Phase Transitions in Bounded Agency

A Geometric Framework for Insight and Collective Intelligence

Alexander Sabine & Nicolas Hinrichs

Royal Society Phil Trans A - March 2026

Royal Society Phil Trans A: Complete Proposal

Special Issue: World/Self-Models, Agency, Reasoning/Planning

Deadline: March 31st, 2026 | **Call:** https://www.adamsafron.com/agency

---

1. Paper Options

Rank	Title	Evidence/Claims	Novelty
1	Phase Transitions in Bounded Agency: Geometric Detection of Insight	**0.75**	High
2	The 16 Nats Threshold: Universal Information Limits	0.65	Very High
3	Geometric Hyperscanning Meets Active Inference	0.70	Medium
4	System 2 Cognition as Dyadic Rupture Synchronization	0.60	High
5	From Cells to Societies: CRR as Scale-Free Grammar	0.55	High
6	Critical Periods and Cognitive Plasticity	0.60	Medium

2. System 1 and System 2: The Core Framework

2.1 Kahneman's Dual-Process Theory

Daniel Kahneman's *Thinking, Fast and Slow* (2011) distinguishes two modes of cognition:

System 1	System 2
Fast, automatic	Slow, deliberate
Effortless	Effortful
Parallel processing	Serial processing
Unconscious	Conscious
Heuristic-based	Rule-based
Associative	Analytical

Key insight: Most cognition is System 1. System 2 is invoked only when System 1 fails or detects novelty.

2.2 CRR Reformulation of System 1/2

The Coherence-Rupture-Regeneration (CRR) framework provides a **mechanistic account** of when and how System 2 engages:

System 1 (CRR)	System 2 (CRR)
Coherence phase (C < Omega)	**Rupture event** (C >= Omega)
Within-model inference	Between-model transition
Gradient descent on free energy	Discrete model switching
Exploitation	Exploration
Amortized inference	Explicit reasoning
Continuous dynamics	Discontinuous dynamics

The key claim: System 2 is not a separate "system" but the **rupture phase** of a unified CRR cycle. Deliberate reasoning emerges when accumulated evidence (coherence) crosses threshold.

2.3 Why This Matters for AI

Current Large Language Models (LLMs) approximate **System 1 only**:

- Autoregressive generation = amortized inference
- No explicit model switching mechanism
- No accumulation-to-threshold dynamics
- Cannot "step back and reason"

CRR prediction: True System 2 in AI requires: 1. Explicit coherence tracking 2. Rupture mechanism at threshold 3. Memory-weighted regeneration 4. Discontinuous transitions

See: `crr_active_reasoning.md` for full formulation.

2.4 System 2 as Dyadic Phenomenon

A radical implication of the CRR-Hyperscanning synthesis:

System 2 cognition may be fundamentally social.

When two agents interact, their coherence dynamics **couple**. Synchronized rupture—when both agents cross threshold together—produces qualitatively different insight than individual reasoning.

Individual System 2	Dyadic System 2
Internal threshold crossing	Synchronized threshold crossing
Monadic insight	Shared understanding
Private reasoning	"Thinking together"
Measured by EEG	Measured by hyperscanning

Geometric Hyperscanning (Hinrichs et al., 2025) provides the measurement methodology: peaks in curvature entropy H(kappa) across the inter-brain network mark **dyadic System 2 events**.

---

3. Recommended Abstract

Phase Transitions in Bounded Agency: Geometric Detection of Insight and Collective Intelligence

Authors: Alexander Sabine & Nicolas Hinrichs

Abstract (250 words)

Agency emerges across scales from cells to societies, yet unified frameworks for its temporal dynamics remain elusive. We synthesize two complementary approaches: **Coherence-Rupture-Regeneration (CRR)**, providing temporal grammar for phase transitions in bounded agents, and **Geometric Hyperscanning**, operationalizing these transitions via Forman-Ricci curvature of inter-brain networks.

Our central thesis: agency is fundamentally *discontinuous*. Agents alternate between coherent exploitation (System 1: gradient descent on free energy) and rupture events (System 2: discrete model switching), with memory-weighted regeneration enabling cumulative learning. We establish mathematical correspondence with the Free Energy Principle, showing CRR provides the missing "between-model" transition structure.

Three key results emerge:

1. **Universal Threshold:** Information capacity before rupture converges on Omega = 16 nats (~23 bits) across biological and cognitive systems—resolving ~10^7 distinguishable states.
2. **Euler Calibration:** At rupture (C = Omega), the memory kernel satisfies exp(C/Omega) = e, providing exact mathematical calibration.
3. **Geometric Operationalization:** Forman-Ricci curvature entropy peaks mark rupture events; this provides a direct neural signature of System 2 engagement.

We address Question 1 (agency taxonomy) by proposing the rigidity parameter Omega characterizes where systems fall on the System 1/System 2 continuum. We address Question 4 (agency across scales) by demonstrating CRR applies uniformly from cellular decisions through individual cognition to collective intelligence.

Testable predictions: (i) EEG precision correlates with coherence; (ii) curvature entropy peaks at insight; (iii) inter-brain coupling accelerates dyadic System 2 reasoning.

Keywords: agency, System 1/System 2, free energy principle, phase transitions, hyperscanning

---

4. Mathematical Framework

4.1 Core CRR Operators

Coherence Accumulation:

$$C(x,t) = \int_0^t L(x,\tau)\, d\tau$$

Rupture Condition (System 2 Trigger):

$$\delta(t - t^*) \quad \text{when} \quad C \geq \Omega$$

Regeneration Operator:

$$R[\phi](x,t) = \int_0^t \phi(x,\tau) \cdot e^{C/\Omega} \cdot \Theta(t-\tau)\, d\tau$$

4.2 FEP-CRR Correspondence

FEP	CRR	System 1/2
Free Energy F(t)	Coherence C(t) = F0 - F	System 1 metric
Precision Pi	(1/Omega)exp(C/Omega)	Confidence
Gradient descent	Coherence phase	System 1
Model switching	Rupture	System 2 trigger

4.3 The 16 Nats Threshold

Empirical convergence:

System	Measured	CRR Prediction
Working memory	17 nats	16 nats
Visual STM	15 nats	16 nats
Conscious bandwidth	16 nats	16 nats
Hyperscanning H(kappa)	15.9 nats	16 nats

Interpretation: 16 nats = log(10^7) corresponds to resolving ~9 million states—the complexity threshold requiring System 2.

4.4 Euler Calibration

At rupture (C = Omega):

$$\exp(C/\Omega)\big|_{C=\Omega} = e \approx 2.718$$

This is **exact**, not fitted. The memory kernel has characteristic value e at the System 1 to System 2 transition.

4.5 Geometric Hyperscanning

Forman-Ricci Curvature:

$$\kappa_F(e) = 4 - d(v_1) - d(v_2)$$

Phase Transition Marker:

$$H(\kappa) = -\sum_i p(\kappa_i) \log p(\kappa_i)$$

Peak in H(kappa) = **dyadic System 2 event**

---

5. Figures

Figure 1: CRR Cycle

File: `diagrams/crr_cycle-1.png`

The Coherence-Rupture-Regeneration cycle. System 1 operates during coherence phase (C < Omega). System 2 triggers at rupture (C >= Omega). Regeneration consolidates learning with memory kernel exp(C/Omega).

Figure 2: Q-Factor Correlation

File: `q_omega_correlation.png`

Empirical relationship between substrate Q-factor and rigidity Omega. Power law fit R^2 = 0.94. High-Q (rigid) substrates = System 1 dominated. Low-Q (dissipative) = System 2 accessible.

Figure 3: Exploration-Exploitation Phase Space

File: `precision_coherence.png`

Phase diagram showing System 1 (exploitation, high precision, low Omega) vs System 2 (exploration, low precision, high Omega) regimes.

Figure 4: Validation

File: `crr_wound_validation_plot.png`

CRR prediction vs empirical wound healing data. R^2 = 0.9989. Demonstrates CRR captures real biological phase transitions.

---

6. Repository Resources

Key Documents

File	Description
`crr_active_reasoning.md`	CRR formulation of active inference and System 2
`fep_crr_integration.md`	Full FEP-CRR mathematical correspondence
`crr_16_nats_hypothesis.md`	16 nats derivation and cross-system validation

File	Description
crr_hyperscanning_16_nats_inductive_test.md	Current hyperscanning analysis
CRR_COMPREHENSIVE_SUMMARY.md	Complete CRR overview
crr_simulation.py	Full simulation framework

Interactive Demos (HTML)

Demo	CRR Concept
16nats_simulation.html	16 nats threshold visualization
fep_crr_dynamics.html	FEP-CRR dynamics
crr-three-phase-visualiser.html	Coherence-Rupture-Regeneration phases
precision_coherence.png	Exploration-exploitation
crr-brain-photorealistic.html	Neural CRR dynamics
fep-crr-game.html	Interactive CRR exploration
child_dev.html	Developmental CRR (critical periods)
Maze.html	Decision-making and System 2
ecosystem.html	Collective CRR dynamics

---

7. Simulation Evidence

Test Results

```
CRR-AHA MOMENT TEST SUITE RESULTS
==================================
Test 1: Bayesian Model Reduction     PASSED
Test 2: Information at Insight       15.9 nats (within 1% of 16)
Test 3: Precision-Coherence          PASSED (exact)
Test 4: Euler Calibration            EXACT (e = 2.718282)
Test 5: Hyperscanning Integration    PASSED

Evidence/Claims Ratio: 0.75
```

Key Quantitative Results

Metric	Value
Q-Omega R^2	0.94
Wound healing R^2	0.999
H(kappa) at rupture	15.9 nats
Euler calibration	2.718282 (exact)

---

8. Research Integration

8.1 Friston & Da Costa: Aha Moments

- **Paper:** "Active Inference, Curiosity and Insight" (2017)
- **Key:** Bayesian model reduction triggers insight
- **CRR:** Rupture = BMR trigger point

8.2 Hinrichs: Geometric Hyperscanning

- **Paper:** "Geometric Hyperscanning of Affect" (2025)
- **Key:** Forman-Ricci curvature tracks phase transitions
- **CRR:** H(kappa) peak = rupture/System 2 event

8.3 Critical Periods (Knudsen, de Villers-Sidani)

- **Key:** Plasticity maximal during critical periods
- **CRR:** Critical period = low Omega (easy rupture)
- **Closure:** Omega increases, System 2 harder to trigger

8.4 Safron: IWMT

- **Key:** Consciousness as integrated world model
- **CRR:** Self-organizing harmonic modes ~ coherent phases
- **Alignment:** Workspace ignition ~ rupture

---

9. References

Core:

- Sabine, A. (2025). CRR: A Memory-Augmented Variational Framework. Working paper.
- Hinrichs, N. et al. (2025). Geometric hyperscanning of affect. arXiv:2506.08599.
- Friston, K. et al. (2017). Active inference, curiosity and insight. Neural Computation, 29(10).
- Kahneman, D. (2011). Thinking, Fast and Slow. FSG.

Supporting:

- Friston, K. (2010). The free-energy principle. Nature Reviews Neuroscience, 11(2).
- Knudsen, E.I. (2004). Sensitive periods. J Cogn Neurosci, 16(8).
- Safron, A. (2020). IWMT of consciousness. Frontiers in AI, 3.
- Cowan, N. (2001). The magical number 4. BBS, 24(1).

---

10. Quick Reference

For Nicolas

CRR provides: - When System 2 engages (C >= Omega) - Universal threshold (16 nats) - FEP integration

Geometric Hyperscanning provides: - How to measure it (curvature entropy) - Neural/dyadic operationalization - Empirical validation pathway

Together: - Complete System 1/2 framework - Testable predictions - Scale-free (cells to societies)

Prepared January 2026 for Royal Society Phil Trans A submission March 2026