Coherent sources and thermal loop correlations in the two-dimensional toric code

S1 source-plane study with `hcphases 0.8.0`

Homological-code phase project

8 November 2026

Abstract

We study equal stabilizer couplings on an $L = 4$ torus using 54 field/temperature groups, four contrasting-start histories per group and 344,064 retained measurement rows. Complete cold and warm source maps are supplemented by refined diagonal, weak-transverse-source and finite-bias profiles. Full spacetime cocycle averaging, shared collective updates and conditional line resampling are used throughout. The data show a weakly polarized cold regime with large connected loops, source-selected polarized regimes with small connected loops, and strong thermal suppression of weak-field correlations. The cold diagonal loop drop is now resolved across several points near the competing-source region. Longer central histories substantially improve magnetization and loop diagnostics. Nevertheless no group passes the complete frozen convergence rule; individual histories, failures and prospective tails are retained without a principal equilibrium phase map. The evidence supports the broad known physical picture but does not locate a thermodynamic critical line or the finite first-order endpoints.

Reading the evidence. This is one combined study, not two independent estimates of continued histories. Colored maps and numerical ranges describe retained trajectories; exact analytic references are labeled separately. No withdrawn scientific rows are included. The accompanying exports retain primitives, errors, masks and diagnostic views.

1 Model, units and the finite-temperature target

We study edge qubits on a periodic $L \times L$ square lattice, with

$$H = -\frac{J}{2}\sum_v (1 + A_v) - \frac{K}{2}\sum_f (1 + B_f) - g\sum_e X_e - h\sum_e Z_e, \tag{1}$$

$$A_v = \prod_{e \ni v} X_e, \qquad B_f = \prod_{e \in \partial f} Z_e. \tag{2}$$

Here J, K multiply stabilizer projectors, while g, h multiply Pauli operators. There is no Pauli-Y field. We fix $J = K = 1$; thus the present study covers representative slices in (g, h, β), not the full space $(K/J, g/J, h/J, \beta J)$. Temperature is $T = 1/\beta$.

At Trotter number M put $\delta = \beta/M$. In the Z representation, write a spacetime edge field as $\sigma_e = (-1)^{\phi_e}$ and its plaquette curvature sign as $F_f = \prod_{e \in \partial f} \sigma_e$. Up to an overall normalization independent of ϕ, the positive weight is

$$w(\phi) = \exp\left[k_s \sum_{f \| xy} F_f + k_t \sum_{f \ni \tau} F_f + H_s \sum_{e \| x,y} \sigma_e + H_t \sum_{e \| \tau} \sigma_e\right], \tag{3}$$

where

$$k_s = \frac{\delta K}{2}, \quad k_t = \frac{1}{2}\log\coth(\delta g), \quad H_s = \delta h, \quad H_t = \frac{1}{2}\log\coth(\delta J/2). \tag{4}$$

This is the binary specialization of the project's spacetime formulation [1]; its normalizations are independently checked below rather than assumed from the working manuscript. These coefficients include the projector-to-Pauli factor of two. For example, the two temporal edge weights before this exponential rewrite are $\cosh(\delta J/2)$ and $\sinh(\delta J/2)$, times the same harmless projector constant. An analogous Fourier transform gives the temporal plaquette weights from g. Such constants cancel in probabilities but *must* be included when differentiating the partition function. Our exact controls include those constants. Equations (3)– (4) describe the finite-M target; they do not remove Trotter error relative to the quantum Gibbs state.

Spatial dualization maps the X representation of the original problem to the same Z machinery on the dual lattice, with

$$(J, K, g, h) \longmapsto (K, J, h, g), \qquad W \longleftrightarrow T. \tag{5}$$

This identity has a separate derivation and matched-runtime audit in the repository[1]. At $J = K$ the reflected source plane therefore supplies $C_T(g, h) = C_W(h, g)$ and $m_X(g, h) = m_Z(h, g)$, with supports mapped as well. In this report reflected values carry the original chain's uncertainty and failure status. They are not additional chains, an independent X-basis validation, or a reason to average two different parameter points. On $g = h$ equality of a datum and its reflection is tautological.

2 Sampler, observables and the meaning of the saved rows

2.1 Exact cocycle averaging and the remaining sampling problem

The engine is the frozen `hcphases 0.8.0` artifact. It uses the shared, direction-symmetric S1 kernel: lazy probability .1, strip .2, coboundary interval .2, coboundary box .2, and local remainder .3, with two conditional cluster updates per sweep. Conditional line updates number $\lceil 9L^2M/(2L + M)\rceil$ per sweep. The allowed direction labels are sampled uniformly; anisotropic couplings and extents still produce different physical scales. This schedule is a coverage heuristic, not a mixing theorem.

For each $\lambda \in \mathbb{F}_2^3$, let c_λ be a combination of the three fixed spacetime cocycle sheets. The algorithm samples

$$\widehat{w}(\phi) = \sum_{\lambda \in \mathbb{F}_2^3} w(\phi + c_\lambda). \tag{6}$$

The representatives have disjoint edge supports. If Q_j is the signed sum on sheet j, the sheet-dependent factor is exactly $\prod_{j=x,y,\tau} 2\cosh(H_j Q_j)$; curvature and all off-sheet terms are unchanged. On the full field space $\sum_\phi \widehat{w}(\phi) = 8Z$. The constant multiplicity cancels in expectations. Fields with nonzero curvature are retained. This construction does not restrict the chain to flat configurations or to a selected quantum ground state. The eight spacetime shifts are not eight quantum ground states.

Exact summation removes the need to tunnel explicitly between these eight representatives. It does not average over every curvature pattern, sheet magnitude, or competing polarization background. Slow invariant modes can remain, even when a winding-loop mean appears stable.

2.2 Joint loop insertions

For spatial orientation $\mu = 0, 1$, choose two parallel winding cycles at transverse positions $0, L/2$, on the common time slice zero. In the Z basis they define Wilson operators W_1, W_2 (products of Z). Their dual counterparts T_1, T_2 are products of X on dual cycles. The desired signed connected quantity is

$$C_O = \langle O_1 O_2\rangle - \langle O_1\rangle\langle O_2\rangle, \qquad O = W \text{ or } T. \tag{7}$$

[1] `reference/xz-dual-spacetime-derivation-2026-11-07.tex`.

Every retained sweep produces a *measurement row*. A complete configuration snapshot is saved only every 128 measured sweeps. Given the background ϕ_i in row i, define

$$p_i(\lambda) = \frac{w(\phi_i + c_\lambda)}{\sum_\eta w(\phi_i + c_\eta)}, \tag{8}$$

$$a_i = \sum_\lambda p_i(\lambda) O_1(\phi_i + c_\lambda), \qquad b_i = \sum_\lambda p_i(\lambda) O_2(\phi_i + c_\lambda), \tag{9}$$

$$d_i = \sum_\lambda p_i(\lambda) O_1(\phi_i + c_\lambda) O_2(\phi_i + c_\lambda), \tag{10}$$

$$\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a_i, \quad \bar{b} = \frac{1}{N}\sum_{i=1}^{N} b_i, \quad \bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i, \qquad \widehat{C} = \bar{d} - \bar{a}\bar{b}. \tag{11}$$

Thus "same background" means the same spacetime field and the same conditional distribution p_i for all three insertions. The full conditional product is evaluated before taking sample means. In general $d_i \neq a_i b_i$, and neither $\overline{d - ab}$ nor an absolute value gives the requested connected observable.

For the homologous loops here, the sheet formula specializes to $a_i = W_1(\phi_i)\tanh(H_\mu Q_\mu)$, $b_i = W_2(\phi_i)\tanh(H_\mu Q_\mu)$, $d_i = W_1(\phi_i)W_2(\phi_i)$. The temporal sheet still enters the sampling weight. Both loop orientations are measured on each background and are correlated. Where a figure uses $C_{\mathrm{av}} = (C_0 + C_1)/2$, its error calculation retains that covariance. Orientations do not increase the number of independent histories.

2.3 Local magnetizations and complementary responses

The native Z-basis quantities are $m_Z = (2L^2)^{-1}\sum_e \langle Z_e\rangle$ and the mean plaquette sign $\langle B\rangle$, measured by `SpatialMean` on spatial slice zero. They are not imaginary-time averages. In contrast, the complementary responses use all temporal edges or plaquettes. Write $N_c = L^2 M$, R_τ for the orbit-conditional temporal edge-sign sum and F_τ for the temporal plaquette-sign sum. The exact finite-M derivative estimators are

$$A_{\mathrm{resp}} = \coth(\delta J) - \frac{R_\tau}{N_c \sinh(\delta J)}, \tag{12}$$

$$m_{X,\mathrm{resp}} = \coth(2\delta g) - \frac{F_\tau}{2N_c \sinh(2\delta g)}. \tag{13}$$

If S_τ is the raw temporal edge sum, then $R_\tau = S_\tau - Q_\tau[1 - \tanh(H_\tau Q_\tau)]$ for the chosen sheet. In the X representation exchange $J \leftrightarrow K$ and $g \leftrightarrow h$. The cancellation in (13) can amplify noise at small δg. Individual response rows need not lie in $[-1, 1]$ and are never clipped; physical ensemble means do obey the bound.

We retain $D = m_{X,\mathrm{resp}} - m_Z$ and $A_{\mathrm{resp}} - B$ as correlated linear-response diagnostics. A histogram of these row estimators is a histogram of the Monte Carlo representation, not the quantum joint distribution of noncommuting operators. Binder moments or phase weights cannot be read from it without another derivation. Nonzero magnetization in an explicit source is likewise not by itself spontaneous symmetry breaking.

3 Exact and analytic reference limits

3.1 Source-free finite torus

At $g = h = 0$ we use the commuting IID API, not an epsilon-field substitute. For $n = L^2$, $t = \tanh(\beta K/2)$, the plaquette signs have the independent Boltzmann law conditioned on even parity, $\prod_f B_f = 1$. For any subset of a plaquettes,

$$\left\langle \prod_{f\in S} B_f \right\rangle = \frac{t^a + t^{n-a}}{1 + t^n}. \tag{14}$$

The strip between the parallel loops has $a = n/2$. A single winding loop averages to zero over logical sectors. Consequently

$$C_W = \frac{2t^{n/2}}{1+t^n}, \qquad \langle B\rangle = \frac{t+t^{n-1}}{1+t^n}, \qquad m_X = m_Z = 0, \tag{15}$$

and $C_T, \langle A\rangle$ follow by $K \to J$. For fixed finite β, $C_W \to 0$ as $L \to \infty$; for fixed L, $C_W \to 1$ as $\beta \to \infty$. The crossover has βK of order $\log L^2$, not a nonzero critical temperature. A cold $L = 4$ loop can therefore be almost one while thermodynamic finite-temperature topological order is absent.

Mazac and Hamma relate source-free thermal topological entropy, confinement and memory across generalized toric codes[3]. The two-dimensional model loses thermodynamic topological order at every positive temperature. Equation (15) is a compatible loop signature, not a computation of their entropy and not a theorem ruling out every finite-temperature transition in the coherently perturbed model.

3.2 Sources and the ground-state comparison

As a post hoc weak-field bulk reference, a single X_e flips two plaquettes with energy cost $2K$. Its perturbative amplitude is $g/(2K)$. To leading order in the ground manifold this gives

$$m_X \simeq g/K, \quad m_Z \simeq h/J, \quad \langle B\rangle \simeq 1 - 2g^2/K^2, \quad \langle A\rangle \simeq 1 - 2h^2/J^2. \tag{16}$$

The two-loop product has $2L$ boundary edges, giving $\langle W_1 W_2\rangle \simeq 1 - Lg^2/K^2$ (and the exchanged formula for T). At $L = 4$, $g = h = .04$, these references are .9968 for stabilizers and .9936 for the loop products. They are neither exact finite-M predictions nor automatically connected correlators: the disconnected contribution must be evaluated separately. Finite temperature, higher orders, finite size and Trotter error remain.

When sources dominate, the independent-spin reference $J = K = 0$ has $r = \sqrt{g^2+h^2}$, $m_X = (g/r)\tanh(\beta r)$ and $m_Z = (h/r)\tanh(\beta r)$, with zero connected correlations between disjoint loops. This explains a limiting trend, not the exact values at the simulated $J = K = 1$.

Wu, Deng and Prokof'ev used unit *Pauli* stabilizer couplings[2]. Their fields and inverse temperature are $h_x = 2g$, $h_z = 2h$ and $\beta_{\text{paper}} = \beta/2$. Their ground-state diagram has a topological region and charge- and vortex-condensed regimes, with a finite self-dual first-order segment approximately $g = h \in [.170, .209]$ in our units. Their finite-size order-parameter analysis is much stronger than a small loop map. The polarized regimes can be connected around the finite endpoint; one should not draw an infinite separating line or infer three distinct adiabatic phases everywhere. These literature locations guide coverage, not our numerical acceptance thresholds or fitted results.

Our $\beta = 8 = 2L$ follows their bulk aspect-ratio choice at $L = 4$. It does not demonstrate temperature saturation. Weak sources also induce very small logical-level splittings on a finite torus; cooling bulk excitations need not select the unique finite-volume logical ground state. Winding correlators depend on this order of limits. Bare-loop suppression with circumference alone does not rule out a dressed topological observable.

4 Campaign design, gates and uncertainty

4.1 Parameter slices and contrasting starts

The paired source maps contain all sixteen combinations of $F = \{.04, .14, .19, .30\}$ at each of $(\beta, M) = (8, 32)$ and $(2, 8)$. Both have $\delta = .25$. Eight additional thermal groups measure the four .04/.30 corners at $(.5, 4)$ and the four diagonal sources in F at $(4, 16)$. Fourteen additional cold groups refine the source plane: eight points on $g = s+\epsilon, h = s-\epsilon$ with $s = .16, .19, .23$ and $\epsilon = 0, \pm.01$ (the ninth, .19 center, is already in the grid), the diagonal points $.175, .21$, the smaller biases $(.185, .195), (.195, .185)$, and the weak-transverse-source points $(.165, .04), (.04, .165)$. Thus there are 54 distinct physical groups. The

fine points appear both in profiles and in the irregular source maps at their actual coordinates. The maps use nearest-point shading and a central magnifier, not a fitted surface or fictitious additional observations.

Every group has four independent update streams: ordered0, ordered1, hot0 and hot1. Ordered histories share their starting configuration, not their random stream; the hot configurations are independently drawn. The combined study has 216 independent histories. Checkpoint continuation, tail views, the two loop orientations and duality reflection do not increase that count. No history is selected because it agrees with theory.

All retained windows were specified before their new draws. The final windows are summarized below; indices are half-open absolute sweep clocks. Measurements are retained every sweep, complete snapshots every 128 measured sweeps, and restart checkpoints every 256 measured sweeps.

Groups	Relaxation	Retained/history
34 unchanged original groups	512	1,024
Five continued cold groups: $(.14,.19),(.19,.14),(.18,.20),(.20,.18),(.23,.23)$	512	3,072
Cold center $(.19,.19)$	512	10,240
Twelve new missing/refinement groups	1,024	1,536
New bias pair $(.185,.195),(.195,.185)$	1,024	3,584

This totals 344,064 scientific rows. They are correlated Monte Carlo measurements, not independent samples. Preserved pilot, control, relaxation and interrupted replay rows are excluded from this total. Original pilot prefixes were followed by additional relaxation to the same fixed window as the other original histories; they were never silently pooled in.

The allocation prioritized the missing map cells, anticipated boundary scales and convergence at the central competing-source point. The final continuation added depth at the center and the smaller bias pair under the same seven-hour deadline. It introduced no new seeds or target law. Prospective suffix screens accompany the complete cumulative windows: the last 4,096 rows at the center and the last 2,048 at the smaller bias pair. Other augmented groups retain their registered suffix lengths. These are sensitivity checks, not independent replications or permission to replace the main window with whichever suffix looks more favorable.

4.2 Frozen convergence rules

A qualified all-history result requires split/rank diagnostics $\widehat{R} \le 1.01$ and bulk ESS ≥ 400 for primary and informative invariant channels, adequate precision, stable block errors, and no physical-bound, within-history drift or between-start incompatibility flags. Connected-loop diagnostics use the influence sequence $d_i - \bar{b}a_i - \bar{a}b_i$. Invariants include action components, curvature-defect fractions and absolute cocycle-sheet sums.

Uncertainty uses joint delete-block jackknife calculations. The original unextended histories retain the block ladder $16, 32, 64, 128$. Augmented windows use $16, 32, 64, 128, 256, 512, 1024$, admitting only lengths that divide the window and yield at least eight blocks per history. All insertions in a block move together; the two-orientation average retains their covariance. The raw reported connected value is Equation (11), with bias corrections retained separately. Precision targets are .01 for magnetizations, .02 for stabilizers, and .03 for each connected loop. The error ladder is unstable if its two largest admissible block errors differ by more than 25%, or its specified increasing trend has not flattened.

When both errors are stable, a contrast is flagged if its size exceeds four times the quadrature MCSE plus a tolerance (.05 for connected loops, .02 for other tracked means). An unstable error makes the comparison unresolved instead; the historical flag `missing_mcse` also covers this case. It does not mean that no numerical error exists. A conditional subset requires a unique compatible set of at least three histories, including both starts, passing the same relevant checks. A favorable majority is insufficient.

Unqualified groups have no principal equilibrium estimate. Their histories and diagnostic all-history summaries remain available separately. Map colors, history ranges and diagnostic error bars describe retained trajectories, not certified equilibrium confidence intervals. A joint gate failure does not prove every observable is biased or the dynamics is nonergodic: short error ladders and rare constant channels can fail conservatively. We distinguish these from actual incompatible contrasts and large rank/ESS failures. The exact self-dual condition $\langle D \rangle = 0$ provides an additional diagnostic, evaluated without forcing symmetry or retroactively changing the acceptance rule.

4.3 Controls, limitations and resource safeguards

Independent tiny-system quantum transfer traces and classical spacetime sums were checked at $L = 2$, $\beta = 1$, $M = 4$, $J = .8$, $K = 1.2$, $g = .3$, $h = .17$. Recorded discrepancies are at most 5.9×10^{-16} in relative partition normalization, 5.6×10^{-16} in joint insertions and 8.5×10^{-15} in responses. Both-basis Monte Carlo controls, source-free commuting IID controls and whole-versus-split checkpoint replay passed their prescribed gates. The continuation also passed both-basis exact replay of payloads, snapshots and persisted random-generator states. These establish tested target, estimator and execution identities, not mixing at the cold points.

The fixed installed `hcphases 0.8.0` wheel and its public APIs were used throughout. Stream first exposures, lineage, absolute clocks, descriptors, receipts and output digests were recorded. Native Windows Jobs contain descendants, enforce six physical cores/twelve logical threads and a 16,724,690,944-byte aggregate memory ceiling, and kill overdue operations. BLAS uses one thread. Five-second native guards support approximately fifteen-minute lightweight watchdog reviews. The extension ended after 6 hours 8 minutes of its first exposure, including controls, pauses and repairs, below the seven-hour hard limit. Read-only sealing, scientific analysis and rendering follow without draws.

No doubled-M, doubled-β, larger-L field check or new independent native-X field history was admitted: the allotted continuation prioritized convergence of the baseline windows. All field evidence remains at $L = 4$. It therefore lacks campaign-specific continuum, temperature-saturation and finite-size extrapolations. Tiny exact controls and tested duality do not substitute for those missing physical checks. A finite-torus crossover or bias response is not by itself a thermodynamic phase boundary.

5 Numerical findings and their evidential status

5.1 Coverage and qualification

The final study contains 54 field/temperature groups, 216 independent histories and 344,064 retained scientific rows. Both sixteen-point core grids are complete; the cold map also includes all fourteen off-grid points. The additional cold diagonal, weak-transverse-source and finite-bias profiles use their precise registered coordinates. The data contain no previous withdrawn scientific campaign.

All 54 groups remain **unresolved under the full frozen rule**: none has a principal equilibrium estimate or a qualified conditional subset. This is not a claim that all plotted means are wrong. Forty-five groups have a numerical rank/ESS failure and three additional groups fail that diagnostic screen through degeneracy or an unexplained constant channel. Six pass that part of the screen but fail other uncertainty or comparison checks. The numerical rank/ESS components pass for m_Z in 38 groups, response m_X in 30, and the two connected-loop influence sequences in 39 and 42. No group has a precise physical-bound violation. Four have actual incompatible contrasts; unstable MCSE flags elsewhere are kept distinct from those contrasts.

All histories consequently remain visible, with failure masks and no equilibrium pooling. In the tables, a range is the *minimum and maximum of four retained-history estimates*, not a confidence interval or a range of instantaneous measurements. Descriptive all-history functionals are labeled explicitly and are never substituted for a principal estimate.

5.2 The completed source maps

At $(g, h) = (.04, .04)$ and $\beta = 8$, the histories give $m_Z = .035$–$.051$, $m_X = .040$–$.044$, and $C_{W,\mathrm{av}} = .9931$–$.9941$. Native $B = .99561$–$.99780$ and the loop products are consistent in scale with weak-field perturbation theory. Numerical rank/ESS checks pass at this point; uncertainty-stability checks remain the limitation. This is useful descriptive evidence for a weakly polarized, strongly correlated cold finite torus, not a thermodynamic topological-order measurement.

The completed outer part of the cold map consistently shows strong source-directed polarization and small connected loops:

g	h	m_Z	$m_{X,\mathrm{resp}}$	$C_{W,\mathrm{av}}$
.30	.04	.026–.043	.820–.829	−.0010–.0264
.30	.14	.120–.131	.812–.824	−.0259–.0339
.30	.19	.174–.190	.796–.805	−.0025–.0280
.04	.30	.800–.820	.032–.104	−0.00002–.0286
.14	.30	.822–.844	.120–.124	−.0007–.0039
.19	.30	.781–.822	.181–.189	−.0018–.0031

The .14/.30 and .19/.30 pairs use 1,536 retained sweeps per history; the other rows use 1,024. Small negative estimates are retained as signed Monte Carlo estimates. They do not establish a negative thermodynamic correlation. The reflected T map transports the corresponding W measurements and their masks without adding independent evidence.

At $\beta = 2$, the same additional points are less polarized. For example, $(.14, .30)$ has $m_Z = .490$–$.530$ and $C_W = .215$–$.240$ (orientation average), while $(.30, .14)$ has $m_X = .505$–$.543$ and $C_W = .048$–$.072$. The completed maps reinforce the broad thermal and polarization trends; they do not supply a qualified thermal transition contour.

5.3 Refined cold profiles: a localized crossover, not a critical line

The diagonal now resolves the loss of large connected correlations across the central source region:

$g = h$	Retained/history	m_Z	$m_{X,\mathrm{resp}}$	$C_{W,\mathrm{av}}$
.14	1,024	.160–.181	.160–.203	.855–.920
.16	1,024	.186–.236	.182–.257	.746–.841
.175	1,536	.219–.257	.224–.276	.714–.781
.19	10,240	.324–.349	.299–.319	.519–.592
.21	1,536	.365–.456	.324–.391	.274–.407
.23	3,072	.497–.527	.393–.432	.184–.212
.30	1,024	.512–.667	.469–.605	.057–.083

The successive four-history loop ranges at $.175, .19, .21, .23$ do not overlap. This makes the broad finite-torus drop much clearer: a large part of the variation lies in roughly $.175 \lesssim g = h \lesssim .23$. That interval describes where the sampled observable changes; it is *not* an uncertainty interval for a thermodynamic critical coupling. Bare winding loops, source selection, thermal rounding, finite size and unqualified mixing all enter the shape.

The weak-transverse-source cuts give a complementary picture. At $(.165, .04)$, $m_X = .222$–$.305$ and $C_W = .732$–$.811$; at $(.19, .04)$ they are .328–.472 and .447–.653. On the reflected cut, $(.04, .165)$ has $m_Z = .216$–$.269$ and $C_W = .937$–$.971$, whereas $(.04, .19)$ has $m_Z = .339$–$.450$ and $C_W = .409$–$.830$. The new point resolves the approach to the rapidly changing region, but the neighboring start dependence prevents a precise boundary estimate. These cuts have a nonzero transverse source .04; the known zero-transverse axis transition cannot be identified with them without an additional limit.

5.4 Finite source bias and the first-order question

For $g = s + \epsilon, h = s - \epsilon$, retain the correlated native difference $D = m_{X,\mathrm{resp}} - m_Z$. At $s = .19$ the final data give

ϵ	Retained/history	Four-history range of D
$-.010$	3,072	$-.226$ to $-.135$
$-.005$	3,584	$-.102$ to $-.040$
0	10,240	$-.048$ to $-.006$
$+.005$	3,584	$-.020$ to .143
$+.010$	3,072	.083 to .144

The larger bias selects the corresponding polarization in all four histories. At half that bias the response is smaller and one positive-bias history remains on the other side of zero. The descriptive all-history secants $[D(\epsilon) - D(-\epsilon)]/(2\epsilon)$ are about 16.0 at $\epsilon = .01$ and 14.4 at .005. These are unqualified finite-volume response estimates, not susceptibility measurements with certified errors. The roughly reduced contrast at smaller bias is compatible with a steep rounded response. It does not demonstrate a nonzero zero-bias jump.

The bias response is weaker at $s = .16$: the $\epsilon = \mp.01$ history ranges are $[-.087, -.023]$ and $[.043, .078]$. At $s = .23$ the old biased points remain much less reliable, with ranges $[-.436, .068]$ and $[-.182, .348]$. These latter data do not show that the literature first-order segment extends beyond its finite endpoint. Disagreeing histories alone cannot distinguish physical metastability from slow algorithmic relaxation. The report therefore retains the literature segment as an external reference; neither endpoint is measured by this campaign.

5.5 Why small connected loops need magnetizations and joint primitives

Small C_W has two different mechanisms in these data. At $(.30, .04)$, individual loop means remain below .015 in magnitude and the two-loop product is small. At $(.04, .30)$, $\bar{a}, \bar{b}$ instead lie between .957 and essentially one, with $\bar{d} = .982$–$.994$: the disconnected term cancels the large product. For orientation zero of hot0, the first point has $(\bar{a}, \bar{b}, \bar{d}) = (.00359, .00551, .06055)$ and $C_W = .06053$; the reflected point has $(.99414, .99804, .99219)$ and $C_W \simeq -8 \times 10^{-6}$. These are illustrative individual histories, not selected equilibrium references. The full table retains every primitive. Magnetization distinguishes source polarization from a small product; a connected-loop map by itself cannot label all phase regimes.

5.6 Temperature dependence and analytic limits

At the weak point $(.04, .04)$ the orientation-averaged loop ranges are

β	Observed four-history range	Exact source-free $L = 4$ value
.5	$-.025$ to $-.011$	.000026
2	.174–.270	.223507
4	.911–.964	.958536
8	.993–.994	.999986

The high-temperature negative estimates are on the scale of their block errors (single-orientation descriptive MCSE about .017 at the weak point), not resolved negative correlations. The exact column is not a fit to the field data. It explains the strong finite-torus thermal suppression, while coherent sources reduce the cold value from its source-free limit.

At $\beta = .5$, corner magnetizations are much smaller: m_Z is about .015–.021 for $h = .04$ and .145–.151 for $h = .30$, close in scale to the leading high-temperature response βh. Corner loop correlations are near zero on their observed noise scale. No high-temperature boundary is resolved. Rare constant temporal-defect channels and a degenerate folded-rank statistic in one binary loop channel explain some hot-point flags; these flags alone are not evidence of trapping.

Cooling need not increase a *connected* loop. At $(.04, .30)$, C_W is near zero at $\beta = .5$, .250–.290 at $\beta = 2$, and near zero to .029 at $\beta = 8$. The cold source almost fixes each loop, so the disconnected subtraction removes its large product. This mechanism is not evidence of a reentrant phase boundary. Large cold loops on this small torus also do not contradict the known absence of thermodynamic source-free topological order at positive temperature.

6 Convergence: what improved and what remains unresolved

6.1 The central point now has stronger observable-specific evidence

At $(.19, .19)$ the final 10,240-row histories give

Channel	$\widehat{R}$	Bulk ESS	Diagnostic all-history MCSE
m_Z	1.00248	1,276.5	.00804
$m_{X,\mathrm{resp}}$	1.00190	481.8	.00808
D	1.00188	578.3	.01391
C_0 influence	1.00158	3,292.1	.01167
C_1 influence	1.00176	3,054.0	.01288
Collapsed log-target	1.00240	383.2	.000834

The primary physical channels pass their numerical rank/ESS components; their descriptive all-history block errors are stable, and the prescribed magnetization and loop precision targets are met. This is a substantive improvement over the sealed 2,048-row central view, whose response m_X had $\widehat{R} = 1.0219$ and ESS about 107.

The complete gate remains closed: the collapsed-action ESS is below 400, and individual/half-window error ladders leave compatibility comparisons unresolved. The prospective last-4,096-row screen has response m_X $\widehat{R} = 1.0242$, ESS 229, and an unstable error ladder; its collapsed-action ESS is 167. A shorter suffix naturally has less information, so these values do not prove permanent bias. They do prevent using a favorable cumulative screen as a blanket equilibrium certificate. Neither this suffix nor a selected start replaces the primary histories.

6.2 Symmetry and actual incompatibilities

At the self-dual center the descriptive native difference is $D = -.03117$ with stable diagnostic MCSE .01391. It does not exceed the frozen 4 MCSE+.02 zero-contrast screen. At .175 the corresponding difference is $-.00194$ with MCSE .01206. At .23, all four primary differences are negative, $[-.134, -.070]$, but the pooled error ladder is unstable; the prospective tail of hot0 changes to $D = .057$. This is a warning about slow common backgrounds, not evidence that exact duality fails. The independent reflected-target magnetization comparisons produce no stable-error incompatibility under the same screen; many remain unresolved. Derived reflected T values are not an independent symmetry test.

Four groups have actual incompatible retained-window contrasts: ordered0 drift in B and C_0 at $(.14, .04)$; between-history polarization, defect and loop differences at $(.22, .24)$; hot0 versus hot1 differences in m_Z, D, a check difference and a sheet magnitude at $(.14, .19)$; and ordered0 magnetization drift at $(.21, .21)$. These histories remain separate. No unique qualified majority exists. Unstable uncertainty in other groups must not be mistaken for demonstrated agreement, nor must a gate failure automatically be called a bad mean.

6.3 Likely causes and the next numerical priority

The data support both finite-run uncertainty and slow orbit-invariant relaxation. The eight cocycle-sheet signs are already summed exactly; remaining problems cannot be attributed solely to failure to flip those signs explicitly. Curvature patterns, sheet magnitudes and collective polarization backgrounds can remain correlated. Temporal derivative estimators also amplify rare-defect noise at small δg. The conditional line kernel substantially improves some modes, but is not a mixing theorem for the entire cold source plane.

First-order metastability is plausible near competing phases, but cannot explain every warning: failures occur beyond the literature segment and at weak-source points too. No new target-law error is established by this campaign. Exact small-system, replay and independent reporting checks passed, and the package runtime was unchanged. The finite histories do not prove nonergodicity or uniquely identify the remaining barriers.

The next useful work is convergence at a few contrasting anchors and separately equilibrated M, temperature and size comparisons. Another large fine grid would not by itself supply those controls. Any future observable-specific acceptance rule should be specified prospectively; this report does not relax its simultaneous gate after seeing the data.

7 Most plausible phase picture

The augmented data give a more detailed, internally checked picture of three useful low-temperature regimes: weak polarization with large finite-torus connected loops, predominantly X polarization, and predominantly Z polarization. The diagonal correlation drop is localized more clearly to the central source range, and small field biases there produce a steep polarization response. Heating suppresses weak-field loop correlations and rounds the source response. The product/disconnected decomposition explains why strongly polarized cold states can also have small connected loops.

These observations are compatible with the known ground-state structure and source-free finite-temperature limit. They neither contradict a known boundary nor independently confirm its precise location. The finite self-dual first-order segment and its endpoints remain literature-based; this $L = 4$, finite-M study does not determine a jump, critical exponents, thermal critical surface or global equilibrium phase diagram. Polarized regimes may connect around the finite endpoint, so the figures do not imply an indefinitely extended separating line. Physical mixing, thermodynamic boundary qualification and full G7/M3 remain open.

8 Graphical evidence atlas

All field figures retain unresolved histories. The source-free curves are exact equilibrium references.

The source maps use physical $g/J, h/J$ coordinates. The cold plane contains the sixteen core points plus fourteen refinements: the diagonal values $.16, .175, .21, .23$; the pairs $(.15, .17)$, $(.18, .20)$, $(.185, .195)$, $(.22, .24)$ and their reflections; and $(.165, .04), (.04, .165)$. All these extra points have $\beta J = 8$. The warm plane at $\beta J = 2$ still has sixteen measured locations. For each point, its Voronoi tile is the part of the drawing closer to it than to another measured point. Its color repeats that one estimate: the polygon is neither a measured parameter bin nor evidence for a boundary. The dashed rectangle and middle row magnify $g/J, h/J \in [.135, .245]$, containing sixteen actual cold sample locations. The same full-plane polygons are clipped for the magnifier, so colors do not change between views. Numerical values and statistical gates are unchanged.

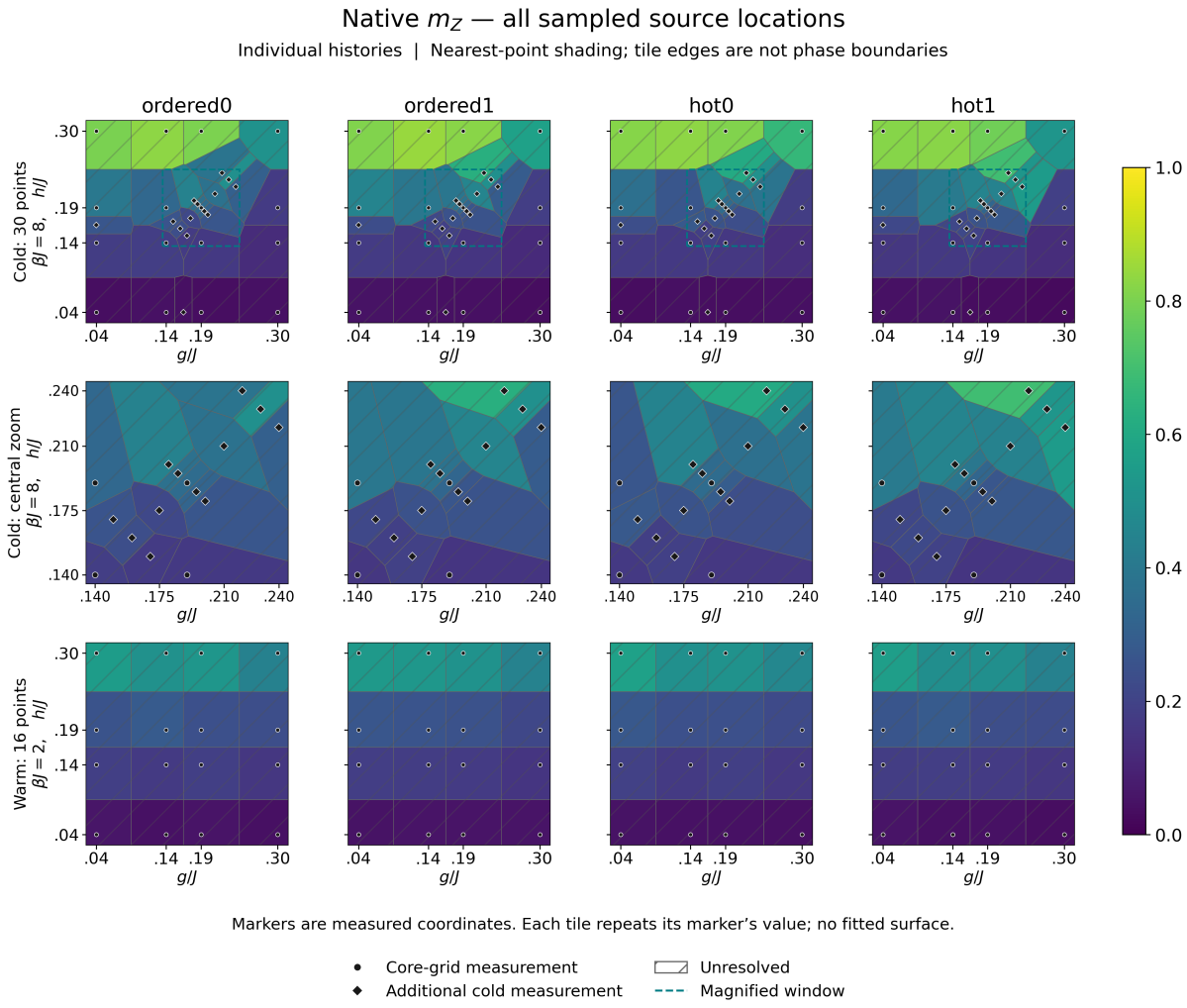


Figure 1: Native m_Z at every sampled location. Rows show the full cold plane (30 points), its magnified center, and the warm plane (16 points); columns retain four independent histories. Circles mark core-grid measurements and diamonds the fourteen additional cold points. Constant-color Voronoi tiles carry the nearest measurement, not a fitted surface. Hatching marks unresolved groups. Tile edges are display boundaries, not phase boundaries.

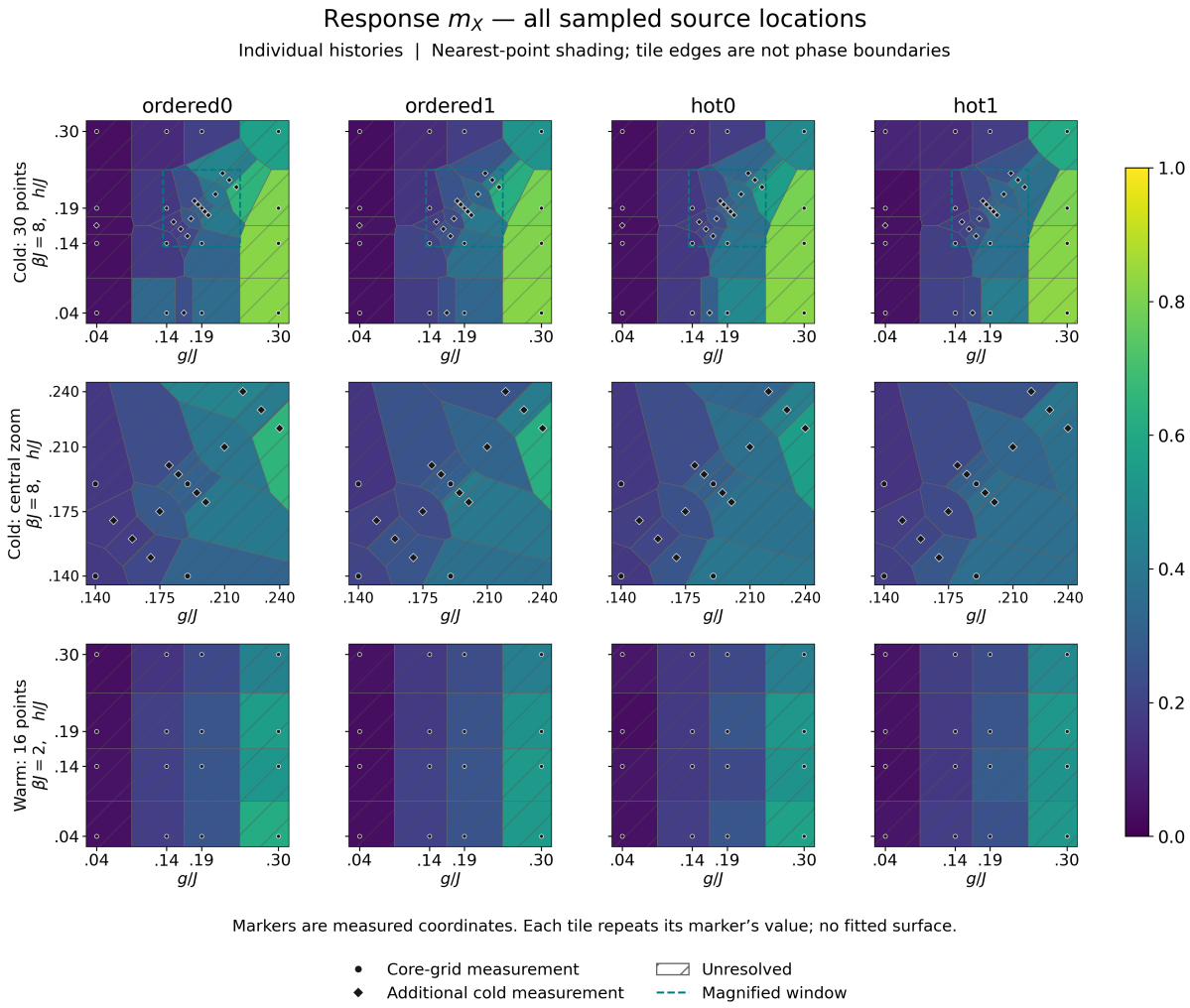


Figure 2: Complementary m_X response on the same full and magnified source planes. All thirty cold points and sixteen warm points are included. Temporal derivatives and native slice insertions can have different statistical efficiency. Markers, nearest-point shading and failure masks have the meaning of Figure 1.

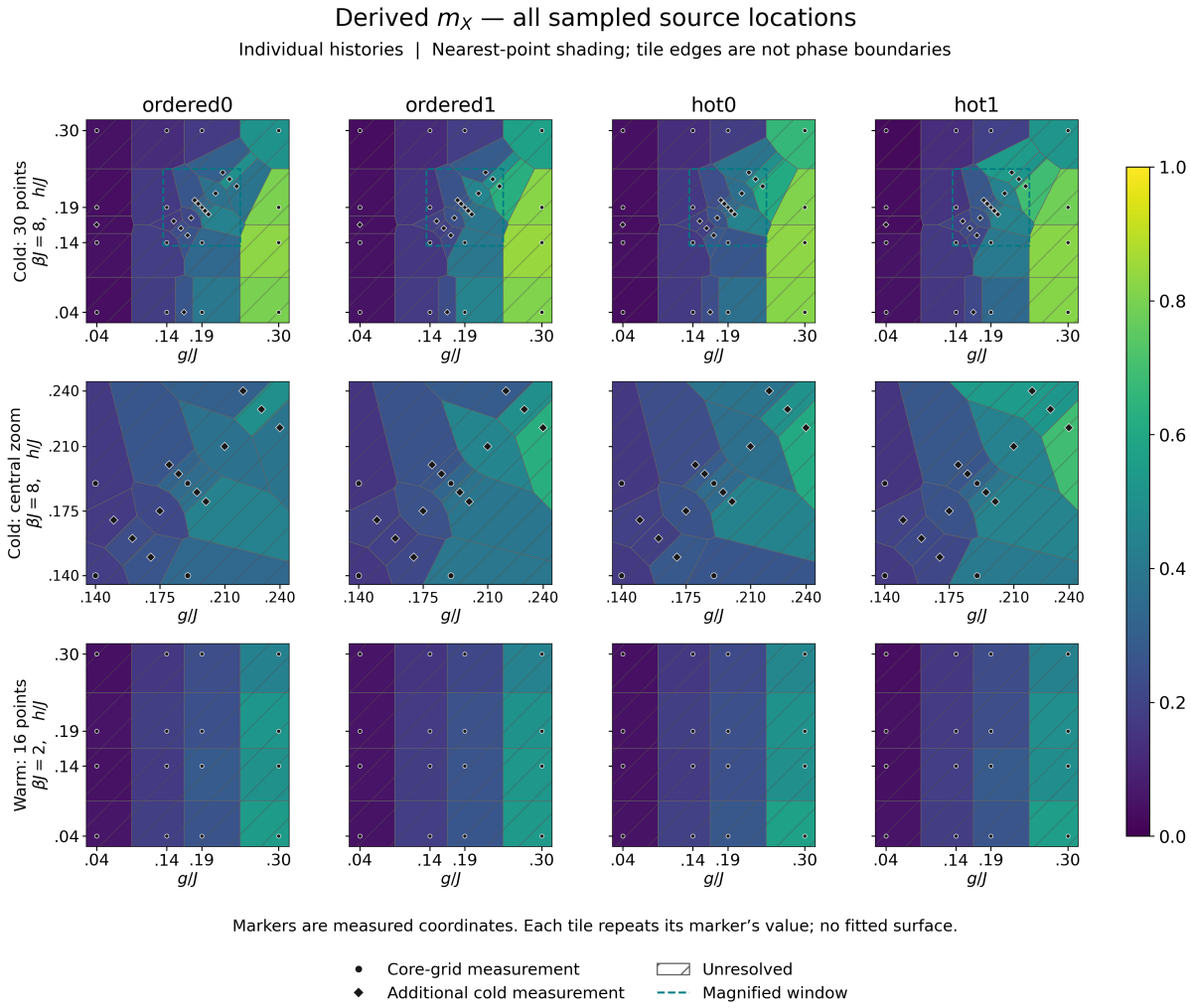


Figure 3: Duality-derived $m_X(g,h) = m_Z(h,g)$ at all sampled locations, including the central refinements. Reflected histories retain their uncertainty and failure status. These transported values are not new independent X-basis histories and are not averaged with the response estimator. The magnifier clips the same nearest-point tiles as the full view.

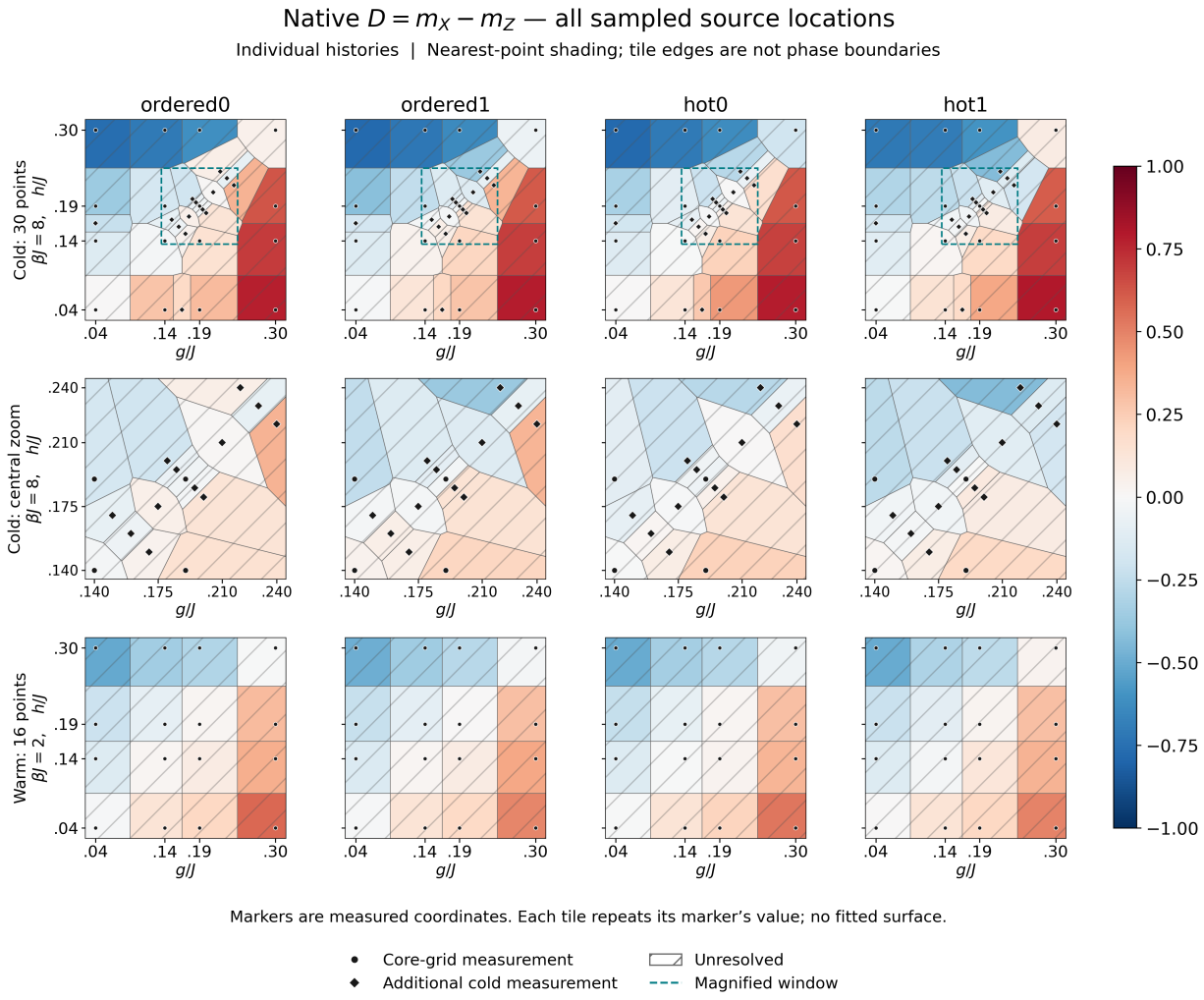


Figure 4: Native correlated difference $D = m_{X,\mathrm{resp}} - m_Z$ with local refinement. The central magnifier resolves the diagonal and nearby finite-bias samples. Values are not forced to vanish on the self-dual line. Every colored tile represents an actual history estimate; its footprint supplies no additional measured parameter values.

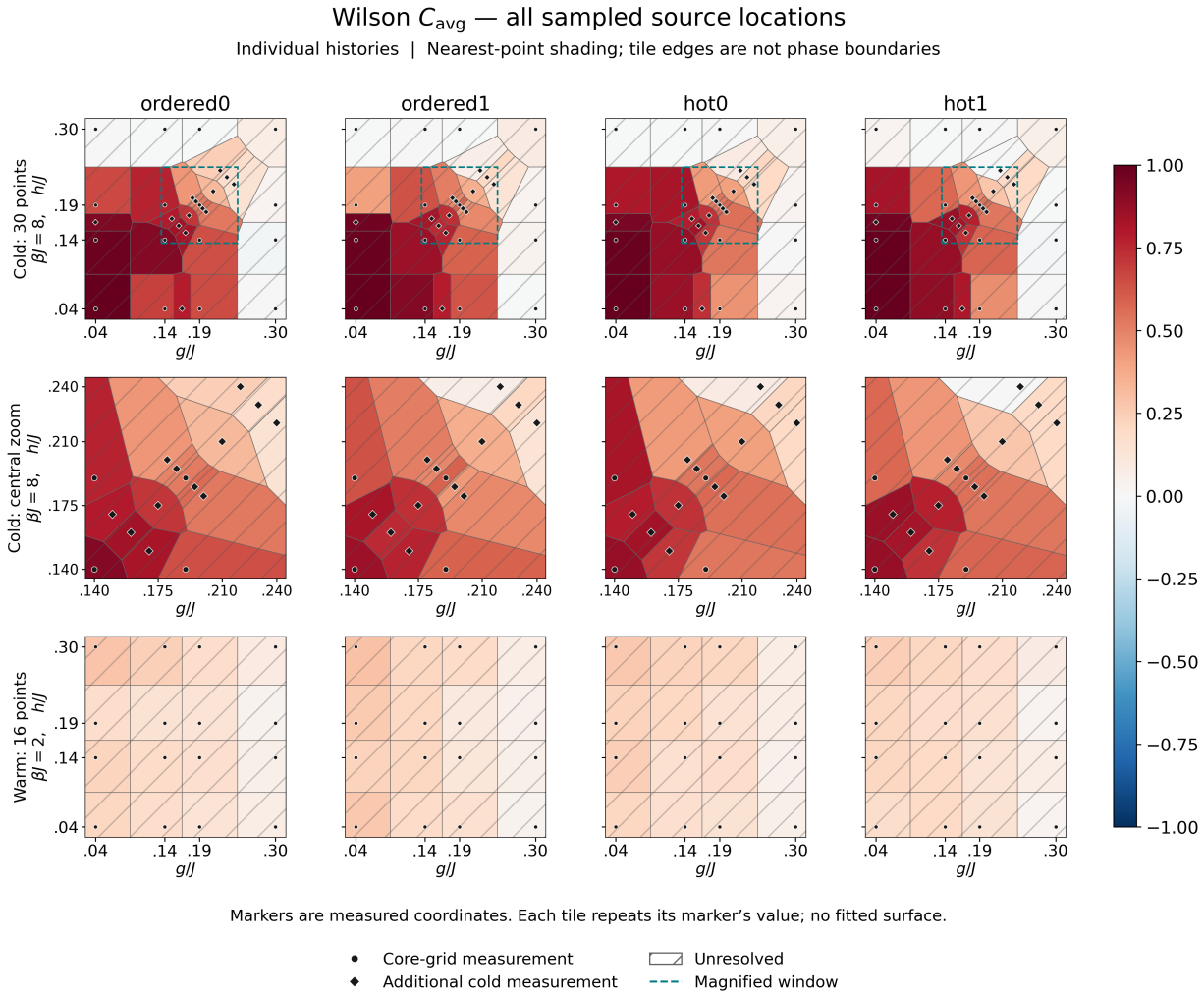


Figure 5: Signed connected Wilson functional averaged over the two correlated orientations within each history. The cold row includes all thirty measured locations, the middle row magnifies its dense central region, and the warm row has sixteen measurements. Diamonds identify additional cold points. Tile edges are nearest-point display boundaries, not measured phase lines. Negative estimates and all unresolved masks are retained.

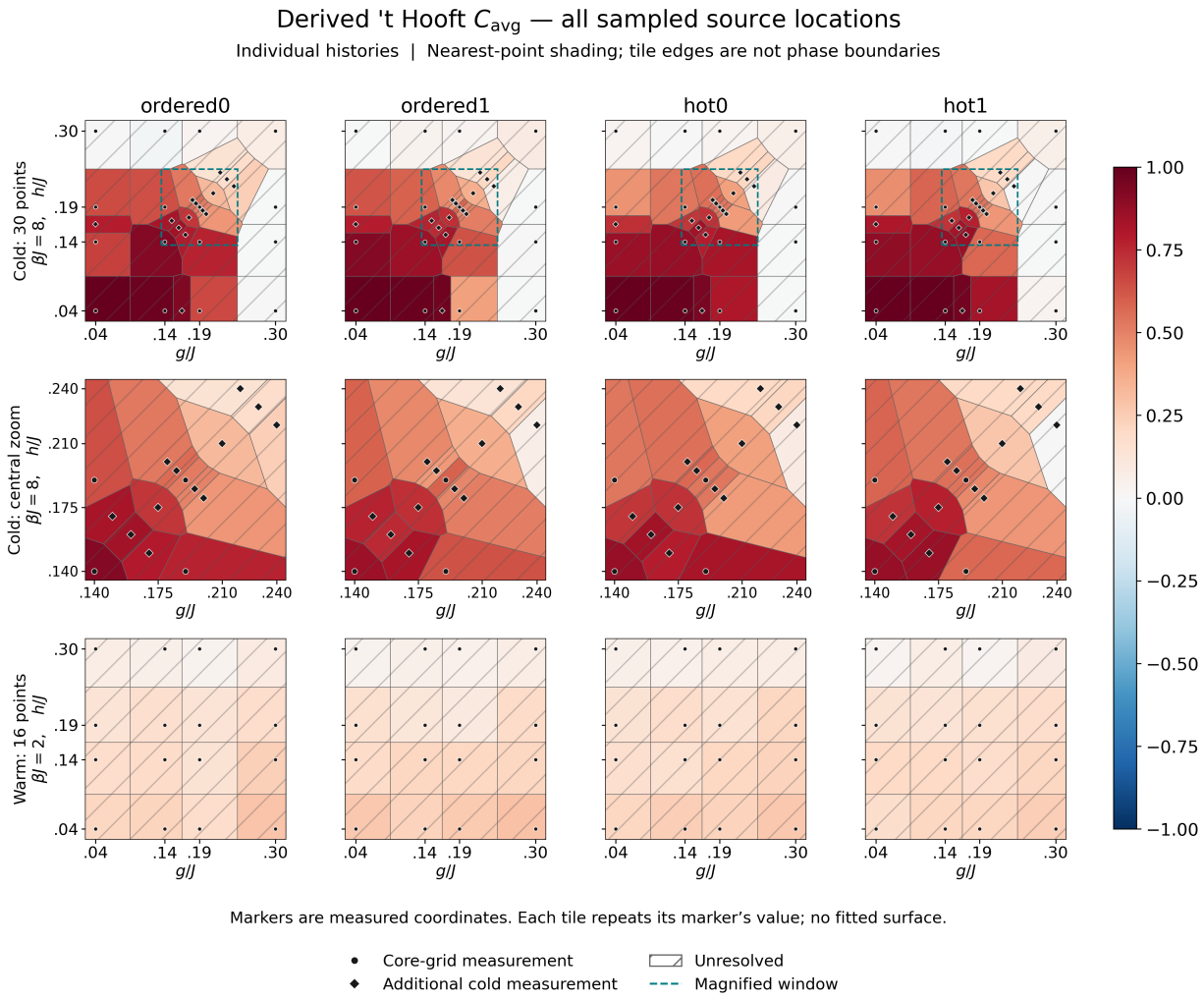


Figure 6: Derived connected T map from reflected W measurements at the full set of source locations. The zoom uses the same cells and colors as the full cold plane. At self-dual points reflection repeats the same datum, not an independent test. Transported failure masks remain visible and no sample count is increased.

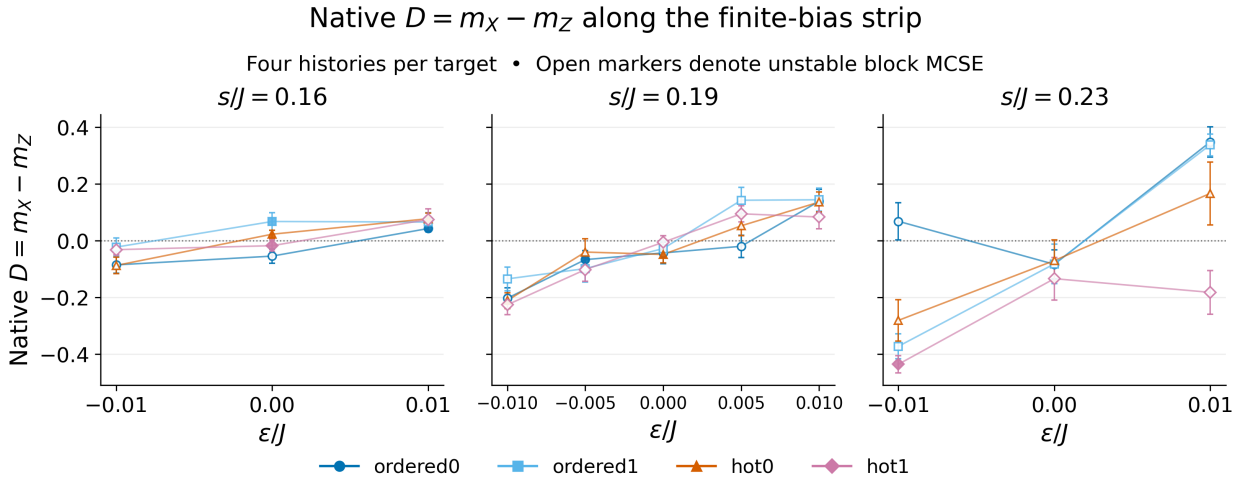


Figure 7: Cold finite-bias profiles $g = s + \epsilon, h = s - \epsilon$. The $s = .19$ panel includes $\epsilon = \pm.005$ and $\pm.01$. Lines connect independently sampled parameter points only to guide the eye. Open markers flag unstable block MCSE; filled markers do not certify the full convergence gate.

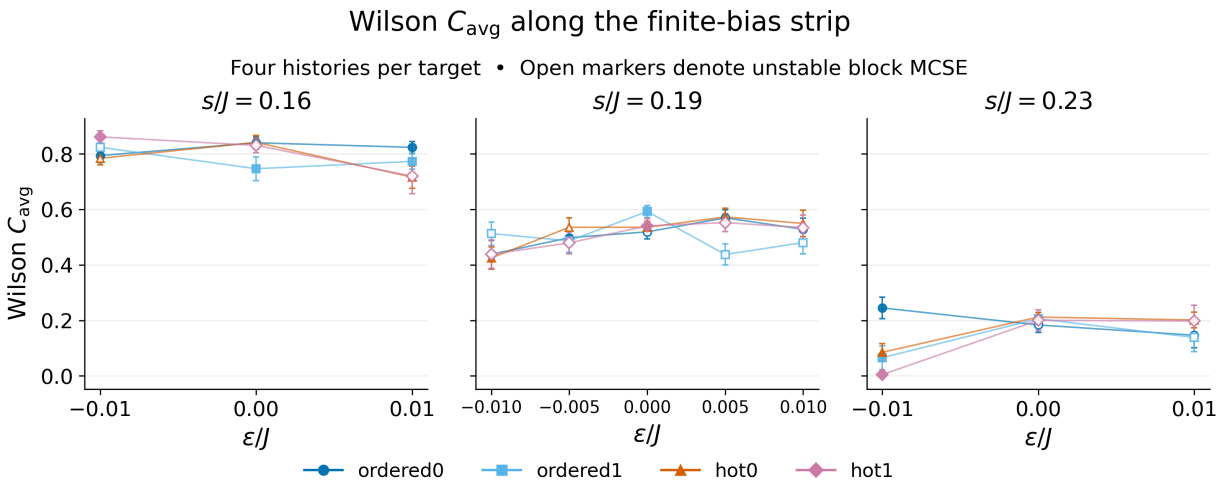


Figure 8: Connected W on the same focused cuts, using final cumulative windows. Errors retain the covariance of the two orientations. These diagnostic error bars are not certified equilibrium confidence intervals; the smaller-bias response does not determine a zero-bias jump.

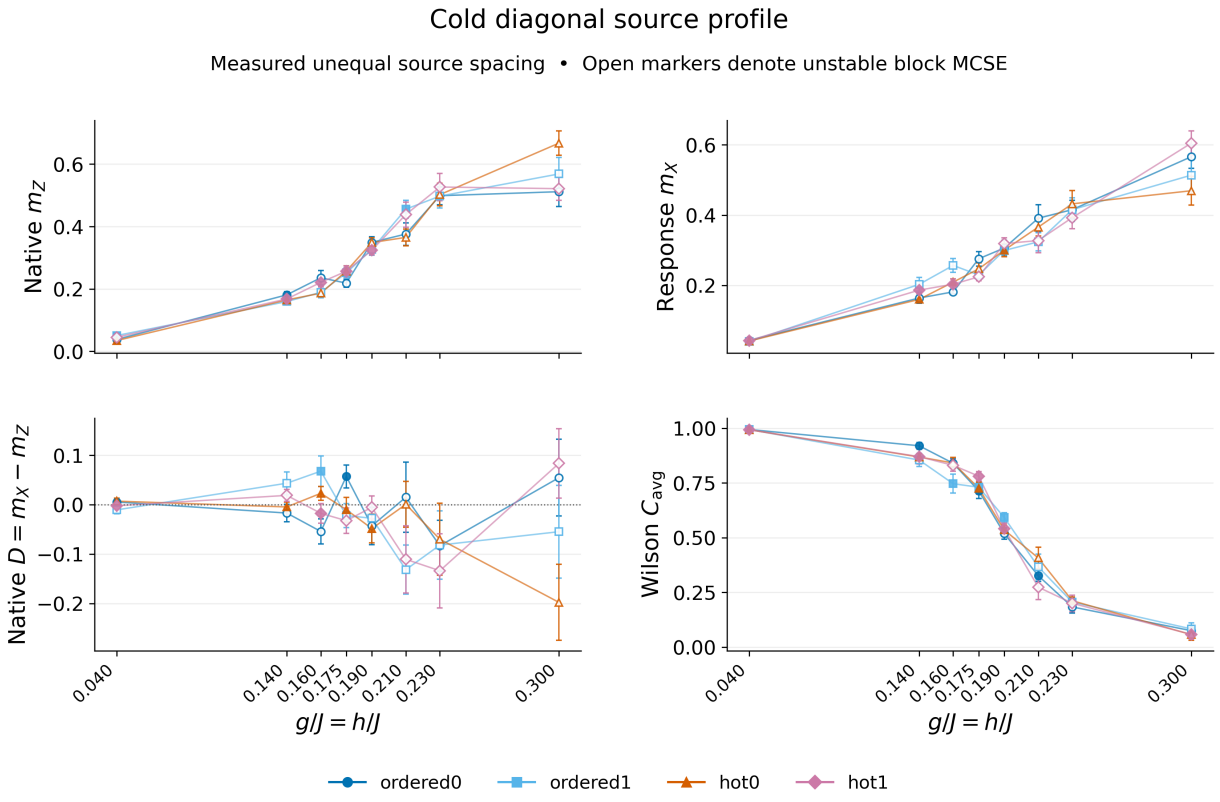


Figure 9: Cold diagonal source profiles at the actual, unequal field spacing. The .175 and .210 refinements resolve a substantial part of the loop drop, while the .190 center has the longest histories. The displayed variation is a finite-torus crossover profile, not a fitted critical line.

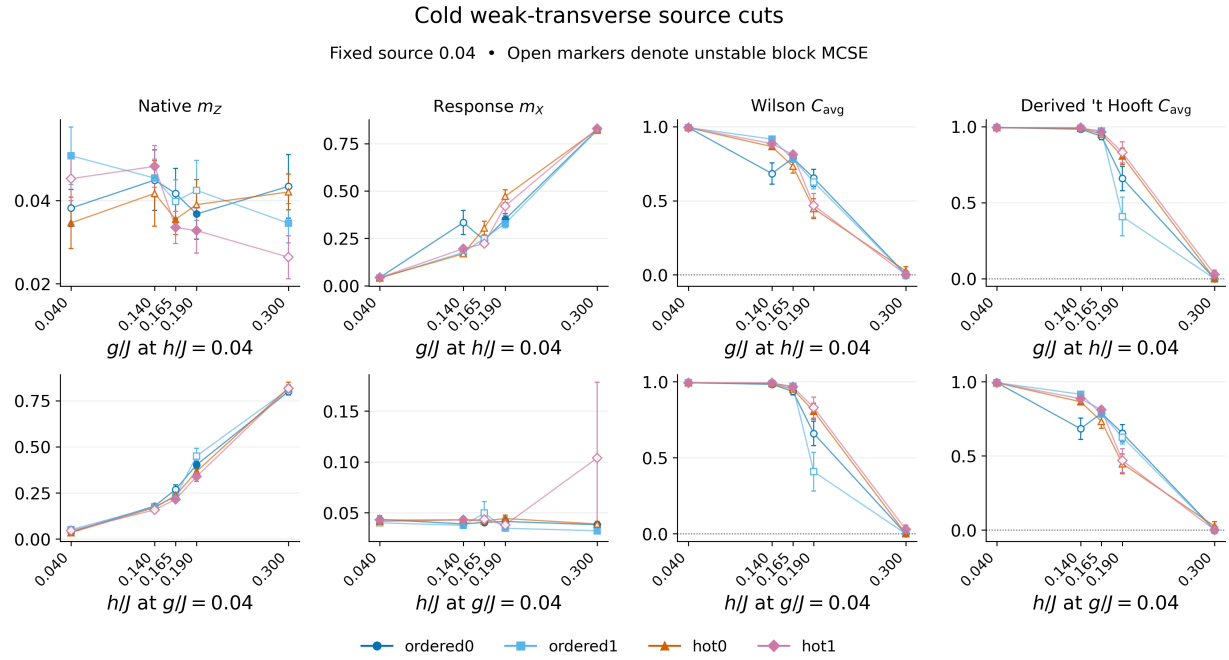


Figure 10: Cold weak-transverse-source cuts with the additional .165 points. Both reflection directions are shown at actual source coordinates. Magnetizations and loops distinguish polarization from product suppression. A transverse field .04 remains nonzero; these are not the exactly single-field axes.

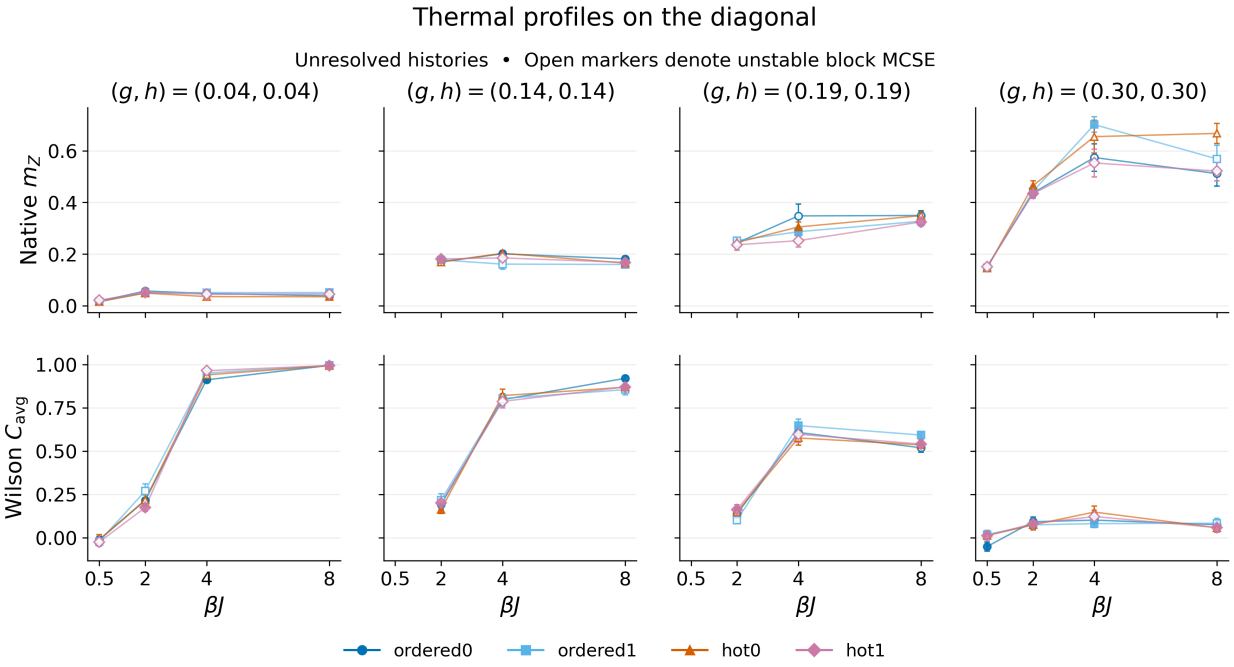


Figure 11: Diagonal temperature profiles with all final cumulative values. Only markers are observations; intermediate-source $\beta = .5$ points were not proposed. Connecting lines do not constitute continuum, cooling or size extrapolations.

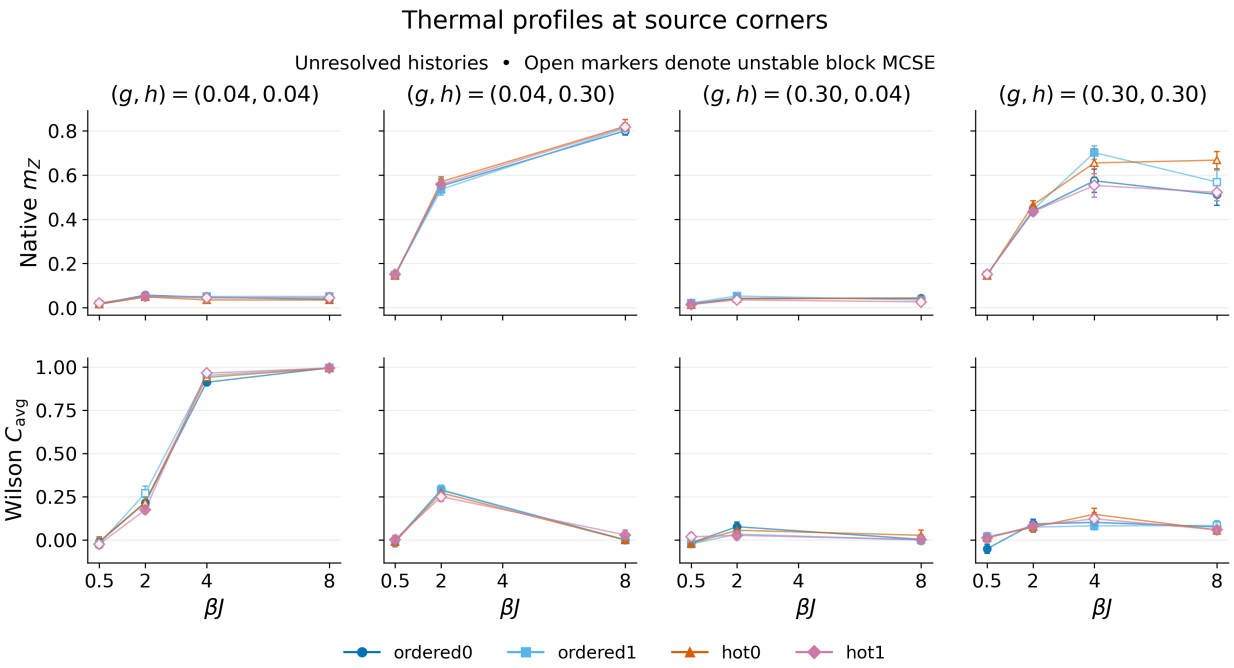


Figure 12: Thermal corner profiles. Weak-source loops grow strongly on cooling, whereas at $(.04, .30)$ cold source selection increases the disconnected term and reduces the connected quantity again. This is not a demonstrated reentrant phase boundary.

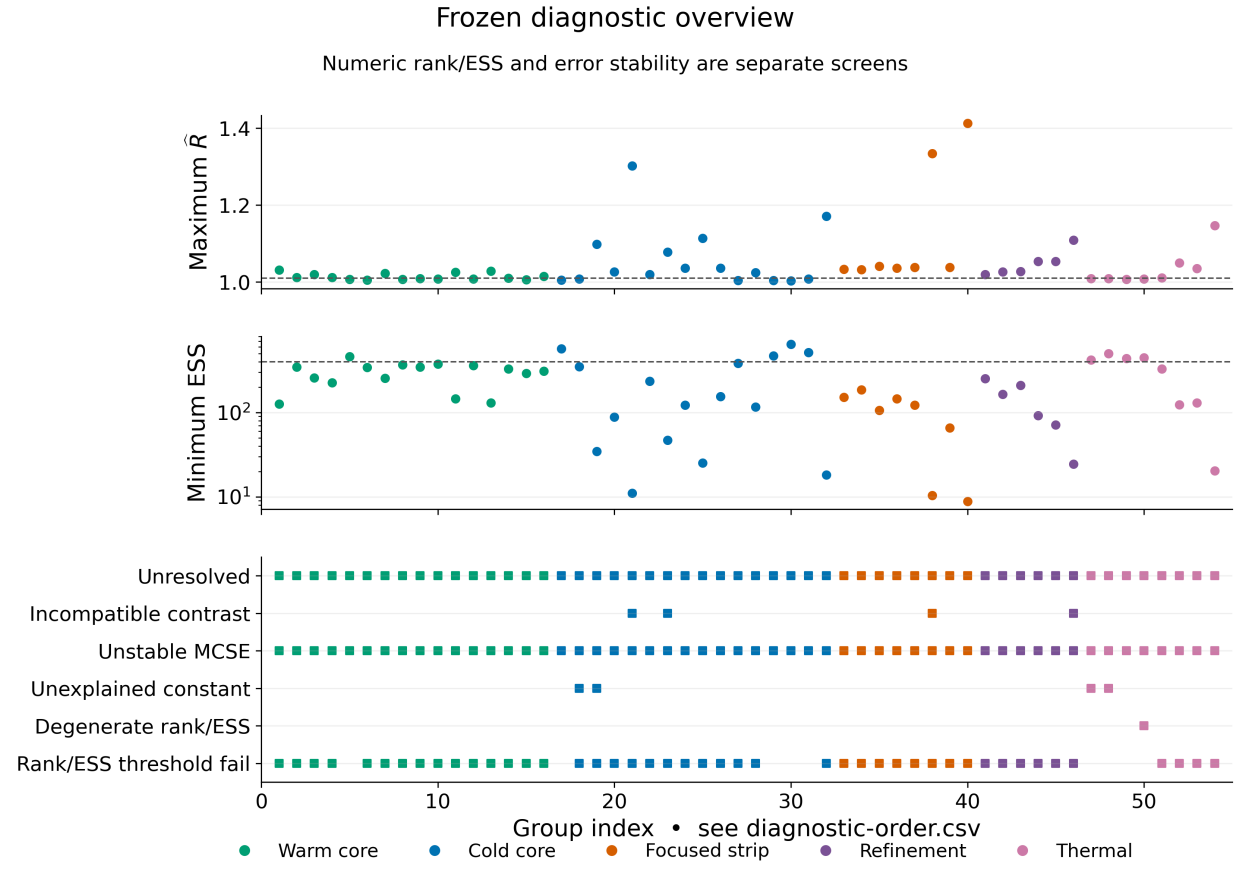


Figure 13: Final diagnostic overview. Extrema range over informative channels with numerical statistics; constant/degenerate channels are tracked separately. Thresholds are $\widehat{R} = 1.01$ and ESS400. Actual incompatible contrasts and unstable error ladders remain distinct. All groups remain unresolved; the full index dictionary follows the atlas.

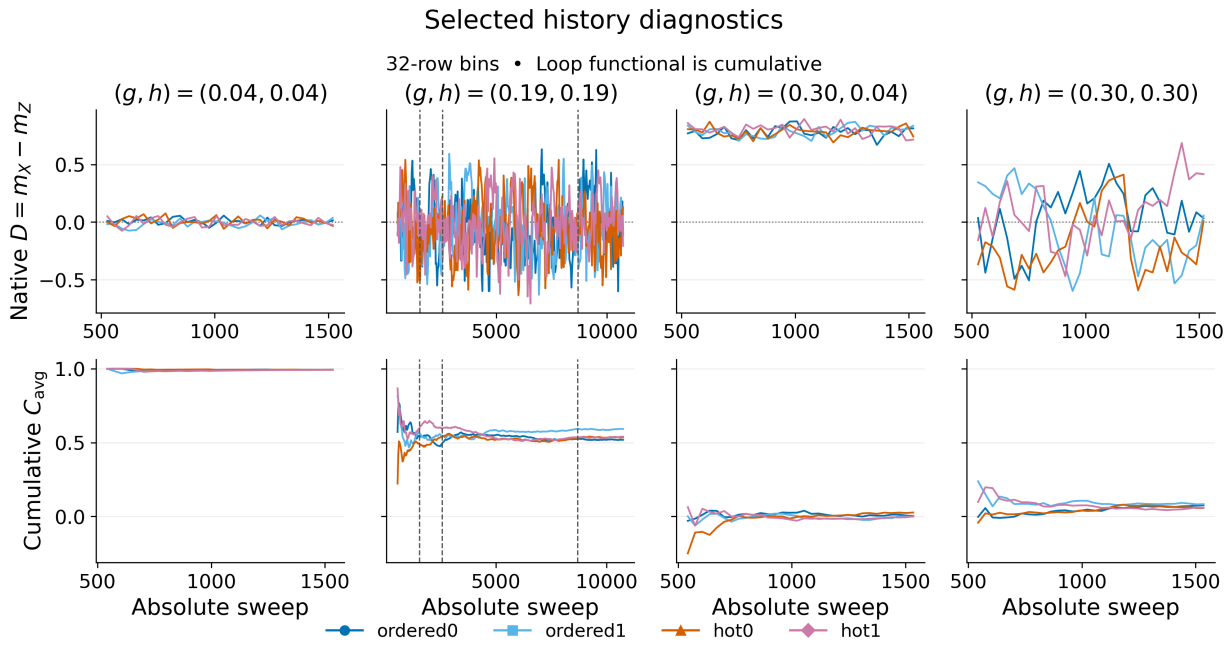


Figure 14: Selected retained histories: fixed 32-row means of native D above, cumulative joint connected functionals below. Vertical markers identify continuations of the same histories. The central point now spans 10,240 retained rows per history. Stable-looking loop curves can coexist with slower invariant or polarization modes.

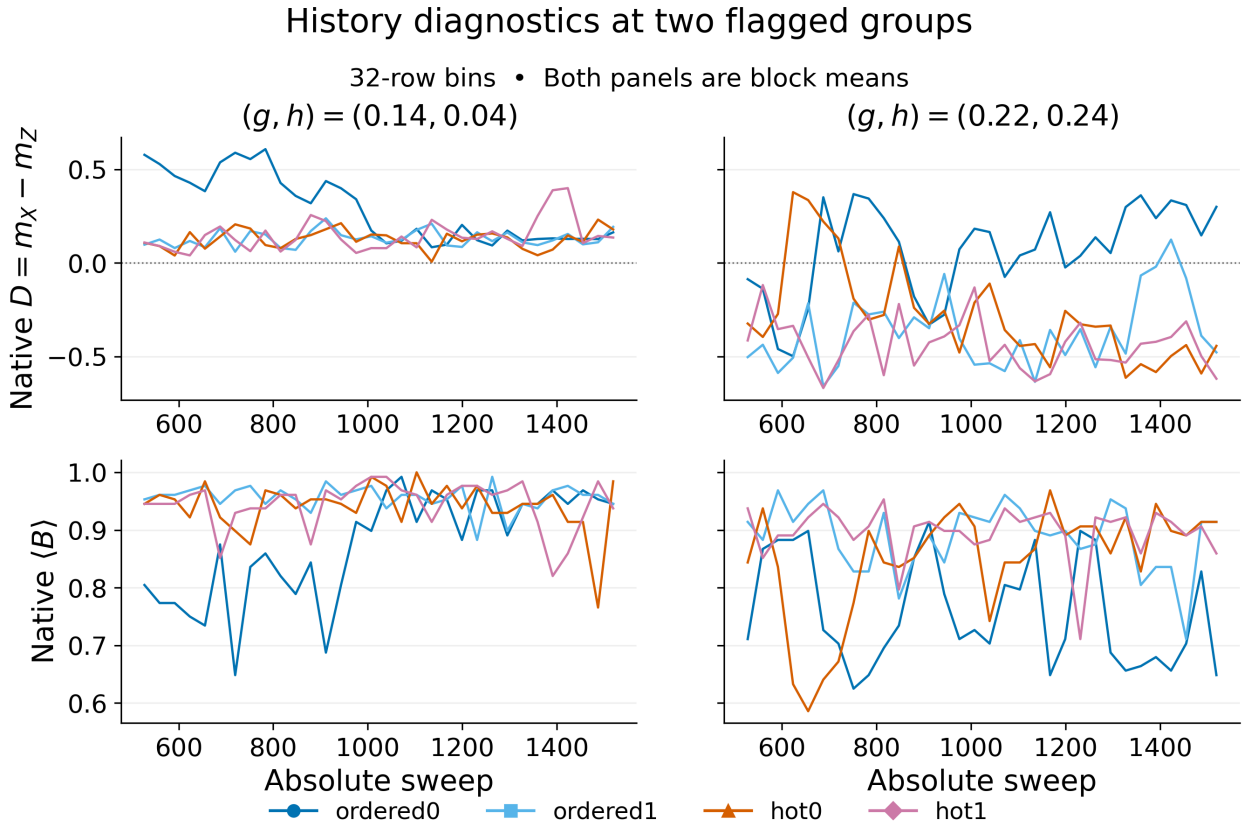


Figure 15: Two retained examples of actual incompatibility. Fixed 32-row bins reveal early ordered0 drift at $(.14, .04)$ and long background differences at $(.22, .24)$. These are Monte Carlo diagnostics, not quantum joint-probability or Binder distributions.

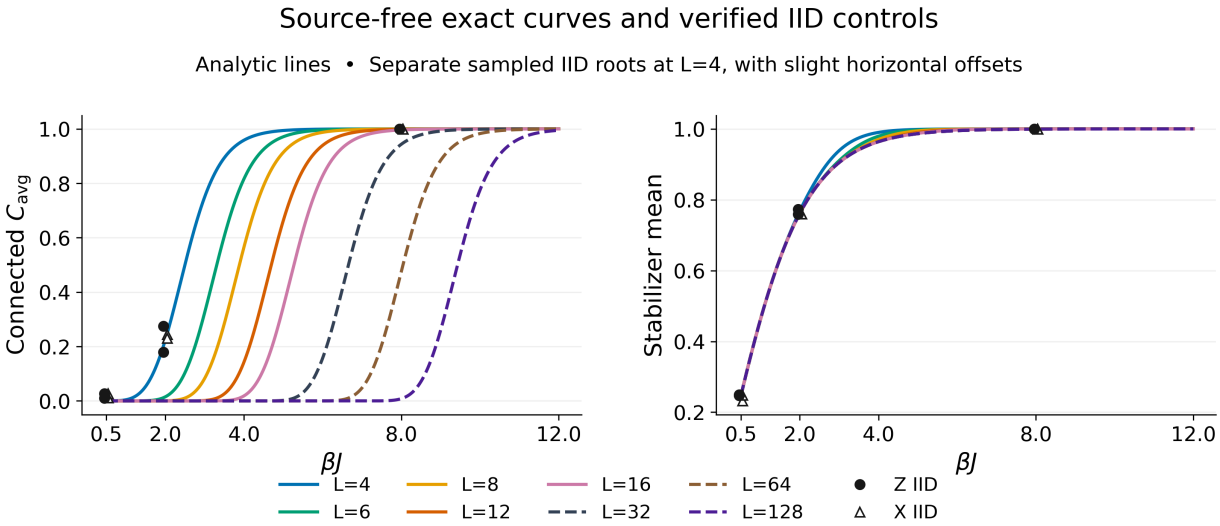


Figure 16: Exact source-free finite-size curves and preserved fresh $L = 4$ commuting IID controls in both bases. Dashed large-L curves are analytic, not extra simulations. Zero observed cold-control variance does not exclude rare events. The limits of size and temperature do not commute.

Diagnostic group dictionary

All rows have $J = K = 1$ and $L = 4$. Indices match Figure 13.

Index	β	M	g	h	Index	β	M	g	h
1	2	8	0.04	0.04	28	8	32	0.19	0.3
2	2	8	0.04	0.14	29	8	32	0.3	0.04
3	2	8	0.04	0.19	30	8	32	0.3	0.14
4	2	8	0.04	0.3	31	8	32	0.3	0.19
5	2	8	0.14	0.04	32	8	32	0.3	0.3
6	2	8	0.14	0.14	33	8	32	0.15	0.17
7	2	8	0.14	0.19	34	8	32	0.16	0.16
8	2	8	0.14	0.3	35	8	32	0.17	0.15
9	2	8	0.19	0.04	36	8	32	0.18	0.2
10	2	8	0.19	0.14	37	8	32	0.2	0.18
11	2	8	0.19	0.19	38	8	32	0.22	0.24
12	2	8	0.19	0.3	39	8	32	0.23	0.23
13	2	8	0.3	0.04	40	8	32	0.24	0.22
14	2	8	0.3	0.14	41	8	32	0.04	0.165
15	2	8	0.3	0.19	42	8	32	0.165	0.04
16	2	8	0.3	0.3	43	8	32	0.175	0.175
17	8	32	0.04	0.04	44	8	32	0.185	0.195
18	8	32	0.04	0.14	45	8	32	0.195	0.185
19	8	32	0.04	0.19	46	8	32	0.21	0.21
20	8	32	0.04	0.3	47	0.5	4	0.04	0.04
21	8	32	0.14	0.04	48	0.5	4	0.04	0.3
22	8	32	0.14	0.14	49	0.5	4	0.3	0.04
23	8	32	0.14	0.19	50	0.5	4	0.3	0.3
24	8	32	0.14	0.3	51	4	16	0.04	0.04
25	8	32	0.19	0.04	52	4	16	0.14	0.14
26	8	32	0.19	0.14	53	4	16	0.19	0.19
27	8	32	0.19	0.19	54	4	16	0.3	0.3

A Execution, failures and reproducibility

A.1 Closed resource accounting

The combined study comprises an original sealed allocation and its approved continuation; those allocations are not independent copies of continued histories. The original first exposure was 7 November 2026, 13:36:24.54 UTC, and its last sampling settlement was 18:36:42.78 UTC. Its conservative monotonic drawing clock was 18,018.80 seconds, with evidence sealed at 18:44:21.89 UTC, below its seven-hour cap. The preserved original accounting contains 15,712.48 union-active wall seconds and 173,219.17 job-seconds.

The extension's first exposure was 7 November 2026, 22:02:19.0952406 UTC. Its last native settlement was 8 November 2026, 04:10:49.013953 UTC, after 22,109.919 seconds (6 hours 8 minutes 29.919 seconds), including controls, pauses, failures and repairs. The user had raised the hard cap from six to seven hours *before these draws*; the unchanged absolute deadline was 05:02:19.0952406 UTC. Optional continuation auto-sealed at 04:10:49.093481 UTC. Main evidence auditing began at 04:11:23.105900 UTC and sealed at 04:35:31.571723 UTC, taking 1,448.466 read-only seconds. The latter interval contains no sampling and is reported separately. All owned native descendants were stopped; the sampling lease was absent. Subsequent export, analysis and rendering used no stochastic draws.

The extension charged 20,484.44 union-active wall seconds and 215,142.839 job-seconds, including technical retries. Job-seconds sum elapsed job durations; they are not measured CPU time. It added 176,128 unique science rows to the original 167,936, giving 344,064. The 151,552-row main extension and 24,576-row final continuation have separate sealed ledgers with the same global clock. Their shared predecessors are counted once. Hardware ceilings and the package runtime remained unchanged.

The extension watchdog retained 27 actual timestamped pulses. The largest recorded gap was 1,575.8 seconds (26 minutes 15.8 seconds), spanning 22:48:04.834–23:14:20.601 UTC during recovery. Later delayed pulses and corrections are explicit; none was invented retrospectively. The coarse observer was therefore approximate, not an exact fifteen-minute scheduler. Native five-second health/resource guards and supervisor checks continued; the final pulse confirms both sealed states and zero active descendants.

A.2 Preserved incidents and recovery

The original allocation retains eight interrupted tiny-oracle operations from a tight wrapper deadline. Its 1,998 durable rows were reproduced bit-for-bit by an explicitly charged technical replay. Original pre-launch memory refusals and their registration repair also remain recorded. These controls and failed prefixes are not scientific rows.

The extension encountered a separate orchestration timeout in wave0011: the cost forecast had also been used as the wave hard wall. All twelve jobs were progressing, with 249–255 durable rows each, when they were stopped. The 3,025 partial rows remain intact. A reviewed controller revision separated the cost forecast from the hard timeout, preserving the original global clock, physics, random streams and committed parents. Replay resumed the preceding valid checkpoints; all twelve partial prefixes matched exactly, including all 3,025 raw rows, arrays, metadata and random states. Those failed rows received no extra independent scientific credit. All failure and replay costs remained charged.

A subsequent RAM admission deferral occurred before reservation or exposure. Under the user's existing authorization, the supervisor trimmed Chrome working-set pages without closing tabs or killing processes. No system setting, dependency or package was changed; reported working-set reductions are not claimed as equivalent physical RAM freed. Sampling resumed only after admission passed. The final continuation had no failures.

A.3 Independent identity and arithmetic checks

Main sealing verified the full original 197,694-file raw baseline and 39 report/archive files, preserving their bytes. It also rechecked the accepted frozen inputs and the package's 54 Python modules. The final main inventory has 598 completed and 610 attempted operations, including the twelve timeout operations; all twenty planned group views completed. The optional inventory adds 24 completed operations and three cumulative views. No runtime change or additional scientific stream followed closure.

The routine reporting export validated all receipt lineages, hashes, retained clocks, means, half windows and block-error/bias calculations. Its largest residual against frozen summaries was 2.68×10^{-13}. Six deterministic fixtures check nonlinear connected functionals, cross-orientation covariance, window/lineage joins, reflection and refined plot coverage. A separate supervisor calculation imported no sampler or reporting helpers: it read 204,800 raw rows across all eighty augmented histories, recomputed signed joint primitives and temporal derivative responses, and compared 160 primary/tail history views. Maximum agreement error was 9.28×10^{-15}. These are reporting/target checks, not proofs of physical equilibration.

All sixteen figure families use the final combined export, in PNG and SVG. The source maps include all off-grid cold points with a central magnifier. Two additional profiles display the precise diagonal and weak-transverse-source refinements. The old report, exports and figures remain in the verified archive. The live report presents the final study coherently; historical and tail views remain separately accessible for audit.

A.4 Artifact entry points

Paths below are relative to the research repository.

- Specification and original parameter catalogue: `simulations/d2-p1-s1/SPECIFICATION.md` and `simulations/d2-p1-s1/configs/prospective-080.json`.
- Combined data, signed primitives, uncertainty and symmetry diagnostics: `simulations/d2-p1-s1/analysis/data/study.json`, `simulations/d2-p1-s1/analysis/data/history-summary.csv` and `simulations/d2-p1-s1/analysis/data/self-dual-D.csv`. Historical/tail/all-history diagnostic views in these files are not extra independent histories or qualified principal estimates.
- Live figures and report: `simulations/d2-p1-s1/reports/`.
- Original frozen evidence: `development/provenance/d2-p1-s1-080-worker-2026-11-07-attempt-01/`.
- Extension registry, receipts, failures, main and optional ledgers: `development/provenance/d2-p1-s1-080-extension-worker-2026-11-08-attempt-01/`. The original report archive is in its `original-report/` subdirectory.
- Root protocol, admissions, independent arithmetic, clocks and verdict: `development/provenance/d2-p1-s1-080-extension-supervisor-2026-11-08/`.
- Staged export, tests, figure manifests and readback: `development/provenance/d2-p1-s1-080-extension-reporting-2026-11-08/`.

Frozen wheel SHA256:
`509e4ddc677ce081b633db38fb4fbd3244120d5cc5a2b867a91484b75535469c`

Main extension registration SHA256:
`990eba002e503cf030a6d2d450b9f1d4ce612cb097a649c8b23de5f19e759e59`

Main sealed ledger SHA256:
`0c10d3f4d45edc073febe00153bca61fef2f838f6174a25b006736aedb8bec86`
Optional sealed ledger SHA256:
`f4b09222b1b1c770cc7dd4ba21ef98936f74aeb252de406d3360cca5286e02d3`
Combined study export SHA256:
`ff245b916485dce6026ebbfd46ad829db33f5395cf8db9d6a4278d28dfe9cfdb`.

References

[1] N. Ishtiaque and S. Chakroborty, *An exact spacetime polymer gas for finite-temperature $\mathbb{Z}_N$ homological quantum code*, working manuscript retained in `reference/draft.tex` and `reference/draft-sections.tex`.

[2] F. Wu, Y. Deng and N. Prokof'ev, *Phase diagram of the toric code model in a parallel magnetic field*, arXiv:1201.6409 (2012).

[3] D. Mazáč and A. Hamma, *Topological order, entanglement, and quantum memory at finite temperature*, arXiv:1112.0947 (2012).