constant $C < \infty$ can satisfy*

$$\left\| \Pi_{\text{PRISM}}(\cdot \mid \mathcal{D}) - \Pi_{\text{PRISM}}(\cdot \mid \mathcal{D}') \right\|_{\text{TV}} \;\le\; C \left\| \mathcal{D} - \mathcal{D}' \right\|$$

for all ρ sufficiently close to 1. The deterministic cascades in strong herding force the total variation distance to jump by 1 when $\mathcal{D}$ crosses the threshold θ.

Proof. (i) From Lemma reference, the likelihood assigns mass to neighborhoods of 0 and 1 with positive probability even when $p_{\text{true}} \in (\theta - \varepsilon, \theta + \varepsilon)$. A finite mixture of Betas cannot track both cascades simultaneously: its posterior mean necessarily lies in a compact subinterval of $(0, 1)$ independent of the data configuration producing the cascades. Thus, the posterior mean cannot converge uniformly to p_{true} as $\rho \to 1$, proving (i).

(ii) Any finite mixture of Betas has unimodal or mildly multimodal densities, but cannot represent a sequence of likelihoods whose mass splits between 0 and 1 in a way that depends discontinuously on the data path. Therefore no finite mixture can uniformly approximate the posterior, yielding (ii).

(iii) Let $\mathcal{D}$ and $\mathcal{D}'$ differ only by whether the empirical mean crosses θ at time t_0. Under strong herding, the posteriors collapse to neighborhoods of 0 and 1 respectively, giving total variation distance equal to 1. Since $\|\mathcal{D} - \mathcal{D}'\|$ can be arbitrarily small (e.g. one flip of a single Bernoulli), no Lipschitz constant can satisfy the inequality uniformly. This proves (iii). □

Remark 20 (Interpretation). The strong–herding regime invalidates the data–generating assumptions required for PRISM's robustness theory. Deterministic cascades destroy mixing, induce bimodal likelihoods, and generate posterior discontinuities that no finite–mixture Beta representation can smooth uniformly. Thus the impossibility theorem above provides a fundamental limitation: PRISM can be consistent and Lipschitz–robust only on subregions of the parameter space where dependence is sufficiently weak and mixing conditions (A7)–(A9) hold.

8.17 Boundary Behavior, Long-Shot Events, and Stabilization

The PRISM posterior for the event probability p is Beta with parameters

$$p \mid \mathcal{I} \sim \mathrm{Beta}(\alpha_n, \beta_n), \qquad \alpha_n = \alpha_0 + y_n^*, \quad \beta_n = \beta_0 + n_n^* - y_n^*,$$

where (y_n^*, n_n^*) are the adjusted counts and (α_0, β_0) are the hybrid prior parameters. In long-shot regimes (p_{true} near 0 or 1) or in very small samples, it is possible for the posterior mass to concentrate near 0 or 1, leading to heavy tails and numerical instability for functions of p (e.g. log-odds or certain risk measures).

We formalize this behavior and describe a simple stabilization based on either truncation or a logit transform.

Lemma 12 (Tail Behavior of Beta Posteriors Near the Boundaries)**.** *Let* Π_n *denote the* $\mathrm{Beta}(\alpha_n, \beta_n)$ *posterior for* p*.*

(a) If $\alpha_n \le 1$*, then the density of* Π_n *behaves like*

$$\pi_n(p) \propto p^{\alpha_n - 1}(1-p)^{\beta_n - 1} \sim p^{\alpha_n - 1} \quad \text{as } p \downarrow 0,$$

so that the left tail near 0 *is heavy whenever* $\alpha_n < 1$*.*

(b) If $\beta_n \le 1$*, then*

$$\pi_n(p) \sim (1-p)^{\beta_n - 1} \quad \text{as } p \uparrow 1,$$

and the right tail near 1 *is heavy whenever* $\beta_n < 1$*.*

(c) If $\alpha_n, \beta_n \ge c > 1$ *uniformly in* n*, then there exists* $\varepsilon > 0$ *such that*

$$\Pi_n\big([0,\varepsilon) \cup (1-\varepsilon, 1]\big) \le C\varepsilon^c,$$

for some constant $C > 0$ *independent of* n*. In particular, the posterior places vanishing mass near* 0 *and* 1 *as* $\varepsilon \downarrow 0$*.*

Proof (Sketch). Parts (a) and (b) follow from the Beta density

$$\pi_n(p) = \frac{1}{\mathrm{B}(\alpha_n, \beta_n)}\, p^{\alpha_n - 1}(1-p)^{\beta_n - 1},$$

and standard asymptotics as $p \to 0$ and $p \to 1$. When $\alpha_n < 1$, the factor $p^{\alpha_n - 1}$ diverges as $p \downarrow 0$, indicating a heavy left tail; an analogous statement holds for $\beta_n < 1$ near 1.

For (c), if $\alpha_n, \beta_n \ge c > 1$ and $p \in (0, \varepsilon)$, then

$$\pi_n(p) \le \frac{1}{\mathrm{B}(\alpha_n, \beta_n)} p^{c-1},$$

and integrating on $(0, \varepsilon)$ gives

$$\Pi_n([0,\varepsilon)) \le C_1 \varepsilon^c$$

for some $C_1 > 0$ that can be chosen uniformly in n due to the compactness of the parameter region. A symmetric argument holds near 1. □

The lemma shows that heavy tails near 0 and 1 arise precisely when the posterior parameters α_n and β_n are small, which can occur under three circumstances: (i) very small effective sample size n_n^*, (ii) extreme long-shot outcomes (e.g. no YES orders in a rare-event market), or (iii) extremely concentrated or misaligned priors.

To mitigate numerical instability and avoid overconfident long-shot posteriors in these regimes, we consider a simple stabilized transform.

Proposition 14 (Stabilized Posterior via Truncation or Logit Transform)**.** *Fix a truncation parameter $\varepsilon \in (0, 1/2)$ and define the truncated interval*

$$I_\varepsilon = [\varepsilon, 1 - \varepsilon].$$

Let Π_n be the $\mathrm{Beta}(\alpha_n, \beta_n)$ posterior and define the truncated posterior Π_n^ε by

$$\Pi_n^\varepsilon(A) = \frac{\Pi_n(A \cap I_\varepsilon)}{\Pi_n(I_\varepsilon)}, \qquad A \subseteq [0,1] \text{ measurable},$$

whenever $\Pi_n(I_\varepsilon) > 0$.

Then:

(a) If $\alpha_n, \beta_n \geq c > 1$ uniformly in n, then

$$\|\Pi_n - \Pi_n^\varepsilon\|_{\mathrm{TV}} = \Pi_n\big([0, \varepsilon) \cup (1 - \varepsilon, 1]\big) \leq C\varepsilon^c,$$

for some $C > 0$ independent of n. Thus, for small ε, the truncated posterior is close in total variation to the original posterior.

(b) Define the logit transform

$$\theta = \log \frac{p}{1-p},$$

and let Λ_n be the induced posterior distribution for θ under Π_n^ε. Then moments of all orders exist for Λ_n, and Λ_n is supported on a compact interval

$$\Theta_\varepsilon = \left[\log \frac{\varepsilon}{1-\varepsilon},\ \log \frac{1-\varepsilon}{\varepsilon}\right].$$

Consequently, functionals of p that are Lipschitz in θ are uniformly bounded and well-behaved under Λ_n.

(c) If $\varepsilon = \varepsilon_n \downarrow 0$ is chosen such that $\varepsilon_n^c \to 0$ and $n_n^* \to \infty$, then the truncated posterior $\Pi_n^{\varepsilon_n}$ remains asymptotically equivalent to Π_n for inference about p_{true} while preventing extreme numerical instability near 0 and 1 at finite n.

Proof (Sketch). For (a), observe that truncation only removes mass near 0 and 1, so the total variation distance equals the probability of the removed regions:

$$\|\Pi_n - \Pi_n^\varepsilon\|_{\mathrm{TV}} = \Pi_n\big([0, \varepsilon) \cup (1 - \varepsilon, 1]\big).$$

The bound then follows directly from part (c) of the lemma, with the same exponent c and an adjusted constant C.

For (b), the truncation ensures that $p \in I_\varepsilon$ almost surely under Π_n^ε. The logit map $p \mapsto \theta = \log(p/(1-p))$ sends I_ε to Θ_ε, a compact interval in $\mathbb{R}$, and is smooth on $(0, 1)$. As a result, all moments of θ under Λ_n exist and are bounded uniformly in n. Any functional of p that can be expressed as a Lipschitz function of θ thus inherits uniform boundedness and stability.

For (c), if $\varepsilon_n^c \to 0$, then the total variation distance between Π_n and $\Pi_n^{\varepsilon_n}$ tends to zero by (a). At the same time, the growing effective sample size n_n^* drives the posterior mass towards p_{true}, and the truncation can be chosen small enough that it does not distort the asymptotic concentration in typical cases (where $p_{\text{true}} \in (0,1)$). Hence the truncated posterior is asymptotically equivalent to the original one for inference while improving finite-sample stability. □

Remark 21 (Practical Guidance for Long-Shot Markets). The analysis above suggests a simple stabilization strategy for PRISM in long-shot or low-liquidity markets:

- Choose a small truncation level ε (e.g. 10^{-4} or 10^{-3}) and work with the truncated posterior Π_n^{ε}.
- When transforming probabilities (e.g. to log-odds), perform the transform on $\theta = \log(p/(1-p))$ under Λ_n rather than directly on p near the boundaries.
- Report both the truncated posterior summaries and the original Beta summaries, especially in cases where α_n or β_n are close to 1 or below.

This keeps the asymptotic properties intact while explicitly addressing the finite-sample instabilities associated with heavy Beta tails near 0 and 1.

8.18 PRISM as a KL Projection: A Formal Information-Theoretic Interpretation

The adjusted PRISM posterior for the event probability p is a Beta distribution of the form

$$p \mid \mathcal{I} \sim \text{Beta}(\alpha_n, \beta_n), \qquad \alpha_n = \alpha_0 + y_n^*, \quad \beta_n = \beta_0 + n_n^* - y_n^*,$$

constructed from the hybrid prior and adjusted parimutuel counts. Up to this point the Beta form has been motivated by conjugacy and interpretability. Here we show that it also admits a fundamental *information-theoretic characterization*: it is the KL-projection of a general posterior onto the Beta family.

Let $\Pi^\star$ denote the (hypothetical) posterior distribution for p that would arise under a fully specified, potentially nonparametric data-generating model with dependence, behavioral distortions, or heterogeneous signals. In a general market this $\Pi^\star$ may not be Beta, and may not be computationally tractable.

PRISM provides a tractable Beta posterior. The following theorem shows that the PRISM posterior coincides with the *I-projection* (information projection) of $\Pi^\star$ onto the Beta family $\mathcal{B} = \{\text{Beta}(a,b) : a, b > 0\}$.

Theorem 25 (KL Projection Theorem for PRISM)**.** *Let $\Pi^\star$ be any posterior distribution on $p \in (0,1)$ with finite mean and finite* log*-moment, and let $\mathcal{B}$ be the family of Beta distributions. Consider the KL divergence*

$$D_{\text{KL}}(\Pi^\star \,\|\, \text{Beta}(a,b)) = \int_0^1 \log \frac{d\Pi^\star}{d\text{Beta}(a,b)}(p)\, d\Pi^\star(p),$$

defined whenever $\Pi^\star$ is absolutely continuous with respect to $\text{Beta}(a,b)$*.*

Define the KL-projection of $\Pi^\star$ onto $\mathcal{B}$ as

$$(a^\dagger, b^\dagger) \in \arg\min_{a,b>0} D_{\text{KL}}(\Pi^\star \,\|\, \text{Beta}(a,b)).$$

Assume that $\Pi^\star$ has mean $m^\star$ and inverse second moment $M^\star = \mathbb{E}_{\Pi^\star}[p^{-1} + (1-p)^{-1}] < \infty$. Then:

(a) *The minimizer* $(a^\dagger, b^\dagger)$ *exists and is unique.*

(b) *The KL-projection satisfies*

$$\frac{a^\dagger}{a^\dagger + b^\dagger} = m^\star, \qquad \frac{a^\dagger b^\dagger}{(a^\dagger + b^\dagger)^2(a^\dagger + b^\dagger + 1)} = \mathrm{Var}_{\Pi^\star}(p),$$

i.e. the minimizing Beta distribution matches the mean and variance of $\Pi^\star$*.*

(c) *If* $\Pi^\star$ *is generated by a hybrid prior and adjusted counts* y_n^*, n_n^* *(with mixing or dependence), then the PRISM posterior* $\mathrm{Beta}(\alpha_n, \beta_n)$ *coincides with the KL-projection:*

$$(\alpha_n, \beta_n) = (a^\dagger, b^\dagger).$$

(d) *In particular, for large effective sample size* n_n^**, PRISM selects, among all Beta distributions, the one closest in KL sense to the ideal but intractable* $\Pi^\star$*.*

Proof (Sketch). Part (a) follows from strict convexity of the KL divergence in (a, b) on the Beta family. For (b), writing the KL divergence explicitly and differentiating under the integral yields two first-order conditions:

$$\frac{\partial}{\partial a} D_{\mathrm{KL}} = 0 \quad \Rightarrow \quad \mathbb{E}_{\Pi^\star}[\log p] = \psi(a) - \psi(a+b),$$

$$\frac{\partial}{\partial b} D_{\mathrm{KL}} = 0 \quad \Rightarrow \quad \mathbb{E}_{\Pi^\star}[\log(1-p)] = \psi(b) - \psi(a+b),$$

where ψ is the digamma function. Using identities for Beta means and variances, these conditions imply the matching of mean and variance stated in (b).

For (c), under a Beta–Binomial model (even with adjusted counts), the exact posterior for p is $\mathrm{Beta}(\alpha_n, \beta_n)$. Thus if $\Pi^\star$ arises from such updating under a fully specified likelihood, $\Pi^\star$ is already in $\mathcal{B}$ and the unique KL minimizer is precisely $\mathrm{Beta}(\alpha_n, \beta_n)$.

When $\Pi^\star$ is more general (due to dependence, behavioral structures, or nonparametric components), the PRISM posterior can be interpreted as the projection of $\Pi^\star$ onto the Beta family using the adjusted empirical mean and variance $(m^\star, \mathrm{Var}_{\Pi^\star})$ induced by the PRISM adjustment rules. This identifies (α_n, β_n) with the unique minimizer $(a^\dagger, b^\dagger)$. □

Remark 22 (Interpretation and Novelty)*.* This theorem provides an information-theoretic justification for the PRISM posterior beyond conjugacy. Even when the true posterior $\Pi^\star$ is nonparametric or analytically intractable, PRISM delivers the *closest Beta distribution in KL divergence.*

Thus the Beta form is not merely a convenient algebraic choice, but the information-projection that preserves the two most important moments of the ideal posterior under the PRISM-adjusted signal process.

This interpretation also helps explain why the hybrid prior and adjustment mechanism remain stable even under dependence or behavioral distortions: PRISM selects the "least distorted" Beta posterior compatible with the adjusted empirical information.

Remark 23 (Limitations of the KL Projection View)*.* Interpreting PRISM as a KL projection onto the Beta family is useful but also restrictive. It guarantees that, among all Beta distributions, the chosen posterior preserves certain moments of a more complex underlying posterior. However, it does not claim that the Beta family is rich enough to capture all features of the true posterior under strong dependence, multimodal priors, or adversarial behavior. In settings where such features are important, the KL-projection perspective should be viewed as an approximation tool rather than a full description of market beliefs, and more flexible models (e.g. MCMC-based or variational) may be warranted.

8.19 Alternatives to Beta Conjugacy under Dependence

While PRISM deliberately uses the Beta–Binomial structure for tractability and interpretability, other approaches can, in principle, accommodate richer dependence at the cost of computational complexity:

- **MCMC with dependent likelihoods.** One can specify an explicit dependent model for the order sequence, such as an Ising or Markov random field for YES/NO decisions, and sample from the posterior for p via Gibbs or Metropolis–Hastings. This yields a more flexible posterior but requires careful tuning and may be slow in large markets.
- **Variational approximations.** Variational Bayes can approximate complex posteriors with factored or low-rank distributions, trading off accuracy for speed. In this setting, one could approximate the joint posterior over $(p, \text{latent states})$ with a product of a Beta distribution for p and a tractable family for the latent dependence structure.
- **Composite likelihoods.** Composite or pseudo-likelihood methods replace the full joint likelihood with products of low-dimensional marginals or conditionals, offering a compromise between full MCMC and the single-sufficient-statistic approach of PRISM.

PRISM chooses Beta conjugacy as a deliberate design decision: it provides closed-form updates, interpretable pseudo-counts, and an information-projection interpretation, while recognizing that more flexible likelihood-based methods are possible when computational resources and data volume permit.

8.20 Asymptotics Under Ergodic Regime Switching and an Impossibility Result

The dynamic bias layer of Section reference models (δ_t, ψ_t) as a hidden Markov chain (S_t) with a finite state space $\mathcal{S} = \{1, \ldots, R\}$ and regime–specific parameters (δ_r, ψ_r). This section establishes (i) a Bernstein–von Mises–type result under ergodic regime switching (mild nonstationarity), and (ii) an impossibility result when regimes switch so quickly that no regime accumulates enough information for consistent learning.

Assumption 10 (Ergodic Regime Switching and Regularity)**.** *Let $(S_t)_{t\geq 1}$ be an irreducible, aperiodic Markov chain on $\mathcal{S} = \{1, \ldots, R\}$ with transition matrix P and unique stationary distribution $\pi = (\pi_1, \ldots, \pi_R)$. Assume:*

*(i) (*Ergodicity*) For each $r, s \in \mathcal{S}$ there exists $k \geq 1$ such that $(P^k)_{rs} > 0$, and the chain is aperiodic. Consequently, for any initial distribution,*

$$\mathbb{P}(S_t = r) \to \pi_r \quad \text{as } t \to \infty.$$

*(ii) (*True parameter vector*) Each regime r has a true parameter $\theta_r^\star = (\delta_r^\star, \psi_r^\star)$ in a compact subset $\Theta_r \subset \mathbb{R}^{d_r}$, and the emission model $p(\mathcal{D}_t \mid S_t = r, \theta_r^\star)$ coincides with the Stage 1/Stage 2 correction structure of PRISM.*

*(iii) (*Identifiability and smoothness*) The mapping $\theta_r \mapsto p(\mathcal{D}_t \mid S_t = r, \theta_r)$ is identifiable and C^2 in a neighborhood of $\theta_r^\star$, with Fisher information matrix $I_r(\theta_r^\star)$ positive definite.*

*(iv) (*Finite moments*) The emission log-likelihoods have finite second moments under the true model.*

We are interested in a smooth functional of the regime parameters, such as the long–run average bias

$$\bar{\delta}^\star := \sum_{r=1}^{R} \pi_r \, \delta_r^\star,$$

or more generally a differentiable functional $\varphi(\theta_1, \dots, \theta_R)$ with $\varphi : \Theta_1 \times \cdots \times \Theta_R \to \mathbb{R}$.

Theorem 26 (Bernstein–von Mises Theorem under Ergodic Regime Switching)**.** *Under Assumption reference, suppose the PRISM posterior over $(\theta_1, \dots, \theta_R)$ is proper and assigns positive prior density in a neighborhood of the true parameter vector $(\theta_1^\star, \dots, \theta_R^\star)$. Let*

$$\hat{\theta}_r(T) \quad (r = 1, \dots, R)$$

denote the (quasi–)maximum likelihood or posterior mean estimator of θ_r based on data $\mathcal{D}_{1:T} = (\mathcal{D}_1, \dots, \mathcal{D}_T)$, and define $\hat{\varphi}(T) := \varphi(\hat{\theta}_1(T), \dots, \hat{\theta}_R(T))$. Then, as $T \to \infty$,

*(i) (*Consistency*)*

$$\hat{\varphi}(T) \xrightarrow{\mathbb{P}} \varphi(\theta_1^\star, \dots, \theta_R^\star).$$

*(ii) (*Asymptotic normality*)*

$$\sqrt{T}\left(\hat{\varphi}(T) - \varphi(\theta_1^\star, \dots, \theta_R^\star)\right) \xrightarrow{d} \mathcal{N}(0, V_\varphi),$$

where V_φ is the asymptotic variance obtained by the delta method applied to the joint asymptotic distribution of $(\hat{\theta}_1(T), \dots, \hat{\theta}_R(T))$.

*(iii) (*Bernstein–von Mises*) The posterior distribution of $\varphi(\theta_1, \dots, \theta_R)$ given $\mathcal{D}_{1:T}$ converges in total variation to the normal law $\mathcal{N}(\hat{\varphi}(T), V_\varphi/T)$:*

$$\left\| \Pi\big(\varphi(\theta_1, \dots, \theta_R) \in \cdot \,\big|\, \mathcal{D}_{1:T}\big) - \mathcal{N}\big(\hat{\varphi}(T), V_\varphi/T\big) \right\|_{\mathrm{TV}} \xrightarrow{T\to\infty} 0.$$

Proof Sketch. Under Assumption reference, the hidden Markov model $(S_t, \mathcal{D}_t)$ is an ergodic HMM with finite state space and regular parametric emission densities. Standard results for HMMs imply that the (quasi–)maximum likelihood estimator of each θ_r is consistent and asymptotically normal, with an information matrix determined by long–run frequencies π_r and the per–regime Fisher information $I_r(\theta_r^\star)$. The joint asymptotic normality of $(\hat{\theta}_1(T), \dots, \hat{\theta}_R(T))$ then yields asymptotic normality of $\hat{\varphi}(T)$ via the delta method. The Bayesian Bernstein–von Mises conclusion follows from general BvM theorems for HMMs with finite state space and regular parametric families, together with the prior positivity condition. Full details follow the usual template for BvM in dependent data models; we omit these for brevity. □

Remark 24 (Application to Dynamic PRISM). In the dynamic PRISM setting of Section reference, the regime–specific parameters $\theta_r^\star$ encode bias corrections $(\delta_r^\star, \psi_r^\star)$ and possibly additional structural quantities. The functional φ may be taken as the long–run average distortion $\bar{\delta}^\star = \sum_r \pi_r \delta_r^\star$, or as a mapping from the collection $(\theta_r^\star)$ to a long–run effective event probability $p_{\mathrm{eff}}^\star$ under dynamic corrections. Theorem reference then ensures that the PRISM posterior for $\bar{\delta}$ (or p_{eff}) concentrates and becomes asymptotically normal despite mild nonstationarity induced by regime switching.

We now show that this positive result has a natural limit: if regime switching is too fast for any regime to accumulate information, no sequential estimator or posterior can remain uniformly well calibrated.

Assumption 11 (Fast Regime Switching with Vanishing Dwell Time)**.** *Let (S_t) be a Markov chain on $\mathcal{S}$ with transition matrix P depending on T such that:*

(i) The minimum expected dwell time in each state is uniformly bounded:

$$\sup_T \max_{r \in \mathcal{S}} \mathbb{E}_T \big[\textit{ time spent in state } r \textit{ up to } T \big] \; < \; C < \infty.$$

(ii) The regime–specific parameters $\theta^\star_r(T)$ are allowed to vary with T, remaining in a compact set, and emissions are generated from $p(\mathcal{D}_t \mid S_t, \theta^\star_{S_t}(T))$.

In words, regimes switch so frequently that the average number of visits to any given state does not grow with T.

Theorem 27 (Impossibility of Uniform Consistency Under Fast Regime Switching)**.** *Under Assumption reference, consider any sequential estimator sequence $\hat{\theta}(T)$ or any Bayesian posterior sequence $\Pi_T(\cdot \mid \mathcal{D}_{1:T})$ for a regime–dependent parameter functional $\varphi(\theta^\star_1(T), \dots, \theta^\star_R(T))$ (e.g. a regime–specific bias $\delta^\star_r(T)$ or a regime–specific event probability $p^\star_r(T)$). Then there exists a choice of parameter sequences $\{\theta^\star_r(T)\}$ such that*

$$\limsup_{T \to \infty} \sup_{\{\theta^\star_r(T)\}} \mathbb{E}\Big[\big| \hat{\varphi}(T) - \varphi(\theta^\star_1(T), \dots, \theta^\star_R(T)) \big| \Big] \; > \; 0,$$

and similarly, no sequence of posteriors Π_T can concentrate around the true functional uniformly over $\{\theta^\star_r(T)\}$. In particular, there is no uniformly consistent dynamic PRISM estimator or posterior in this fast switching regime.

Proof Sketch. Under Assumption reference, the expected number of observations generated in any fixed regime r up to time T is uniformly bounded by C. Therefore, for any estimator or posterior targeting a regime–specific parameter (or a functional that depends nontrivially on the per–regime values), the effective sample size per regime does not grow with T. By standard parametric lower bounds, no estimator can achieve vanishing risk uniformly over the parameter sequences $\{\theta^\star_r(T)\}$ when each regime is observed only $O(1)$ times. One can construct pairs of parameter sequences that are indistinguishable from the data but induce separated values of the functional φ, forcing a nonzero lower bound on the estimation error. The same argument applies to Bayesian posteriors: with bounded regime information, the posterior cannot concentrate uniformly on the true functional. $\square$

Remark 25 (Interpretation for Dynamic PRISM). Theorem reference and Theorem reference identify a *feasible* and an *infeasible* nonstationary regime for PRISM. Under ergodic regime switching with growing effective sample size per regime, dynamic PRISM remains consistent and asymptotically normal for smooth functionals of the regime parameters. When regime switching becomes so fast that no regime accumulates information, no sequential estimator or posterior can be uniformly well calibrated. In practice, PRISM should therefore treat fast, high–frequency structural breaks as a regime where the bias layer and posterior must be explicitly flagged as fragile, and any pricing output should carry a warning about nonstationary uncertainty that cannot be statistically resolved.

References

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.