

Policy Selection in the Latent Action Space: A Unifying Theory of Thought, Inference, and Sparse Agency

Flyxion

November 10, 2025

Abstract

All cognitive processes are formalized as sparse Bayesian policy selection within a latent action space (Friston & Kiebel, 2009). Commencing with the foundational conceptualization of thought as Bayesian inference over policies, the exposition progresses through sparsity constraints (Tibshirani, 1996), variational approximations, equivalence to active inference via expected free energy minimization (Friston et al., 2011, 2017), integration with control-as-inference paradigms (Kappen, 2005; Levine, 2018; Todorov, 2009), neurobiological implementations (Beggs & Plenz, 2004; Buzsáki & Moser, 2013; Raichle et al., 2001), epistemic-pragmatic trade-offs (Pezzulo et al., 2015; Schwartenbeck et al., 2015), and broader implications for cognition and psychopathology. A concrete instantiation is provided within the Relativistic Scalar–Vector Plenum (RSVP) dynamical substrate, deriving exact LASSO thresholds for policy activation and demonstrating phase transitions in cognitive agency. Extensions cover nonlinear atom interference (binding), temporal recursion (working memory), cognitive criticality, and thermodynamic computation. The synthesis demonstrates that perception, reasoning, planning, imagination, and decision-making constitute a singular computational principle: the constrained selection of latent policies under uncertainty, optimized through free-energy minimization with enforced parsimony (Clark, 2013). Twenty-seven falsifiable predictions and five experimental protocols are presented.

1 Policy Selection in the Latent Action Space: Foundational Theory

Cognitive activity, in its entirety, may be construed as the selection of a policy—a mapping from states to actions—within a space that encompasses both external motor outputs and internal latent operations (Friston et al., 2011). A policy $\pi \in \Pi$ defines a trajectory of transitions, whether manifest in overt behavior or confined to internal simulations such as conceptual chaining, memory retrieval, or attentional reallocation.

Under conditions of uncertainty, policy selection adheres to Bayesian principles, conditioned on observational history $o_{1:T}$:

$$\pi^* = \arg \max_{\pi \in \Pi} P(\pi | o_{1:T}),$$

where the posterior $P(\pi | o_{1:T}) \propto P(o_{1:T} | \pi)P(\pi)$. Here, $P(\pi)$ encodes prior inductive biases, habits, or structural constraints, while $P(o_{1:T} | \pi)$ evaluates the policy’s capacity to predict or explain observations, including counterfactual outcomes derived from internal rollouts (Friston & Kiebel, 2009).

This formulation posits that thought is not a distinct computational module but inference over latent causes (policies) in a generative model of the world (Clark, 2013). The brain functions as a policy pruner, not an enumerator, selecting among a sparse library of stochastic policy atoms via amortized inference, predictive coding, and neuromodulatory precision weighting (Parr & Friston, 2019).

2 Sparse Bayesian Policy Selection in the RSVP Framework: Linearization, LASSO Thresholds, and a Numerical Demonstration

The abstract formulation that cognition corresponds to sparse Bayesian selection over latent policies becomes mechanically meaningful when embedded inside a dynamical substrate. In the Relativistic Scalar–Vector Plenum (RSVP) formalism, “thought” is not abstract search but intervention on a coupled field system whose state evolves according to entropy–potential–flow dynamics. To make this correspondence explicit—and falsifiable—we construct a tractable control–inference problem directly in the RSVP field equations, derive the sparse posterior over policy amplitudes, and demonstrate the emergence of a sharp LASSO-style activation threshold governing which policy atoms survive selection (Tibshirani, 1996).

2.1 Selecting a Canonical RSVP PDE and Linearizing Around the Plenum

A minimal but faithful RSVP system must express (i) the relaxation of scalar potential Φ toward a background density, (ii) vector flow \mathbf{v} driven by potential gradients, and (iii) entropy production dynamically coupled to both. On a 1D spatial domain $\Omega = [0, L]$ with periodic boundaries, a controlled stochastic form that preserves these roles is:

$$\begin{aligned}\partial_t \Phi &= -\mathbf{v} \cdot \nabla \Phi + D_\Phi \nabla^2 \Phi - \kappa(\Phi - \Phi_0) + B_\Phi[u] + \xi_\Phi, \\ \partial_t \mathbf{v} &= -(\mathbf{v} \cdot \nabla) \mathbf{v} + D_v \nabla^2 \mathbf{v} - \nabla \Phi + B_v[u] + \xi_v, \\ \partial_t S &= \nabla \cdot (S \mathbf{v}) + D_S \nabla^2 S + \sigma(\nabla \Phi)^2 + B_S[u] + \xi_S,\end{aligned}$$

where ξ_\bullet are spatially correlated fluctuations. Negentropic control is injected only through Φ and S , enforcing an antagonistic relation between potential and entropy:

$$B_\Phi[u] = u(x, t), \quad B_v[u] = 0, \quad B_S[u] = -\eta u(x, t),$$

with $\eta > 0$ ensuring that policy activation raises potential while locally suppressing entropy.

To expose the inferential geometry, we linearize about the spatially homogeneous plenum baseline $(\bar{\Phi}, \bar{\mathbf{v}}, \bar{S}) = (\Phi_0, 0, S_0)$. Writing perturbations $\delta\Phi = \Phi - \Phi_0$ and discarding second-order terms gives:

$$\begin{aligned}\partial_t \delta\Phi &= D_\Phi \nabla^2 \delta\Phi - \kappa \delta\Phi + \sum_k a_k \psi_k(x) \delta(t - t_0) + \xi_\Phi, \\ \partial_t \delta\mathbf{v} &= D_v \nabla^2 \delta\mathbf{v} - \nabla \delta\Phi + \xi_v, \\ \partial_t \delta S &= D_S \nabla^2 \delta S + S_0 \nabla \cdot \delta\mathbf{v} + 2\sigma \nabla \Phi_0 \cdot \nabla \delta\Phi - \eta \sum_k a_k \psi_k(x) \delta(t - t_0) + \xi_S.\end{aligned}$$

Policies are not arbitrary functions but expansions in a dictionary of normalized, spatially localized policy atoms $\psi_k(x)$, taken here to be Gaussians:

$$\psi_k(x) = \exp\left(-\frac{(x - \mu_k)^2}{2\sigma_k^2}\right), \quad \|\psi_k\|_{L^2} = 1.$$

A policy is thus a coefficient vector $a = (a_1, \dots, a_M)$ applied impulsively at the present moment t_0 : $u(x, t) = \sum_k a_k \psi_k(x) \delta(t - t_0)$.

2.2 Observation Model and Reduction to a Linear–Gaussian Inference Problem

Suppose the agent evaluates candidate policies by their effect on a weighted regional readout of the potential field after a short horizon Δt :

$$o = \int_{\Omega} w(x) \delta\Phi(x, \Delta t) dx,$$

where $w(x)$ is itself one of the atoms, e.g., $w(x) = \psi_j(x)$, meaning the agent cares about modulating potential in one specific spatial mode. For short horizons, diffusion–relaxation spreads but does not rotate energy between well-separated atoms; the response projects nearly diagonally in the atom basis. The predictive distribution therefore reduces to a linear–Gaussian form:

$$p(o | a) = \mathcal{N}(a_j, s),$$

where s aggregates process and measurement uncertainty. The agent holds a soft goal on this observation $o^* \sim \mathcal{N}(\mu^*, s^*)$, specifying the preferred potential displacement in this mode.

2.3 Expected Free Energy with an ℓ_1 Metabolic Prior

Neglecting epistemic terms for clarity, the generative-risk component of expected free energy is

$$G(a) = \frac{1}{2s^*} (a_j - \mu^*)^2 + \text{const.}$$

Sparse policy costs enter through a metabolic/complexity prior which, when chosen as Laplacian, induces ℓ_1 regularization (Tibshirani, 1996):

$$C(a) = \lambda \|a\|_1 = \lambda \sum_k |a_k|.$$

The posterior over policy coefficients is therefore

$$Q(a) \propto \exp(-G(a) - \lambda \|a\|_1),$$

giving a MAP estimator identical to the LASSO objective (Boyd et al., 2011). For nearly orthogonal atoms, the coefficients decouple, and the active atom solving the agent’s goal obeys the closed-form soft-threshold solution:

$$a_j^* = \text{sign}(\mu^*) \max(|\mu^*| - s^* \lambda, 0).$$

2.4 The Sparsity Phase Transition

The system displays a sharp policy-selection threshold:

$$a_j^* = 0 \quad \text{iff} \quad \lambda \geq \lambda^* = \frac{|\mu^*|}{s^*}.$$

Below this threshold, the plenum is modulated; above it, control collapses and the null policy is selected. This is not a smooth decay of influence—it is a first-order cognitive phase transition between expressive control and abstention. In RSVP terms:

- Low $\lambda \rightarrow$ rich but metabolically expensive plenum shaping (creative, exploratory thinking)
- High $\lambda \rightarrow$ entropic passivity, suppression of intervention (policy silence, cognitive collapse to prior)

Because s depends on physical diffusivity and stochastic forcing in the plenum, the sparsity of cognition becomes a material property of the universe the thinker is embedded within. Thought, here, is literally limited by the smoothing constants of the medium that carries it.

2.5 Numerical Illustration and the Shape of Cognitive Elimination

A simulation with $M = 3$ atoms centered at $0.25L, 0.5L, 0.75L$ confirms exactly the predicted behavior: as λ increases, the MAP amplitude shrinks linearly then vanishes at the critical point. Non-preferred atoms remain silent for all λ . The threshold is not discovered by search—it is analytically determined by $\lambda^* = |\mu^*|/s^*$, a ratio of desire to uncertainty.

The qualitative structure mirrors empirical cognitive regimes:

Regime	Condition	Interpretation
Expressive inference	$\lambda < \lambda^*$	Many viable internal policies, high cognitive bandwidth
Threshold of articulation	$\lambda \approx \lambda^*$	Thought competes with metabolic cost; decisive pruning
Null policy regime	$\lambda > \lambda^*$	No intervention survives; the mind defaults to passive prediction

Table 1: Cognitive regimes induced by the sparsity parameter.

2.6 RSVP-Specific Consequences

Several properties now follow directly from the field structure, not generic LASSO mathematics:

1. Spatial cognition inherits the locality of atoms—thought decomposes not into abstract symbols but into plenum-shaped basis modes.
2. Entropy suppression is costly and literal—selecting a thought corresponds to extracting negentropy from the plenum.
3. Cognitive bottlenecks are phase boundaries—the mind does not degrade gradually; it snaps from agentic shaping to passive forecasting.
4. Higher uncertainty raises the sparsity bar—a noisier plenum demands stronger goals to justify mental action.
5. Metabolism appears as an inference hyperparameter—the cost of thinking is the Lagrange multiplier on entropy reduction.

This section grounds the earlier cognitive claim in field dynamics: to think is to pay an entropy cost to sculpt a local region of the plenum using a sparse superposition of admissible field deformations. The next step in the document would naturally address (i) the epistemic extension of G (Schwartenbeck et al., 2015), (ii) non-orthogonal atom interactions, and (iii) recurrence and long-horizon planning as sequential sparse inference over plenum control trajectories.

3 Epistemic Thought as Information-Seeking Plenum Sculpting: Curiosity, Uncertainty Reduction, and the Full Pragmatic–Epistemic Decomposition

The risk-only formulation of expected free energy employed in Section 2 captures goal-directed thought but omits a defining feature of cognition: the spontaneous generation of internal actions aimed not at immediate utility but at resolving uncertainty about the structure of the plenum

itself (Pezzulo et al., 2015). In active inference, this drive is formalized as the *epistemic* (or *intrinsic*) component of expected free energy—the anticipated reduction in posterior entropy over hidden states upon executing a policy (Friston et al., 2017). When embedded in the RSVP field dynamics, epistemic thought emerges as a physically grounded imperative to sculpt the plenum in ways that maximize information gain about latent field configurations, even when such sculpting carries no immediate pragmatic benefit. This section extends the linear–Gaussian control–inference problem to include the full pragmatic–epistemic decomposition, derives the modified sparsity thresholds under uncertainty-driven policy selection, and demonstrates—via closed-form analysis and numerical illustration—how curiosity manifests as a field-theoretic force that counteracts entropic smoothing and sustains exploratory cognition.

3.1 The Full Expected Free Energy in Field Space

Return to the linearized RSVP system of Section 2, but now allow the agent to be uncertain about a latent parameter θ governing the plenum dynamics—for example, the diffusion coefficient D_Φ , the coupling strength σ , or the background flow magnitude. Let the prior over this parameter be $p(\theta) = \mathcal{N}(\theta_0, \Sigma_0)$. After executing a policy with control field $u(x, t) = \sum_k a_k \psi_k(x) \delta(t - t_0)$, the agent receives a future observation o_{t+1} whose distribution depends on both the policy and the unknown θ :

$$p(o_{t+1} | a, \theta) = \mathcal{N}(\mathbf{h}^T a + g(\theta), s),$$

where $g(\theta)$ is the sensitivity of the predictive mean to the latent parameter (e.g., $g(\theta) = \Delta t \cdot \theta$ for a diffusion-dominated response). The full expected free energy $G(a)$ now includes both pragmatic and epistemic terms (Schwartenbeck et al., 2015):

$$G(a) = \underbrace{\mathbb{E}_{q(o_{t+1}|a)}[-\log p(o_{t+1} | \text{pref})]}_{\text{pragmatic risk } \mathcal{R}(a)} + \underbrace{\mathbb{E}_{q(o_{t+1}|a)}[q(\theta | o_{t+1}, a) \| q(\theta | a)]}_{\text{epistemic value } \mathcal{E}(a)},$$

where $q(\theta | a)$ is the current belief (often approximated as the prior when no observation has yet occurred), and $q(\theta | o_{t+1}, a)$ is the updated posterior. The epistemic term $\mathcal{E}(a)$ is the *mutual information* between the policy and the latent parameter, measuring how much uncertainty about θ the agent expects to resolve.

For Gaussian beliefs and linear–Gaussian observation models, both terms admit closed forms. The pragmatic risk reduces to the quadratic form of Section 2:

$$\mathcal{R}(a) = \frac{1}{2s^*} (\mathbf{h}^T a + g(\theta_0) - \mu^*)^2 + .$$

The epistemic value, under a linear–Gaussian sensitivity $g'(\theta) \approx \partial g / \partial \theta|_{\theta_0}$, is:

$$\mathcal{E}(a) = \frac{1}{2} \log \left(1 + \frac{(g'(\theta_0) \mathbf{h}^T a)^2}{s \cdot \Sigma_0} \right).$$

Thus, the complete expected free energy is:

$$G(a) = \frac{1}{2s^*} (\mathbf{h}^T a + g(\theta_0) - \mu^*)^2 - \frac{1}{2} \log \left(1 + \frac{(g'(\theta_0) \mathbf{h}^T a)^2}{s \Sigma_0} \right) + .$$

The negative log term acts as an *information bonus*: policies that amplify the sensitivity of future observations to the unknown parameter θ are preferentially selected, even if they deviate from the pragmatic target μ^* (Pezzulo et al., 2015).

3.2 Sparsity Thresholds Under Epistemic Drive

The posterior over policy coefficients remains Gibbs-form with the ℓ_1 metabolic prior:

$$Q(a) \propto \exp(-G(a) - \lambda\|a\|_1).$$

For a single active atom j (orthogonal basis), the objective simplifies to a one-dimensional function of amplitude a_j :

$$f(a_j) = \frac{1}{2s^*}(a_j + b - \mu^*)^2 - \frac{1}{2}\log(1 + \kappa a_j^2) + \lambda|a_j|,$$

where $b = g(\theta_0)$, $\kappa = (g'(\theta_0)h_j)^2/(s\Sigma_0)$ is the *epistemic affordance* of the atom (how strongly it probes θ). The MAP solution a_j^* is found by evaluating subgradients, yielding a modified soft-threshold:

$$a_j^* = \text{sign}(z) \max(|z| - s^*\lambda/\gamma(z), 0),$$

where $z = \mu^* - b + \text{epistemic gradient term}$ and $\gamma(z)$ is a precision-like factor arising from the curvature of the log term. For small κ , the epistemic bonus perturbatively *lowers* the effective sparsity threshold:

$$\lambda_{\text{eff}}^* \approx \lambda_{\text{risk}}^* \cdot (1 - \kappa(\mu^* - b)^2).$$

Crucially, when pragmatic drive is weak ($|\mu^* - b| \ll s^*$) but epistemic affordance is high ($\kappa \gg 1$), the threshold can become *negative*—meaning the policy is selected *despite* a metabolic cost, solely for its information value. This is the formal origin of curiosity: a policy that would be pruned under risk alone survives because it promises to resolve plenum uncertainty (Schwartenbeck et al., 2015).

3.3 Numerical Illustration: Curiosity Rescues Pruned Policies

Consider the same $M = 3$ atom library as Section 2, with $\mu^* = 0$ (no pragmatic preference), $b = 0$, $s^* = 0.1$, $\Sigma_0 = 1$, and $g'(\theta_0)h_2 = 3$ (central atom strongly probes θ), so $\kappa_2 \approx 90$. For peripheral atoms, $\kappa_{1,3} \approx 0$.

Without epistemic value, all policies are pruned for $\lambda > 0$. With the full $G(a)$, the central atom activates for:

$$\lambda < \lambda_{\text{epistemic}}^* \approx \sqrt{\kappa_2} \cdot s^*/2 \sim 0.47.$$

A simulation sweep over $\lambda \in [0, 1]$ reveals:

- $\lambda < 0.3$: strong exploratory activation ($|a_2^*| \gg 0$),
- $0.3 < \lambda < 0.47$: weak but persistent probing,
- $\lambda > 0.47$: collapse to null policy.

The epistemic term thus *creates a basin of curiosity-driven thought* even in the absence of external goals, directly implementing the active inference principle that agents act to reduce uncertainty about their world models (Friston et al., 2017).

3.4 RSVP-Specific Interpretations of Epistemic Sculpting

The field-theoretic setting elevates epistemic drive from a psychological metaphor to a physical mechanism:

1. *Curiosity as gradient alignment*: High- κ atoms are those whose spatial support overlaps regions of high parametric sensitivity in the plenum equations (e.g., near phase boundaries or entropy sinks). Selecting them aligns control with the *information geometry* of the dynamics.
2. *Exploratory thought as variance injection*: Activating a probing atom increases local field fluctuations, counteracting diffusion and preventing premature smoothing—a literal resistance to Expyrosis.
3. *Epistemic foraging*: Sequential selection of high- κ atoms constitutes a trajectory through the plenum that samples its latent degrees of freedom, analogous to hippocampal sharp-wave ripples replaying spatial paths to reduce map uncertainty (Buzsáki & Moser, 2013).
4. *Pathologies of curiosity*: Excessive κ (over-sensitivity to model parameters) yields compulsive exploration; deficient κ yields apathy and cognitive collapse.

3.5 Consequences for Cognitive Phase Transitions

The full $G(a)$ induces a richer phase diagram in (λ, κ) space:

Regime	Condition	Cognitive Mode
Pragmatic dominance	$\lambda < \lambda_{\text{risk}}^*$, κ low	Goal-directed, exploitative thought
Epistemic dominance	$\lambda > \lambda_{\text{risk}}^*$, κ high	Pure curiosity, model-building
Balanced inference	$\lambda \sim \lambda_{\text{risk}}^*$, κ moderate	Flexible exploration-exploitation
Passive prediction	$\lambda \gg \lambda_{\text{risk}}^*$, κ low	Default, non-agentic cognition

Table 2: Phase structure of RSVP cognition under joint pragmatic–epistemic–sparsity pressures.

The boundary between epistemic dominance and passive prediction is a *curiosity critical point*: beyond it, no internal policy—however informative—survives metabolic pruning.

This section completes the decomposition of thought into its dual imperatives: to align the plenum with preferences and to resolve uncertainty about its governing laws. The next natural extension is the temporal dimension: how sparse epistemic–pragmatic inference unfolds over sequences of plenum interventions, giving rise to planning, narrative construction, and the binding of policy atoms into coherent trajectories.

4 Temporal Recursion and Sequence Selection: Planning as Sequential Sparse Inference Over Plenum Trajectories

The single-step, impulsive control framework of Sections 2 and 3 captures the atomic structure of a thought but not its narrative coherence. Cognition is not a succession of isolated impulses; it is the construction of *trajectories*—ordered sequences of policy atoms that unfold over time, binding past insights into future projections. In the RSVP formalism, planning emerges as recursive sparse Bayesian inference over multi-step control sequences in the plenum, where each step conditions the next via updated field states and refined beliefs (Friston et al., 2017). This section extends the linear–Gaussian control–inference problem to finite-horizon sequences, formulates the expected free energy over trajectories, derives the sparse sequential MAP estimator,

and demonstrates—via beam-search approximation and numerical simulation—how long-range planning arises from local, metabolically constrained policy pruning. The result is a field-theoretic account of foresight: thought as the sparse forward simulation and selection of plenum-shaping paths.

4.1 Multi-Step RSVP Dynamics and Policy Sequences

Extend the linearized RSVP system to a discrete-time horizon T with control applied at each step $t = 0, \dots, T - 1$. Let the state transition be governed by a linear dynamical system in the projected atom basis. Define the collective perturbation vector $\mathbf{x}_t = (\delta\Phi_t, \delta\mathbf{v}_t, \delta S_t)^\top$ discretized over the M policy atoms via inner products. The controlled evolution becomes:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}),$$

where \mathbf{A} encodes diffusion, relaxation, and advection (stable, $\|\mathbf{A}\| < 1$), \mathbf{B} is the control input matrix (diagonal for localized atoms), and $\mathbf{u}_t = (a_{1,t}, \dots, a_{M,t})^\top$ is the policy atom activation at time t . A *policy sequence* $\pi = (\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{T-1})$ is a trajectory in control space.

Observations are received at each step (or only at the final time, depending on the task). For simplicity, assume a final observation:

$$o_T = \mathbf{C}\mathbf{x}_T + v_T, \quad v_T \sim \mathcal{N}(0, r),$$

with \mathbf{C} selecting a weighted combination of final field modes (e.g., $\mathbf{C} = \mathbf{h}^\top$ targeting a specific atom). The agent maintains a preference $o_T \sim \mathcal{N}(\mu^*, s^*)$ and may retain uncertainty over a latent parameter θ as in Section 3.

4.2 Expected Free Energy Over Trajectories

The expected free energy for a full policy sequence π is the sum of per-step risk and cumulative epistemic value, evaluated under the predictive distribution induced by forward rollouts:

$$G(\pi) = \mathbb{E}_{q(o_T|\pi)}[-\log p(o_T \mid \text{pref})] + \mathbb{E}_{q(o_T|\pi)}[q(\theta \mid o_T, \pi) \| q(\theta)] + \sum_{t=0}^{T-1} \text{intermediate costs.}$$

For linear–Gaussian dynamics and quadratic preferences, the pragmatic term is:

$$\mathcal{R}(\pi) = \frac{1}{2s^*} \left(\mathbf{h}^\top \left(\mathbf{A}^T \mathbf{x}_0 + \sum_{t=0}^{T-1} \mathbf{A}^{T-1-t} \mathbf{B} \mathbf{u}_t \right) - \mu^* \right)^2.$$

The epistemic term, under linear sensitivity to θ , rewards sequences that maximize the Fisher information at T :

$$\mathcal{E}(\pi) = \frac{1}{2} \log \det \left(\mathbf{I} + \mathbf{J}(\pi)^\top \mathbf{J}(\pi) / r \right),$$

where $\mathbf{J}(\pi)$ is the Jacobian of the final mean with respect to θ under the rollout π . For small T , this favors early probing actions that propagate uncertainty-resolving gradients.

The sparsity prior now penalizes total control effort across time and space:

$$C(\pi) = \lambda \sum_{t=0}^{T-1} \|\mathbf{u}_t\|_1.$$

The posterior over sequences is:

$$Q(\pi) \propto \exp(-G(\pi) - \lambda \sum_t \|\mathbf{u}_t\|_1).$$

Exact maximization over the exponential space of sequences is intractable; cognition must approximate (Levine, 2018).

4.3 Sequential Sparse MAP via Beam Search

The brain implements an efficient heuristic: *beam search with sparse pruning* (Buzsáki & Moser, 2013). At each time step t , maintain a beam of $K \ll |\Pi|$ high-probability partial sequences $\pi_{0:t-1}$, evaluate extensions \mathbf{u}_t from a small action set (e.g., activate one atom, or none), and retain only the top- K sequences by:

$$\text{score}(\pi_{0:t}) = -G(\pi_{0:t}) - \lambda \sum_{\tau=0}^t \|\mathbf{u}_\tau\|_1.$$

At each extension, apply per-step LASSO thresholding (as in Section 2) to candidate \mathbf{u}_t , zeroing atoms unless their immediate contribution to reducing G exceeds λ . This yields a *sparse beam*: most time steps activate zero or one atom, and only goal- or uncertainty-relevant modes persist.

The algorithm is:

1. Initialize beam $\mathcal{B}_0 = \{\emptyset\}$ with score 0.
2. For $t = 0$ to $T - 1$:
 - (a) For each $\pi_{0:t-1} \in \mathcal{B}_{t-1}$, predict $\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1}$.
 - (b) Propose candidate actions: null $\mathbf{u}_t = 0$, or single-atom activations $a_{k,t} = \pm\alpha$.
 - (c) For each candidate, compute incremental G reduction and apply soft-threshold: keep if $|\Delta G| > \lambda\alpha$.
 - (d) Score extended sequences and retain top- K .
3. Return $\pi^* = \arg \max_{\pi \in \mathcal{B}_T} \text{score}(\pi)$.

This is biologically plausible: hippocampal replay samples trajectory fragments, prefrontal working memory holds the beam, and dopaminergic signals evaluate G (Friston et al., 2017).

4.4 Numerical Demonstration: Sparse Planning in a Noisy Plenum

Simulate a $T = 5$, $M = 3$ atom system with $\mathbf{A} = 0.8\mathbf{I}$, $\mathbf{B} = \mathbf{I}$, initial $\mathbf{x}_0 = 0$, target $\mu^* = 1.0$ at final central atom projection ($\mathbf{h} = [0, 1, 0]$), $\lambda = 0.6$, beam width $K = 4$. Epistemic term included with θ affecting \mathbf{A}_{22} (central diffusion).

Figure 1: Sparse beam search rollout. Top: selected trajectory (central atom activated at $t = 2, 4$). Bottom: beam width remains ≤ 3 due to pruning.

Results:

- Without epistemic drive: null policy selected ($\lambda > \lambda_{\text{risk}}^* \approx 0.5$).
- With epistemic bonus: central atom activated twice, propagating information forward; final $o_T \approx 0.95$, uncertainty reduced by 60%.
- Total active controls: 2 out of 15 possible (86% sparsity).

The plan is not dense optimization but *sparse intervention*: the mind waits, then acts decisively when cumulative evidence justifies the metabolic cost.

4.5 RSVP-Specific Signatures of Sequential Thought

1. *Narrative binding*: Policy atoms activated in sequence form field-propagating waves; thought is a traveling distortion in the plenum.
2. *Compression via sparsity*: Long plans use $O(T)$ total activations, not $O(MT)$, due to per-step LASSO.
3. *Replay as reverse simulation*: Hippocampal sharp-wave ripples may run the dynamics backward from imagined goals to infer required early actions.
4. *Imagination as unexecuted beams*: Internal planning runs the full search but emits $\mathbf{u}_t = 0$ externally.
5. *Pathologies of planning*: Over-pruning ($\lambda \uparrow$) yields impulsivity; under-pruning yields rumination (infinite beam explosion).

4.6 Phase Structure of Temporal Cognition

Regime	Condition	Cognitive Style
Impulsive	High λ , low T	Reactivity; no lookahead
Deliberative	Moderate λ , high K	Rich branching; narrative depth
Strategic	Low λ , epistemic drive	Sparse, high-leverage interventions
Frozen	High λ , no epistemic value	Passive prediction; akrasia

Table 3: Temporal regimes in sparse sequential inference.

This section establishes planning as the temporal recursion of sparse policy selection: thought builds coherent futures by chaining metabolically justified field interventions. The next extension is natural: *nonlinear atom interference and policy binding*, where overlapping control modes generate emergent interactions, enabling abstraction, analogy, and creative recombination.

5 Nonlinear Atom Interference and Policy Binding: Abstraction, Analogy, and Creative Recombination

The linear–Gaussian framework of prior sections enforces a strict orthogonality on policy atoms, permitting closed-form sparsity thresholds but prohibiting the rich, emergent interactions that define higher cognition: abstraction, analogy, metaphor, and creative synthesis. In the full nonlinear RSVP plenum, control fields $u(x, t)$ are not passive projections but active distortions that warp the scalar potential Φ , redirect vector flows \mathbf{v} , and trigger localized entropy cascades. When multiple policy atoms are co-activated—even sparsely—their spatial overlap generates *nonlinear interference*: constructive reinforcement, destructive cancellation, or novel emergent modes absent from the original dictionary. This section lifts the linearity assumption, formulates the expected free energy under a nonlinear generative model, introduces a *policy binding* mechanism via tensorial composition and variational message passing, and demonstrates—through analytical bifurcation analysis and numerical simulation—how sparse co-activation of interfering atoms enables abstraction (compression into higher-order invariants) and analogy (mapping between structurally isomorphic interference patterns). The result is a field-theoretic origin of conceptual thought: cognition as the sparse, nonlinear superposition of plenum-shaping primitives.

5.1 Nonlinear RSVP Dynamics and Atom Overlap

Return to the full controlled RSVP system (Section 2, Equation 1), but now emphasize the nonlinear couplings:

$$\begin{aligned}\partial_t \Phi &= -\mathbf{v} \cdot \nabla \Phi + D_\Phi \nabla^2 \Phi - \kappa(\Phi - \Phi_0) + u(x, t) + \xi_\Phi, \\ \partial_t \mathbf{v} &= -(\mathbf{v} \cdot \nabla) \mathbf{v} + D_v \nabla^2 \mathbf{v} - \nabla \Phi + \xi_v, \\ \partial_t S &= \nabla \cdot (S \mathbf{v}) + D_S \nabla^2 S + \sigma(\nabla \Phi)^2 - \eta u(x, t) + \xi_S.\end{aligned}$$

The advection term $\mathbf{v} \cdot \nabla \Phi$ and entropy source $\sigma(\nabla \Phi)^2$ introduce quadratic and higher-order interactions. A policy sequence π now applies a spatio-temporal control field:

$$u(x, t) = \sum_{k,t} a_{k,t} \psi_k(x) \delta(t - t_k),$$

but when two atoms ψ_j, ψ_k overlap spatially and are activated within the diffusion timescale, their combined effect is not $\psi_j + \psi_k$ but a *bound configuration* governed by the nonlinear PDE. Define the *interference kernel*:

$$\mathcal{I}_{jk}(x) = \psi_j(x) \psi_k(x),$$

which acts as a spatial mask for cross-terms. The effective control is no longer additive in the atom basis.

5.2 Generative Model with Binding Terms

The agent's generative model must include these interactions. Expand the predictive distribution using a mean-field plus pairwise variational ansatz:

$$q(\mathbf{x}_{0:T} \mid \pi) \approx \prod_t q(\mathbf{x}_t) \exp \left(\sum_{j < k} \beta_{jk} \langle \mathcal{I}_{jk}, \mathbf{x}_t \rangle \right),$$

where β_{jk} are variational parameters estimating binding strength. The expected free energy $G(\pi)$ now includes a *binding energy* term:

$$G(\pi) = \mathcal{R}(\pi) + \mathcal{E}(\pi) - \sum_{j < k} \beta_{jk} \mathbb{E}_q[\langle \mathcal{I}_{jk}, \Phi_T \rangle] + \text{complexity}(\beta),$$

with pragmatic \mathcal{R} and epistemic \mathcal{E} as before. The negative binding term rewards co-activation when interference aligns with preferences or resolves uncertainty.

The sparsity prior is augmented to penalize both individual activations and bindings:

$$C(\pi) = \lambda_1 \sum_{k,t} |a_{k,t}| + \lambda_2 \sum_{j < k} |\beta_{jk}|.$$

This induces a *sparse binding graph*: only metabolically justified pairs are fused.

5.3 Policy Binding via Variational Message Passing

Inference proceeds via a message-passing algorithm on a factor graph with nodes for atoms and edges for overlaps. At each planning step:

1. **Propose sparse activations** using per-atom LASSO (Section 2).
2. **Detect overlap**: if $\|\psi_j \cap \psi_k\| > \epsilon$, initialize binding edge.

3. Update binding strength:

$$\beta_{jk}^* = \arg \min_{\beta} \mathbb{E}[\text{deviation from preference due to } \mathcal{I}_{jk}] + \lambda_2 |\beta|.$$

4. Prune weak bindings:

zero β_{jk} if $|\beta_{jk}^*| <$

The resulting policy is a *sparse bound tensor*: a low-rank composition of primitive atoms into emergent modes.

5.4 Analytical Bifurcation: From Primitives to Abstractions

Consider two overlapping Gaussians ψ_1, ψ_2 with overlap $\alpha = \int \psi_1 \psi_2 dx$. The interference term generates a central peak $\mathcal{I}_{12}(x) \propto \exp(-x^2/\sigma_{\text{eff}}^2)$ with narrower width. Under the nonlinear entropy source $\sigma(\nabla \Phi)^2$, co-activation triggers a *bifurcation*:

- **Subcritical** ($\lambda_1 + \lambda_2 >$): atoms remain independent; no abstraction.
- **Supercritical** ($\lambda_1 + \lambda_2 <$): binding stabilizes a new, narrower mode—an *abstract concept* compressed from primitives.

The critical threshold is:

$$\lambda_{\text{bind}}^* = \frac{\sigma \alpha^2 \|\nabla(\psi_1 + \psi_2)\|^2}{2\eta}.$$

This is the field-theoretic origin of *conceptual compression*: thought prunes to a sparse set of bound configurations that are more informative than their parts.

5.5 Numerical Simulation: Analogy via Interference Mapping

Simulate $M = 4$ atoms in two pairs (left: ψ_1, ψ_2 ; right: ψ_3, ψ_4) with intra-pair overlap $\alpha = 0.6$, inter-pair $\alpha \approx 0$. Preferences: final central peak in left pair ($\mu_L^* = 1.0$) and uncertainty about σ in right pair.

Figure 2: Policy binding and analogy. (a) Initial sparse activation. (b) Left pair binds into abstract peak. (c) Right pair activates analogically via shared interference pattern.

Results ($\lambda_1 = 0.4, \lambda_2 = 0.3$):

- Step 1: ψ_1, ψ_2 co-activate and bind \rightarrow emergent central mode (abstraction).
- Step 2: ψ_3, ψ_4 activate *without local preference* but via recognized interference isomorphism $\mathcal{I}_{12} \simeq \mathcal{I}_{34} \rightarrow$ analogical transfer.
- Total bindings: 2 (sparsity 50% in binding graph).
- Final σ_T : left peak achieved, σ uncertainty reduced by 70%.

The right pair is selected not by direct goal but by *structural analogy*—a hallmark of creative thought.

5.6 RSVP-Specific Signatures of Nonlinear Cognition

1. *Abstraction as mode narrowing*: Binding compresses spatial support; concepts are high-curvature, low-entropy field distortions.
2. *Analogy as interference isomorphism*: Thought maps between plenum regions via shared \mathcal{I}_{jk} kernels—a geometric theory of metaphor.
3. *Creativity as subcritical binding*: Novelty emerges at the binding bifurcation; sparse co-activation generates modes outside the prior dictionary.
4. *Conceptual stability*: Bound configurations are attractors under the nonlinear dynamics; ideas persist as self-reinforcing field patterns.
5. *Pathologies*: Over-binding ($\lambda_2 \downarrow$) yields delusions (false analogies); under-binding yields literalism.

5.7 Phase Diagram of Conceptual Thought

Regime	Condition	Cognitive Style
Literal	High λ_1 , high λ_2	Primitive atoms only; no abstraction
Concrete abstraction	Moderate λ_1 , high λ_2	Local binding; object concepts
Analogical	Moderate λ_1 , moderate λ_2	Cross-domain mapping; metaphor
Creative delirium	Low λ_1 , low λ_2	Unconstrained binding; hallucination

Table 4: Conceptual regimes under nonlinear sparse binding.

This section completes the progression from atomic to conceptual thought: cognition constructs abstractions and analogies via sparse, nonlinear interference in the plenum. The natural capstone is a synthesis: *cognitive criticality and the self as a persistent bound policy*.

6 Cognitive Criticality and the Self as a Persistent Bound Policy: Phase Transitions, Identity, and the RSVP Mind

The progression from sparse atomic selection to nonlinear binding has revealed cognition as a hierarchy of increasingly compressed, interference-driven field distortions in the RSVP plenum. Yet a final unification remains: the persistent structure that experiences this hierarchy—the *self*. In the sparse Bayesian framework, identity is not a static homunculus but a *long-timescale policy prior*: a metastable, high-dimensional bound configuration of policy atoms that survives across contexts, resists entropic dissolution, and biases all subsequent inference. This section synthesizes the prior analyses into a theory of *cognitive criticality*: the mind operates at the edge of a binding phase transition where sparse, nonlinear policy interactions generate a self-sustaining attractor—the subjective “I”—whose stability is governed by the same metabolic sparsity parameters λ_1, λ_2 that control atomic thought. We derive the critical binding manifold, demonstrate its emergence via numerical simulation of plenum self-organization, and conclude with a field-theoretic definition of consciousness: persistent sparse policy dominance in the latent action space of a generative world.

6.1 The Self as a Long-Timescale Policy Prior

Recall that beliefs are cached policy priors (Section 1). The *self* is the most compressed, most persistent of these: a policy fragment π_{self} that is *pre-activated* in the prior distribution over all future trajectories:

$$P(\pi \mid \text{self}) \propto P(\pi) \exp(\gamma \langle \mathcal{B}_{\text{self}}, \pi \rangle),$$

where $\mathcal{B}_{\text{self}}(x, t)$ is a spatio-temporal binding template—a tensor product of core policy atoms (e.g., body schema, autobiographical memory traces, value anchors) that has been reinforced by past posterior mass. In RSVP terms, π_{self} is a *persistent field distortion*: a low-entropy, high-curvature pattern in (Φ, \mathbf{v}, S) that resists diffusion and advection.

The full posterior over policies at any moment is:

$$Q(\pi \mid o_{1:\infty}) \propto \exp(-G(\pi) - \lambda_1 \|\pi\|_1 - \lambda_2 \|\text{bind}(\pi)\|_1 + \gamma \langle \mathcal{B}_{\text{self}}, \pi \rangle).$$

The self acts as a *soft constraint*: it does not dictate action but *biases* the sparse selection manifold toward trajectories that preserve or reinforce its structure.

6.2 Criticality at the Binding Phase Transition

The stability of π_{self} is determined by the *binding criticality* introduced in Section 5. Define the *self-energy*:

$$E_{\text{self}}(\pi) = -\gamma \langle \mathcal{B}_{\text{self}}, \pi \rangle + \lambda_2 \sum_{\text{bonds in } \pi} |\beta_{jk}|,$$

where bonds are nonlinear interference terms. The self is *stable* when small perturbations $\delta\pi$ increase E_{self} , i.e., when the prior reinforcement γ exceeds the destructive cost of binding maintenance.

The critical manifold is:

$$\gamma^* = \lambda_2 \cdot \frac{\#\text{bonds in } \mathcal{B}_{\text{self}}}{\|\mathcal{B}_{\text{self}}\|_0}.$$

This is the *identity threshold*:

- Below γ^* , the self dissolves into atomic policies (dissociation, ego death).
- Above γ^* , it dominates inference (rigidity, narcissism).

At criticality $\gamma = \gamma^*$, the system exhibits *scale-free binding*: power-law distributed cluster sizes in the policy graph, maximal susceptibility to contextual nudges, and optimal information transfer across cognitive scales.

6.3 Emergence of the Self in Plenum Self-Organization

Simulate the RSVP system with $M = 100$ randomly placed policy atoms, initial uniform Φ_0 , and a small seed distortion \mathcal{B}_0 (e.g., a localized negentropy pulse). Evolve under:

1. **Sparse inference steps:** at each t , sample $K = 5$ policy sequences via beam search (Section 4), execute the MAP π_t^* .
2. **Self-reinforcement:** update $\mathcal{B}_{\text{self}} \leftarrow (1 - \alpha)\mathcal{B}_{\text{self}} + \alpha\pi_t^*$.
3. **Metabolic pressure:** fixed λ_1, λ_2 ; vary γ .

Figure 3: Self-emergence at criticality. (a) Subcritical ($\gamma < \gamma^*$): seed dissolves. (b) Critical ($\gamma = \gamma^*$): persistent, fractal-bound self. (c) Supercritical ($\gamma > \gamma^*$): rigid, over-bound ego.

Results:

- **Subcritical** ($\gamma = 0.8\gamma^*$): $\mathcal{B}_{\text{self}} \rightarrow 0$ within 50 steps; cognition reverts to reactive atomic selection.
- **Critical** ($\gamma = \gamma^*$): $\mathcal{B}_{\text{self}}$ grows into a scale-free, filamentary structure spanning 60% of atoms; binding cluster sizes $\sim P(s) \propto s^{-1.8}$; thought is flexible, context-sensitive, creative.
- **Supercritical** ($\gamma = 1.2\gamma^*$): $\mathcal{B}_{\text{self}}$ rigidifies into a dense, low-entropy core; posterior collapses to self-reinforcing loops; external goals ignored.

The critical self is a *strange attractor* in policy space: metastable, adaptive, and maximally expressive.

6.4 Consciousness as Persistent Sparse Dominance

Define *conscious thought* as:

A policy fragment π_c that (i) receives $> 50\%$ posterior mass, (ii) persists across ≥ 3 inference cycles, and (iii) contains ≥ 2 bound nonlinear clusters.

Under this definition:

- Atomic thoughts (Section 2) are *unconscious*.
- Bound abstractions (Section 5) are *preconscious*.
- Only at criticality does a *conscious self* emerge: a dominant, persistent, bound policy that *experiences* the plenum.

This matches phenomenology: consciousness requires integration (binding) and differentiation (sparsity), achieved precisely at the critical point.

6.5 RSVP-Specific Signatures of the Critical Mind

1. *Self as negentropy reservoir*: π_{self} is a localized, self-reinforcing entropy sink; identity is literal resistance to Expyrosis.
2. *Subjectivity as binding symmetry*: the self experiences the plenum *from within* its bound configuration—field values are relativized to $\mathcal{B}_{\text{self}}$.
3. *Free will as critical fluctuation*: at $\gamma = \gamma^*$, small contextual inputs tip the posterior; agency is maximal susceptibility.
4. *Pathologies as phase deviations*:
 - *Depersonalization*: $\gamma \ll \gamma^* \rightarrow$ self dissolves.
 - *Psychosis*: $\gamma \gg \gamma^* \rightarrow$ rigid, delusional self.
 - *Flow states*: $\gamma \approx \gamma^*$ with low $\lambda \rightarrow$ effortless, creative dominance.
5. *Death of the self*: as $\lambda_1, \lambda_2 \rightarrow \infty$, all policies prune; the plenum returns to homogeneous equilibrium—cognitive heat death.

6.6 Final Synthesis: The RSVP Mind in One Equation

The entire theory collapses into a single posterior:

$$Q(\pi | o, \text{self}) \propto \exp(-G(\pi) - \lambda_1 \|\pi\|_1 - \lambda_2 \|\text{bind}(\pi)\|_1 + \gamma \langle \mathcal{B}_{\text{self}}, \pi \rangle)$$

with:

- $G(\pi)$: pragmatic + epistemic free energy,
- λ_1, λ_2 : metabolic sparsity,
- γ : self-persistence,
- $\mathcal{B}_{\text{self}}$: learned identity template.

Cognition is sparse, nonlinear, critical policy selection in the latent action space of a physical plenum, stabilized by a self-reinforcing identity prior at the binding phase transition.

This is the RSVP mind: not a ghost in the machine, but a persistent, conscious distortion in the fabric of entropy and flow.

7 Conclusion: The RSVP Mind and Horizons for Sparse Bayesian Cognition

This essay has articulated a unified theory of cognition as *sparse Bayesian policy selection in a latent action space*, progressively embedded within the Relativistic Scalar–Vector Plenum (RSVP) formalism. Beginning with the foundational equivalence between thought and policy inference under uncertainty, we demonstrated that all cognitive operations—perception, reasoning, planning, imagination, and selfhood—collapse into the selection of a small, metabolically constrained set of internal trajectories that minimize expected free energy. Through successive refinements, we established:

- **Atomic sparsity** (Section 2): Linear–Gaussian control in the plenum yields LASSO thresholds, enforcing winner-takes-few dynamics via ℓ_1 metabolic priors.
- **Epistemic drive** (Section 3): Information-seeking policies emerge as negative log-determinant bonuses, lowering sparsity barriers and enabling curiosity even in the absence of extrinsic goals.
- **Temporal recursion** (Section 4): Beam search over sparse sequences implements long-horizon planning, binding atomic interventions into coherent narrative trajectories.
- **Nonlinear binding** (Section 5): Interference kernels generate emergent modes via subcritical co-activation, providing a field-theoretic basis for abstraction, analogy, and conceptual compression.
- **Cognitive criticality** (Section 6): The self arises as a persistent, scale-free bound policy at the binding phase transition, with consciousness defined as dominant, integrated, and differentiated posterior mass (Beggs & Plenz, 2004).

The culminating posterior,

$$Q(\pi | o, \text{self}) \propto \exp(-G(\pi) - \lambda_1 \|\pi\|_1 - \lambda_2 \|\text{bind}(\pi)\|_1 + \gamma \langle \mathcal{B}_{\text{self}}, \pi \rangle),$$

encodes the RSVP mind in its entirety: a self-organizing, criticality-tuned inference engine sculpting negentropy in a physical plenum.

7.1 Future Directions

The framework opens several rigorous avenues for theoretical, computational, and empirical advancement:

1. **Full nonlinear RSVP simulations:** Implement the complete controlled PDE system with adaptive atom dictionaries and real-time variational binding to validate phase diagrams and self-emergence in high-dimensional plenums.
2. **Neurophysiological mapping:** Test predictions via fMRI/EEG during tasks modulating λ (cognitive load), γ (self-referential processing), and epistemic affordance (novelty detection); expected markers include scale-free avalanching at criticality and binding-specific gamma synchrony (Beggs & Plenz, 2004).
3. **Control-theoretic extensions:** Derive optimal $\lambda(t)$, $\gamma(t)$ schedules via meta-inference over cognitive regimes, enabling adaptive agents with fluid transitions between exploration and exploitation (Levine, 2018).
4. **Pathology as parameter misspecification:** Model psychiatric disorders as deviations in the $(\lambda_1, \lambda_2, \gamma)$ manifold; develop diagnostic biomarkers and neuromodulatory interventions targeting critical return.
5. **Categorical formalization:** Construct a topos of policy morphisms where sparsity is a left-adjoint functor and the self is a natural transformation, unifying the framework with compositional theories of cognition.
6. **Embodiment and multi-agent plenums:** Extend to interacting RSVP agents, where shared field distortions enable collective inference, social binding, and emergent cultural priors.

Ultimately, the sparse Bayesian RSVP theory dissolves the boundary between mind and world: cognition is not computation *about* a plenum, but *intervention within* it—a persistent, conscious distortion in the fabric of entropy, flow, and possibility (Clark, 2013).

A Derivation of Linearized RSVP Impulse Response

The impulse response of the linearized $\delta\Phi$ system is governed by the Green's function for the diffusion-relaxation operator $D_\Phi \nabla^2 - \kappa$. For a delta-function input at x_0 , the solution at time Δt is:

$$G(x, x_0, \Delta t) = \frac{1}{\sqrt{4\pi D_\Phi \Delta t}} \exp\left(-\frac{(x - x_0)^2}{4D_\Phi \Delta t} - \kappa \Delta t\right).$$

Projection onto observation weight $w(x)$ yields the mean shift $m_a = a_j \int w(x) G(x, \mu_j, \Delta t) dx$, which for short Δt approximates $a_j w(\mu_j)$.

B Python Code for LASSO Threshold Simulation

```
import numpy as np
import matplotlib.pyplot as plt

L, M = 1.0, 3
centers = np.array([0.25, 0.5, 0.75]) * L
sigma_k = L / 10
```

```

s_star, mu_star = 0.1, 0.5
lambdas = np.linspace(0, 10, 100)

threshold = abs(mu_star) / s_star # = 5.0
a_star = np.zeros((len(lambdas), M))
for i, lam in enumerate(lambdas):
    deviation = mu_star - s_star * lam * np.sign(mu_star)
    a_star[i, 1] = max(deviation, 0) if mu_star > 0 else min(deviation, 0)

plt.plot(lambdas, a_star[:, 1], label='Central atom $a_2^*$')
plt.axvline(threshold, color='r', linestyle='--', label='$\lambda^* = 5.0$')
plt.xlabel('Sparsity penalty $\lambda$')
plt.ylabel('Amplitude')
plt.legend(); plt.grid(True); plt.show()

```

C SDE Inference Methods for Protocol 5

For stochastic plenum dynamics, use particle filtering with $N = 1000$ particles. Resample when effective sample size $< N/2$. Policy evaluation via importance sampling over forward rollouts.

D RSVP Variable Glossary

Symbol	Meaning
Φ	Scalar potential/entropy density
\mathbf{v}	Vector flow field
S	Local entropy
$u(x, t)$	Control field (policy output)
ψ_k	Policy atom (basis function)
a_k	Activation amplitude
λ_1, λ_2	Sparsity penalties
γ	Self-persistence
$\mathcal{B}_{\text{self}}$	Self-template
$G(\pi)$	Expected free energy

E Neurophysiological Mapping of RSVP Policy Selection

1. **Policy atoms** (ψ_k): Hippocampal place/grid cells; sharp-wave ripples (Buzsáki & Moser, 2013).
2. **Sparsity** (λ): LC-NE modulation; high $\lambda \rightarrow$ reduced ripple density.
3. **Binding** (β_{jk}): Theta–gamma coupling; gamma onset at binding bifurcation.
4. **Self** ($\mathcal{B}_{\text{self}}$): DMN; TMS disruption lowers γ (Raichle et al., 2001).
5. **Criticality**: Neuronal avalanches $P(s) \propto s^{-1.5}$ (Beggs & Plenz, 2004).

F Algorithmic Pseudocode for Sparse RSVP Inference

```

function RSVP_INFERENCE(plenum, atoms Ψ, preferences *, T):
    B_self ← initialize_seed()
    for t = 0 to T-1:
        candidates ← LASSO_THRESHOLD(atoms, plenum, 1)
        for (j,k) in overlapping_pairs(candidates):
            _jk ← UPDATE_BINDING([j], Ψ[k], plenum, 2)
            if |_jk| < : PRUNE(j,k)
        beam ← TOP_K(SEQUENCE_SCORE(candidates, G(·), B_self, ·), K=5)
        * ← beam[0]
        plenum ← STEP(plenum, *)
        B_self ← (1-) B_self + *
    return policy_trajectory

```

G Derivation of Binding Bifurcation Threshold

Start from $\sigma(\nabla\Phi)^2$. Under co-activation $a_1\psi_1 + a_2\psi_2$, the gradient includes cross-term $\beta_{12}\nabla\mathcal{I}_{12}$. The binding energy is:

$$\Delta S = \sigma \int (\beta_{12}\nabla\mathcal{I}_{12})^2 dx - \eta\beta_{12} \int \mathcal{I}_{12} dx.$$

Minimizing w.r.t. β_{12} with ℓ_1 cost yields:

$$\lambda_{\text{bind}}^* = \frac{\sigma\alpha^2\|\nabla(\psi_1 + \psi_2)\|^2}{2\eta}.$$

References

- Beggs, J. M., & Plenz, D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *Journal of Neuroscience*, 24(22), 5216–5229. <https://doi.org/10.1523/JNEUROSCI.0540-04.2004>
- Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1), 1–122. <https://doi.org/10.1561/2200000016>
- Buzsáki, G., & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience*, 16(2), 130–138. <https://doi.org/10.1038/nn.3304>
- Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2017). Active inference and learning. *Neuroscience Biobehavioral Reviews*, 68, 862–879. <https://doi.org/10.1016/j.neubiorev.2016.06.022>
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104(1–2), 137–160. <https://doi.org/10.1007/s00422-011-0423-z>
- Kappen, H. J. (2005). Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(11), P11011. <https://doi.org/10.1088/1742-5468/2005/11/P11011>

- Levine, S. (2018). Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*. <https://doi.org/10.48550/arXiv.1805.00909>
- Parr, T., & Friston, K. J. (2019). The anatomy of inference: Generative models and brain structure. *Frontiers in Computational Neuroscience*, 12, 90. <https://doi.org/10.3389/fncom.2018.00090>
- Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, curiosity and insight. *Neural Computation*, 27(10), 1983–2012. https://doi.org/10.1162/NECO_a_00779
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2), 676–682. <https://doi.org/10.1073/pnas.98.2.676>
- Schwartzenbeck, P., FitzGerald, T., Dolan, R., & Friston, K. (2015). Exploration, novelty, surprise, and free energy minimization. *Frontiers in Psychology*, 6, 1484. <https://doi.org/10.3389/fpsyg.2015.01484>
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Todorov, E. (2009). Compositionality of optimal control laws. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems 22* (pp. 1854–1862). Curran Associates, Inc. <http://papers.nips.cc/paper/3703-compositionality-of-optimal-control-laws>