Measuring Orthogonality Between Paired Reservoirs: Duplicate Twins Are Reparametrizations, Independent Twins Are Nearly Free

A Preprint

Caitlyn Meeks
AuxiLab Tenerife — research@auxi.cafe

August 11, 2026 · amended August 13, 2026

registered 2026-08-11 | internal run log #18, #19, #25, #50 | 3 seeds, replicated claims 3/3
replicated before publication | promoted 2026-08-11 | amended 2026-08-13 after audit

Amendment notice, 2026 August 13. The first of the two findings below — the duplicate twin as a pure reparametrization — was independently audited on 2026-08-13 and was an outright **pass**. No number changed; the audit found the tie holds in more places than this preprint claimed, and it verified by direct measurement a premise the study had only asserted. One qualification does attach, and it is about how the evidence should be weighed rather than whether it holds: the audit showed the tie is *algebraically forced*, so the three seeds are three instances of an identity rather than three independent confirmations of an uncertain proposition. What the replication count certifies here is the measurement pipeline, not the effect. §4.1 reports this in full. The second finding — that an independent twin beats the duplicate on every instrument and nearly matches a full-width reservoir — has not yet been audited.

Abstract

An open-ended design brief called for two reservoirs, one "orthogonal" to the other, leaving open what orthogonal should mean; this study makes it a measured quantity. Two half-width reservoirs (N=512 each) run in parallel on the identical text stream, states concatenated to width 1024; the second reservoir — the twin — is either a *bit-identical duplicate* of the first or an *independent draw*; orthogonality is measured by canonical correlation between the two 512-dimensional blocks. Two findings, each obtained on all three seeds, and one registered methodological observation. (1) A duplicate twin is a pure reparametrization, **exactly**: it ties the lone half-width reservoir on ridge bpc and decode depth to the last recorded digit, three seeds in a row — a tie a subsequent independent audit showed to be algebraically forced rather than empirically lucky, so the three seeds certify the measurement pipeline rather than an uncertain effect (§4.1). (2) An independent twin beats the duplicate on **every** instrument — ridge, logistic, and decode depth — and lands within 0.0008–0.0049 bits of a full-width 1024-unit reservoir built as a single unit: near-free doubling. (3) The observation: the naive reading of the orthogonality measure is misleading — the two reservoirs' *single best-aligned* directions correlate above 0.999 whether or not the wiring was copied, because shared input dominates the top of the spectrum; only the *full-spectrum* mean canonical correlation (≈0.90 duplicate vs ≈0.45 independent at full budget, recorded on two of the three seeds) distinguishes a copy from a genuinely independent reservoir, and predicts which one pays downstream.

Keywords reservoir computing · canonical correlation analysis · redundancy · ensembles · feature width · pre-registration

§1 Introduction

When does a second reservoir add information to a first one? Folk intuition says: when it is "orthogonal" — differently wired, independently random. But two reservoirs driven by the same input are not independent systems; both spend most of their capacity linearly tracking the same highly-recoverable recent-character history. The brief left the interpretation of orthogonality open; the interpretation adopted here is an *operationalized* one — pick a measure, then check whether the measure predicts the downstream payoff of concatenating the two reservoirs' states.

§2 Method

Two half-width reservoirs built from the program's reference sparse-mixing tanh construction (N=512 each, spectral radius ρ=0.95, fanin 10, the standard input map, leak 1.0) [1] run in parallel on the identical text8 stream [3], 2M train characters; their states are concatenated to readout width 1024, matched to the standard full-width reservoir of the addressed-memory ESN study so no feature-width confound enters. Three cells per seed: **single512**, one half-width reservoir alone; **twin_redundant**, in which the twin — the paired second reservoir — is built from the first reservoir's own seed: identical wiring, bit-identical states every step, the maximally non-orthogonal case; **twin_orthogonal**, in which the twin is an independent draw. Orthogonality is computed, not assumed: canonical correlation analysis [2] between the two 512-dim state blocks on subsampled validation rows (≤100k, ridge-regularized whitening), summarized two ways — `top_canon_corr`, the single best-aligned direction, and `mean_all_canon_corr`, the average over all 512 canonical directions. The full-width comparator (ridge 3.1472, logistic 2.6727, depth 10.11) is the archived reference cell, cited from the archive and not rerun. Instruments: ridge and logistic validation bits per character (bpc), and U3 decode depth — a linear decode-depth probe, so designated in the program's internal instrument numbering.

§3 Pre-registered predictions

The pre-registration pilot run held a genuine surprise, disclosed in the registration before the full run: *independent draws do not decorrelate the top of the spectrum.* The predictions were written with that disclosed:

- **O1/O2** (manipulation checks): the full-spectrum mean separates the twins by a wide margin, while the top canonical correlation stays ≥ 0.99 in *both* twin arms — with the named alternative that more data at full budget resolves the saturation away, in which case it is an artifact of the pilot run. (Both hit; the saturation is real at 20× the data.)
- **O3** (reparametrization null): the duplicate twin ties the lone half-width reservoir on ridge bpc (|gap| ≤ 0.01) and decode depth (|gap| ≤ 0.3). (Hit — exactly: ridge gap 0.0000, depth gap 0.00.)
- **O4** (the headline): the independent twin beats the duplicate on all three of ridge, logistic, and depth. Named alternative: any single reversal = the pilot run's pattern was small-sample noise. (Hit, 3/3.)

- **O5** (near-free doubling): the independent twin lands closer to the full-width comparator than the lone half-width reservoir does, on both bpc instruments. (Hit.)

Quantitative bands Q1–Q4 were registered separately; two hit and two missed *in the beat-expectations direction* — the 100k-character pilot run undersold how far the logistic readout converges at 2M characters, so both logistic bands' actuals came in below their registered floors. Recorded as misses, consistent with this program's observed pattern of quantitative bands missing. The seed-1 rerun registered O3/O4/O5 unbundled (each promoting alone); a third seed was subsequently commissioned and run under the identical registered wording.

---

§4 Results

Table 1. All three seeds (internal run log #18 / #19 / #25; values seed 0 / 1 / 2). Full-width comparator, cited from the archive and not rerun: ridge 3.1472, logistic 2.6727, depth 10.11. Zero consistency-gate failures, and no activation of the divergence tripwire (the program's fit-invalidating divergence rule).

cell	ridge val bpc	logistic val bpc	U3 depth
single512	3.3142 / 3.3100 / 3.3013	2.8356 / 2.8318 / 2.8266	8.66 / 8.57 / 8.18
twin_redundant	3.3142 / 3.3100 / 3.3013	2.8330 / 2.8289 / 2.8229	8.66 / 8.57 / 8.18
twin_orthogonal	**3.1442 / 3.1487 / 3.1423**	**2.6713 / 2.6735 / 2.6714**	**10.08 / 10.18 / 10.03**

The duplicate is inert, to the digit. On every seed, twin_redundant's ridge bpc and decode depth equal single512's exactly as recorded — 3.3142/8.66, 3.3100/8.57, 3.3013/8.18 — an exact tie at every seed, for the reason given in §4.1. A duplicated readout column adds no information the ridge instrument or the decode-depth probe can use; this is the 2×-redundancy edge case of the program's measured result that widening a readout in this way is a reparametrization, now with an exact copy instead of a permuted view.

The independent twin wins on every instrument, three times. Ridge 3.1442/3.1487/3.1423 vs the duplicate's 3.3142/3.3100/3.3013; logistic 2.6713/2.6735/2.6714 vs 2.8330/2.8289/2.8229; depth 10.08/10.18/10.03 vs 8.66/8.57/8.18. And it sits 0.0030/0.0015/0.0049 ridge-bits and 0.0014/0.0008/0.0013 logistic-bits from the full-width comparator, where the lone half-width reservoir sits 0.15–0.17 bits away on both — at least thirty times closer, on every seed and instrument. Two independently drawn half-width reservoirs recover almost everything a single full-width reservoir obtains.

The top-direction summary saturates. top_canon_corr: 0.9999 (duplicate, every seed) vs 0.9995/0.9994/0.9995 (independent) — indistinguishable in practice. The full-spectrum mean separates them decisively: 0.9060/0.9022 vs 0.4477/0.4441 on the two seeds where the full-budget run recorded it (the third seed's verdict table recorded the top-direction summary only; its pilot run showed the same shape, 0.8967 vs 0.4562). Registered as a methodological observation and carried in the program's intake notes rather than its confirmed-findings record: it is a point about how to *measure* orthogonality between co-driven reservoirs, not an effect size.

Why the duplicate's full-spectrum mean is 0.906 and not 1.0. A reader may reasonably expect any whole-spectrum correlation summary between an exactly identical pair to read 1.0. For this regularized construction it does not, and the shortfall is closed-form rather than empirical. The estimator adds a ridge term reg·I, here 10^{-3}, to both auto-covariances before whitening, but not to the cross-covariance. For a bit-identical pair the three covariances are the same matrix C, so the whitened matrix is $(C + reg \cdot I)^{-1/2}\, C\, (C + reg \cdot I)^{-1/2}$, which shares C's eigenvectors and has eigenvalues $\lambda_i/(\lambda_i + reg)$ — strictly below 1 for every finite λ_i. A follow-up (internal run log #50) confirmed the pair *is* bit-identical by direct array comparison at seed 0 (maximum absolute difference exactly 0.0) and checked the prediction numerically: the predicted eigenvalue ratios match the estimator's actual singular values to 4.4×10^{-14}, and the predicted mean reproduces the recorded 0.9060 to all four decimals. The raw covariance spectrum runs from 0.00117 to 14.15, with 212 of 512 eigenvalues below 10^{-2}, ten times the ridge term — the regime where the suppression bites — which is consistent with a 512-unit reservoir driven by a 27-symbol alphabet carrying a long tail of low-variance, near-redundant directions. Sweeping the ridge term down confirms the limit: the predicted mean rises to 0.999885 at reg=1e-6 and to 1.000000, at displayed precision, at reg=1e-9. The shortfall from 1.0 is entirely and only a function of the regularization, not evidence of any difference between the two blocks. This paper makes no claim that the statistic *should* read 1.0; it is used throughout as a discriminating quantity (0.906 against 0.448) and that use is unaffected. The check described here was an internal validity check; the independent audit reported in §4.1 came later and covers the first of this paper's two findings, not this statistic.

---

§4.1 What the independent audit found, and what it changes

The tie re-derives, and runs deeper than this paper claimed. The audit of 2026-08-13 recomputed all six quoted values from the raw result files with freshly written code and matched them exactly. It then checked places this paper does not report, and the duplicate and the lone half-width reservoir agree in all of them: the complete eleven-lag decode-accuracy ladder (33 of 33 lag accuracies bit-equal across the three seeds), the untouched test split (3.3196/3.3131/3.3050, identical in both arms), and even the fitted calibration temperature (0.1005/0.1002/0.1004). The trained logistic readout correctly does *not* tie — the duplicate comes in 0.0026/0.0029/0.0037 bits below the lone half-width reservoir — which is the expected conditioning-rather-than-information signature and serves as a negative control: had logistic tied as well, the two arms would be under suspicion of being accidentally the same object.

An asserted premise was measured. This study's construction code *asserts* that a same-seed twin is bit-identical to its original at every step and nowhere tests it. The audit tested it. At all three seeds the two reservoirs match bit-for-bit in mixing matrix (5,120 nonzeros), input map and bias; stepped in lockstep for 400 steps using the auditor's own update loop — written from the documented rule rather than reusing this program's stepper, precisely so that a bug in the stepper could not hide itself — the maximum absolute state difference is exactly 0.000, while the independent-draw arm diverges (maximum 1.97, mean 0.65). That run also cross-validates the program's own batch stepper against the auditor's loop, again at exactly 0.000.

The qualification: the tie is a theorem, so the seed count means something narrower than it appears. Ridge regression on a duplicated design — the design matrix X placed beside an exact copy of itself — with penalty λ produces predictions identical to ridge on X alone with penalty $\lambda/2$: the duplicated weight splits evenly between the two copies; the audit verified this numerically to a maximum prediction difference of 1.0×10^{-9}. Because this program's ridge penalty is absolute rather than per-sample, $\lambda=10^{-2}$ accumulated against a two-million-row normal matrix is a relative perturbation of order 10^{-8}, so λ against λ/2 is invisible far below the fourth decimal place, and the decode probe's penalty is smaller still. The exact tie is therefore forced by algebra. This paper's wording for the finding — "a pure reparametrization", "adds no information a linear instrument can use" — is exactly that theorem, correctly stated, which is why the audit passed it without amendment. But it follows that the three seeds are three instances of an identity, not three independent confirmations of an uncertain proposition. The right reading of this finding is: *the pipeline reproduces an algebraic identity to the last digit, three times* — a real and reassuring thing to know about the instrument stack — and not: *an uncertain empirical effect replicated three times*. The second finding of this paper, the independent twin's near-free doubling, carries no such identity behind it and is a genuinely empirical result; it has not yet been audited.

---

§5 Discussion

Duplication is exactly worthless; independence is nearly free. If a second reservoir is to add value it must be drawn independently; doing so buys almost the whole doubling: two half-width reservoirs recover a full-width reservoir to within thousandths of a bit, three seeds in a row. Copying the wiring buys exactly nothing, to the recorded decimal — an algebraic identity the pipeline reproduced exactly at all three seeds (§4.1). The result composes with the partitioned sub-reservoir ensemble finding from the same day's program — there, 64 blocks of a single draw at matched width never cost and usually helped; here, two fully independent draws (own mixing matrix *and* own input map) nearly match a monolithic reservoir.

"Orthogonal" is a measurement, not a construction. Two reservoirs with independent wiring *sound* orthogonal, but their best-aligned directions correlate above 0.999, because the shared input dominates what both reservoirs most strongly represent. Independence only shows up — and only pays — across the rest of the spectrum. Any pipeline that certifies diversity by construction ("different seeds, therefore decorrelated") is measuring the wrong end of the canonical spectrum.

---

§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. H. Hotelling, "Relations between two sets of variates," *Biometrika* 28, 1936.
3. M. Mahoney, "Large text compression benchmark" (text8), mattmahoney.net/dc/textdata.

---

§7 Provenance

- **c660a423** — the initiating design brief; 3 cells, seed 0 (internal run log #18). Registered before running; the pilot run's saturation surprise disclosed in the registration. Headline hit 3/3 ⇒ provisional, replication auto-queued before any confirmed-findings entry.
- **142d0104** — seed 1, fresh independent draws of both reservoirs (internal run log #19). All three unbundled claims hit; promoted.

- **8879faae** — seed 2, commissioned as a direct third-seed confirmation; identical registered wording (internal run log #25). All three claims hit again; the both findings entries annotated 3/3 seeds.
- **48c99128** — a read-only validity check prompted by a reader objection to this page's 0.906 figure (internal run log #50). Bit-identity verified directly; the estimator's behaviour derived in closed form. No published number moved and no confirmed-findings entry changed; the methodological paragraph in §4 was added as a result.

Audit status: the duplicate-twin finding was independently audited 2026-08-13 and passed; the independent-twin finding is still audit pending. This program requires every promoted finding to be independently audited by a party that did not run the science, re-deriving its numbers from the raw result files with freshly written code. For the first finding that has now happened: an AUDIT PASS. The six quoted values were re-derived from the raw result files with freshly written code, and the study's unmeasured bit-identity premise was re-measured at bench scale and matched; no full-budget rerun was performed. No number moved, and one recorded consequence attaches to how the evidence should be weighed — reported in full in §4.1. For the second finding — that an independent twin beats the duplicate on every instrument and nearly matches a full-width reservoir — it has not. The 2026-08-12 validity check recorded above is narrower still and was run inside the program; it is not an audit of either finding and does not stand in for one. Replication and audit are different guarantees, and this page now carries the second for one of them, and the first for both — with the caveat of §4.1 that for the duplicate-twin finding the replication certifies the pipeline rather than the effect.

All claims judged strictly against the registered wording, misses included above; 3 seeds; replicated before publication, and the first finding audited afterward. Internal designation: orthogonal reservoir.
