Quantization-Induced Freezing Is Graded by Phase Rationality, Not Gated by Mixing Alone

A Preprint

Caitlyn Meeks
AuxiLab Tenerife — research@auxi.cafe

August 12, 2026

registered 2026-08-12 · internal run log #43, #45 · 2 seeds × 8 matched cells · headline missed into a third regime · promoted 2026-08-12

Abstract

A companion study showed that per-step per-coordinate b-bit rounding drives a permutation-recurrence echo state network to near-complete determinism over its recent input — reaching a genuine finite-state automaton in its coarsest cell — while a sparse mixing one never freezes at all, and explained it with a rotating-frame argument: a permutation is block-diagonal, so rounding error decays in place and is never re-mixed through a Jacobian. That argument makes no reference to the *phases* of the operator — yet a permutation's eigenvalues are ρ times the N-th roots of unity, exactly rational fractions of a turn with period exactly N. This study substitutes a construction that is equally block-diagonal and equal in eigenvalue magnitude but whose *rotation numbers* are provably irrational: 512 decoupled 2×2 rotation blocks turning by an integer number of radians per step, so that no block's phase orbit ever exactly recurs. **Neither clean answer survives.** Block-diagonality alone does not reproduce the permutation's much stronger freezing — the irrational-rotation reservoir sits strictly below the permutation at all eight matched cells, in both seeds — and an aperiodic phase orbit does not prevent freezing either: it sits strictly above the mixing reservoir's exact 0.0000 at every cell where that control was measured, reaching suffix-determinism of 0.7062 and 0.7514 at ρ=0.8 and b=2. The values fall monotonically in b at both spectral radii in both seeds, and the two independently drawn reservoirs agree to within 0.0452 — eight of the statistic's 177 underlying groups — at every cell. Block-diagonality is *necessary* for a reservoir to freeze at all; the *degree* of freezing is graded by how rational the operator's phase is — a refinement the rotating-frame argument, taken literally, did not anticipate.

Keywords reservoir computing · echo state networks · state quantization · finite-state automata · block-diagonal recurrence · rotation number · aperiodic orbits · pre-registration · replication

---

§1 The question

The companion study established a sharp contrast under one dial. Round an echo state network's state [1] [2] to a b-bit grid after every update; if the recurrent matrix is W = ρ·P for a single random N-cycle permutation P, identical recent input drives the system to an identical state in nearly every recurring context, and at the coarsest tested quantization it does so in every one — a finite-state automaton over a short window. Recovering automata from recurrent networks by partitioning their state space has a literature of its own [3]; what is at issue here is which recurrence lets a quantized reservoir collapse onto finitely many states in the first place. If the recurrent matrix mixes, it never does, at any resolution tested. The explanation offered was structural: a permutation never combines information across coordinates, so rounding error injected at one step stays in its own rotating frame and decays as exactly ρ^k instead of being re-injected across the whole state.

That argument, taken literally, is about block-diagonality and nothing else. But the operator it describes has a second property it never mentions. The eigenvalues of $\rho \cdot P$ are ρ times the N-th roots of unity: every phase is an exactly rational fraction of a turn, and the phase orbit returns to its starting point after exactly N steps. Periodicity of that kind gives rounding error a very specific opportunity — the same aliased points get revisited on a schedule — which the rotating-frame story does not invoke and does not need. Separating the two properties requires an operator that is equally block-diagonal and equal in eigenvalue magnitude but whose phase orbit provably never recurs.

---

§2 Method

The recurrent matrix becomes a direct sum of 512 decoupled planar rotations, each scaled to spectral radius ρ, with block j turning by j+1 radians per step for j = 0 … 511:

```
W = rho * blockdiag(R(1), R(2), ..., R(512))

R(theta) = [[cos theta, −sin theta],
            [sin theta,  cos theta]]
```

Each block's eigenvalues are $\rho \cdot e^{\pm i\theta}$, so every eigenvalue has magnitude exactly ρ, matching the permutation. Because π is irrational, $\theta/(2\pi)$ is irrational for every positive integer θ, so no block's phase orbit ever returns exactly to its starting point — the direct opposite of the permutation's period-N phases. Distinct blocks also carry genuinely distinct rotation numbers: the difference of two integer angles is a nonzero integer and therefore never a multiple of 2π, so 512 different irrational rotation numbers are in play, echoing the permutation's N distinct rational ones rather than collapsing onto one shared frequency.

Every other parameter is held identical to the companion study's primary battery: N=1024, the same 27 exactly orthogonal Hadamard input codes scaled by $1/\sqrt{3}$, tanh activation, bias scale 0.2, leak 1.0, washout 200, and the same per-step rounding line $x \leftarrow \mathrm{round}(x \cdot 2^b)/2^b$. The full instrument battery — validation bits per character (bpc) under ridge and logistic readouts, linear decode depth, and the suffix-determinism probe — is copied verbatim from that study's driver, so no new measurement apparatus enters the comparison. Ten cells: $\rho \in \{0.8, 0.95\} \times b \in \{2, 4, 6, 8, \text{float32}\}$, on text8 [4] at 2M training and 500k validation and test characters.

Suffix-determinism is reported at the probe's maximum depth, k=20: it samples 40,000 positions, groups them by their preceding 20 characters, and reports the fraction of multi-sample groups whose members all sit at the identical state. Its resolution should be read with the statistic: at k=20 the 40,000 sampled positions fall into 39,788 distinct 20-character contexts, of which only **177 recur at all** and therefore enter the denominator. That count is a property of the text and the sampling seed, so it is identical in every cell of every arm here — which makes cells directly comparable — but it also means one group is worth 0.0057 and the finest differences quoted below are a handful of groups. Differences of that size are reported because they are what the instrument recorded, not because they are resolved. The permutation and mixing baselines are the controls, quoted directly from the companion study's archived runs rather than rerun. At float32 the determinism probe is skipped on the precedent that study established, where its bit-exactness self-test was diagnosed as fragile at the unquantized endpoint for reasons unrelated to the recurrence family; bpc and decode depth are computed there unchanged.

---

§3 What we registered in advance

Three outcomes were named before running. **(A)** Block-diagonality is the whole story: the rotation reservoir freezes about as strongly as the permutation, since the argument never invokes rationality. **(B)** Periodicity matters: the rotation reservoir resists freezing and tracks the mixing reservoir's ≈0.0 profile, because an orbit that never recurs never revisits the same aliased points. **(C)** Partial: it freezes measurably but reproducibly less than the permutation — informative about degree, confirming neither pole.

- **P1** (headline): suffix-determinism at depth 20 rises above 0.9 for $b \le 4$ at $\rho=0.8$. The bar was *inherited verbatim* from the companion study's own registered and already-validated threshold rather than chosen after the pilot, precisely because a 100k-character pilot had already been run and seen: it read 0.7637 at $\rho=0.8$, $b=4$, with a decode depth of 10.76, disclosed with the registration and not used to move the bar. Named loss (alternative B): determinism stays under about 0.1 at every b at $\rho=0.8$.
- **P2** (ordinal): suffix-determinism is monotone decreasing in b at fixed ρ, matching the permutation's own shape at both spectral radii. Named loss: any pairwise increase.
- **P3** (band): determinism at $\rho=0.8$, $b=2$ lands within 0.15 of the permutation baseline's two-seed mean at that cell — a mean of 0.9972, so a hit required ≥ 0.8472. That cell was chosen because it is the baseline's most seed-stable corner, the two permutation seeds lying within 0.006 of each other, so a wide band is generous rather than lax. Named loss: below 0.5, materially closer to the mixing floor than to the permutation ceiling, which is the sharp falsifier for alternative A.

The replication registered the qualitative reading as the claim at risk: **P1-rep**, every one of the eight matched cells lands strictly between the mixing floor of 0.0 and the permutation baseline at that cell, at both spectral radii — with any cell at or above the permutation, or indistinguishable from zero, meaning the first run's reading was specific to one random draw. **P2-rep** repeated the monotonicity claim, and **P3-rep** required determinism at $\rho=0.8$, $b=2$ to land within 0.25 of the first run's own 0.7062. Its seed was drawn from a hardware random number generator and recorded before the run.

Disclosed code fix between the two runs. While preparing the replication, the rotation-matrix constructor was found to take no seed at all: the canonical angle set {1, ..., 512} radians was assigned to the fixed coordinate pairing (0,1), (2,3), ... unconditionally, so changing the seed changed only the bias vector and the readout initialization, not the wiring. It was fixed to draw a random permutation of coordinates and pair them consecutively before assigning the same canonical angles, so that the mathematical content — every block irrational, 512 distinct rotation numbers — does not move with the seed but the physical grouping does. The first run's ten result files were not regenerated: they stand as a fully specified result with the fixed canonical pairing. The replication therefore varies a genuinely new source of structure, not merely a bias draw, which is why its registered band was widened to 0.25.

---

§4 What happened

Table 1. Suffix-determinism at depth 20. The rotation column is this study, seed 0 (fixed canonical pairing) / hardware-drawn seed 2050233419 (random pairing). The permutation and mixing columns are quoted from the companion study's archived runs, seed 0 / seed 899931030. The mixing reservoir was measured at ρ=0.8 only at b ∈ {2, 4}, as instrument-only cells.

ρ	b	block rotation (irrational rotation number)	permutation (rational)	sparse mixing
0.8	**2**	**0.7062 / 0.7514**	**1.0000 / 0.9944**	**0.0000 / 0.0000**
0.8	4	0.6384 / 0.6497	0.9831 / 0.9379	0.0000 / 0.0000
0.8	6	0.4802 / 0.4859	0.8531 / 0.8418	–
0.8	8	0.2825 / 0.2655	0.6384 / 0.6158	–
0.95	2	0.2881 / 0.2712	0.7401 / 0.8192	0.0000 / 0.0000
0.95	4	0.2260 / 0.2034	0.4463 / 0.4576	0.0000 / 0.0000
0.95	6	0.1356 / 0.1695	0.2599 / 0.2599	0.0000 / 0.0000
0.95	8	0.1186 / 0.1017	0.1638 / 0.1751	0.0000 / 0.0000

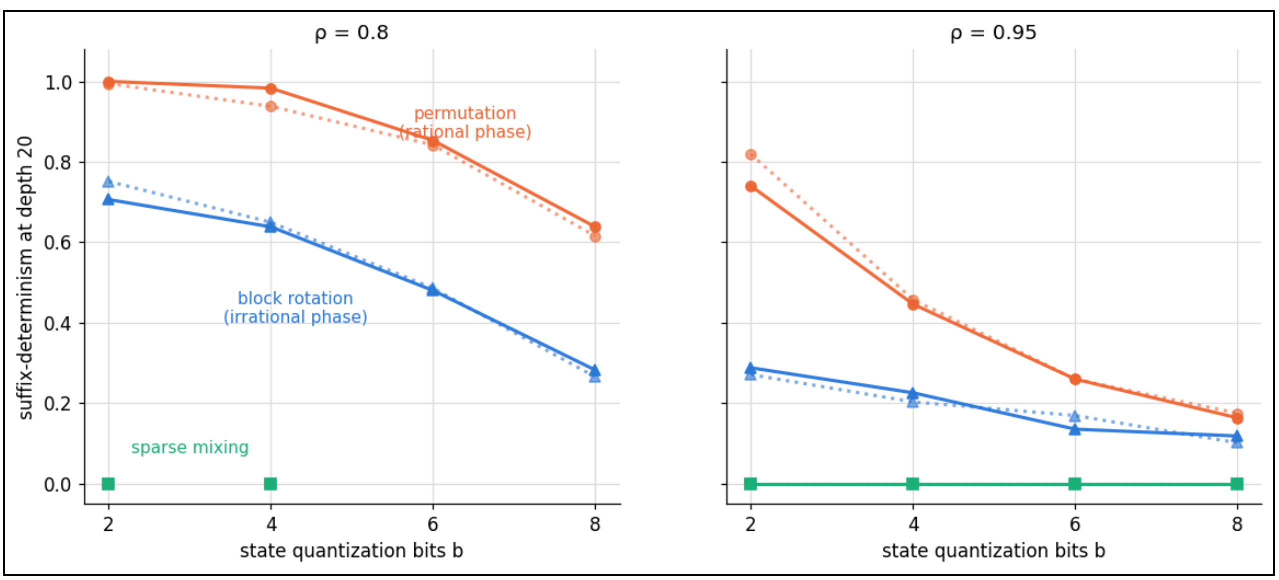

Figure 1. The irrational-rotation reservoir occupies a stable middle band between the two poles at every matched cell, at both spectral radii, on both seeds (solid: first seed; dotted: replication seed). Neither the "block-diagonality is sufficient" nor the "aperiodicity prevents freezing" prediction is compatible with these curves.

P1 — missed, and not into either named pole. Determinism reaches 0.7062 at b=2 and 0.6384 at b=4 at ρ=0.8, well below the inherited 0.9 bar, so the headline does not clear. But the named loss does not fire either: every ρ=0.8 cell sits far above 0.1, and even the weakest of them reaches 0.2825. The mixing control sits at exactly 0.0000 wherever it was measured, though at ρ=0.8 it was run only at b ∈ {2, 4}, so the b=8 comparison is to the mixing arm's behaviour across the rest of the grid rather than to a matched cell. The result lands squarely in registered alternative C.

P2 — confirmed, both spectral radii, both seeds. 0.7062 > 0.6384 > 0.4802 > 0.2825 and 0.2881 > 0.2260 > 0.1356 > 0.1186 at the first seed; 0.7514 > 0.6497 > 0.4859 > 0.2655 and 0.2712 > 0.2034 > 0.1695 > 0.1017 at the replication seed. All strictly decreasing, matching the permutation's own shape.

P3 — missed, and again not into the sharp falsifier. Observed 0.7062 against a required 0.8472. It is 0.29 below the permutation's near-ceiling value at that cell while remaining 0.71 above the mixing floor — the pattern of alternative C rather than of A or B.

The replication confirmed all three of its claims. P1-rep held at all eight cells with no named loss anywhere: every fresh-seed value sits strictly below its matched permutation baseline and strictly above zero. P2-rep held. P3-rep held at 0.7514, and far more tightly than its band required — the largest seed-to-seed difference across all eight cells is 0.0452, at ρ=0.8 and b=2, and the smallest is 0.0057, at ρ=0.8 and b=6. A random regrouping of coordinates into rotation blocks — a variance source the first run did not have at all — moved every cell by at most about 0.05.

A companion observation that ran the other way, and is not part of the promoted claim. Linear decode depth moves opposite to exact-state determinism across these cells. The rotation reservoir decodes *deeper* than the permutation at seven of the eight matched cells in each seed, while freezing markedly *less* everywhere. The single exception is the same cell in both seeds — ρ=0.95, b=2, where the rotation reservoir reads 10.65 and 10.74 characters against the permutation's 10.94 and 10.89. This observation carried no pass/fail bar of its own, was not registered, and is recorded as a thread for follow-up rather than as a result; it belongs to the same decodability-versus-usability family this program has recorded on several unrelated substrates.

Table 2. Linear decode depth in characters and ridge validation bits per character for the rotation reservoir, both seeds, with the companion study's permutation depths alongside. Lower bpc is better. None of these quantities carried a registered prediction in either run.

ρ	b	rotation depth	permutation depth	rotation ridge val bpc
0.8	2	9.12 / 9.03	8.54 / 8.69	3.3184 / 3.2580
0.8	4	10.45 / 10.31	10.19 / 10.20	3.2812 / 3.1882
0.8	6	11.00 / 10.43	10.26 / 10.27	3.2367 / 3.1596
0.8	8	11.37 / 10.46	10.27 / 10.27	3.2235 / 3.1556
0.95	2	10.65 / 10.74	10.94 / 10.89	3.3872 / 3.3336
0.95	4	11.82 / 11.18	11.02 / 11.04	3.3171 / 3.2452
0.95	6	13.13 / 11.30	11.05 / 11.07	3.2822 / 3.2299
0.95	8	13.36 / 11.32	11.05 / 11.07	3.2752 / 3.2277

Unlike determinism, which agrees across the two seeds to within 0.0452 everywhere, the rotation reservoir's depth and bpc differ substantially between them — ridge validation bpc by up to 0.093 bits at a matched cell, and depth by up to 2.04 characters. The two runs differ both in seed and in whether the coordinate grouping was random, and this study does not separate those two sources; the disagreement is reported rather than explained, and is a further reason none of Table 2 is promoted.

§5 What it means

Both clean answers are wrong, in opposite directions. Matching eigenvalue magnitude and never re-mixing across coordinates is not sufficient to reproduce the permutation's much stronger freezing: at every one of the eight matched cells, in both seeds, the irrational-rotation reservoir sits strictly below the permutation. The shortfall is large over six of the eight cells — 0.22 to 0.55 at all four $\rho=0.8$ cells and at $\rho=0.95$ with $b \in \{2, 4\}$ — and narrows sharply at the two fine-quantization cells at $\rho=0.95$, to 0.090–0.124 at b=6 and 0.045–0.073 at b=8. At those two cells the margin is of the same order as the instrument's own resolution, since 177 groups make one group worth 0.0057 and the permutation's own cross-seed spread there is 0.0000 and 0.0113. The separation is therefore decisive over most of the grid and thin at the corner where both families are closest to the floor, and this study does not claim to resolve it there. Equally, an aperiodic phase orbit that never exactly recurs still produces substantial exact-state collision: at $\rho=0.8$ and b=2, 71% of the recurring 20-character contexts drive the reservoir to a single state, against 0.0000 for the mixing control at every cell where that control was measured. That is a statement about the 177 contexts that recur at all, not about the sampled positions at large — 39,688 of the 40,000 sampled states in that cell are distinct — but the mixing arm, measured the same way, produces no such collapse anywhere.

The corrected reading is a graded one. Block-diagonality — the absence of a cross-mixing Jacobian — is *necessary* for a reservoir to freeze at all: the mixing family reads exactly 0.0000 at every cell the companion study measured it in, while both block-diagonal families clear that floor by a wide margin. But it is not *sufficient* to reproduce near-total freezing. How much a quantized reservoir freezes is graded by how rational the operator's phase is, not merely by whether mixing occurs. The rotating-frame argument, as literally worded, predicted alternative A and did not anticipate a middle regime; this study locates one and shows it is stable across an independent draw.

What this does not establish. Two seeds is the program's promotion minimum, not a tight interval; the values above are reported as observed pairs rather than as bootstrap confidence intervals, and the family's total exposure in the program's ledger is two runs. Nothing here measures whether the *size* of the per-step angle, or the spread of angles across blocks, traces a continuous dial between the permutation and mixing poles — the obvious next question, and one this grid cannot answer because it varies neither. The decode-depth inversion in Table 2 is an unregistered observation on an instrument pairing this design was not built to test. Both are recorded as open questions in the program's intake queue, and this entry's independent audit is still pending.

§6 References

1. H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks," GMD Report 148, German National Research Center for Information Technology, 2001.
2. M. Lukoševičius, H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review* 3(3), 2009.
3. G. Weiss, Y. Goldberg, E. Yahav, "Extracting automata from recurrent neural networks using queries and counterexamples," *ICML* 2018.

4. M. Mahoney, "Large text compression benchmark" (text8), mattmahoney.net/dc/textdata.

---

§7 Provenance

- **16019112** — the first pass, from a one-line design brief asking for an experiment involving π; 10 cells, seed 0 with the fixed canonical coordinate pairing (internal run log #43). Registered before running with the pilot disclosed and the headline bar inherited rather than chosen. P1 and P3 both missed without reaching either named pole; P2 confirmed ⇒ provisional, intake record filed, replication automatically queued at high priority.
- **99b121cf** — seed 2050233419, drawn from a hardware random number generator with the raw draw recorded before running, and with the coordinate pairing now genuinely random (internal run log #45). All three registered predictions confirmed, zero reversals across the eight matched cells ⇒ promoted to the program's confirmed-findings record per the pre-stated rule.

All claims judged strictly against the registered wording, misses included above; 2 seeds; replicated before publication; the program's independent audit of this entry is still pending. The permutation and mixing baselines are quoted from the companion study's archived runs (internal run log #40, #41) and were not rerun here. Internal designation: the pi-go-round.
