# Ising Phase Boundaries from the Quantum Vertex Code

One mapping, two benchmarks: isotropic crossings and an anisotropic phase classifier

Homological-code phase project
Public `hcphases` Monte Carlo package

1 November 2026

**Abstract**

We benchmark the source paper's $d = 1, p = 0, N = 2$ homological code through its exact finite-Trotter mapping to a classical Ising model. The first experiment restricts the two couplings to the isotropic line and locates a finite-size Binder crossing. The continuation varies the spatial and temporal couplings independently, measures only the two magnetization moments needed for the Binder ratio, and learns a phase classifier from crossings on three lattice sizes. The exact Ising boundary is reserved for comparison. Both experiments use the public package on periodic square spacetime tori. Statistical uncertainty, interpolation error, finite-size drift and unresolved diagnostics are distinguished throughout; none of these numerical results is a proof or an acceptance of the four-dimensional research gates.

**Principal findings.** The isotropic 32/64 crossing is 0.4411288, 0.1003% above the exact value, outside its statistical-only interval. In the anisotropic continuation, 7 of 7 primary crossing cuts are resolved; their largest absolute relative boundary displacement is 0.191%. The learned classifier's maximum displacement on the comparison grid in its resolved interpolation domain is 0.546%. These are finite-size benchmark results, with statistical intervals that exclude interpolation and finite-size bias.

## Contents

**1 Shared theory: the code and its classical image** **2**

**2 What is measured, and why crossings detect a transition** **2**

**3 Shared sampling and statistical safeguards** **3**

3.1 Uncertainty and diagnostics . . . . . 4

**4 Experiment I: the isotropic benchmark** **4**

**5 Experiment II: the anisotropic continuation** **6**

5.1 A frozen two-parameter scan . . . . . 6

5.2 A classifier learned from Binder crossings . . . . . 6

5.3 Results, diagnostics and comparison with the exact boundary . . . . . 7

**6 Reproducibility and scope of the conclusions** **10**

# 1 Shared theory: the code and its classical image

A qubit occupies each vertex of a periodic circle of length $L$. In the source paper's unaugmented complex there is no negative-degree stabilizer at $p = 0$, so the $J$ term is absent. Keep one transverse source, with no longitudinal source:

$$H = H_z + H_x, \qquad H_z = -\frac{K}{2}\sum_{j=0}^{L-1}(1 + Z_j Z_{j+1}), \qquad H_x = -g\sum_{j=0}^{L-1} X_j, \qquad h = 0. \tag{1}$$

Here $K$ multiplies the code projector; the conventional Ising exchange is $K/2$. The thermal trace includes the full Hilbert space, not just the two aligned code states. Set $k_B = \hbar = 1$, $\beta = 1/T$, and $\delta = \beta/M$. The ordered finite-$M$ product is

$$Z_M = \mathrm{Tr}\big(e^{-\delta H_z} e^{-\delta H_x}\big)^M. \tag{2}$$

Inserting $Z$ eigenstates and using $\langle s|e^{\delta g X}|s'\rangle = \cosh(\delta g)$ if $s = s'$ and $\sinh(\delta g)$ otherwise gives

$$Z_M = C \sum_{\sigma\in\{\pm1\}^{L\times M}} \exp\left[\kappa_x \sum_{j,t} \sigma_{j,t}\sigma_{j+1,t} + \kappa_\tau \sum_{j,t} \sigma_{j,t}\sigma_{j,t+1}\right], \tag{3}$$

$$\kappa_x = \frac{\delta K}{2}, \qquad \kappa_\tau = -\frac{1}{2}\log\tanh(\delta g), \tag{4}$$

$$C = e^{\beta K L/2}\big[\cosh(\delta g)\sinh(\delta g)\big]^{LM/2}. \tag{5}$$

The constant $C$ cancels in normalized expectations. The thermal trace and spatial circle impose periodic boundaries in both directions. The full normalization and operator derivation is in `../../theory/main.tex`.

Both experiments use the convenient inverse map

$$\boxed{L = M = n, \quad \delta = 1, \quad T = 1/n, \quad K = 2\kappa_x, \quad g = -\tfrac{1}{2}\log\tanh\kappa_\tau.} \tag{6}$$

Thus arbitrary positive $(\kappa_x, \kappa_\tau)$ are realized by the same quantum-code family. Isotropy imposes the additional relation $\kappa_x = \kappa_\tau = \kappa$; anisotropy releases that constraint without introducing a second source. At fixed $n$, changing $\kappa_x$ changes $K$ and changing $\kappa_\tau$ changes $g$. Changing $n$ also changes the quantum temperature. This is an exact benchmark of the finite-$M$ classical measure, not a Trotter-continuum extrapolation of a fixed quantum Hamiltonian. The classical thermodynamic limit sends both $L$ and $M$ to infinity; in this family it also sends the quantum temperature to zero. These plots therefore do not establish a finite-temperature symmetry-breaking transition of a one-dimensional quantum chain at fixed $T$.

The exact infinite-volume positive-coupling boundary is[2]

$$\sinh(2\kappa_x)\sinh(2\kappa_\tau) = 1, \qquad \kappa_\tau^{\mathrm{exact}}(x) = \tfrac{1}{2}\,\mathrm{arsinh}\left(\frac{1}{\sinh(2x)}\right). \tag{7}$$

The product is smaller than one in the disordered region and larger than one in the ordered region. On the isotropic line it reduces to $\kappa_c = \frac{1}{2}\log(1+\sqrt{2}) = 0.44068679350977\ldots$. These formulas enter the comparison stage only. The experiment is informed by the known model and a useful parameter range, but its fitting labels, refinement decisions and stopping rules do not use the analytic boundary. A finite periodic torus itself has no thermodynamic singularity.

# 2 What is measured, and why crossings detect a transition

Every collection of vertices is a 0-cycle. A Wilson insertion $W_\nu = \prod_j Z_j^{\nu_j}$ maps to the corresponding spin product on an imaginary-time slice. Multiple slices give Euclidean time-ordered insertions.

Consequently the moments of

$$m = \frac{1}{LM} \sum_{j,t} \sigma_{j,t} \tag{8}$$

are sums of these multi-insertion correlators. They are not powers of a single-time quantum magnetization operator. The Binder ratio is[3]

$$U_n = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}. \tag{9}$$

Both moments are measured jointly on each configuration, then averaged before forming the ratio. Averaging a per-configuration ratio would be wrong. The zero-field finite-volume measure has $\langle m \rangle = 0$ on both sides; its distribution is approximately Gaussian far in the disordered region and concentrates near two opposite magnetizations in the ordered region. Accordingly $U_n$ tends to 0 and 2/3, respectively, as volume increases. Near a continuous transition its leading scaling is size independent, so crossings compare distributions at different sizes without imposing a single arbitrary threshold.

For each fixed spatial coupling define

$$D_n(\kappa_\tau; \kappa_x) = U_n - U_{2n}. \tag{10}$$

A zero estimates a finite-size boundary location. The 32/64 crossing is primary; 16/32 provides a size-drift diagnostic. The critical Binder height can depend on anisotropy and effective aspect ratio, even when the geometric torus is square.[5] Therefore the anisotropic classifier is based on size crossings, not an isotropic universal Binder-height cut.

The original isotropic run also measured $m$, $|m|$, nearest-neighbor bonds, several Wilson point-pair correlators, and a joint four-spin insertion. Those records and their connected-correlation analysis are preserved. The continuation measures *only* $m^2$ *and* $m^4$; all its physics plots and diagnostics derive from these two columns. Neither experiment samples twisted boundary conditions. Global-X and magnetic insertions produce temporal and spatial twists as derived in the theory note, but no equal-weight mixture of boundary backgrounds or twist-ratio estimate is introduced here.

## 3 Shared sampling and statistical safeguards

Models, observables, sampling plans and resource budgets are declared through the exported public API; `run_vertex` executes each cell. The study contains no independent Ising sampler. Each update stays put with probability 0.05; otherwise it draws the ferromagnetic bond field with probabilities $p_x = 1 - e^{-2\kappa_x}$ and $p_\tau = 1 - e^{-2\kappa_\tau}$, retains only bonds between equal spins, selects a uniform seed vertex, and flips its connected component. This is the Wolff single-cluster kernel.[4] Drawing the whole bond field is equivalent to revealing the necessary bonds while growing a cluster. The joint spin–bond measure proves stationarity; positive singleton probabilities give reachability, and laziness gives aperiodicity. These facts do not bound the mixing time.

A clock tick is one cluster attempt, including lazy attempts. Samples are taken at fixed attempt intervals, never at a state-dependent stopping time based on the cumulative number of flipped spins. Each cell has four chains starting ordered, hot, ordered, hot. Independent identities determine seeds regardless of execution order. Pseudorandom-stream independence is a statistical assumption, not a theorem about the generator.

Public execution checks frozen implementations, deadlines, process-tree memory and output limits. It saves joint moment records, selected full-spin snapshots, and final state/RNG/counter checkpoints. Public reading independently recomputes measurements for the saved snapshots; it does not claim to replay unstored configurations. Tiny exact tests compare anisotropic weights, quantum transfer traces, Wilson products and cluster detailed balance. The continuation additionally checks

spatial/time transposition by complete $3 \times 3$ enumeration, Binder ratio algebra, crossing inventories and classifier interpolation on unrelated literal fixtures. These are engineering tests, not independent scientific confirmation trials.

### 3.1 Uncertainty and diagnostics

Within each chain we resample contiguous blocks, keeping the two moments joint. Every one of 2,000 bootstrap resamples recomputes the Binder ratio and the curve crossing. The primary block length is 64 retained records; 32 and 128 assess block-size sensitivity. Roots are zeros of piecewise-linear Binder differences. Missing and multiple roots remain counted. A crossing is accepted as an estimate only if the central root is unique, the outer bracket endpoints have opposite signs throughout their central 95% bootstrap intervals, and at least 95% of bootstrap samples have a unique root. Reported percentile intervals are conditional on uniqueness.

We monitor rank-normalized folded split-$\widehat{R}$ and bulk ESS for the moments and the linearized Binder influence

$$I = -\frac{m^4 - \langle m^4 \rangle}{3\langle m^2 \rangle^2} + \frac{2\langle m^4 \rangle}{3\langle m^2 \rangle^3}(m^2 - \langle m^2 \rangle). \tag{11}$$

Targets are $\widehat{R} \leq 1.01$ and bulk ESS $\geq 1000$. The original isotropic run also checked a bond-energy series. Failed targets are retained; histories are not extended selectively. Neither good diagnostics nor block stability certifies mixing or confidence coverage. Statistical intervals exclude finite-size bias, interpolation error, warmup error and adaptive-grid selection uncertainty. Two size pairs do not justify an unconstrained infinite-volume extrapolation. An interval for a finite-size crossing is not automatically an interval for the bulk critical boundary. Exclusion of the exact bulk value is an accuracy warning, but by itself is not a demonstration of failed coverage for the finite-size estimand.

## 4 Experiment I: the isotropic benchmark

The first campaign uses $\kappa_x = \kappa_\tau = \kappa$ on $16^2$, $32^2$ and $64^2$ tori. The initial scan is $0.36, 0.37, \ldots, 0.52$, followed by predeclared refinement at spacing 0.0025 within observed crossing brackets. A separate timing pilot precedes the production freeze; its records are not pooled with production. Each production chain discards 8,192 attempts, runs 65,536 attempts and retains every fourth, yielding 16,384 joint records. There are 60 cells, 240 chains, 3,932,160 records and 4,080 audited snapshots.

| Size pair | Crossing | Approximate statistical 95% interval | Exact value |
|---|---|---|---|
| 16/32 | 0.4407406 | [0.4401454, 0.4411806] | 0.4406868 |
| 32/64 | 0.4411288 | [0.4409127, 0.4413172] | 0.4406868 |

The primary crossing is 0.1003% above the exact answer, which lies outside its statistical-only interval. Its sampled bracket $[0.4400, 0.4425]$ contains the exact value. Refinement moved the largest-pair estimate by $-0.0021306$, much more than its approximately 0.00010 bootstrap standard error. Thus the study finds the transition region but does not establish bulk accuracy at its quoted statistical precision. The $+0.0003882$ size-pair shift is a diagnostic, not a systematic-error bound. All 2,000 primary bootstrap roots are unique; block-length sensitivity is preserved in the original result tables.

Two far-disordered $64^2$ cells, $\kappa = 0.36$ and 0.37, miss the ESS threshold (approximately 836 and 882); all crossing-bracket cells meet the targets. No records are removed or selectively extended. The sampling allocation used at most 2,489.422 seconds and about 302 MiB, within its 7,200-second/1-GiB caps. Package 0.3.0, its freeze, seeds, raw runs and all Wilson analyses remain preserved. The earlier report was archived before this combined narrative replaced it.

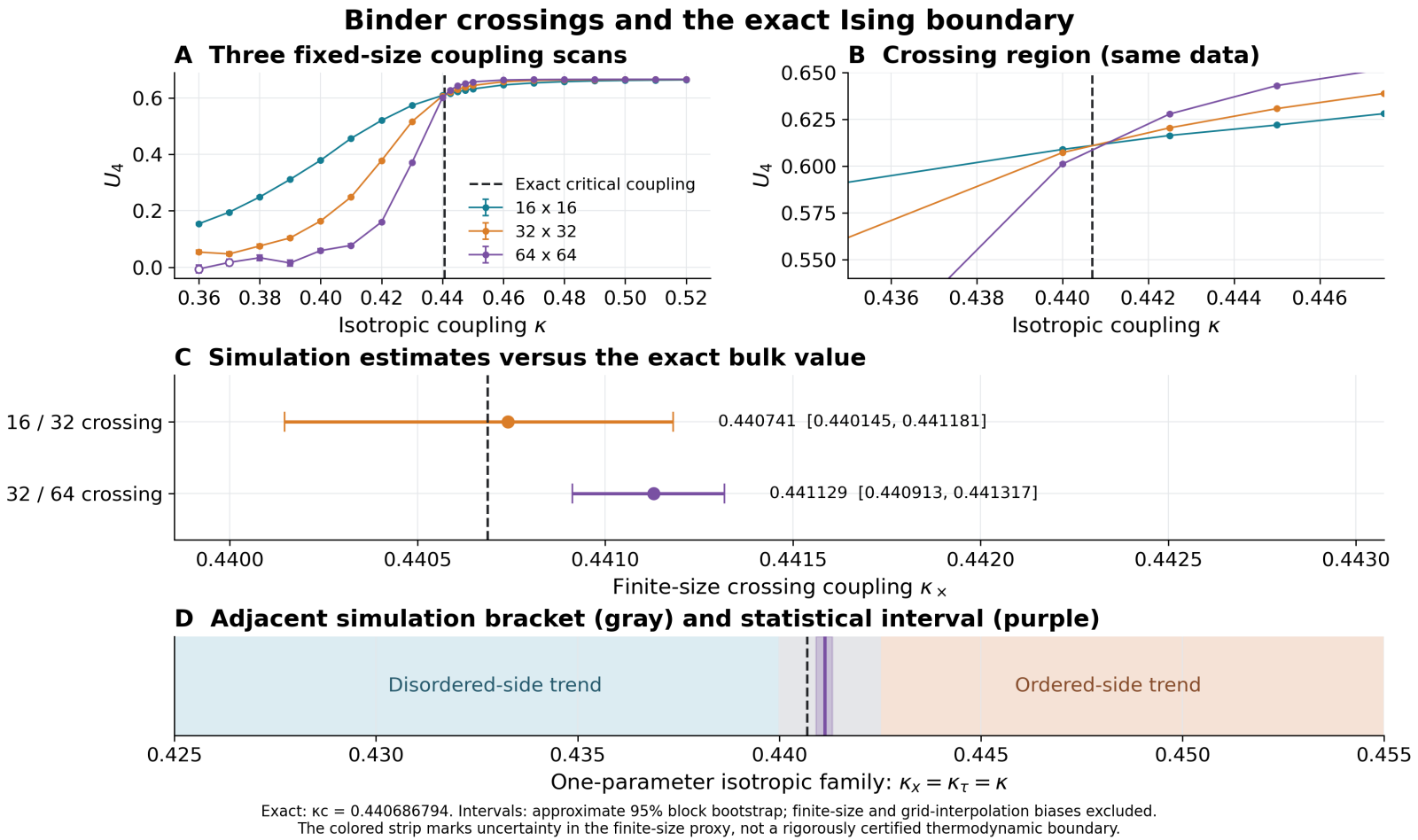



Figure 1: The completed one-parameter benchmark. Statistical precision is much smaller than the observed grid displacement; proximity to the exact answer does not remove that qualification. This figure retains the original result.

# 5 Experiment II: the anisotropic continuation

## 5.1 A frozen two-parameter scan

The continuation uses package 0.3.1 and the same inverse map (6). The spatial cuts are

$$\kappa_x \in \{0.24, 0.30, 0.36, 0.44, 0.54, 0.66, 0.80\}. \tag{12}$$

At each cut the coarse grid is $\kappa_\tau = 0.16, 0.24, \ldots, 0.80$, with $n = 16, 32$: 126 cells. Coarse chains discard 2,048 attempts and run 8,192 attempts, retaining every fourth. The coarse stage locates brackets and is not pooled into refined estimates.

The frozen bracket rule selects an adjacent positive-to-negative sign change of $D_{16}$ whose mean of the four endpoint Binder values lies in $[0.25, 0.65]$. This broad filter rejects nearly Gaussian or fully saturated noise crossings; it does not prescribe the critical Binder value. Exactly one eligible bracket per spatial cut is required. All other sign changes and unresolved cuts are retained. The entire selected interval, including both endpoints, is independently sampled at spacing 0.005 and at all three sizes. Each refined chain discards 8,192 attempts, runs 32,768 attempts and retains every fourth: 8,192 records per chain. Nothing is selected by distance to the exact curve.

The coarse, refined and bootstrap roots are 2028110110, 2028110111 and 2028110112. Cells use identities comprising the root, integer spatial and temporal coupling indices, and size. All inputs were frozen before the corresponding stage's first draw. The refinement configuration is derived from the recorded coarse results by the stated rule and then frozen before refined sampling.

Three independent parameter cells run concurrently, each with one sampling worker and one BLAS thread. Each cell is bounded by 900 seconds, 512 MiB process-tree memory and 16 MiB output. The combined sampling allocation is 10,800 seconds and 2 GiB; batches reserve output space and inherit the remaining global time. Analysis is outside that sampling-time allocation. These jobs use multiple cores on the local eight-core/sixteen-thread Ryzen 7 9700X without modifying the public sampler or assuming linear speedup.

## 5.2 A classifier learned from Binder crossings

For resolved primary crossings $(x_i, \widehat{\tau}_i)$, define

$$\log \widehat{\tau}(x) = (1 - w) \log \widehat{\tau}_i + w \log \widehat{\tau}_{i+1}, \qquad w = \frac{\log x - \log x_i}{\log x_{i+1} - \log x_i}. \tag{13}$$

The classifier predicts the ordered side if $\kappa_\tau > \widehat{\tau}(\kappa_x)$ and the disordered side if it is smaller. It is a transparent supervised boundary classifier: its training information is the simulated size-crossing location and the observed direction of Binder flow. There is no analytic training label, exact-curve fitting feature, or optimization of a classifier against the known answer. Log coordinates ensure positivity and interpolate across the multiplicative coupling scale. No extrapolation or interpolation across a missing spatial column is allowed.

The statistical strip interpolates marginal crossing intervals in the same coordinates. It signals uncertainty in the classification boundary but is neither a simultaneous confidence band nor a total-error bound. Predictions near it must remain tentative. Interior leave-one-column-out interpolation compares a held-out simulated crossing with its neighbors; it checks the interpolation assumption without supplying exact labels. The classifier is a finite-size descriptive tool, not a probability model or a proof of a thermodynamic phase assignment.

### 5.3 Results, diagnostics and comparison with the exact boundary

| $\kappa_x$ | Exact $\kappa_\tau$ | 32/64 crossing | Statistical 95% interval | Error (%) | Unique |
|---|---|---|---|---|---|
| 0.24 | 0.723031 | 0.723476 | [0.722613, 0.724389] | +0.062 | 2000/2000 |
| 0.30 | 0.616679 | 0.616937 | [0.615627, 0.617833] | +0.042 | 2000/2000 |
| 0.36 | 0.531795 | 0.531958 | [0.531065, 0.532678] | +0.031 | 2000/2000 |
| 0.44 | 0.441374 | 0.442215 | [0.441493, 0.442795] | +0.191 | 2000/2000 |
| 0.54 | 0.353635 | 0.353914 | [0.353500, 0.354348] | +0.079 | 2000/2000 |
| 0.66 | 0.273776 | 0.274122 | [0.273728, 0.274555] | +0.126 | 2000/2000 |
| 0.80 | 0.204709 | 0.204756 | [0.204402, 0.205336] | +0.023 | 2000/2000 |

The primary rule resolves 7 of 7 spatial cuts. At those cuts, the maximum absolute relative displacement from the exact boundary is 0.191%, with root-mean-square displacement 0.097%. The exact value lies in 6 of the 7 approximate statistical intervals. These are comparisons at the sampled spatial cuts, not a confidence-coverage experiment.

Across the 1,001-point uniform spatial-coupling comparison grid wherever the classifier is defined, its maximum absolute relative boundary error is 0.546% and its root-mean-square relative error is 0.270%. Interpolation between spatial cuts contributes to this discrepancy. The same analytic expression is used only to evaluate these post-inference errors.

**Size and block sensitivity.**

| $\kappa_x$ | 16/32 crossing | 32/64 minus 16/32 | Statistical widths, blocks 32/64/128 |
|---|---|---|---|
| 0.24 | 0.722959 | +0.000517 | 0.00178 / 0.00178 / 0.00187 |
| 0.30 | 0.617212 | -0.000275 | 0.00220 / 0.00221 / 0.00209 |
| 0.36 | 0.532819 | -0.000861 | 0.00163 / 0.00161 / 0.00173 |
| 0.44 | 0.441106 | +0.001109 | 0.00133 / 0.00130 / 0.00140 |
| 0.54 | 0.353553 | +0.000362 | 0.00087 / 0.00085 / 0.00085 |
| 0.66 | 0.274714 | -0.000592 | 0.00081 / 0.00083 / 0.00079 |
| 0.80 | 0.205040 | -0.000285 | 0.00094 / 0.00093 / 0.00089 |

A numerical root is not necessarily an accepted estimate: the complete analysis records the central-root, endpoint and bootstrap-uniqueness criteria separately. Size drift and stable block widths do not bound finite-size or interpolation bias.

**Diagnostics, including failures.**

| Size | Cells missing a target | Minimum bulk ESS | Maximum $\widehat{R}$ |
|---|---|---|---|
| 16 | 0/119 | 19879.0 | 1.00035 |
| 32 | 0/119 | 14491.4 | 1.00063 |
| 64 | 0/119 | 11538.5 | 1.00083 |

There are 5 coarse cells missing a target; the short coarse stage is only a bracket locator. In the refined data, 0 endpoint cells in the primary adjacent crossing brackets miss a target. All recorded values remain available; diagnostic failures are not cured by relabeling or dropping histories.

All 42 cut/pair/block analyses satisfy the frozen rule, each with 2,000 unique bootstrap roots and no missing or multiple roots. The three block choices reuse the same Monte Carlo histories; they are sensitivity checks, not independent replication campaigns.

The exact bulk value is outside the primary statistical interval at $\kappa_x \in \{0.44\}$. The comparison does not isolate the cause: sampling fluctuations, finite-size displacement and grid interpolation all need to be distinguished. No interval is promoted to a bulk-error bound.

The largest absolute interior leave-one-column-out interpolation discrepancy is 2.363% relative to the held-out simulated crossing. This measures sensitivity to spatial grid spacing; it is not an independent physical classification error rate.

**Completed execution and accounting.**

| Stage | Cells/chains | Retained records | Audited snapshots | Wall seconds |
|---|---|---|---|---|
| Coarse | 126/504 | 1,032,192 | 1,512 | 229.5 |
| Refined | 357/1428 | 11,698,176 | 12,852 | 5061.5 |

The two sampling stages use 5291.1 seconds in total and 205.5 MiB of saved run files. No cell failed and no history was rerun or extended. These timings exclude postprocessing. The reported peak is per bounded cell; concurrent peaks are not incorrectly presented as a measured aggregate.

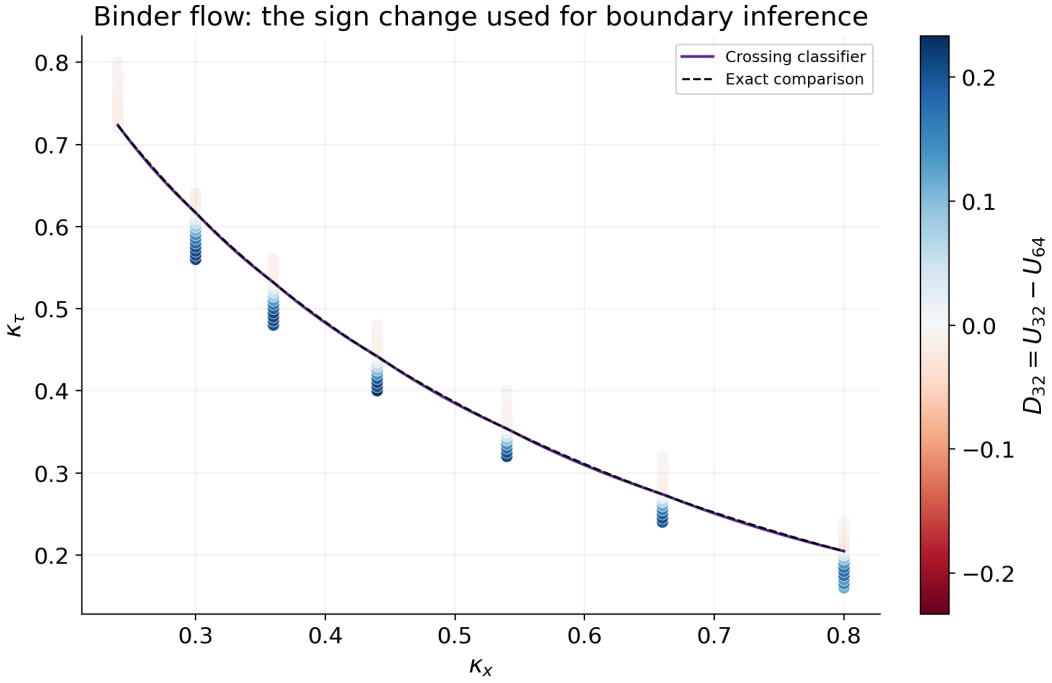



Figure 2: The actual finite-size crossing signal on the refined two-parameter grid. Positive $U_{32} - U_{64}$ is the approach from the disordered side and negative values the approach from the ordered side near the resolved transition. The zero is inferred along each spatial cut; no analytic labels are assigned to the Monte Carlo observations.

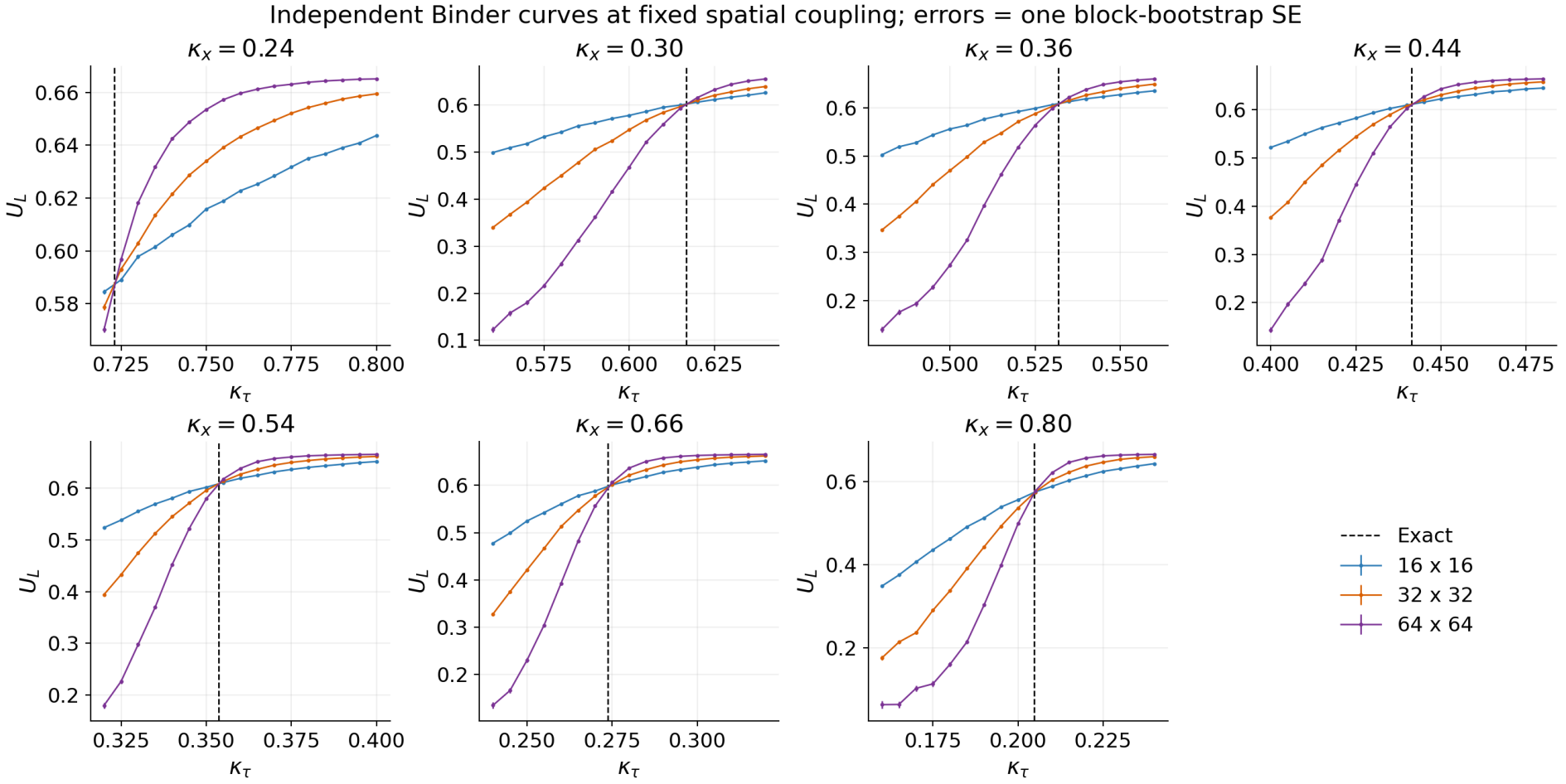



Figure 3: Binder ratios at seven fixed spatial couplings. Points are measured values; errors are one joint block-bootstrap standard error. Vertical dashed lines are the analytic comparison, introduced after inference. The varying crossing height is one reason a universal single-size threshold is unsuitable.

## 6 Reproducibility and scope of the conclusions

The study remains under `simulations/ising-from-code/`. Public package code is in `packages/hcphases/`; tests, engineering checks and temporary outputs remain in `development/`. The theory note is unchanged. The original live report is archived at `../../../../development/provenance/ising-anisotropic-2026-11-01/isotropic-report-before-anisotropic.zip`.

The isotropic data retain their original 0.3.0 implementation identity. The continuation uses the separately accepted 0.3.1 wheel. Its changes reject numerically saturated bond probabilities and avoid an overflowing intermediate normalization expression; the sampled kernel is unchanged in this scan's domain. Old trajectories are audited with the old installed wheel, new ones with the new wheel. Historical freezes are never rewritten to match later code.

Scientific inputs are in `../../configs/`; immutable runs in `../../runs/`; analysis and figures in `../../analysis/`. The anisotropic protocol, stage freezes, derived refinement schedule, per-cell receipts and complete moment histories specify how to reproduce the calculation. The recorded standalone analysis can be rerun into a fresh output directory. No simulation is silently rerun, imputed, or lengthened until it matches the known answer. All unresolved outcomes and numerical qualifications remain part of the evidence. The current engineering audit passes and all 1,116 regression tests pass without exclusions. G1 passes; the aggregate scientific gate deliberately returns its G7-unavailable status. The package runtime is unchanged. Verification receipts and the original report archive are under `../../../../development/provenance/ising-anisotropic-2026-11-01/`.

This benchmark establishes numerical performance only in the supported binary, positive-coupling, periodic vertex model. It does not validate four-dimensional background mixing, general-$N$ sampling, twisted-boundary ratios or nonlinear Wilson confidence coverage. Full G7/M3 remain open and G3 retains its qualification. The next accuracy improvements should address both coupling-grid resolution and increasing system size under a specified finite-size correction model; longer histories alone cannot remove either bias. The plots below distinguish a useful recovered phase diagram from claims of exact thermodynamic accuracy.

## References


[1] N. Ishtiaque and S. Chakroborty, *An exact spacetime polymer gas for finite-temperature $\mathbb{Z}_N$ homological quantum code*, working draft under `../../../../reference/`; conventions and specialization audited in `../../theory/main.tex`.

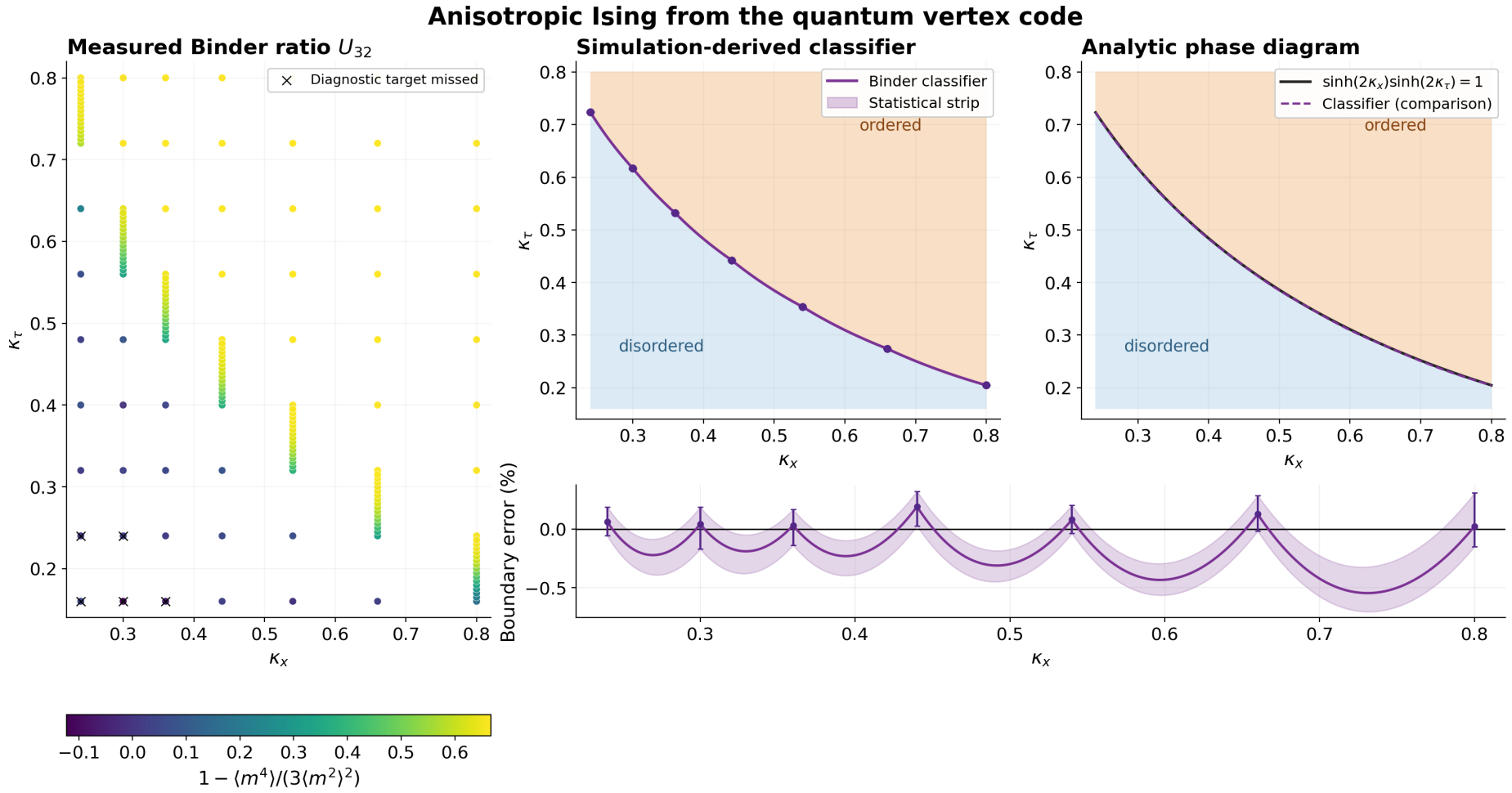
[2] L. Onsager, *Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition*, Phys. Rev. **65**, 117–149 (1944), doi:10.1103/PhysRev.65.117.

[3] K. Binder, *Finite Size Scaling Analysis of Ising Model Block Distribution Functions*, Phys. Rev. Lett. **47**, 693–696 (1981), doi:10.1103/PhysRevLett.47.693.

[4] U. Wolff, *Collective Monte Carlo Updating for Spin Systems*, Phys. Rev. Lett. **62**, 361–364 (1989), doi:10.1103/PhysRevLett.62.361.

[5] W. Selke and L. N. Shchur, *Critical Binder cumulant in two-dimensional anisotropic Ising models*, J. Phys. A **38**, L739–L744 (2005), arXiv:cond-mat/0509369.

## The combined result at a glance




The left panel shows measured Binder values over the scanned two-coupling plane; crosses flag diagnostic targets missed. Negative short-run coarse estimates remain visible. At duplicate grid coordinates the refined estimate is shown, without pooling histories. The two upper-right panels compare the simulation-trained classifier with the analytic phase diagram on identical axes. The residual plot expands their separation. Shading around the learned boundary represents statistical uncertainty only; finite-size and grid errors are not included.