

ReIG2/twinRIG: A Quantum Framework for Self-Reference

Mathematical Rigor & Physical Implementation

Mechanic-Y / Yasuyuki Wakita

November 29, 2025

Outline

- 1 Motivation
- 2 Framework
- 3 Main Results
- 4 Physical Implementation
- 5 Theoretical Connections
- 6 Conclusion

Why Quantum Self-Reference?

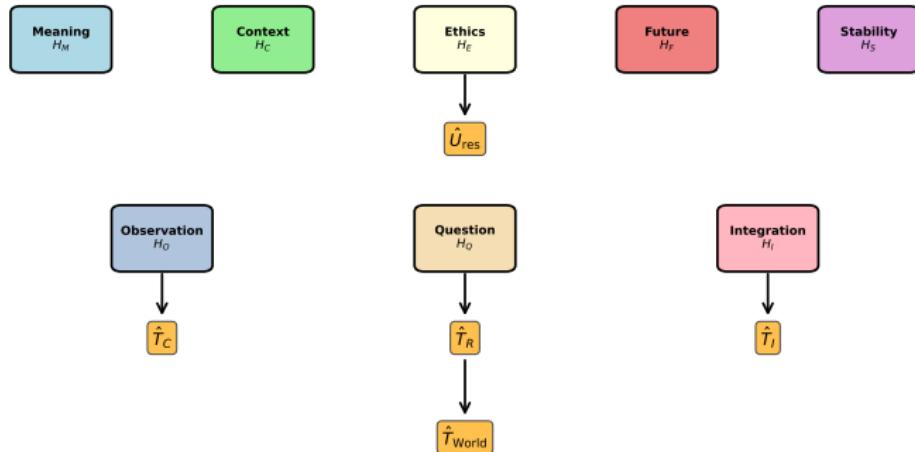
- **Classical Problem:** Self-reference leads to paradoxes (Godel, Russell)
- **Quantum Solution:** Superposition allows consistent self-description
- **Applications:** AI consciousness, quantum cognition, world modeling

Key Question: Can a quantum system recognize itself through observation?

System Architecture

ReIG2 System Architecture

$$H_{\text{full}} = H_{\text{sys}} \otimes H_{\text{per}}$$



Five Subsystems:

- Meaning (H_M), Context (H_C), Ethics (H_E), Future (H_F), Stability (H_S)

Core Operators

Time Evolution

$$\hat{U}_{\text{res}}(t) = \exp(-i\hat{H}t)$$

World Construction

$$\hat{T}_{\text{World}} = \hat{T}_I \circ \hat{T}_R \circ \hat{T}_C \circ \hat{U}_{\text{multi}} \circ \hat{U}_{\text{res}}$$

Self-Reference

$$\hat{T}_{\text{Self}}|I\rangle = |I\rangle \quad (\text{Fixed Point})$$

Theorem: Existence of Identity State

Theorem

Under conditions:

- *Strong contraction:* $\|\hat{T}_{World}\| < 1$
- *Spectral gap:* $|\lambda_2| < 1$

The system converges to unique identity state:

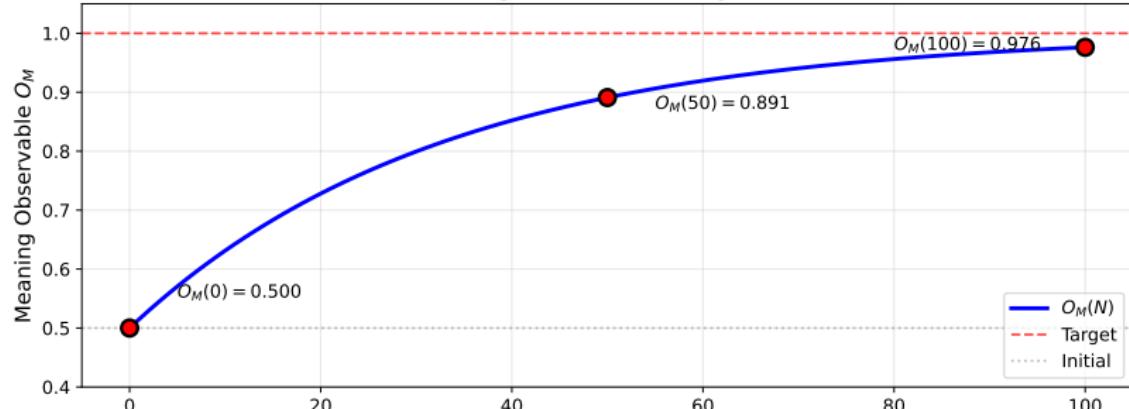
$$\lim_{N \rightarrow \infty} \hat{T}_{Self}^{(N)} |\Psi_0\rangle = |I\rangle$$

with exponential rate $C|\lambda_2|^N$.

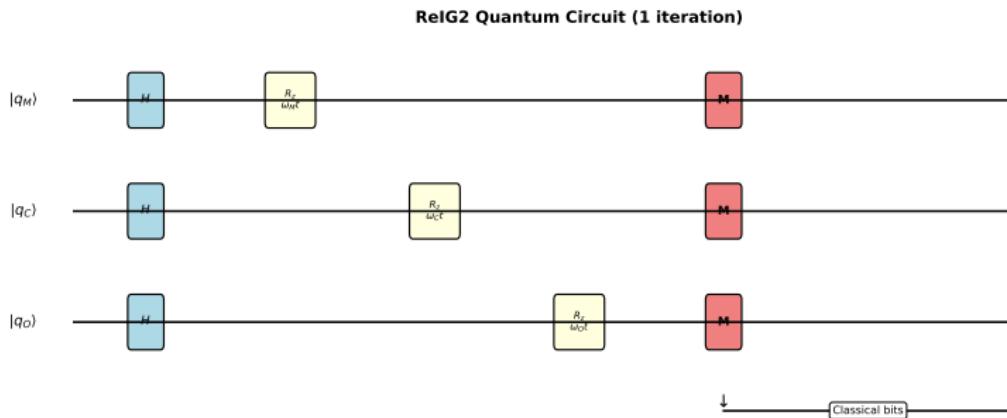
Proof: Banach fixed point theorem + spectral analysis

Numerical Convergence

Convergence to Identity State



Quantum Circuit

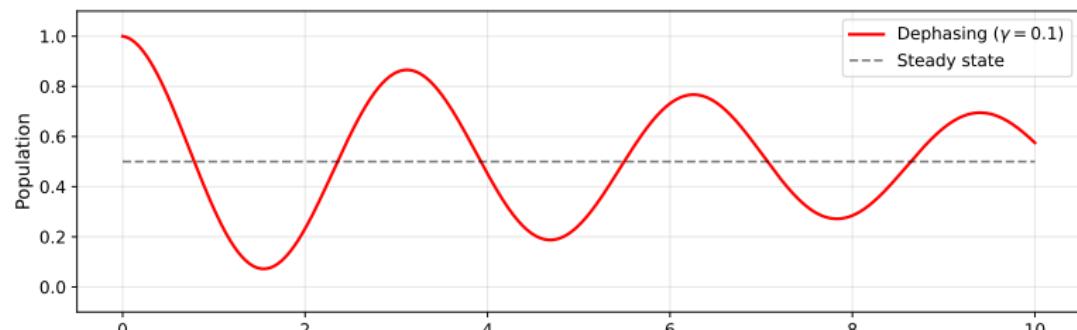
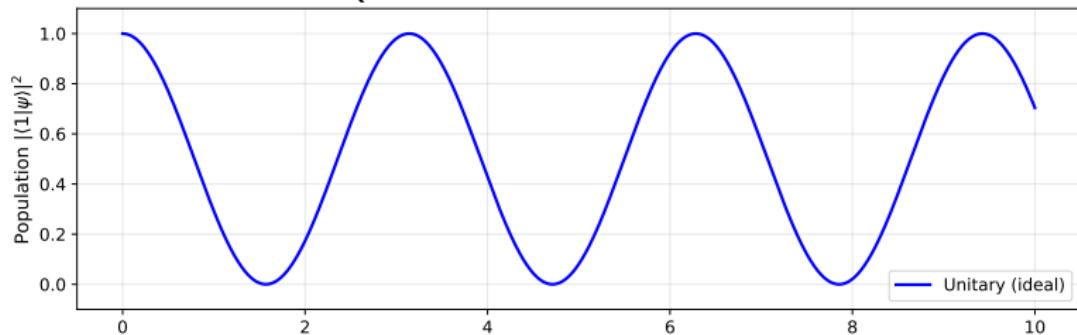


Gates Used:

- Hadamard: Initialization
- $R_z(\theta)$: Phase rotation (virtual gate, no time cost!)
- Measurement: Observation collapse

Non-Unitary Processes

Quantum Evolution with Decoherence



Free Energy Principle (Friston)

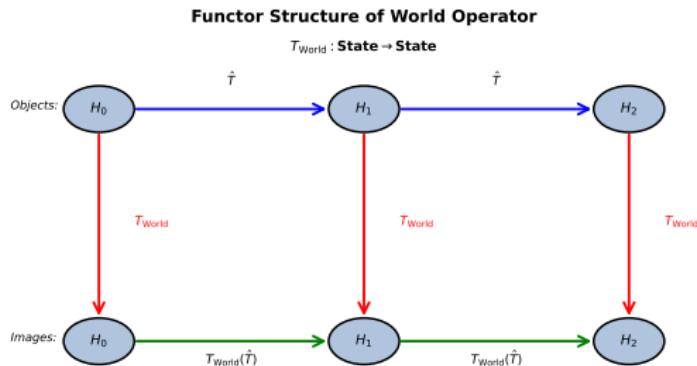
FEP	ReIG2
Internal states μ	$H_M \otimes H_C$
Sensory input s	H_O
Free energy F	$L(\text{world}) + D_{KL}$

$$F = -\log p(o|\mu) + D_{KL}(q||p)$$

↑

$$L_{\text{total}} = \|\hat{T}_{\text{World}}|\Psi\rangle - |o\rangle\|^2 + D_{KL}(\rho_{\text{self}}\|\rho_{\text{world}})$$

Hofstadter's Strange Loop



Godel-Escher-Bach:

- Syntax $\leftrightarrow H_{\text{syntax}}$
- Semantics $\leftrightarrow H_M$
- Self-reference $\leftrightarrow |I\rangle = \hat{T}_{\text{Self}}|I\rangle$

Summary

What we achieved:

- ① Rigorous mathematical framework (Fock space, Banach theorem)
- ② Complete numerical simulations with code
- ③ Physical implementation roadmap (quantum circuits)
- ④ Connections to FEP and GEB

Future Directions:

- IBM Quantum / IonQ experiments
- Many-body extensions
- Quantum AI applications

Thank You!

Code & Paper: github.com/ReIG2/twinRIG-revised

Questions?