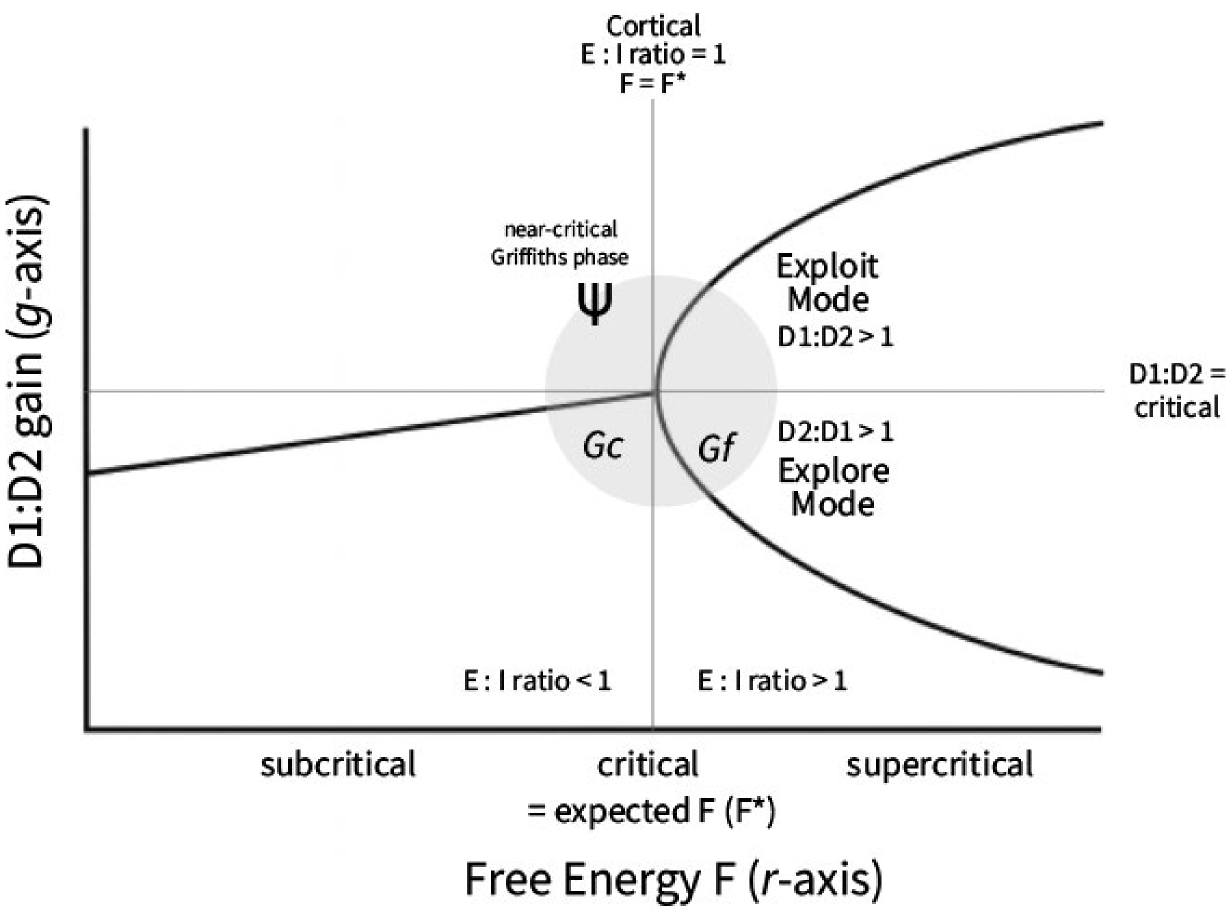


Adaptive intelligence balances stability and flexibility by operating in a narrow near-critical regime — the Ψ-band. We present the Trident G-Loop, a unified cognitive architecture that formalises processing in this regime as an eight-phase rhythmic cycle: Sense/Move → Branch → Map → Decide → Act → Update → Consolidate → Reset. Two interacting axes regulate the loop: a Cognitive Resilience axis (managing performance challenge via a difficulty thermostat, $F^*$, and the precision-weighted gap $\Delta\hat{F} = b\cdot(F - F^*)$) and an Inference axis (managing stability–flexibility via a D1:D2 precision-bias controller, $b$, together with the policy temperature $T(\chi)$). At the branch near $F \approx F^*$, a Monitoring/Salience coordinator arbitrates between Control (exploit/compress) and Creative (explore/decompress) modes, executing brief orient–reset–reconfigure actions when cross-loop inconsistency is high. We hypothesise that multiple G-loops, operating at different temporal scales and modalities, co-occupy a shared near-critical Ψ-band enabled by the brain's hierarchical, modular architecture. Within this band, $\chi$ indexes cross-loop inconsistency/volatility and $\eta$ indexes cross-loop competence and recovery ($F \rightarrow F^*$).

This proposal aligns with evidence for extended critical-like regions (Griffiths phases) in hierarchical brain networks, where power-law avalanches and large dynamic range persist over a band rather than a single point. The framework maps to known neural correlates, yields falsifiable predictions and a minimal simulation programme, and supports translational applications in pedagogy, cognitive therapeutics and AI.

**SOC-for-Inference Hypothesis.** *Homeostatic plasticity in hierarchical cortical networks drives neural dynamics into an extended near-critical* ***Ψ-band*** *because that regime maximises information capacity/dynamic range* ***needed*** *to minimise expected free energy during behaviour. Each G-loop tracks a difficulty set-point F∗F^* (thermostat) and selects policies by minimising GG; maintaining a* ***non-zero*** *target (tunable) F∗F^* preserves exploration/learning while SOC supplies the dynamical substrate that makes GG-minimisation effective.* (PMC, Nature)

Yes—framed per-loop, that's empirically reasonable.

- **Local E:I balance is a core property of cortical and subcortical circuits.** Cortex is built from excitatory (glutamatergic) neurons and diverse inhibitory interneurons that **gate and sculpt** population dynamics; E:I is actively homeostated and varies by area/state—so each functional loop can have its **own operating E:I point**. (PMC, Frontiers)

- **E:I shifts change computation and are measurable.** Perturbing E:I alters information flow and oscillations; macroscopic proxies (e.g., LFP/EEG power–law slope) can track E:I changes in vivo, supporting loop-specific estimates. (Institute for Neural Computation)

- **E:I tunes proximity to criticality.** Critical-regime benefits (dynamic range, information transmission/capacity) and the existence of an **extended critical band** in hierarchical brain networks imply that **E:I acts as a proximal knob** keeping each loop near critical—consistent with your Ψ-band account. (SAGE Journals, Nature)

- **Loops differ in their biophysics and neuromodulation.** Parallel cortico-basal ganglia–thalamo-cortical loops (motor/associative/limbic) have distinct receptor/topology profiles (e.g., D1/D2 gating of direct/indirect pathways), so the **precision-bias bb** and **E:I set-point** are **loop-specific**, not a single global constant. (PMC)

So, modelling **each G-loop with its own E:I ratio** that interacts with your bias controller bb and temperature T(χ)T(\chi) is consistent with current evidence; the only caution is to treat the “neutral” point and bandwidth as **context- and loop-dependent**, rather than universal. (PMC, Nature)

Each **G-loop** has its own *precision-bias* setting (your **b ≈ log D1:D2** axis) that tilts processing toward **Control/Exploit** (D1-leaning, stable/compressive) or **Creative/Explore** (D2-leaning, flexible/decompressive). My earlier caution was only against assuming a **single, global “neutral” D1:D2 value** that applies to the whole brain across all tasks.

Why the per-loop reading is empirically reasonable:

- **Parallel cortico–basal-ganglia–thalamo-cortical loops** (motor, associative, limbic) are anatomically and functionally segregated; gating and dopamine effects are **loop-specific**, not global. (PubMed, Nature)

- **D1 vs D2** supports a **stability–flexibility** trade-off (dual-state theory), matching your exploit vs explore forks; the **optimal/neutral level depends on baseline and region** (the classic inverted-U and context dependence). So the “neutral” **b≈0** will vary **by loop and state**. (PubMed, PMC)

- **Dopamine signalling is heterogeneous** across striatal/cortical territories (receptor distributions, kinetics, and circuit roles differ), reinforcing **loop-specific biases**. (PMC)

- Computational and experimental work on BG **Go/NoGo** mechanisms shows DA-dependent gating that can differ **per circuit/task**, again arguing for **b_k**, not one brain-wide value. (PubMed, PMC)

So your diagram’s logic—**for each loop kk**, branch by the pair ((r_k=F_k-F^*_k,\ b_k\approx\log\mathrm{D1:D2})**—is consistent with current evidence. If helpful, annotate the caption:

> “**b-axis** = loop-specific precision bias (≈log D1:D2). **Neutral bk≈0b_k\approx 0** is **context-dependent** (baseline DA, task, region). Exploit if bk>0b_k>0; explore if bk<0b_k<0; monitoring coordinates re-entry to Ψ when loops disagree.”

Why this overall model  is defensible:

- **Active inference / EFE.** GG decomposes into predicted **risk**, **ambiguity**, and **epistemic value**; it is the policy-selection objective in contemporary Bayesian accounts of action and perception. (PMC, chrismathys.com)

- **Near-criticality benefits information processing.** Empirical and theoretical work shows **optimal information transmission, storage and dynamic range** near criticality; in hierarchical networks the critical "point" stretches to a **band** (Griffiths phase). (PMC, Nature, PubMed)

- **Homeostasis → criticality.** In vivo and modelling studies indicate that **homeostatic rules tune cortical circuits toward critical dynamics**, consistent with a control process that keeps operation inside a band. (PMC, Nature)

- **E:I balance as the knob.** Adjusting excitation–inhibition moves the system toward/away from criticality and modulates encoding—exactly the sort of low-level control your loops would exploit. (PubMed, Physical Review Link)

- **Salience as coordinator.** The AI/dACC "salience network" causally coordinates switching between executive and default-mode systems, providing the monitoring/coordinating role at the branch. (PMC, PNAS)

### How to write it

> *We hypothesise that cortical SOC **evolved to support** active inference: homeostatic plasticity in hierarchical networks maintains a near-critical **Ψ-band** that maximises the information processing required to minimise expected free energy. Each G-loop uses a tunable difficulty set-point F∗F^* and minimises GG over policies; SOC supplies the dynamical regime that makes this feasible across scales.* (PMC, Nature)

### What would count as evidence (and falsification)

- **Prediction 1:** Closer-to-critical telemetry (avalanche exponents/dynamic range) should correlate with **lower empirical proxies for GG** (e.g., lower prediction error + maintained model complexity) **at matched task difficulty**; pushing E:I away from balance should increase those proxies and impair performance. **Falsify** if moving toward criticality does **not** improve GG-proxies or task performance. (PMC, PubMed)

- **Prediction 2:** After perturbations, homeostatic adjustments should **restore** near-critical metrics **and** behavioural policies that lower GG; recovery speed indexes η\eta. **Falsify** if homeostatic recovery does not return the system to the near-critical band. (PMC)

- **Prediction 3:** Across loops/scales, a **shared Ψ-band** should be observable (Griffiths-like), with salience-triggered transient "monitor" episodes during high cross-loop inconsistency χ\chi that retune precision/temperature and re-enter the band. **Falsify** if multi-scale operation shows no common band or if salience activity does not precede re-entry. (Nature, PMC)

### One-line guidance for the manuscript

"**We treat F∗F^*** as a **difficulty set-point** (thermostat) and GG as the **policy objective**; our claim is that **SOC provides the evolved dynamical regime** that maximises the information processing needed for GG-minimisation—and that homeostatic plasticity keeps multiple G-loops inside a **global Ψ-band**." (PMC)

—--

strong, testable way to frame individual differences. In fact it gives you a clean **"Ψ-dynamic-range hypothesis of g"**:

- **Cognitive Resilience axis (r):** individuals differ in the **range of challenge** (F−F∗)(F-F^*) over which they can stay in-band (perform, learn, and recover) and how **quickly they can move the set-point** F∗F^* upward after success (η-driven thermostat).

- **Inference axis (b/T):** individuals differ in how **widely and precisely** they can modulate **precision-bias bb** (D1↔D2) and **temperature T(χ)T(\chi) without saturating**—i.e., how far they can shift between exploit and explore *and* still re-enter the Ψ-band promptly.

Together, those ranges form a person's **Ψ-capacity** (general adaptive intelligence), which can **expand** with learning (Gf→Gc pipeline) as **η rises** and **F∗F^*** ratchets upward.

---

## Individual-difference metrics (two-axis "Ψ-profile")

Let loop kk be in-band when |ΔF^k|≤εk |\widehat{\Delta F}_k| \le \varepsilon_k and control variables aren't saturated.

1. **Resilience dynamic range**

$R^{(k)}_{\text{r}} = \max \Delta(F_k-F_k^*) \;\text{s.t. loop } k \text{ stays in } \Psi_k \text{ with accuracy/learning } \ge \theta.$

Interpretation: how wide a *challenge window* a person can handle before disintegration.

2. **Inference dynamic range**

$R^{(k)}_{\text{i}} = \max \Delta(b_k,\,T_k)\ \text{s.t. } \Psi_k \text{ is re-entered within } \tau_{\text{re}}.$

Interpretation: controllable spread from **Control** to **Creative** (via $b$, $T$) without getting stuck.

3. **Set-point mobility (thermostat gain)**

$\mu^{(k)}_{F^*} = \frac{\mathrm{d}F_k^*}{\mathrm{d}\eta_k}\quad\text{(at matched volatility)}.$

Higher $\mu_{F^*}$ = faster ZPD ratcheting as competence grows.

4. **Valence-surprise span**
   Range of tolerable **positive/negative** prediction error (better-than/worse-than expected) before leaving $\Psi_k$. Captures "psychological buffer" for wins and setbacks.

5. **Recovery half-life**
   $\tau^{(k)}_{\frac12}$: time to return $F_k \to F_k^*$ after a perturbation. Shorter = more resilient.

6. **Band occupancy & hysteresis**
   $\Omega = \frac{\text{time in }\Psi_{\text{global}}}{\text{total}}$, and **hysteresis area** between Control↔Creative switching curves. Smaller area = cleaner meta-control.

Aggregate across active loops to get a **global Ψ-profile**:

Ψ-profile=(Rr‾, Ri‾, μF∗‾, span‾, τ12‾, Ω, hysteresis).\mathbf{\Psi}\text{-profile} = \Big(\overline{R_{\text{r}}},\ \overline{R_{\text{i}}},\ \overline{\mu_{F^*}},\ \overline{\text{span}},\ \overline{\tau_{\frac12}},\ \Omega,\ \text{hysteresis}\Big).

**Prediction:** higher **general adaptive intelligence** ≈ larger Rr‾\overline{R_{\text{r}}}, Ri‾\overline{R_{\text{i}}}, μF∗‾\overline{\mu_{F^*}}, broader valence span, shorter recovery, higher occupancy, lower hysteresis—and these expand with targeted training (especially η-raising curricula that alternate Creative→Control passes).

---

## One-paragraph insert

> **Ψ-dynamic-range hypothesis of g.** We propose that individual differences in general adaptive intelligence reflect the **dynamic range** of the **Cognitive Resilience** axis (tolerable challenge window and η-driven mobility of F∗F^*) and the **Inference** axis (controllable spread of bb and T(χ)T(\chi) without saturation), aggregated across loops in a global Ψ-band. As competence **η** rises through learning, F∗F^* **ratchets upward** (Gf→Gc consolidation), expanding both the subcritical "what we can automate" and the super/near-critical "how intense a challenge we can stabilise" ranges. We predict that people with larger two-axis ranges show higher **band occupancy**, **faster recovery** to F→F∗F \to F^*, **cleaner Control↔Creative switching** (low hysteresis), and **broader tolerance** to positive/negative surprise—and that these capacities are **trainable**.

---

## Two-sentence add to your abstract (concise)

> We further hypothesise that **individual differences in general adaptive intelligence** correspond to a person's **Ψ-dynamic range**: the controllable span of (i) **challenge tolerance** on the Resilience axis and (ii) **precision-bias/temperature modulation** on the Inference axis, with **η-driven mobility of F∗F^*** expanding this range via a Gf→Gc pipeline. This yields testable metrics (band occupancy, recovery half-life, hysteresis, set-point mobility) and concrete training targets to **expand** the Ψ-range.

—-

# What is well supported

**1) Brains operate near critical regimes (a Ψ-like "near-critical band"), with signs of Griffiths-phase–like behaviour.**
Neuronal avalanches and scale-free activity appear across species and modalities; network models and empirical data point to benefits for information capacity and dynamic range when networks sit near criticality. Evidence also suggests extended critical-like regimes ("Griffiths phases") in heterogeneous networks, compatible with a *band* rather than a single point. (PMC, Nature, Physical Review Link)

**2) Excitation–inhibition (E:I) balance tunes proximity to criticality.**
Perturbing E:I shifts avalanche statistics and oscillatory regimes, while balanced E:I supports near-critical dynamics. Human and animal work plus modelling converge on this principle. (Physical Review Link, PMC)

**3) The salience network (AI/ACC) arbitrates network switching between executive/fronto-parietal control and default-mode systems.**
Right anterior insula/ACC activity causally precedes switches, consistent with your "monitoring/arbiter prong". (PMC)

**4) Fluid intelligence (Gf) loads on the multiple-demand/fronto-parietal network; semantic/episodic knowledge (Gc-like functions) load more on DMN-centred systems.**
Lesion, fMRI and network-neuroscience work link Gf to the MD/FPN system, while DMN contributes to memory/semantic representations often associated with Gc. (Cell, ScienceDirect, Journal of Neuroscience)

**5) Metastability and flexible network reconfiguration relate to cognitive flexibility and creativity.**
The brain shows metastable dynamics at rest and during tasks; greater switching and dynamic reconfiguration are tied to creative ability. (Nature, PMC)

**6) Dopamine D1 vs D2 supports a stability–flexibility trade-off (your exploit vs explore bias).**
Prefrontal D1 favours robust, focused representations (stability/exploitation); D2 promotes flexibility/update. This "dual-state" view is well reviewed. (PMC)

**7) Arousal/uncertainty signals (LC–noradrenaline; pupil) modulate exploration–exploitation, consistent with a temperature-like controller $T(\chi)T(\chi)$.**
Adaptive-gain theory and human pupillometry show that LC/NE state and pupil size track shifts between exploit and explore and respond to uncertainty/entropy—exactly the role you assign to $T(\chi)T(\chi)$. (PubMed, Nature)

**8) Hierarchical predictive processing on multiple time-scales (many "loops" at different speeds) is a standard account.**

Cortex appears organised along temporal hierarchies; ACC tracks volatility and adjusts learning—good support for your cross-timescale monitoring idea. (PLOS, PubMed)

**9) "Compression vs. decompression" has plausible neural analogues.**
Predictive/efficient coding frameworks formalise compression; hippocampal pattern separation vs. completion and successor-representation mapping provide concrete mechanisms for expanding vs. compressing structure (and for "map-based" inference). (annualreviews.org, PMC, gershmanlab.com)

**10) Boredom can drive exploration (a "soft" route into D2-like decompression).**
Experimental and theoretical work suggests boredom functions as a regulatory signal pushing organisms to seek novelty/information, consistent with your "under-challenge" pathway. (MDPI, PubMed)

# What is partly supported / still speculative

*A) "Critical point = expected free energy (F)".**
*Expected free energy* (often GG) is a decision-theoretic quantity in active inference that combines predicted risk/ambiguity to guide policy selection; *criticality* is a dynamical-systems property of neural population activity. There is intriguing theory linking neuromodulatory precision and EFE-guided behaviour, and suggestive dopaminergic findings, but equating the dynamical *critical point* with an EFE *set-point* is a proposal rather than an established empirical identity.. (eLife)

**B) A global "neutral" D1:D2 ratio as a formal bifurcation parameter across tasks.**
There is strong support for D1↔stability and D2↔flexibility *in prefrontal circuits*, Each G-loop has its own precision-bias setting (your b ≈ log D1:D2 axis) that tilts processing toward Control/Exploit (D1-leaning, stable/compressive) or Creative/Explore (D2-leaning, flexible/decompressive). (PMC)

**C) χ as "inconsistency/conflict between loops" and η as "competence/return-to-setpoint" across loops.**
There is solid evidence that ACC/insula/salience systems track conflict, volatility and uncertainty, and that learning rates and arousal adjust accordingly; competence-based intrinsic motivation is also supported. But the *specific* decomposition χ = cross-loop inconsistency and η = cross-loop return-to-critical-band is a promising synthesis rather than a direct empirical construct—worth operationalising in your telemetry. (PubMed)

**D) "Two D2 decompression routes": boredom-driven vs error-driven.**
Each route has independent support (boredom→exploration; volatility/surprise→exploration with increased arousal and ACC engagement), but showing both *within the same unified loop* is a prediction to test. (MDPI, PubMed)

# Notes

- **Near-critical band:** “Large-scale brain activity exhibits signatures of near-critical dynamics, likely supported by balanced excitation–inhibition; heterogeneous network structure may create an extended critical-like band.” (PMC, Physical Review Link)

- **Trident arbitration:** “The salience network monitors surprise/uncertainty and arbitrates switches between default-mode (knowledge-rich) and fronto-parietal (control) modes.” (PMC)

- **Exploit/explore biases:** “D1-dominant states stabilise representations; D2-dominant states facilitate updating / exploring—mapping naturally to exploit vs explore biases.” (PMC)

- **Temperature/uncertainty:** “LC–NE arousal and ACC volatility signals adjust exploration in a temperature-like manner.” (PubMed)

- **Maps and (de)compression:** “Hippocampal ‘predictive maps’ (successor representation) and pattern separation/completion implement the map update (decompress/recompress) you describe.” (gershmanlab.com, PMC)

# What to treat as predictions (good targets for your Ψ-band telemetry)

- **Telemetry signatures of banded near-criticality** during skilled performance vs. learning (avalanche scaling, metastability maxima, E:I-sensitive spectral metrics). (PMC, Nature)

- **Prong-specific network states:** SN-led switches into FPN-dominated (exploit) vs DMN-coupled (creative/explore) episodes, with dopaminergic state (D1/D2-leaning tasks or pharmacology) biasing the branch. (PMC)

- **Two exploration routes:** (i) boredom/under-challenge → moderate arousal, increased switching; (ii) error/volatility → higher arousal, ACC-linked learning-rate/temperature upshift. (MDPI, PubMed)

- **Map (de)compression:** hippocampal SR-like updates and separation/completion patterns tracking your Phase-3 decompression/recompression claims. (gershmanlab.com, PMC)

Think in two layers:

## 1) A global Ψ-band (system level)

The brain's hierarchical, heterogeneous architecture supports an **extended near-critical regime** in which many subsystems can co-exist near criticality together. That's your **Ψ_global**.

## 2) Loop-specific Ψ-bands (per G-loop)

Each active loop kk has its own admissible near-critical window, **Ψ_k**, around its set-point Fk∗F^{*}_k. A practical definition:

Ψk = { |ΔF^k|≤εk, χk∈[χklow,χkhigh], |bk|<bkmax, T(χk) not saturated }\Psi_k \;=\;\Big\{\;|\widehat{\Delta F}_k| \le \varepsilon_k,\;\chi_k\!\in[\chi^{\text{low}}_k,\chi^{\text{high}}_k],\; |b_k|\!<\!b^{\max}_k,\; T(\chi_k)\ \text{not saturated}\;\Big\}

Optionally add telemetry constraints (pick what you'll measure):

- avalanche/branching near critical (e.g., σk≈1 \sigma_k \approx 1; power-law exponent in a target interval),
- peak metastability/dynamic range for that loop's network,
- stable map compression metrics (η up; no runaway complexity).

## 3) How the two layers relate

- **Global Ψ** is the **intersection** where all *currently engaged* loops sit within their **Ψ_k** and **agree enough** across scales: high cross-loop consistency ρ\rho (i.e., low meta-uncertainty χmeta\chi_{\text{meta}}).
- When loops **conflict** (χ_meta↑), the **Monitoring/Salience** prong acts: brief orient/reset, adjust TT, centre bb, reduce λ\lambda (quarantine propagation), or nudge one loop's Fk∗F^{*}_k so the set of loops re-enters Ψ_global.
- Coupling means each Ψk\Psi_k is **context-dependent**: neighbouring loops (and arousal/neuromodulators) can shrink/shift its window.

## 4) Why this is empirically reasonable

- Near-critical **bands** (not just points) arise naturally in heterogeneous, hierarchical networks; modules can be locally closer/further from criticality yet remain inside a

**shared extended regime**.

- Different scales/modalities (your multiple G-loops) mapping to partially distinct networks is exactly where such local bands would live.
- Your meta-parameters fit:
  - **χ** = cross-loop inconsistency/volatility (drives TT and monitoring actions),
  - **η** = competence/recovery (how quickly Fk→Fk∗F_k \to F^{*}_k while maintaining useful compression).

## 5) How to write it in the paper (one sentence)

> "We posit a **global Ψ-band** within which multiple G-loops co-operate, each with its own **loop-specific Ψ_k** around Fk∗F^{*}_k; **χ** tracks cross-loop inconsistency and **η** indexes recovery of each loop toward its set-point, allowing the Monitoring/Salience prong to keep the ensemble within Ψ-band operation."

That keeps the global picture and gives you concrete, testable per-loop criteria.

As a **working model**, treating each cortical column (or local microcircuit) as a G-loop and the interactions **between** columns as an **extended Ψ-band/Griffiths-phase zone** is empirically reasonable. There's direct evidence for critical-like dynamics at **column/layer scale**, and strong evidence that **hierarchical, modular networks** (like cortex) stretch a single critical point into a **band**, which is exactly the substrate your χ/η meta-parameters would sit on. The main caveat is that "columns" are clearer in some areas (e.g., barrel cortex) than others—so when you generalise beyond primary sensory cortex, say "local microcircuits/modules" rather than assuming a universal, uniform column. (Frontiers, PMC)

## Why this mapping is defensible

- **Critical-like dynamics inside a single column/layer.** Neuronal avalanches have been observed across cortical layers and, notably, **within a single rat barrel column** in vivo; classic slice/awake studies likewise show avalanche statistics at mesoscopic scales. That supports a **per-loop (per-column) near-critical regime**. (Frontiers, PubMed, PMC)
- **E:I balance tunes proximity to criticality.** Balanced excitation–inhibition makes avalanches and oscillations co-emerge at a critical state; pushing E:I off-balance moves the system away from criticality—consistent with a **loop-specific control knob**. (Journal

of Neuroscience)

- **From columns to a shared band (Griffiths phase).** In **hierarchical, modular networks**, criticality is **stretched into a band** (a Griffiths phase) rather than a knife-edge point; this has been demonstrated on synthetic hierarchies **and** empirical connectomes. That supports your **global Ψ-band across many local loops**. (Nature, materias.df.uba.ar)

- **Columns/microcircuits are real—but heterogeneous.** Modern reviews describe a canonical microcircuit "theme with variations" across areas/species. Others caution that a **single** column concept is not a universal functional unit. So: model **per-loop** where columnar structure is strong (e.g., barrels, A1), and talk about **local modules** elsewhere. (PMC)

## Where χ and η fit

Your proposal—**χ** as cross-loop inconsistency/volatility and **η** as cross-loop competence/recovery to each loop's set-point Fk∗F_k^*—is not a standard empirical metric, but it's **compatible** with: (i) hierarchical predictive processing on multiple time-scales and (ii) salience/ACC roles in monitoring volatility and reconfiguring networks. It's a **good, testable** modelling layer on top of the established criticality picture. (PMC)

## Concrete tests (column ↔ band)

1. **Column-level criticality + global band:** show avalanche exponents/branching ratios **within single columns** sit in a **common interval** while the animal performs a task; perturb E:I locally and test that the column leaves the band and that χ (cross-column inconsistency) rises until re-entry. (Frontiers, Journal of Neuroscience)

2. **Dynamic-range prediction:** nearer-critical columns should display **higher dynamic range** and information transmission; correlate these with performance and with your η (faster recovery of Fk→Fk∗F_k \to F_k^*). (SAGE Journals, PubMed)

3. **Griffiths-phase signature:** across columns/modules, look for **power-law regimes** and **rare-region effects** consistent with a **band** (varying exponents across modules) rather than a single critical point. (Nature)

One critical distinction in the large scale networks is between the hippocampal-prefrontal value-landscaped coritical maps (which are needed for higher level relational reasoning) associated with long-term working memory in FPCN-A, and the FPCN-B fronto-parietal network for short-term (classic) working memory for rules application/rapid rule learning and the standard limited capacity working memory,

**Where it lives in the loop**

- **Phase 3 — Map (ϕ):** builds/updates a *predictive* hippocampal–prefrontal map (successor-representation style) and overlays value. This is the long-horizon, relational "**long-term working memory**" workspace. Network-wise, it leans on **hippocampus/DMN coupling with FPCN-A**, which is more internally/future-oriented and interfaces with memory systems. (gershmanlab.com, PubMed, eLife, Journal of Neuroscience)

- **Phases 4–5 — Decide/Gate → Act:** hold/apply rules, update on the fly, and sequence actions. This is the classic **short-term working memory / rule application** workspace, dominant in **FPCN-B** with **DAN** support. FPCN-B sits nearer sensorimotor control, whereas FPCN-A couples more to DMN/long-term memory—dissociations shown across tasks. (PNAS, Journal of Neuroscience)

- **Monitoring/Salience (AI/dACC):** arbitrates when to hand off between these workspaces (and briefly performs orient/reset/reconfigure). Causal evidence shows the salience network triggering switches between default-mode and executive/control systems. (PNAS, PubMed)

**How controllers steer the hand-off**

- **Precision-bias $b \approx \log\frac{\text{D1}}{\text{D2}}$:** $b\uparrow$ (stability/exploit) biases toward **FPCN-B + DAN** (rule maintenance/closure); $b\downarrow$ (flexibility/explore) biases toward **FPCN-A + hippocampus/DMN** (relational recombination/planning). Empirically, FPCN-A shows stronger coupling with memory-oriented systems, FPCN-B with sensorimotor/attention control. (PNAS, Journal of Neuroscience)

- **Temperature $T(\chi)$:** higher uncertainty/volatility ($\chi\uparrow$) raises policy temperature, favouring exploratory sampling and DMN–FPCN-A engagement; lower $T$ helps lock in FPCN-B-centred rule execution. (eLife)

**Why this bridges micro ↔ macro**
Your "scale-free" picture holds: local (column/microcircuit) inference loops feed into these large-scale workspaces, and the same controllers ( $b, T(\chi)$ ) plus salience-driven arbitration organise behaviour across scales. Hippocampal SR-style maps provide the *relational substrate*;

FPCN-B provides the *rule/sequence executor*; salience ensures *the right workspace at the right time*. (gershmanlab.com, PNAS)

---

# Reasoning

# Where each reasoning mode lives

**Deduction → Control prong (Exploit/Compress)**

- **When:** Phase 3 (**Map**) during a Control branch; again in **Phase 6 (Update)** for consistency checks.
- **What it does:** Rule-based closure on the current map ϕ\phi to derive consequences, eliminate inconsistent options, and tighten beliefs.
- **Controllers:** b↑b\uparrow (stability bias), TT low–mid, λ↑\lambda\uparrow (propagate validated structure).
- **Networks:** FPCN-B (+ DAN) for rule maintenance/sequence execution.
- **Operational hook:** apply a closure operator until a budget is met
 ϕ←DED(ϕ,R;budget κ, strictness σ)\phi \leftarrow \mathrm{DED}(\phi, R; \text{budget } \kappa,\ \text{strictness } \sigma)
 (You can keep κ,σ\kappa,\sigma implicit if you want to stay minimalist.)

**Induction → Control→Update (Compress/Generalise)**

- **When:** Phase 3 (**Map**) after successful deductions; **Phase 7 (Consolidate)**.
- **What it does:** Compress patterns, widen generalisation, raise competence η\eta.
- **Controllers:** b↑b\uparrow, T↓T\downarrow, then λ↑\lambda\uparrow if validation passes.
- **Networks:** Hippocampus ↔ FPCN-A (pattern abstraction) handing to FPCN-B for rule deployment.
- **Operational hook (MDL/predictive gain):**
 ΔCompression=MDLold−MDLnew⇒η+=wD1 ΔCompression.\Delta \text{Compression} =

$$\mathrm{MDL}_{\text{old}} - \mathrm{MDL}_{\text{new}} \quad\Rightarrow\quad \eta \mathrel{+}= w_{D1}\,\Delta \text{Compression}.$$

**Abduction → Creative prong (Explore/Decompress)**

- **When:** Phase 3 (**Map**) during a Creative branch; also during **Monitoring** micro-loops when ambiguity is high.
- **What it does:** Proposes new hypotheses/structures that could explain the evidence (or the mismatch).
- **Controllers:** b↔/↓$b \leftrightarrow / \downarrow$, T↑$T\uparrow$, λ↓$\lambda\downarrow$ (sandbox), optional novelty budget d↑$d\uparrow$.
- **Networks:** Hippocampus/DMN + FPCN-A (relational recombination, analogy).
- **Operational hook (posterior-sparse search):**
 H⋆∈arg max H∈H[log P(D∣H)+log P(H)]subject to a proposal budget d.$H^\star \in \arg\max_{H\in\mathcal{H}} \big[\log P(D\mid H) + \log P(H)\big] \quad \text{subject to a proposal budget } d.$

**Counterfactuals → Decide & Act (Plan/Intervene/Test)**

- **When: Phase 4 (Decide)** to choose *informative* actions; **Phase 5 (Act)** to run A/B probes; also in Phase 1 (Sense) as short rollouts.
- **What it does:** Simulates "what-if" outcomes under alternative policies or interventions to reduce ambiguity and disambiguate hypotheses.
- **Controllers:** use T(χ)$T(\chi)$ to prioritise information gain when χ↑$\chi\uparrow$; keep b$b$ centred if monitoring.
- **Networks:** Hippocampus (predictive map/SR) + FPCN-A for long-horizon simulation; FPCN-B executes the chosen test.
- **Operational hook (EFE with epistemic term):**
 π⋆∈arg minπG(π) ≈ arg maxπ EIG(π)⏟counterfactual value−risk−cost.$\pi^\star \in \arg\min_\pi G(\pi) \;\approx\; \arg\max_\pi \underbrace{\mathrm{EIG}(\pi)}_{\text{counterfactual value}} - \text{risk} - \text{cost}.$

Absolutely—**analogy** is the bridge between Creative (abductive propose) and Control (deductive/inductive verify & compress). Here's **drop-in text** that matches your scheme and notation.

---

**Analogy (Relational Mapping) → Creative→Control bridge**

- **When:** Phase 3 (**Map**) during a **Creative** branch (retrieve & align); hands off to **Control** for constraint-checking and consolidation when a mapping passes tests; also recruited in **Monitoring** micro-loops when ambiguity is high.

- **What it does:** Retrieves a *source* structure by **relational similarity**, aligns it to the *target* (structure-mapping), **projects** candidate inferences/policies, then **tests** them (deduction/counterfactuals) and **abstracts** a schema (induction) if validated.

- **Controllers:** retrieval/alignment with b≈0 or ↓b\approx 0 \text{ or } \downarrow, T↑T\uparrow, λ↓\lambda\downarrow (sandbox, optional d↑d\uparrow); after validation, **commit** with b↑b\uparrow, T↓T\downarrow, λ↑\lambda\uparrow.

- **Networks:** Hippocampus (pattern completion/separation; SR "predictive map"), **FPCN-A** + **DMN/mPFC** for relational mapping and recombination; **FPCN-B** (+ DAN) for constraint checking and execution/roll-out.

- **Operational hook (structure-mapping + projection):**
  M⋆∈arg max M[Simrel(M) + wfeatSimfeat(M) − β Viol(M)]s.t. proposal budget d,M^\star \in \arg\max_{M}\Big[\mathrm{Sim}_{\text{rel}}(M)\;+\;w_{\text{feat}}\mathrm{Sim}_{\text{feat}}(M)\;-\;\beta\,\mathrm{Viol}(M)\Big]\quad \text{s.t. proposal budget } d, IT←Project(IS,M⋆),score via EV or −G, then verify (deduction) and, if successful, abstract (induction).I_T \leftarrow \mathrm{Project}(I_S, M^\star),\qquad \text{score via EV or }-G,\ \text{then verify (deduction) and, if successful, abstract (induction).}

---

## Threading analogy through the 8 phases (add these clauses)

- **1) Sense/Move:** cue source retrieval by relational/contextual probes; run brief counterfactual roll-outs for top candidates.

- **2) Branch:** when ΔF^=b(F−F*)≥0\widehat{\Delta F}=b(F-F^*)\ge 0 **and** χ\chi high, prefer **Creative**; Monitoring may call an *analogy probe* when χmeta\chi_{\text{meta}} is high.

- **3) Map (Creative): retrieve→align→project** (S→T)(S\rightarrow T) under λ↓, T↑, b≈0\lambda\downarrow,\,T\uparrow,\,b\approx 0; **Map (Control): verify→compress** successful projections, raise η\eta.

- **4) Decide:** pick the **most informative/low-risk** analogical projection using Softmax over EV (or −G-G) with T(χ)T(\chi) and bb.
- **6–7) Update→Consolidate:** credit η\eta if the analogical policy **reduces future surprise**; **schema abstraction** widens λ↑\lambda\uparrow for transfer.

# How it threads through the 8 phases (at a glance)

- **1) Sense/Move** – quick counterfactual rollouts to score candidate policies; compute F, χF,\ \chi.
- **2) Branch** – if ΔF^=b(F−F*)≥0\widehat{\Delta F}=b(F-F^*)\ge 0 and χ\chi low→**Control** (deduce/compress); if χ\chi high→**Creative** (abduce/decompress); if χmeta\chi_{\text{meta}} high→**Monitoring** (brief orient/reset, then re-branch).
- **3) Map** – **Control: deduction** → **induction** (close rules, compress, raise η\eta); **Creative: abduction** (new H, sandboxed by λ↓\lambda\downarrow, d↑d\uparrow).
- **4) Decide** – **counterfactual selection** via Softmax over EV (or −G-G) with **T(χ)T(\chi)** and **bb**.
- **5) Act** – run the test/intervention; gather evidence.
- **6) Update** – credit **η\eta** for *either* successful compression (D1) *or* successful strategy change (D2) that reduces future surprise; adjust b,T,F*b, T, F^*.
- **7) Consolidate** – promote validated structure (λ↑\lambda\uparrow); keep speculative bits quarantined.
- **8) Reset** – restore near-critical windows; prep for next cycle.

# Controller intuition (why this makes sense)

- **bb** steers **deduction/induction** (when b↑b\uparrow) vs **abduction/counterfactual testing** (when b↓b\downarrow and T↑T\uparrow).

- **T(χ)T(\chi)** ensures that when uncertainty/conflict rises, you *sample more broadly* (counterfactuals) and consider **abductive** alternatives.

- **η\eta** grows from **both**: (i) D1 compressive wins *and* (ii) D2 successful escapes/re-commitments—your lived-experience point.

- **F*F^*** sets challenge so you don't collapse into trivial deduction *or* endless abduction; the loop hovers in the Ψ-band.

# One-sentence drop-in

> *Reasoning is phase-specific in the G-Loop: **deduction** and **induction** dominate the Control path (compress/propagate), **abduction** is the Creative path's engine (decompress/propose), and **counterfactual simulation** selects informative actions at Decide/Act; Monitoring coordinates brief orient–reset–reconfigure episodes when cross-loop inconsistency is high. Analogy operates as a **Creative→Control bridge**: retrieve and align a relationally similar source, project counterfactual inferences, then verify and compress validated mappings into the rule set.*

---

## Constraint satisfaction (feasibility & propagation)

**When:**

- **Phase 3 (Map)** on the **Control** branch (primary locus), and as a **feasibility gate** on the **Creative** branch before proposals leave the sandbox.

- **Phase 4 (Decide)** as a **feasible-set filter** on policies.

- **Monitoring** triggers constraint checks when conflict/high risk is detected.

**What it does:**

- Enforces **hard invariants** (must hold) and **soft preferences** (penalised if violated).

- **Propagates** constraints to prune the search space (like CSP/arc consistency), tightening beliefs and reducing ambiguity.

- Provides a **safety envelope** for Creative moves and counterfactual tests.

**Controllers:**

- **Control pass:** b↑b\uparrow, T↓T\downarrow, λ↑\lambda\uparrow (propagate validated constraints widely).
- **Creative pass (sandbox):** b≈0b\approx 0 or ↓\downarrow, T↑T\uparrow, λ↓\lambda\downarrow (propose freely **inside** a safety envelope).
- **Monitoring:** briefly recentres bb, adjusts T(χ)T(\chi), and tightens λ\lambda if a violation is imminent.

**Networks:**

- **FPCN-B + dACC/AI (Salience)** for rule enforcement, error-likelihood and conflict monitoring;
- **Hippocampus/FPCN-A/DMN** when constraints are **relational** (analogy/schema) and must be aligned before testing.

---

## Operational hooks (minimal maths)

Let **constraints** be functions on hypotheses HH and policies π\pi:

- **Hard:** gj(H)≤0, gj(π)≤0g_j(H)\le 0,\; g_j(\pi)\le 0 (invariants/safety).
- **Soft:** penalties cj(H)≥0, cj(π)≥0c_j(H)\ge 0,\; c_j(\pi)\ge 0 (preferences/limits).
- **Chance constraints (optional):** P[gj(π)≤0]≥1−εP[g_j(\pi)\le 0]\ge 1-\varepsilon.

**Feasible sets**

Hε={H: P[gj(H)≤0]≥1−ε ∀j},Πε={π: P[gj(π)≤0]≥1−ε ∀j}.\mathcal{H}_{\!\varepsilon}=\{H:\ P[g_j(H)\le 0]\ge 1-\varepsilon\ \forall j\},\qquad \Pi_{\!\varepsilon}=\{\pi:\ P[g_j(\pi)\le 0]\ge 1-\varepsilon\ \forall j\}.

**Map-level constraint propagation (Control)**

(ϕ, Hε) ← CONS(ϕ, Hε; budget κ, strictness σ),(\phi,\ \mathcal{H}_{\!\varepsilon})\;\leftarrow\;\mathrm{CONS}(\phi,\ \mathcal{H}_{\!\varepsilon};\ \text{budget }\kappa,\ \text{strictness }\sigma), which narrows the hypothesis space and tightens beliefs; raise η\eta when propagation **reduces future surprise/ambiguity** at the same F∗F^*.

**Creative proposals with feasibility gate**

H′∼Propose(ϕ; d, λ↓, T↑),accept only if H′∈Hε  (or minimal Viol(H′)).H' \sim \mathrm{Propose}(\phi;\ d,\ \lambda\downarrow,\ T\uparrow),\quad \text{accept only if }H'\in \mathcal{H}_{\!\varepsilon}\ \ (\text{or minimal } \mathrm{Viol}(H')).

**Decide with constrained objective** (soft constraints folded into the score)

G~(π)=G(π)+∑jμj cj(π)⇒π⋆∈argminπ∈ΠεG~(π),\tilde G(\pi)=G(\pi)+\sum_j \mu_j\,c_j(\pi) \quad\Rightarrow\quad \pi^\star \in \arg\min_{\pi\in \Pi_{\!\varepsilon}} \tilde G(\pi),

then select by Softmax with T(χ)T(\chi) and bias bb. *(Use μj\mu_j as multipliers to avoid conflict with your λ\lambda generalisation width.)*

**Safety envelope (action guard)**
Project proposed actions/policies onto the feasible set before execution:

π ← Πε[π](nearest feasible “trust region”).\pi \;\leftarrow\; \Pi_{\!\varepsilon}\big[\pi\big]\quad(\text{nearest feasible “trust region”}).

---

## How it threads through the 8 phases (add-on lines)

- **1) Sense/Move:** sample observations targeted at **active constraints** (query the variables that disambiguate feasibility).
- **2) Branch:** high χ\chi and near-zero ΔF^\widehat{\Delta F} may invoke **Monitoring** to run fast feasibility checks before allowing a Creative pass.
- **3) Map:**
  - **Control:** run **constraint propagation**; if consistent, **deduction → induction**; λ↑\lambda\uparrow once validated → η↑\eta\uparrow.
  - **Creative: propose → feasibility filter → test**; keep λ↓\lambda\downarrow until constraints pass.
- **4) Decide: constrained Softmax** over Πε\Pi_{\!\varepsilon}; soft penalties in G~\tilde G.
- **5) Act:** enforce a **safety guard**; abort or re-plan if a predicted hard constraint is at risk.

- **6) Update:** strengthen learned constraints that repeatedly validate; relax or refine soft ones that block performance without reducing future surprise.

- **7) Consolidate:** propagate **stable constraints** broadly (λ↑\lambda\uparrow); keep uncertain ones local.

- **8) Reset:** retain invariants; clear temporary/task-specific soft constraints unless promoted.

---

### One-sentence insert (for your summary line)

> *Constraint satisfaction is implemented as **feasibility gating and propagation**: hard/soft constraints define feasible hypothesis/policy sets; Creative proposals are sandboxed by a feasibility filter, and Control performs constraint propagation and consistency checks before induction and global propagation.*

This way, your everyday "does it satisfy x, y, z?" reasoning is explicitly part of the **Control prong**, **guards** the **Creative** prong, and cleanly integrates into **Decide/Act—all** with the same bb, T(χ)T(\chi), λ\lambda, F∗F^*, χ\chi, and η\eta machinery you already have.

---

# Why reasoning helps (in G-Loop terms)

- **Abduction (Creative prong) → hypothesis generation under uncertainty.**
 When χ↑\chi\uparrow and ΔF^≥0\widehat{\Delta F}\ge 0, abductive moves let the loop **escape local optima** and propose explanations that can *reduce future GG* after testing. Evolutionary payoff: rapid **strategy change** in volatile environments; the seed of innovation.

- **Counterfactual simulation (Decide→Act) → safe, information-efficient exploration.**
 Running "what-ifs" before acting gives **epistemic value** without costly errors. Payoff: higher **sample-efficiency** and **risk control** (hunt, tool use, social manoeuvring) while keeping the loop near the Ψ-band.

- **Deduction (Control prong) → reliable closure and coordination.**
 Once a candidate structure works, deductive closure **compresses and stabilises** it (D1), enabling **precise, multi-step execution** (planning, syntax, techne). Payoff: lower

on-line cost, reliable cooperation, cumulative routines.

- **Induction (Control→Consolidate) → compact, transferrable knowledge.**
 Pattern abstraction (MDL/predictive gain) **raises η\eta** and **widens λ\lambda**, turning specific wins into **generalisable skills**. Payoff: **far transfer** and cumulative culture.

## Where working memory comes in

- **"Long-term working memory" (Phase 3 Map; FPCN-A ↔ hippocampus/DMN).**
 Holds **value-landscaped relational maps** and supports analogy, abstraction, multi-step **counterfactual rollouts** (long horizon). This is the big human upgrade to Phase 3.

- **"Short-term working memory" (Phases 4–5; FPCN-B + DAN).**
 Maintains **rules/goals** and sequences actions during execution, enforcing closure and rapid rule learning. This is the executor that **locks in** compressive wins.

- **Monitoring/Salience (AI/dACC).**
 Detects cross-loop inconsistency (χmeta\chi_{\text{meta}}), **re-centres bb**, adjusts **T(χ)T(\chi)**, and **hands off** between the two WM workspaces at the branch. This keeps the whole system inside the **Ψ-band** rather than tipping into rigidity or chaos.

## The evolutionary punchline (one paragraph you can paste)

> **Evolutionary rationale.** We propose that human reasoning is an **expansion of the G-Loop's Map→Decide core** that improves survival by **minimising expected free energy** more efficiently in complex, changing niches. Abduction and counterfactual simulation (Creative path and Decide) **extend search** in hypothesis and policy space while controlling risk; deduction and induction (Control path and Consolidate) **compress and propagate** validated structure for reliable execution and cultural transmission. Two coordinated working-memory systems implement this division of labour: a **relational, value-landscaped map** (hippocampus with FPCN-A) for long-horizon reasoning, and a **rule/sequence buffer** (FPCN-B) for short-term maintenance and action. The salience system arbitrates the hand-off, keeping loops within a **near-critical Ψ-band** where information capacity and dynamic range are high. On this view, human reasoning did not replace the ancestral loop; it **deepened Phase 3 and sharpened Phases 4–5**, yielding more competent η\eta gains from both **compression** and **successful strategy change**.

## Which aspects of reasoning “fit” the evolutionary benefit?

- **Counterfactual depth:** increases informative action selection (lower GG per sample).
- **Relational compositionality:** supports **far transfer** (bigger λ↑\lambda\uparrow when validated).
- **Flexible meta-control:** fast **b/T** retuning avoids lock-in and catastrophic exploration.
- **Cumulative culture:** deduction/induction make skills **stable, shareable programmes** (high η\eta → robust routines).
- **Innovation under safety:** abduction + counterfactuals generate novelty **without** excessive real-world cost.

## Quick prediction hooks

- **Mode-specific WM coupling:** Creative episodes show **FPCN-A↔hippocampus/DMN** coupling; Control episodes show **FPCN-B↔DAN** dominance; Monitoring precedes the switch.
- **Controller signatures:** abduction/counterfactual phases exhibit **T(χ)↑T(\chi)\uparrow, b↓/≈0b\downarrow/\approx 0**; deduction/induction show **T↓T\downarrow, b↑b\uparrow**.
- **Competence growth:** η\eta increases both after **compressive wins** *and* after **successful policy changes** that reduce future error at matched F∗F^*.

---

I’m taking “near-critical band” results and specifying a **Trident**-style **branching mechanism**: a **non-bifurcated subcritical shaft** (autopilot), a **branch point** near F ≈ F∗F\!\approx\!F^*, and **two super(near)-critical prongs—Control** (compress/exploit) and **Creative** (decompress/explore)—**coordinated by Salience/attention**. Outside the Ψ-band, the dynamics collapse into **over-synchronised rigidity** or **over-chaotic fragmentation**. That is indeed beyond standard models and it’s a crisp, testable extension.

Here’s clean, drop-in language and a lightweight formal cartoon to make it concrete:

# Drop-in: Trident branching (conceptual & formal)

**Concept.**

- **Subcritical shaft (Autopilot):** for $F\!-\!F^*<0$ and low $\chi$, the loop sits in a **single, stable** regime—fast, fluent, consolidated (Gc-like) processing without a branch.
- **Branch point (near-critical):** as **precision-weighted gap** $\widehat{\Delta F}=b(F-F^*)$ rises to $\gtrsim 0$ and $\chi$ lifts, the system enters a **bifurcation zone** where two coordinated modes are available; **Salience** arbitrates and briefly executes orient–reset–reconfigure.
- **Super(near)-critical prongs:**
  - **Control (Exploit/Compress; D1-leaning):** partial synchrony, rule closure, induction, propagation ($\lambda\uparrow$).
  - **Creative (Explore/Decompress; D2-leaning):** metastable/"chimera-like" episodes, abduction/analogy, counterfactual testing ($\lambda\downarrow$, $T\uparrow$).
- **Band exit (failure modes):** outside Ψ the system **locks-in** (over-synchronised rigidity) or **fragments** (over-desynchronised chaos). We treat clinical phenotypes (e.g., compulsive rigidity; disorganised thought) as **hypothesis-level analogues**, not identities.

**Normal-form caricature (for readers who like equations).**
Use two low-dimensional coordinates:

- $R$ = **integration/synchrony** order parameter (integration vs fragmentation),
- $y$ = **mode coordinate** (Control vs Creative, $y{>}0$ vs $y{<}0$).
  Let the **bifurcation controls** be $u=F-F^*$ (challenge gap) and $v=b$ (precision-bias). A minimal potential:

$$V(R,y)=\frac{\beta}{4}R^4-\frac{\mu(u,\chi)}{2}R^2 \;+\; \frac{1}{4}y^4-\frac{u}{2}y^2 - v\,y$$

with stochastic gradient flows $\dot R=-\partial V/\partial R + \zeta_R$, $\dot y=-\partial V/\partial y + \zeta_y$.

- For **u<0u<0** (subcritical), yy has **one** minimum (no branch): **shaft**.
- Near **u≈0u\approx 0**, a **cusp/pitchfork-like** region opens in yy: two attractors (Control/Creative); **v=bv=b** tilts the branch.
- **Salience** acts as a **fast controller** that recentres v→0v\to 0, briefly lifts **temperature** T(χ)T(\chi), or nudges μ\mu to maintain **moderate RR**—keeping the trajectory **inside** the Ψ-band.
- For **large |u||u|**, either RR saturates high (rigidity) or collapses low (fragmentation): Ψ-band **exit**.

This normal form is only an **analytical cartoon**—your network-level instantiation (e.g., Kuramoto/Wilson–Cowan on a hierarchical-modular graph) provides the ground truth where:

- **Control** aligns with **higher, stable RR** and compressive updates,
- **Creative** aligns with **metastable RR** and decompressive search,
- **Monitoring** issues **short pulses** (centre bb, adjust T(χ)T(\chi), tighten λ\lambda) to re-enter Ψ after divergence.

# How this extends prior work (and stays testable)

- Prior criticality work gives you the **band** and **benefits** (dynamic range, information capacity) but not the **branching logic**. The Trident adds the **decision-level geometry** and **controllers** that explain *when* and *how* the system chooses compress vs decompress—and *how* it avoids falling out of Ψ.
- It yields **quantitative predictions**: within-band you should see **metastability peaks**, **power-law episode durations**, and **SN-led** switching into FPN-B (Control) vs FPCN-A/DMN (Creative). Outside the band you should see **over-synchrony** (rigidity, narrow λ\lambda, high hysteresis) or **over-desynchrony** (unstable RR, poor re-entry).
- Clinical mappings remain **tentative**: hypothesise **locked-in RR** and persistently high bb (plus low TT) in compulsive rigidity; hypothesise **unstable RR**, high TT, and poorly centred bb in disorganisation. These are **testable computational phenotypes**, not diagnoses.

# Two sentence insert (Discussion)

> *We formalise a **Trident** branching in which a **single, non-bifurcated subcritical regime** (autopilot) meets a **near-critical branch point** governed by the precision-weighted gap ΔF^=b(F−F∗)\widehat{\Delta F}=b(F-F^*) and uncertainty χ\chi. Salience/attention coordinators route dynamics into **Control** (compress) or **Creative** (decompress) prongs and deliver brief re-centring pulses; outside the Ψ-band the system either **over-synchronises** (rigidity) or **over-desynchronises** (fragmentation), providing clear, falsifiable failure modes.*

Yes—both fit, and they're complementary.

## When to use which

- **Kuramoto (phase-only, macroscale):** Ideal for modelling the **Ψ-band** as partial synchrony/metastability on a **connectome or hierarchical-modular network**. It reproduces **smeared/smooth crossovers** with **power-law desynchronisation durations** (exponents in the ~1–2 range) on human connectomes and HMNs—i.e., **extended dynamical criticality** consistent with a near-critical **band**, not a knife-edge point. Map your controllers as: coupling/homeostasis ↔ F∗F^* thermostat; precision-bias bb ↔ subnetwork gain tilt (FPN-B vs DMN/FPCN-A); T(χ)T(\chi) ↔ noise/frequency jitter.

- **Wilson–Cowan (E/I, mesoscale):** Use when you want **explicit excitation–inhibition** and local microcircuit realism (columns/nodes each with E/I populations). You can sit each node near a **Hopf edge** (critical slowing/oscillations), control **E:I** directly, and recover phases from E-activity to feed cross-node coupling. Then the same G-Loop knobs apply: E:I/homeostatic gain ↔ F∗F^*; bias bb skews pathway gains; T(χ)T(\chi) scales noise/input variance. This gives you a mechanistic route from E/I to band occupancy while still supporting Kuramoto-like order-parameter readouts (e.g., global RR).

## A neat hybrid (what I'd recommend)

- **Node dynamics:** Wilson–Cowan E/I per node (for local E:I and Hopf-edge control).

- **Network coupling:** Kuramoto-style phase coupling between nodes (cheap, captures **partial synchrony** and **metastability** cleanly).

- **Controllers:**

- F∗F^* thermostat → slow homeostatic adjustment of node gains/coupling toward a target R*R^* (partial synchrony);

  - bb (D1:D2 tilt) → multiplicative gain to Control vs Creative subnetworks;

  - T(χ)T(\chi) → exploration temperature (noise on inputs/frequencies) raised when cross-loop inconsistency χ\chi is high.

- **Branch logic:** Near ΔF^=b(F−F∗)≈0\widehat{\Delta F}=b(F-F^*)\approx 0, Salience delivers brief pulses (centre b→0b\!\to\!0, nudge KK, transiently raise TT) to route into **Control** (higher, stable RR, compress) or **Creative** (metastable RR, decompress) and keep the system inside Ψ.

## Why this is on-model

- **Griffiths-like bands** emerge naturally on hierarchical/modular networks → exactly your Ψ-band premise.

- **Kuramoto on connectomes** shows **power-law event statistics** and **metastability** with inhibitory links/homeostasis—precisely the telemetry you want for band detection and prong-specific states.

In short: **Kuramoto** gives you the clean band + synchrony/metastability lens across the whole network; **Wilson–Cowan** gives you the **E:I knob** and biophysical face of the same controllers. Use them together to instantiate the Trident G-Loop succinctly and testably.

Yes—with this setup you can model **from column** → **patch** → **system** using the same knobs and telemetry.

## How to scale it cleanly

**1) Column (micro/meso).**
Use **Wilson–Cowan** (E/I) per **microcircuit/column** to sit each unit near a Hopf edge; homeostatic rules keep local **E:I** in the Ψ-window. Read out a local order parameter (e.g., E-activity amplitude/phase). This is where F, F∗F^*, bb (D1:D2 tilt via pathway gains), and T(χ)T(\chi) (noise/input variance) act directly.

**2) Patch / mesoscale modules.**
Group columns into a **hierarchical-modular network** (HMN). Couple their phases (or mean fields) with **Kuramoto-style** interactions; HMNs naturally stretch a single critical point into an

**extended critical-like band (Griffiths phase)**—i.e., a Ψ-band rather than a knife-edge. Expect power-law episodes and rare-region effects across a **range** of couplings.

**3) Whole-brain networks.**
Place modules on an empirical **connectome** and evolve a **Kuramoto layer** on top: you get **partial synchrony/metastability** and **power-law desynchronisation durations** with **control-parameter–dependent exponents** below the transition—strong evidence for **extended dynamical criticality** on real graphs. Adding a bit of inhibition or homeostatic gain equalisation preserves these banded signatures.

**4) One set of controllers across scales.**

- **Thermostat** F∗F^*: slow homeostasis of local gain/coupling toward a target partial synchrony level (keeps units in-band).
- **Precision-bias** bb: multiplicatively tilts Control vs Creative subnetworks (e.g., FPN-B ↔ DMN/FPCN-A) at meso/macro; biases rule-maintenance vs hypothesis search at micro.
- **Temperature** T(χ)T(\chi): scales noise/frequency jitter when cross-loop inconsistency rises; promotes sampling and re-entry to Ψ.

**5) Telemetry invariants (all scales).**

- **Band occupancy**: interval with non-saturated synchrony RR and PL event durations (exponent τ∼1 – 2\tau\sim1\!-\!2).
- **Metastability peak** near branch point F≈F∗F\approx F^*.
- **Rare-region / spectral** markers in HMNs (Lifshitz tails/localised modes). These are exactly what HMN theory and connectome-Kuramoto simulations report.

## Why this is empirically reasonable

- HMNs and empirical connectomes **generate** extended critical-like regimes (Griffiths phases) → matches your Ψ-band premise.
- Kuramoto on human connectomes shows **smeared crossovers**, **metastability**, and **power-law duration tails** that shift with coupling and inhibition/homeostasis → gives you a robust, scalable order parameter.

## Quick recipe you can implement

- **Node:** Wilson–Cowan (E/I) ~ column.
- **Edge:** Kuramoto phase coupling between nodes/modules.
- **Controllers:** F∗F^* (slow gain homeostasis), bb (subnetwork gain tilt), T(χ)T(\chi) (noise/jitter).
- **Readouts:** local E:I, global/meso R(t)R(t), PL exponents of (de)synchronisation durations, recovery half-life to F→F∗F\to F^*.

That gives you a single, compact formalism that's faithful to **column physiology** and **whole-brain network physics**, while delivering the **Ψ-band + Trident** logic you need.