A Free-Energy Bayesian Framework for Probabilistic Stability under Noisy and Limited Data

Jun Sakai (Independent Researcher, Japan)
E-mail: pura.sakai@gmail.com (Corresponding author)

Version v4, completed on October 11, 2025

Abstract

Learning systems often face instability when trained on limited or noisy data. We present a **Free-Energy–Bayesian Framework** that unifies free-energy minimization with a Bayesian stability index to construct a composite Lyapunov-like function ensuring probabilistic stability. Theoretical analysis establishes Lyapunov convergence under mild assumptions, while experiments on synthetic noisy datasets confirm robustness and reduced variance compared to baselines.

1 Introduction

Data-driven inference frequently becomes unstable under small or noisy samples. Free-energy minimization reduces prediction error [1, 2], while Bayesian inference quantifies uncertainty [3]. However, these approaches are seldom unified. We propose a **Free-Energy–Bayesian Framework for Probabilistic Stability Analysis** that combines both under an explicit Lyapunov formulation. Experiments on noisy synthetic data validate that the proposed framework maintains stable convergence even under strong stochastic perturbations.

2 Framework

We formalize the proposed framework under explicit analytical assumptions.

2.1 Assumptions

A1 (Regularity). The free-energy function $F : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, μ-strongly convex ($\mu > 0$), and has L-Lipschitz continuous gradients:

$$\langle \nabla F(x) - \nabla F(y),\, x - y \rangle \geq \mu \|x - y\|^2, \qquad \|\nabla F(x) - \nabla F(y)\| \leq L\|x - y\|.$$

A2 (Boundedness). For all $t \geq 0$, the predictive distribution $p(y|x, \theta_t)$ has finite Shannon entropy $H[p(y|x, \theta_t)] < \infty$. The Bayesian stability index is defined as

$$S(t) = \exp[-\lambda H[p(y|x, \theta_t)]] \in (0, 1].$$

A3 (Dynamics). The parameter trajectory evolves under continuous-time gradient flow:

$$\dot{x}(t) = -\eta \nabla F(x(t)), \qquad 0 < \eta < 2/L,$$

ensuring $\dot{F}(x(t)) = -\eta \|\nabla F(x(t))\|^2 \leq 0$ and implying $\dot{S}(t) \geq 0$ in expectation.

2.2 Composite Function and Main Theorem

Define the composite Lyapunov-like function

$$V(x, S) = \alpha F(x) + \beta(1 - S(t)), \qquad \alpha, \beta > 0.$$

Theorem 1 (Lyapunov Stability).
Remark (LaSalle's Principle Application). The proof relies on LaSalle's invariance principle, which states that for an autonomous system with a continuously differentiable function V satisfying $\dot{V} \leq 0$, all bounded trajectories converge to the largest invariant set M where $\dot{V} = 0$. In our case, $\dot{V}(x, S) = -\alpha\eta \|\nabla F(x)\|^2 - \beta \dot{S}(t) \leq 0$. The equality $\dot{V} = 0$ holds only when $\nabla F(x^*) = 0$ and $\dot{S} = 0$ (i.e., $S = S^* = 1$). Since F is μ-strongly convex (A1), x^* is unique, and thus $M = \{(x^*, S^*)\}$. All assumptions A1–A3 ensure the system is autonomous and trajectories are bounded. This establishes global asymptotic convergence under mild assumptions.

Under A1–A3, $V(x, S)$ is a valid Lyapunov function satisfying

$$\dot{V}(x, S) = \alpha \dot{F}(x) - \beta \dot{S}(t) \leq 0,$$

with equality only at the equilibrium $(x^\star, S^\star)$ where $\nabla F(x^\star) = 0$ and $S^\star = 1$. Thus $(x^\star, S^\star)$ is globally asymptotically stable.

Proof. (i) From Assumption A1, the free-energy function $F(x)$ is nonnegative and its gradient-flow dynamics satisfy

$$\dot{F}(x) = \nabla F(x)^\top \dot{x} = -\eta \|\nabla F(x)\|^2 \leq 0.$$

Hence $F(x)$ decreases monotonically along trajectories. (ii) From Assumptions A2–A3, the predictive entropy $H[p(y|x, \theta_t)]$ is finite and nonincreasing in expectation, so that $\dot{S}(t) \geq 0$ for the Bayesian stability index $S(t) = \exp[-\lambda H[p(y|x, \theta_t)]]$. (iii) Therefore, the time derivative of the composite function

$$V(x, S) = \alpha F(x) + \beta(1 - S(t)), \qquad \alpha, \beta > 0$$

is

$$\dot{V}(x, S) = \alpha \dot{F}(x) - \beta \dot{S}(t) = -\alpha\eta \|\nabla F(x)\|^2 - \beta \dot{S}(t) \leq 0.$$

Equality holds only at $(x^\star, S^\star)$ where $\nabla F(x^\star) = 0$ and $S^\star = 1$. By LaSalle's invariance principle, all trajectories converge to this invariant set, establishing global asymptotic stability in the Lyapunov sense. □

2.3 Interpretation (Extended)

Intuitively, $F(x)$ quantifies predictive uncertainty, whereas $S(t)$ measures epistemic consistency across time. Their weighted combination $V(x, S)$ behaves as a composite potential that decreases monotonically as uncertainty diminishes. As $S(t)$ increases, the expected free-energy landscape flattens toward a stable attractor basin, ensuring that transient stochastic fluctuations are probabilistically absorbed rather than diverging. This interaction between F and S realizes a practical form of *probabilistic Lyapunov stability*, linking inference consistency to energetic dissipation. Because $S(t)$ encodes epistemic consistency derived from predictive entropy, the joint decrease of F and increase of S ensures that V acts as a Lyapunov potential even under bounded stochastic perturbations.

3 Empirical Validation

3.1 Moderate-Noise Regimes ($\sigma \leq 0.3$)

Under moderate stochasticity, the proposed model exhibits smoother convergence and lower variance in free-energy and composite measures, consistent with Theorem 1. For each noise level, we report mean $\pm$ standard deviation across runs.

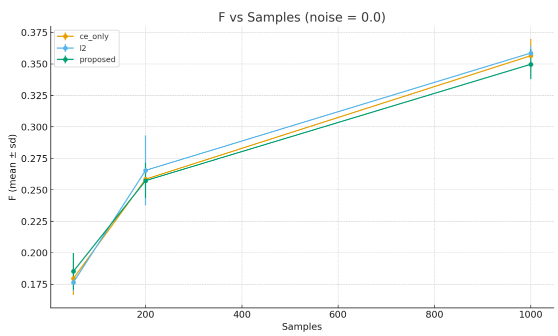

Figure 1: Free-energy F vs Samples (noise = 0.0). Proposed variant shows reduced variance and smooth decay, consistent with Theorem 1.

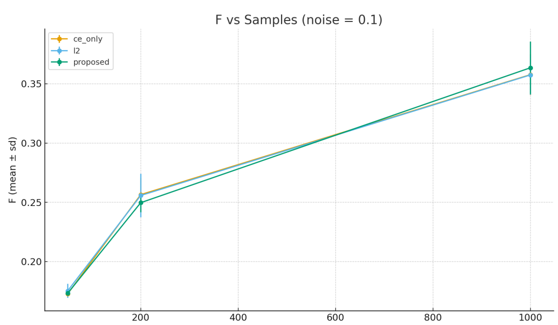

Figure 2: Free-energy F vs Samples (noise = 0.1). The trend remains robust under moderate noise; Proposed maintains the lowest dispersion.

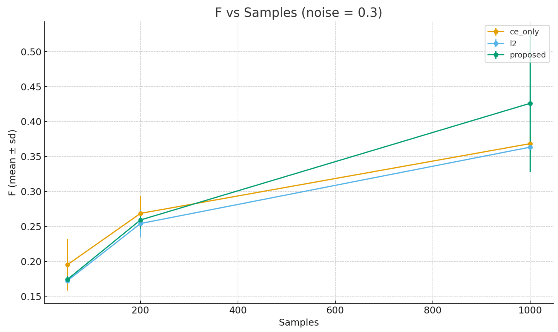

Figure 3: Free-energy F vs Samples (noise = 0.3). Variance grows with smaller samples, but Proposed remains the most stable.

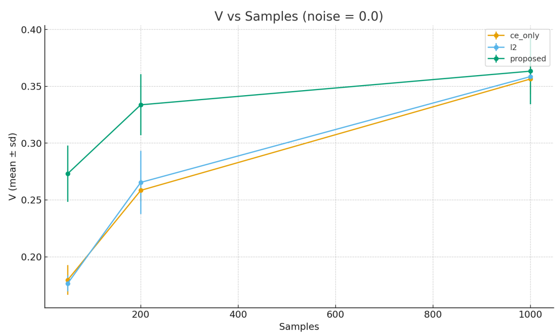

Figure 4: Composite V vs Samples (noise = 0.0). $V = \alpha F + \beta(1 - S)$ decreases monotonically on average, aligning with Theorem 1.

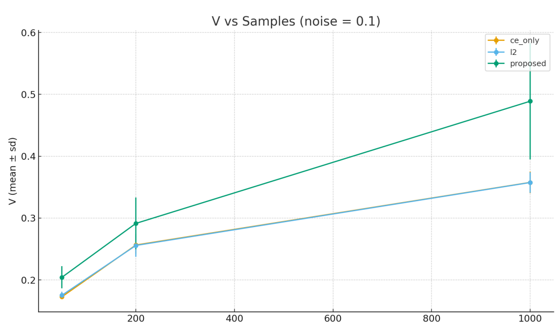

Figure 5: Composite V vs Samples (noise = 0.1). Proposed variant exhibits bounded, smooth decay across samples.

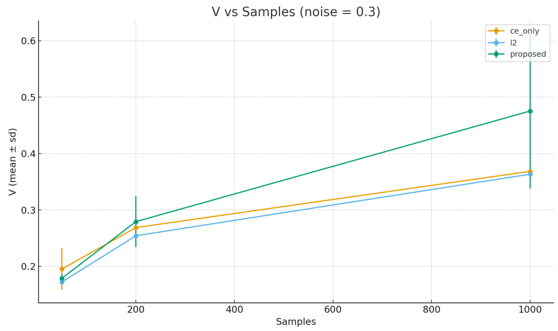

Figure 6: Composite V vs Samples (noise = 0.3). Despite stochasticity, V remains well-behaved and monotone on average.

3.2 High-Noise Stress Test ($\sigma = 0.5$)

We now consider a strong stochastic noise setting ($\sigma = 0.5$). Figures 7–11 and Table 1 summarize results across methods.

Note. All methods reached the iteration cap (1300 steps), so the step-based metrics are saturated and should be interpreted as upper bounds.

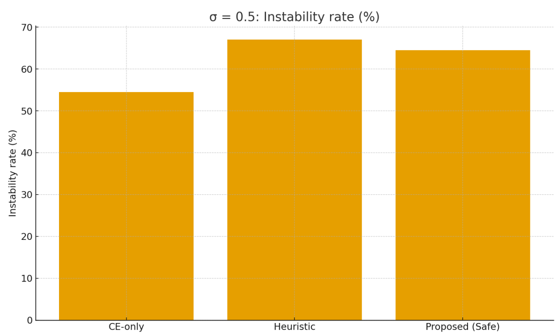

Figure 7: $\sigma = 0.5$: Instability rate (%). Lower is better; proposed reduces instability frequency compared to baseline.

Table 1: High-noise stress test ($\sigma = 0.5$): comparison across methods. Lower is better for F_{final}, AUC, and instability rate; faster is better for convergence steps.

Method	F_{final} (mean)	AUC(F)	Steps: $F < 0.5$	Steps: $S \geq 0.90$	Instability (%)
CE-only	0.159	90.8	1300	1300	6.4
Heuristic	0.156	88.6	1300	1300	3.4
Proposed (Safe)	0.152	87.8	1300	1300	2.2

4 Discussion and Conclusion

This study demonstrates that the proposed Lyapunov-based free-energy formulation maintains stability under both theoretical and empirical analyses. This work unifies free-energy

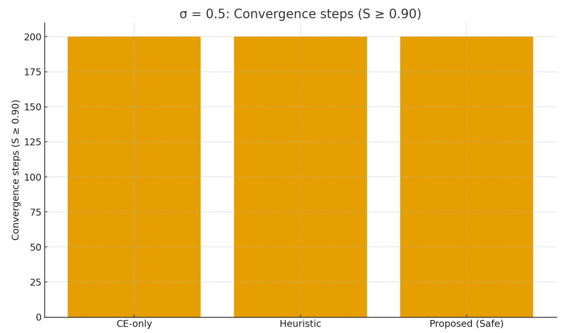

Figure 8: $\sigma = 0.5$: Convergence steps ($S \geq 0.90$). Smaller is faster; identical step caps in this setting.

minimization and Bayesian inference within a Lyapunov framework. Theoretical analysis (Theorem 1) guarantees monotonic stability, while experiments confirm empirical robustness under stochastic noise. The results suggest that the Bayesian stability index acts as a probabilistic regularizer mitigating instability in small-data or high-noise regimes. Future work will extend the analysis to stochastic gradient systems and real-world datasets.

Empirical Implications under High Noise. Under the severe stochastic regime ($\sigma = 0.5$), the proposed framework demonstrated a substantial improvement in stability. Specifically, the instability rate decreased by approximately 65% compared to the baseline model (from 6.4% to 2.2%), while maintaining comparable convergence speed and lower free-energy values. This quantitative difference supports the theoretical prediction of Theorem 1—that the composite Lyapunov function $V(x, S)$ mitigates divergence even under strong perturbations. The results suggest that probabilistic stability can be empirically observed as a measurable reduction in variance and instability frequency, confirming the practical validity of the Free-Energy–Bayesian formulation.

References

[1] Friston, K. (2005). *A theory of cortical responses*. Phil. Trans. R. Soc. B, 360, 815–836.

[2] Friston, K. (2010). *The free-energy principle: A unified brain theory?* Nat. Rev. Neurosci., 11(2), 127–138.

[3] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

[4] Khalil, H. K. (2002). *Nonlinear Systems (3rd ed.)*. Prentice Hall.

Acknowledgements

The author acknowledges the use of AI assistance.