Unified Probability Theory (UPT) – A walkthrough for the simple minded and willfully ignorant. Definitions, Dynamics, and Proofs

Jon Poplett

August 27, 2025

Purpose of this supplement

This note makes UPT maximally explicit: (1) Every operator (potential, emergence, resonance, cascade) is defined using familiar constructs from probability and dynamical systems; (2) The evolution is a well-posed ODE on the probability simplex with conservation and positivity; (3) Bayes' rule appears as a special case (impulsive updates) of the same formalism.

1 Standing assumptions and notation

Fix a finite measurable partition $\mathcal{A} = \{A_1, \dots, A_K\} \subset \mathcal{F}$ of Ω (atoms). Write the time-dependent probabilities

$$\mathbf{P}(t) = (P_1(t), \dots, P_K(t)) \in \Delta^{K-1} \quad \text{with} \quad P_k(t) = P(A_k, t), \quad \sum_{k=1}^{K} P_k(t) = 1, \quad P_k(t) \geq 0.$$

Let $\boldsymbol{z} \in \Delta^{K-1}$ be an *anchor* (baseline distribution, e.g. a model prior or baseline predictor). Denote $\mathbf{1} = (1, \dots, 1)^\top$. The *probability simplex* is $\Delta^{K-1} = \{\mathbf{p} \in \mathbb{R}^K : \; p_k \geq 0, \; \sum_k p_k = 1\}$.

Remark 1 (Why dynamics? Bayes as impulses)**.** Bayes' rule updates $\mathbf{P}$ when new evidence arrives at discrete times. UPT adds *continuous-time* dynamics between evidence arrivals to model endogenous reorganization (resonance, cascades). Section 4 embeds Bayes updates as impulses superimposed on the flow.

2 Operators: definitions grounded in classical terms

2.1 Probability potential and gradient field

Definition 1 (Potential)**.** Define the scalar potential

$$\mathcal{V}(\mathbf{P}) \;=\; \sum_{k=1}^{K} \tfrac{1}{2}\, \phi_a \, (P_k - z_k)^2,$$

with strength $\phi_a > 0$. Its coordinate gradient is $g_k(\mathbf{P}) = \partial \mathcal{V} / \partial P_k = \phi_a \, (P_k - z_k)$.

Remark 2 (Force on probabilities = vector field)**.** "Force" is shorthand for the right-hand side of an ODE on Δ^{K-1}. To make it conservative, we use the zero-sum projection

$$\Pi(\mathbf{v}) \;=\; \mathbf{v} - \frac{1}{K}\, (\mathbf{1}^\top \mathbf{v})\, \mathbf{1},$$

so that $\mathbf{1}^\top \Pi(\mathbf{v}) = 0$. Potential descent uses $-\Pi(\nabla \mathcal{V})$; potential ascent uses $+\Pi(\nabla \mathcal{V})$. The demo uses the same potential shape; theory below is agnostic in the sign.

2.2 Emergence (viability) operator

Definition 2 (Emergence / Viability)**.** For each A_k, define a viability nonlinearity

$$\Lambda_k(\mathbf{P}) \;=\; \beta \,\tanh\!\left(\frac{V_k(\mathbf{P}) - V_{\text{thresh}}}{\sigma}\right), \quad V_k(\mathbf{P}) = \tfrac{1}{2}\,\phi_a\,(P_k - z_k)^2,$$

with parameters $\beta \in \mathbb{R}$, $V_{\text{thresh}} \in \mathbb{R}$, $\sigma > 0$. To preserve total probability we use the zero-sum form

$$\widehat{\Lambda}_k(\mathbf{P}) \;=\; \Lambda_k(\mathbf{P}) \;-\; \frac{1}{K}\sum_{i=1}^{K} \Lambda_i(\mathbf{P}).$$

Remark 3. This is exactly the stabilizer used in the demo (turned small to isolate resonance), now recentered so the operator conserves total mass by construction.

2.3 Resonance: pairwise conservative coupling

Definition 3 (Resonance kernel and flux)**.** Let $S_{ij}(\mathbf{P})$ be a symmetric pairwise kernel encoding co-activation strength between A_i and A_j. A canonical choice aligned with the paper is

$$S_{ij}(\mathbf{P}) \;=\; A_{\text{res}}\,\tanh\!\left(\frac{\widehat{V}_{ij}(\mathbf{P}) - V_c}{\tau_{\text{res}}}\right), \qquad \widehat{V}_{ij}(\mathbf{P}) = \tfrac{1}{2}\big(V_i(\mathbf{P}) + V_j(\mathbf{P})\big),$$

where $A_{\text{res}}, V_c, \tau_{\text{res}} > 0$ and $S_{ij} = S_{ji}$. Define conservative resonance fluxes

$$r_{i\to j}(\mathbf{P}) \;=\; \gamma\, S_{ij}(\mathbf{P})\, P_i, \qquad \gamma > 0,$$

and the net resonance term per coordinate

$$R_k(\mathbf{P}) \;=\; \sum_{j\neq k} \Big(r_{j\to k}(\mathbf{P}) - r_{k\to j}(\mathbf{P})\Big).$$

Lemma 1 (Resonance conserves total probability)**.** $\sum_{k=1}^{K} R_k(\mathbf{P}) = 0$ *for all* $\mathbf{P}$.

Proof. Sum telescopes pairwise since each $r_{i\to j}$ appears once with $+$ and once with $-$. □

2.4 Cascade: threshold-gated, antisymmetric reallocation

Definition 4 (Cascade potential)**.** Define the cascade potential

$$\Phi(\mathbf{P}) \;=\; \max_{i\neq j}\, \Big(S_{ij}(\mathbf{P}) \cdot E_{ij}(\mathbf{P})\Big), \quad E_{ij}(\mathbf{P}) = \beta\,\tanh\!\left(\frac{\widehat{V}_{ij}(\mathbf{P}) - V_{\text{thresh}}}{\sigma}\right).$$

Definition 5 (Cascade flux and operator)**.** Let $[x]_+ = \max\{x, 0\}$. For each ordered pair (i, j) with $i \neq j$, define gated weights

$$W_{i\to j}(\mathbf{P}) \;=\; \mathbf{1}\{\Phi(\mathbf{P}) \geq \Phi_{\text{critical}}\}\, \frac{\big[S_{ij}(\mathbf{P})\, E_{ij}(\mathbf{P})\big]_+}{\sum_{\ell\neq i} \big[S_{i\ell}(\mathbf{P})\, E_{i\ell}(\mathbf{P})\big]_+ + \varepsilon}, \quad \varepsilon > 0.$$

Set cascade fluxes $c_{i\to j}(\mathbf{P}) = W_{i\to j}(\mathbf{P})\, P_i$, and the net cascade term

$$\mathcal{C}_k(\mathbf{P}) \;=\; \sum_{j\neq k} \Big(c_{j\to k}(\mathbf{P}) - c_{k\to j}(\mathbf{P})\Big).$$

Lemma 2 (Cascade conserves total probability)**.** $\sum_{k=1}^{K} \mathcal{C}_k(\mathbf{P}) = 0$ *for all* $\mathbf{P}$*.*

Proof. As for resonance, the pairwise antisymmetry cancels in the sum. $\square$

3 Unified evolution on the simplex

Definition 6 (UPT vector field)**.** Let $\sigma_V \in \{-1, +1\}$ choose descent (-1) or ascent $(+1)$ on $\mathcal{V}$. The UPT field is

$$F(\mathbf{P}, t) \;=\; \underbrace{\sigma_V \, \Pi\big(\nabla \mathcal{V}(\mathbf{P})\big)}_{\text{potential}} \;+\; \underbrace{\widehat{\Lambda}(\mathbf{P})}_{\text{emergence}} \;+\; \underbrace{R(\mathbf{P})}_{\text{resonance}} \;+\; \underbrace{\alpha(t)\, \mathcal{C}(\mathbf{P})}_{\text{cascade}},$$

with $\alpha(t) \geq 0$ the cascade schedule (e.g. your doubling law).

Theorem 1 (Well-posedness and simplex invariance)**.** *Assume $\alpha(\cdot)$ is locally bounded and piecewise continuous. Then the ODE $\dot{\mathbf{P}} = F(\mathbf{P}, t)$ has (Carathéodory) solutions on $[0, \infty)$ for any $\mathbf{P}(0) \in \Delta^{K-1}$. Moreover,*

$$\sum_{k=1}^{K} \dot{P}_k(t) = 0, \qquad P_k(t) \geq 0, \qquad \sum_{k=1}^{K} P_k(t) = 1, \quad \forall t \geq 0.$$

Sketch. All components are locally Lipschitz in $\mathbf{P}$ (tanh is smooth; the rational gating has $\varepsilon > 0$), so existence/uniqueness holds piecewise in t. By construction $\mathbf{1}^\top F(\mathbf{P}, t) = 0$ (potential term is projected; others are pairwise antisymmetric), hence $\sum_k P_k$ is constant. Nonnegativity follows from standard inward-pointing arguments on the boundary of Δ^{K-1}: when $P_k = 0$, all outward fluxes $r_{k\to j}, c_{k\to j}$ vanish and the potential/emergence projections do not decrease P_k below 0 for sufficiently small time; see Nagumo's theorem on viability. $\square$

Corollary 1 (Probability conservation)**.** $\int_\Omega dP(\cdot, t) = \sum_k P_k(t) = 1$ *for all* t*.*

Theorem 2 (Entropy bound)**.** *Let $S(\mathbf{P}) = -\sum_k P_k \log P_k$. Then $S(\mathbf{P}(t)) \leq \log K$ for all t, with equality iff $\mathbf{P}(t)$ is uniform.*

Proof. Standard Shannon bound on the simplex; dynamics need not be entropy-monotone, but the upper bound always holds on Δ^{K-1}. $\square$

4 Bayes as a special case (impulsive updates)

Let evidence events E_n arrive at times t_n. Let $L^{(n)} = (L_1^{(n)}, \ldots, L_K^{(n)})$ with $L_k^{(n)} = P(E_n \mid A_k)$ (likelihoods). Define the Bayes jump map

$$\mathcal{B}_{E_n}(\mathbf{P}) \;=\; \frac{L^{(n)} \odot \mathbf{P}}{\mathbf{1}^\top\big(L^{(n)} \odot \mathbf{P}\big)}.$$

UPT with impulses:

$$\dot{\mathbf{P}} = F(\mathbf{P}, t) \;+\; \sum_n \delta(t - t_n) \Big(\mathcal{B}_{E_n}(\mathbf{P}^-) - \mathbf{P}^-\Big).$$

If $F \equiv 0$ one recovers pure Bayes: piecewise-constant $\mathbf{P}$ with jump $\mathbf{P}^+ = \mathcal{B}_{E_n}(\mathbf{P}^-)$ at each t_n.

5 Resonance/cascade phases and horizons

Definition 7 (Phases)**.** Write $\Phi(\mathbf{P})$ as above. Then:

- Stable: $\Phi(\mathbf{P}) < \Phi_{\text{critical}}$ (no cascade flux).
- Resonant transition: S_{ij} near its steep band (tanh argument near V_c); small amplification via R.
- Cascade: $\Phi(\mathbf{P}) \geq \Phi_{\text{critical}}$ and $\alpha(t) > 0$; rapid reallocation via $\mathcal{C}$.

Definition 8 (Event graph and horizon)**.** Let $G = (V, E)$ be a graph on $\{1, \ldots, K\}$ encoding structural proximity (e.g. ordinal neighbors). Let $d_G(i)$ be graph distance from a reference set. Suppose the cascade kernel $W_{i \to j}(\mathbf{P})$ has compact support $d_G(i, j) \leq r$. If $\alpha(t) \geq \alpha_{\text{horizon}}$ for $t \geq T$, then nodes with $d_G(i) > r$ receive no cascade inflow and can be driven to vanishing mass by sustained outflow (model-dependent). We call r the probability horizon radius.

Remark 4. This formalizes the horizon statement as locality of cascade support on an event graph. Concrete decay rates follow from bounds on $S_{ij}E_{ij}$ beyond radius r.

6 Alignment with the reference implementation

- **Potential:** $\mathcal{V} = \sum_k \frac{1}{2}\phi_a(P_k - z_k)^2$; code uses $g_k = \phi_a(P_k - z_k)$.
- **Emergence:** $\beta \tanh((V_k - V_{\text{thresh}})/\sigma)$ (recentered here to zero-sum).
- **Resonance:** neighbor-only S_{ij} with $\widehat{V}_{ij} = \frac{1}{2}(V_i + V_j)$ for $|i - j| = 1$.
- **Cascade:** present as a stub in the demo; the conservative flux form above is a drop-in replacement that preserves $\sum_k P_k$ and matches the paper's gating ($\Phi \geq \Phi_{\text{critical}}$) and schedule $\alpha(t)$ (e.g. doubling).
- **Simplex invariance:** the demo projects to the simplex after each step; theory uses zero-sum projection Π and inward-pointing arguments to guarantee invariance in continuous time.

7 Plain-language glossary (for the avoidance of doubt)

- **Force on probabilities:** the vector field (right-hand side) of the ODE driving $\mathbf{P}(t)$ on the simplex; not literal energy.
- **Viability / Emergence:** a smooth switch that turns on when an event's potential clears a threshold; implemented by tanh.
- **Resonance:** pairwise co-activation making linked events exchange probability mass (conservatively).
- **Cascade:** when (resonance $\times$ emergence) crosses a threshold, a structured, antisymmetric flow rapidly reallocates mass across many events.
- **Horizon:** locality radius of cascade influence on an event graph; beyond it, events can be suppressed to near zero under sustained cascades.
- **Bayes vs UPT:** Bayes = jumps at evidence times; UPT = continuous flow between jumps, with Bayes embedded as impulses.

Interpretation metaphor (why this matters)

In many real settings (scientific reading, model selection, multi-hypothesis reasoning), belief does not update as isolated jumps only. Local co-activations build until a few connections cross a threshold; then a *cascade* reorganizes the whole distribution in a burst. UPT provides the minimal conservative mechanics on the probability simplex to describe that trajectory: potentials bias, viability gates, resonance couples, and once criticality is reached, cascades rapidly reallocate probability—all while conserving total mass and remaining compatible with Bayes at evidence times.