From Structural Resonance to Adaptive Validity:
A Unified View of RBD, HHA, MIT, ARH, and Dual-Horizon Resonance

A Conceptual Companion Note

Régis Rigaud
RQZ Prospective
regis.rigaud@rqz-prospective.fr

December 2025

Abstract

This note provides a unified conceptual overview of seven interconnected papers: Resonance-Based Detection (RBD), Homeostatic Hamiltonian Agent (HHA), Model Invalidity Test (MIT), Adaptive Resonance Horizon (ARH), Dual-Horizon Resonance (DH-σ and DH-Δ), and Conflict Resolution (CR). Together, they form a coherent framework for adaptive inference under non-stationarity. Rather than summarizing each paper, this companion note explains *why* each component exists, *what question* it answers, and *how* they integrate into a unified control loop. The goal is to provide readers with a conceptual map before engaging with the technical details.

1 Motivation: Why Model Validity Is Not Binary

Most adaptive systems implicitly assume that a learned model is either *valid* or *broken*. This binary view leads to brittle behavior: systems either trust their model completely or discard it entirely.

In non-stationary environments, this assumption fails. Validity is not binary—it is *gradual*, *temporal*, and *contextual*:

- **Gradual:** A model can be partially valid, accurate in some regimes but not others.
- **Temporal:** Validity changes over time as the environment evolves.
- **Contextual:** What counts as "valid enough" depends on the stakes of the current decision.

The framework presented here treats model validity as an *object of control*, not a static property. The agent continuously monitors, regulates, and adapts its epistemic state—not just its predictions.

2 RBD: Structural Coherence as a Prior

The **Resonance-Based Detection (RBD)** component addresses a foundational question:

> *"What kind of world am I allowed to believe in?"*

RBD imposes structural constraints on the learned dynamics. Rather than allowing the model to fit any pattern in the data, RBD requires that learned transitions remain compatible with physical invariants (e.g., energy conservation, symplectic structure).

This is not about accuracy—it is about *coherence*. A model that violates fundamental structure may achieve low prediction error temporarily, but it will generalize poorly and fail catastrophically when the environment shifts.

Role in the framework: RBD defines the space of *plausible* world models. It acts as a structural prior that prevents pathological learning.

3 HHA: Stress as a Continuous Regulatory Signal

The **Homeostatic Hamiltonian Agent (HHA)** addresses the question:

> *"How hard should I push my current model right now?"*

HHA introduces *stress* as a continuous signal derived from inter-temporal energy surprise. Stress is not the same as prediction error: it measures the *tension* between what the model expects and what it observes, accumulated over time.

Crucially, HHA does not make discrete decisions based on stress. Instead, it uses stress to *regulate* inference dynamics—adjusting the "friction" of belief updates. High stress increases caution; low stress permits confident extrapolation.

Role in the framework: HHA provides continuous regulation. It modulates *how aggressively* the agent should rely on its current model, without yet deciding whether the model should be abandoned.

4 MIT: When Adaptation Becomes Necessary

The **Model Invalidity Test (MIT)** addresses the discrete decision:

> *"Should I keep trusting this model, or rebuild from scratch?"*

While HHA regulates continuously, MIT provides a *circuit breaker*—a principled trigger for structural reconstruction. MIT fires only when three conditions are jointly satisfied:

1. **Persistence (Pillar 1):** Stress has accumulated beyond a threshold.
2. **Capacity (Pillar 2):** The system has sufficient resources to reconstruct.
3. **Validity (Pillar 3):** The model has lost temporal coherence (resonance collapse).

This three-pillar design prevents premature reconstruction (which wastes resources) and delayed reconstruction (which causes cascading failures).

Role in the framework: MIT arbitrates the *discrete* decision to rebuild. It converts continuous stress and resonance signals into a binary action when necessary.

5 The Missing Piece: Temporal Evidence

MIT's Pillar 3 (validity) depends on *resonance*—a measure of temporal coherence computed over a rollout horizon N. But the original MIT formulation used a fixed horizon, creating an inherent tension:

- **Short N:** Reactive but noise-sensitive. Transient perturbations trigger false alarms.
- **Long N:** Robust but slow. Genuine failures are detected too late.

The heuristic $N = \max(6, W/3)$ worked empirically but lacked theoretical justification. This gap motivated the development of ARH.

6 ARH: Validity as Temporal Stability

The **Adaptive Resonance Horizon (ARH)** reframes the validity question:

> *"Over what temporal depth should I trust my own predictions?"*

ARH introduces a key conceptual shift: *the system does not detect rupture; it measures stability.* Rupture is not a signal to be found—it is the *absence* of confirmed stability.

This creates an asymmetry:

- **Confirming stability** requires sustained evidence (slow, cumulative).
- **Losing stability** can happen quickly (fast, phasic).

ARH implements this asymmetry by modulating the rollout horizon N based on an internal confidence signal $C(t)$:

- High confidence $\rightarrow$ long horizon $\rightarrow$ tolerant, slow to change.
- Low confidence $\rightarrow$ short horizon $\rightarrow$ vigilant, fast to react.

Role in the framework: ARH upgrades MIT's resonance estimator (Pillar 3) with a principled, state-dependent mechanism. It does not replace MIT; it refines how temporal evidence is gathered.

7 The Remaining Gap: Noise vs. Drift

ARH modulates temporal depth based on confidence, but it cannot distinguish *why* confidence collapsed. Consider two scenarios that both cause resonance to drop:

- **Noise burst:** Sensor variance increases temporarily. The underlying dynamics are unchanged. *Correct response: wait.*
- **Structural drift:** The dynamics have shifted. The model is now systematically wrong. *Correct response: rebuild.*

ARH treats both identically—confidence collapses, horizon shortens, vigilance increases. But the appropriate responses differ fundamentally. This gap motivated the development of Dual-Horizon Resonance (DH).

8 DH-σ: Sign Structure Diagnostic

The **DH-σ** component addresses the diagnostic question:

> *"How is my model failing—noise or drift?"*

The key insight is that noise and drift produce fundamentally different *sign patterns* in prediction errors:

- **Noise:** Errors alternate signs randomly. The sign change rate $\rho \approx 0.5$.
- **Drift:** Errors persist with consistent sign (all positive or all negative). The sign change rate $\rho \approx 0$.

This distinction is robust: it depends only on the *sign* of errors, not their magnitude. A noise burst with large errors still produces alternating signs; a small drift still produces consistent signs.

DH-σ measures the sign change rate over a sliding window and uses it as a *gate* on confidence:

- High ρ (noise-like) $\rightarrow$ protect confidence from collapse.
- Low ρ (drift-like) $\rightarrow$ allow confidence to collapse normally.

Role in the framework: DH-σ integrates directly with ARH as a confidence gate. It provides noise tolerance without drift blindness—the system can weather transient perturbations without losing sensitivity to structural failure.

9 DH-Δ: Position-Velocity Detection

The **DH-Δ** component addresses a different question:

> *"Is something changing in my environment?"*

While DH-σ diagnoses failure *type*, DH-Δ detects regime *transitions*. The insight comes from a physical analogy: combining resonance R with its trend dR/dt captures both *position* and *velocity* in performance space:

$$P(t) = w_R \cdot (1 - R_t) + |\text{trend}_t|$$

Here R measures *where* the model is (position), and trend measures *where it's going* (velocity). These signals are orthogonal: their linear combination is the natural first-order approximation.

Position-Velocity detection achieves 81% discrimination on real-world change point benchmarks—it identifies elevated disturbance near annotated regime transitions.

Role in the framework: DH-Δ provides an input signal for MIT Pillar 1 (persistence). The detection score $P(t)$ accumulates like stress; when its integral exceeds a threshold, it contributes to the reconstruction trigger. This allows MIT to incorporate explicit change detection alongside its existing stress-based criteria.

10 CR: Conflict Resolution

The **Conflict Resolution (CR)** component addresses the meta-question:

> *"What do I do when my validators disagree?"*

Multi-component systems inevitably face situations where different validators issue conflicting recommendations. For example, DH-σ may identify noise ($\rho \approx 0.5$) while MIT triggers a rebuild due to accumulated stress. Both are "correct" given their evidence—yet their recommendations conflict.

CR introduces the insight that **conflict is informative**. Rather than forcing arbitrary hierarchies ("DH always overrides MIT") or simple voting, CR:

- Recognizes conflict as evidence of ambiguity
- Introduces an explicit INVESTIGATE state between NOISE and DRIFT
- Uses the *Temporal Center of Gravity* (CoG) of sign changes to partition decisions geometrically

The CoG mechanism naturally produces three zones:

- Recent sign changes ⇒ low CoG ⇒ NOISE (recent burst)
- Distant sign changes ⇒ high CoG ⇒ DRIFT (past event, now stable)
- Spread sign changes ⇒ intermediate CoG ⇒ INVESTIGATE (accumulate evidence)

Role in the framework: CR provides the arbitration layer when components disagree. It transforms conflict from a failure mode into an actionable signal.

11 The Unified Loop

The seven components integrate into a coherent control loop:

RBD defines what is plausible.
HHA regulates how hard to infer.
ARH defines how long to trust.
DH-σ reads the signature of failure (noise vs. drift).
DH-Δ detects regime transitions.
CR arbitrates when validators disagree.
MIT decides when to rebuild.

The signal flow:

From	$\rightarrow$	**To**
RBD (structure)	$\rightarrow$	HHA (defines what is "surprising")
HHA (stress)	$\rightarrow$	MIT Pillar 1 (persistence)
ARH (confidence)	$\rightarrow$	MIT Pillar 3 (validity depth)
DH-σ (sign rate)	$\rightarrow$	ARH (confidence gate)
DH-Δ (detection score)	$\rightarrow$	MIT Pillar 1 (change signal)
CR (conflict signal)	$\rightarrow$	DH/MIT (arbitration when disagreement)
MIT (rebuild)	$\rightarrow$	RBD (resets structure)

More precisely:

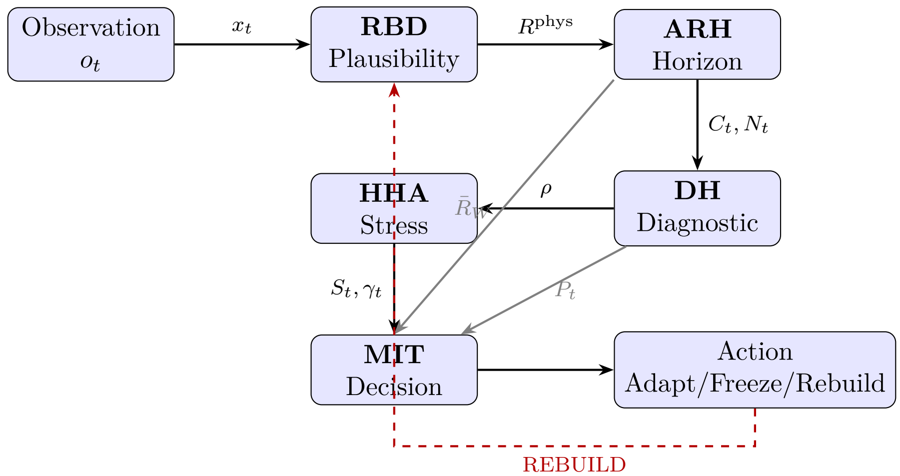
1. **RBD** constrains the hypothesis space, ensuring that learned models respect structural invariants.
2. **HHA** continuously monitors stress and adjusts inference dynamics in real time.
3. **ARH** modulates the temporal scale over which model validity is assessed, based on internal confidence.
4. **DH-σ** gates ARH confidence: high sign change rate (noise) protects confidence; low rate (drift) allows collapse.
5. **DH-Δ** provides MIT with an explicit change detection signal via position-velocity detection.
6. **CR** arbitrates when components disagree, using the Temporal Center of Gravity to partition decisions into NOISE, DRIFT, or INVESTIGATE.
7. **MIT** triggers reconstruction when sustained invalidity is confirmed across all three pillars.

This is not a pipeline—it is a *loop*. Each component feeds into the others:

- RBD's structural constraints affect what HHA considers "surprising."
- HHA's stress signal informs MIT's persistence criterion (Pillar 1).
- DH-σ's sign structure gates ARH's confidence update.
- ARH's confidence modulates the temporal depth of MIT's validity assessment (Pillar 3).
- DH-Δ's detection score provides MIT with change detection evidence (Pillar 1).
- MIT's reconstruction decision resets the loop, allowing RBD to re-impose structure.

12 System Architecture Diagram

The following diagram visualizes the complete signal flow through all components:



Legend:

- Solid arrows: Primary signal flow
- Gray arrows: Secondary inputs to MIT
- Dashed red arrow: Rebuild loop (resets RBD structure)

13 Consolidated Equations

The following table summarizes the core equations across all components:

Component	**Equation**	**Output**	**Role**
RBD	$R^{\text{phys}} = \exp\left(-\frac{\|f_\theta(x)-\Phi(x)\|^2}{\sigma^2}\right)$ $\hat{y} = R \cdot f_\theta(x) + (1-R)\cdot \Phi(x)$	$R \in [0,1]$ $\hat{y}$	Plausibility score R-weighted blend
HHA	$S_t = w_\Delta \lvert \Delta E_t \rvert + w_c[E(z_{t-1}, o_t) - E_{t-1}]^+$ $\gamma_t = \beta\gamma_{t-1} + (1-\beta)\text{clip}(\gamma_{t-1} + \alpha(S_t - S^*))$	S_t γ_t	Stress signal Friction coefficient
MIT	$\text{REBUILD} = (\Sigma_S > \Theta) \wedge (\bar{E}_W > \varepsilon) \wedge (\bar{R}_W < \delta)$	Boolean	3-pillar decision
ARH	$N_t = N_{\min} + (N_{\max} - N_{\min}) \cdot C_t$	N_t	Adaptive horizon
DH	$\rho_t = \frac{1}{W-1}\sum_{i=1}^{W-1} \mathbf{1}[\text{sign}(e_i) \neq \text{sign}(e_{i+1})]$ $P_t = w_R(1-R_t) + \lvert\text{trend}_t\rvert$	$\rho \in [0,1]$ P_t	Noise/drift diagnostic Detection score

Key thresholds:

- $\rho > 0.4 \Rightarrow$ NOISE (alternating signs, wait it out)
- $\rho < 0.15 \Rightarrow$ DRIFT (persistent sign, adapt or rebuild)
- MIT fires only when *all three pillars* are satisfied simultaneously

14 Positioning and Scope

This framework makes no claim to be a theory of truth or a path to general intelligence. Its scope is deliberately narrow:

- **Agent-centric:** All validity assessments are internal to the agent. No oracle access to ground truth.

- **Model-relative:** Validity means temporal self-consistency, not correspondence to external reality.
- **Non-stationary focus:** The framework assumes that environments change, and that adaptation is necessary.

The framework is compatible with—but does not depend on—Active Inference, Model Predictive Control, and model-based Reinforcement Learning. It can be seen as providing an *epistemic control layer* that sits above any specific learning or planning algorithm.

15 Notation Glossary

To ensure consistency across the framework, we define the following notation:

Symbol	Definition	Source
R^{phys}	Structural resonance (alignment with Φ)	RBD
R^{temp}	Temporal resonance (rollout stability)	ARH
C	Confidence (EMA of R^{temp})	ARH, DH
S	Stress (inter-temporal energy surprise)	HHA, MIT
ρ	Sign change rate of prediction errors	DH-σ
P	Detection score (position-velocity)	DH-Δ
CoG	Temporal center of gravity of sign changes	CR
σ_{noise}	Observation noise standard deviation	All
σ_R	Resonance kernel scale (exponentiation)	RBD, ARH
σ_{CoG}	Gaussian width for CoG weighting	CR
N	Rollout horizon (adaptive)	ARH, MIT
W	Sliding window length	DH, MIT

Key distinction. The framework uses two types of "resonance":

- R^{phys} (RBD): Does the model respect physical invariants?
- R^{temp} (ARH): Is the model self-consistent over time?

Both converge to 1 for valid models, but capture orthogonal aspects of validity.

16 Why This Matters

Three core insights emerge from this unified view:

1. **Validity $\neq$ performance.** A model can have low prediction error and still be invalid (e.g., overfitting to a transient regime). Validity is about temporal coherence, not instantaneous accuracy.

2. **Adaptation $\neq$ permanent reconstruction.** Rebuilding a model is costly. The framework provides mechanisms to *regulate* without *replacing*—to adjust confidence and inference dynamics before resorting to reconstruction.

3. **Confidence is an object of control.** The agent's belief in its own model is not a passive byproduct of learning. It is an internal state that can and should be actively managed.

The goal is not to build a model that never fails.
It is to build an agent that knows when to trust, when to doubt, and when to rebuild.

References

[1] R. Rigaud. *Resonance-Based Distribution Shift Detection via Model-Physics Disagreement*. 2025.

[2] R. Rigaud. *Homeostatic Hamiltonian Agent: Inter-temporal Energy Surprise as a Minimal Principle for Adaptive Inference Under Non-Stationarity*. 2025.

[3] R. Rigaud. *Structural Uncertainty as a Control Problem: When to Adapt, Freeze, or Reconstruct World Models*. 2025.

[4] R. Rigaud. *Adaptive Resonance Horizon: Confidence-Modulated Stability Measurement for World Model Validation*. 2025.

[5] R. Rigaud. *Resonance of Resonances: Multi-Scale Temporal Coherence for Decoupling Observation Noise from Model Error*. 2025.

[6] R. Rigaud. *The Conflict Signal: Resolving Component Disagreement in Multi-Validator Adaptive Inference*. 2025.