Spontaneous magnetization and thermal decay are undecidable at positive temperature

Seth Douglas Nidhal Mghirbi

Version 1.0.0, November 2026

Abstract

Spontaneous magnetization is undecidable at a fixed positive temperature for translation-invariant finite-range classical lattice models with strictly positive rational Boltzmann weights. A computable family on 2304 local states has maximal visible nine-colour Potts magnetization exactly zero or a fixed positive constant, attained in a translation-invariant Gibbs state. Its pressure in a uniform field is uniformly computable, but differentiability at zero field is undecidable with a fixed derivative-jump gap. Rational weights are asserted at zero field.

We also prove undecidability of exponential decay of one specified thermal spin correlation, in every Gibbs state, on 512 local states. Both constructions couple critical visible spins to cellular-automaton replicas. Surjectivity cancels the collision corrections; nonsurjectivity gives strict subcritical bond domination and exponential decay of the specified correlation uniformly over states. The interaction range depends on the input, and the branches are fine-tuned at their critical or coexistence points. The results transfer to diagonal quantum interactions and survive a prescribed output-reset noise family, the same construction with reduced retention. Nonsurjective full-state uniqueness and exponential clustering of arbitrary local observables remain open.

AI systems assisted with this work. The authors take responsibility for its claims; see "How this work was done." This preprint has not been peer reviewed.

Contents

1 Introduction **3**
1.1 Prior work . . . 5
1.2 Proof route . . . 5

2 The local positive-weight model **6**
2.1 The auxiliary bonds . . . 7

3 Collision identities and surjectivity **7**

4 Witness completion and strict stochastic comparison **9**
4.1 An attractive comparator . . . 10
4.2 Covariance interpolation . . . 10

5 The comparison in every Gibbs state **11**
5.1 Uniform conditional contraction of the hidden layer . . . 12
5.2 Conditional signs and marginal bond odds . . . 13
5.3 Wired cores and arbitrary exterior conditions . . . 14

6 The two branches and completion of the magnetic proof **14**

7 Fixing the cellular-automaton alphabet **16**

8 Exact diagonal quantum transfer **18**

9 Noise and effective impossibility corollaries **19**
9.1 A prescribed output-reset noise family . . . 19
9.2 Effective compactness and no total positive decay modulus . . . 20
9.3 A general arithmetic upper bound . . . 21

10 Potts order with a fixed gap and a computable pressure **22**
10.1 The local model and the unchanged ingredients . . . 22
10.2 The changed comparison statements . . . 23
10.3 The two Potts branches . . . 25
10.4 Computable pressure, undecidable differentiability . . . 26

11 Auxiliary uniqueness routes **27**
11.1 What attractiveness does prove for the L-model . . . 27

12 Open problems **29**

13 Discussion **31**

How this work was done **33**

A Reproducibility **33**

1 Introduction

Can a finite description determine whether a positive-temperature lattice model has spontaneous magnetization? We construct a family for which the answer is undecidable even with a fixed gap in the standard Potts order parameter. The temperature is fixed, the local weights are strictly positive rationals, and the maximal magnetization is either zero or the same positive constant for every ordered input. The pressure as a function of a uniform field is computable, yet its differentiability at zero field is undecidable. Theorem E states this result first; its proof reuses the thermal-decay machinery developed for Theorem A.

Prominent many-body undecidability results concern low-energy properties. These include ground-state observables of Gu, Weedbrook, Perales and Nielsen [27], the spectral gap of Cubitt, Pérez-García and Wolf [13] and the phase diagrams of Bausch, Cubitt and Watson [4]. Purcell, Li and Cubitt construct a quantum ground-state phase transition at an uncomputable parameter [38]. Perales-Eceiza and collaborators ask, "given a translational invariant Hamiltonian", whether exponential spatial decay of the associated thermal correlations is decidable [37, Section 5.1.4, arXiv v2 numbering]. We resolve a specified-observable formulation of this question: exponential decay of a distinguished visible spin correlation is undecidable; the formulation asking for exponential clustering of all local observables remains open.

The mechanism for both results starts with a cellular automaton whose surjectivity is undecidable. Two independent copies produce collision weights. If the automaton is onto, every nonconstant collision coefficient vanishes and the visible spins remain exactly at the chosen critical point: discontinuous nine-colour Potts for magnetization, or anisotropic Ising for decay. Otherwise a finite positive coefficient changes the bond distribution. A stochastic comparison pushes it strictly into the subcritical regime. Conditional contraction of the hidden variables makes this argument work in every Gibbs state.

An interaction is given by finitely many local weight tables and all their translates. At inverse temperature $\beta = 1$, a table $b_X : S^X \to \mathbb{Q}_{>0}$ represents the energy $\Phi_X = -\log b_X$. Write $\mathcal{G}(\Phi)$ for all DLR Gibbs states. The first coordinate is visible: a Potts colour in Theorem E, and an Ising spin in Theorem A. Distances below use the ℓ^1 norm.

Theorem E (spontaneous magnetization with a fixed gap). *There is a computable family of translation-invariant finite-range classical interactions Ψ_t on $\mathbb{Z}^2$, with strictly positive rational local weights at $\beta = 1$ and on-site space*

$$S_{\mathrm{P}} = \{1, \dots, 9\} \times (\{0,1\}^2)^4, \qquad |S_{\mathrm{P}}| = 2304,$$

for which existence of spontaneous visible magnetization is undecidable with a fixed gap. More precisely, let

$$X_x = (9\mathbf{1}_{\{\sigma_x = 1\}} - 1)/8, \qquad m_9 = \phi^1_{3/4,9}(0 \longleftrightarrow \infty) > 0.$$

The index t specifies a CA F_t from a surjectivity-undecidable family.

- *(i) If F_t is surjective, a translation-invariant Gibbs state has $\mu(X_0) = m_9$, and $\sup_{\mu \in \mathcal{G}(\Psi_t)} \mu(X_0) = m_9$.*
- *(ii) If F_t is nonsurjective, every Gibbs state has only finite clusters in its canonical visible bond extension, with independent uniform cluster colours. Every state is invariant under visible S_9 permutations and has $\mu(X_0) = 0$. The excess equal-colour correlation decays exponentially, uniformly over states and positions.*

Thus $\sup_{\mu \in \mathcal{G}(\Psi_t)} \mu(X_0) \in \{0, m_9\}$ exactly, also with the supremum restricted to translation-invariant states. Any fixed rational $m_* \in (0, m_9]$ supplies a promise gap. The interaction range depends on the input.

Corollary 10.7 proves the pressure assertion: the one-sided derivatives are exactly m_9 and $-m_9/8$ in the surjective branch, and both zero otherwise. The derivative jump is therefore

$9m_9/8$ or zero. The field weights need not be rational away from zero. The ordered branch is fine-tuned at coexistence; no arbitrary-perturbation robustness is asserted.

For the separate thermal-decay result, the visible coordinate is $\sigma_x \in \{-1, +1\}$.

Theorem A (thermal spin correlation). *There is a computable family of translation-invariant finite-range interactions Φ_t on*

$$S_* = \{-1, +1\} \times (\{0, 1\}^2)^4, \qquad |S_*| = 512,$$

with strictly positive rational local Boltzmann weights at $\beta = 1$, for which it is undecidable whether

$$\forall \mu \in \mathcal{G}(\Phi_t) \quad \exists C_\mu < \infty,\ a_\mu > 0 \quad \forall x \in \mathbb{Z}^2 : \quad |\mu(\sigma_0 \sigma_x)| \le C_\mu e^{-a_\mu \|x\|_1}. \tag{1.1}$$

Every Gibbs state has zero visible magnetization. The family is constructed from a surjectivity-undecidable family of cellular automata, with this dichotomy:

(i) For a nonsurjective input, there are $C_t < \infty$ and $a_t > 0$ such that

$$\forall \mu \in \mathcal{G}(\Phi_t)\ \forall x, y \in \mathbb{Z}^2 : \quad 0 \le \mu(\sigma_x \sigma_y) \le C_t e^{-a_t \|x-y\|_1}. \tag{1.2}$$

(ii) For a surjective input, the full Gibbs state is unique. Its visible marginal is the critical zero-field anisotropic Ising state with

$$e^{2K_H} = 3, \qquad e^{2K_V} = 2.$$

On this family, (1.1) is equivalent to the version with constants uniform over states and positions, and to either version restricted to translation-invariant Gibbs states. The interaction range depends on the input.

The indexing is $t \mapsto F_t \mapsto \widetilde{G}_t \mapsto \Phi_{\widetilde{G}_t} = \Phi_t$: Section 7 encodes each hard CA F_t as a four-symbol CA $\widetilde{G}_t$, to which Section 2 applies.

The identified critical Ising correlation has no exponential envelope, by infinite critical susceptibility [16, Proposition 2.2]. Zero magnetization makes the displayed correlations equal to connected correlations. We first construct the interaction for an arbitrary CA alphabet A, on $\{-1, +1\} \times A^4$. Section 7 reduces the CA alphabet to four symbols.

Theorem B (diagonal quantum interaction). *Theorem A holds for the corresponding diagonal interactions on the full quasilocal matrix algebra with on-site space $\mathbb{C}^{512}$, with Gibbs states replaced by $\beta = 1$ KMS states and σ_x replaced by $Z_x = \sigma_x^z \otimes I_{256}$.*

The quantum consequence uses the standard diagonal Gibbs–KMS correspondence, recalled with a proof in Section 8. The Hamiltonian entries are logarithms of rational numbers, as prescribed by the weight-table input convention.

Theorem C (output-reset noise). *For every fixed rational $\eta \in (0, 1]$, independently resetting each CA output in each replica to a uniform symbol with probability $1 - \eta$, before averaging the collision weights, preserves Theorems A and B. This is exactly the base construction with $\varepsilon \mapsto \eta\varepsilon$; arbitrary-perturbation robustness is not asserted.*

Theorem D (no total decay modulus). *There is no total algorithm that returns positive rational C_t, a_t on every input of the constructed family and guarantees (1.2) on every nonsurjective input. The same holds for effectively described positive computable-real outputs.*

The last statement concerns a procedure required to halt on both branches. A search for a finite collision witness does halt on each nonsurjective input. Section 9 proves both statements.

1.1 Prior work

Undecidability of PCA ergodicity, due to Kurdyumov and developed by Toom (2000) and Bušić–Mairesse–Marcovici (2013) [42, 8], concerns the long-time dynamics of cellular automata, rather than equilibrium phase transitions of the spatial Gibbs models considered here.

Kari proved undecidability of two-dimensional CA surjectivity [31]. The balance equivalence originates with Hedlund in one dimension [29] and Maruoka–Kimura in higher dimensions [36]. The connection between surjectivity, uniform Bernoulli measure and equilibrium measures is developed by Kari and Taati [32]. The correspondence between probabilistic cellular automata and equilibrium systems goes back to Goldstein, Kuik, Lebowitz and Maes [25] and Lebowitz, Maes and Speer [34]. Their space-time representation is useful context; our proof works directly with a spatial interaction on the square lattice.

Marcovici and Taati are the closest neighbor [35]. Their Theorem 3 shows that, for $d \geq 2$, PCA with positive rational local rules admitting a stationary uniform Bernoulli measure and those admitting no stationary Bernoulli measure are computably inseparable; the reduction from Kari perturbs a CA by uniform-replacement noise ($\kappa = 1/2$ in their notation). Their rapid-forgetting theorem concerns time evolution. We credit that noisy-CA ingredient here; the collision tilt and its comparison with critical and subcritical spatial Ising correlations supply the additional argument.

Chazottes–Hochman [9] and Chazottes–Shinoda [10] construct nonconvergent zero-temperature limits; Gayral, Sablik and Taati study ground-state sets and chaotic zero-temperature limits [23]. Those limiting questions differ from the fixed positive temperature used here. Fawzi, Fawzi and Scalet give certified algorithms for local equilibrium expectations [19]. Computability of a correlation at each fixed distance is compatible with undecidability of an asymptotic decay envelope. Likewise, computability of pressure [23, Proposition 14] does not decide its differentiability. The Potts inputs are critical-point sharpness [5] and discontinuity for more than four colours [17], used separately in Section 10.3.

The Edwards–Sokal coupling [18], Harris and FKG inequalities [28, 21], Holley's comparison principle [30], and Dobrushin's contraction criterion [14, 39] are established tools. So are the Ising critical-point and sharpness inputs cited in Section 6. The new comparison combines a finite collision witness, a positive closed-pattern reward and a covariance interpolation. The reward acts as an essential diminishment, in the sense motivating [2, 6, 3]; our anchor-covariance interpolation replaces pivotal surgery. It does not require positive association of the actual collision bond measure.

To our knowledge, Theorem E is the first fixed-positive-temperature undecidability result for spontaneous magnetization with strictly positive rational local weights and a fixed order-parameter gap. Theorem A is, to our knowledge, the first undecidability result for exponential decay of a specified thermal spin correlation at a fixed positive temperature with the state quantifier ranging over every Gibbs state. These provisional priority statements rest on the review, the cited neighboring constructions and targeted literature searches.

1.2 Proof route

The load-bearing chain is

$$3.1 \longrightarrow 4.1 \longrightarrow 4.2 \longrightarrow 4.4 \longrightarrow 5.2/5.4 \longrightarrow 5.5 \longrightarrow 6.1/6.3 \longrightarrow 7.2.$$

Lemma 4.3 supplies the interpolation step, and Lemmas 5.1 and 5.3 supply the all-state conditional laws. Quantum transfer and the corollaries follow in Sections 8–9. Section 10 transfers the collision and hidden-layer lemmas, states each changed Q-version, and adds the Potts branches and pressure argument. The later arithmetic upper bound is proved

independently; the discussion's algebraic obstruction concerns fixed templates. Neither asserts uniqueness of the nonsurjective collision model.

2 The local positive-weight model

Let $F : A^{\mathbb{Z}^2} \to A^{\mathbb{Z}^2}$ be a cellular automaton, specified by a local rule $f : A^N \to A$, where $N \subset \mathbb{Z}^2$ is finite and nonempty. Put

$$q = |A| \geq 2, \qquad m = |N|, \qquad \varepsilon = \frac{1}{100q(m+1)^2}. \tag{2.1}$$

At each site place a visible spin $\sigma_i \in \{-1, +1\}$ and four hidden variables

$$X_i^H,\ X_i'^H,\ X_i^V,\ X_i'^V \in A. \tag{2.2}$$

The two orientations use distinct replica pairs. Set $e_H = (1, 0)$, $e_V = (0, 1)$, and

$$p_H = \frac{2}{3}, \qquad p_V = \frac{1}{2}, \qquad h(a, b) = q\mathbf{1}_{\{a=b\}} - 1. \tag{2.3}$$

For $r \in \{H, V\}$, define

$$w_i^r = 1 + \varepsilon^2 h\big(F(X^r)_i, F(X'^r)_i\big). \tag{2.4}$$

Equivalently, if $\widetilde{Y}_i, \widetilde{Y}_i'$ are the two deterministic outputs independently passed through uniform-replacement channels of retention ε, then

$$w_i^r = q\ \Pr[\widetilde{Y}_i = \widetilde{Y}_i'].$$

Two independent channels explain ε^2; uniform replacement makes the collision probability strictly positive. Composing each channel with retention η gives $\varepsilon \mapsto \eta\varepsilon$, precisely Theorem C, and locates the noisy-CA ingredient of [35]. Now set

$$B_{i,r} = p_r \mathbf{1}_{\{\sigma_i = \sigma_{i+e_r}\}} + (1 - p_r) w_i^r, \tag{2.5}$$

$$\Phi_{i,r} = -\log B_{i,r}. \tag{2.6}$$

These are finite-support functions, translated identically at every lattice site.

Lemma 2.1 (positivity and effective description)**.** *The interaction Φ_F defined above is translation invariant and finite range. Every local Boltzmann weight is a strictly positive rational. Moreover,*

$$w_{\min} = 1 - \varepsilon^2 \leq w_i^r \leq 1 + (q-1)\varepsilon^2 = w_{\max}, \tag{2.7}$$

and $\mathcal{G}(\Phi_F)$ is nonempty.

Proof. The kernel h takes the integer values -1 and $q - 1$, and $0 < \varepsilon < 1$. Thus $w_i^r > 0$, and $B_{i,r} \geq (1-p_r)w_{\min} > 0$. All entries are rational. The support is contained in $\{i, i+e_r\} \cup (i+N)$, with the relevant hidden components at $i + N$. The finite rule table, the two rational probabilities, and ε determine every entry by finite arithmetic. Existence of a Gibbs state follows by taking a weakly convergent subsequence of finite-volume Gibbs measures; the configuration space is compact and the finite-range, positive specification is continuous. □

With counting a-priori measure on every copy of $S = \{-1, +1\} \times A^4$, the full finite-volume Hamiltonian is

$$H_\Lambda(\xi \mid \bar{\xi}) = - \sum_{\substack{i \in \mathbb{Z}^2,\ r \in \{H,V\} \\ (\{i, i+e_r\} \cup (i+N)) \cap \Lambda \neq \varnothing}} \log\Big[p_r \mathbf{1}_{\{\sigma_i = \sigma_{i+e_r}\}} + (1 - p_r) w_i^r\Big].$$

Each factor uses ξ on Λ and $\bar{\xi}$ outside. There are no hidden one-site energy terms. Normalizing counting measure to uniform measure only changes normalization constants.

2.1 The auxiliary bonds

Introduce a bond $\omega_{i,r} \in \{0,1\}$ for each nearest-neighbor edge $e = (i, i+e_r)$. Its joint local factor is

$$\widetilde{B}_e = \begin{cases} p_r \mathbf{1}_{\{\sigma_i = \sigma_{i+e_r}\}}, & \omega_e = 1, \\ (1-p_r) w_i^r, & \omega_e = 0. \end{cases} \tag{2.8}$$

For a finite set of affected factors $\mathcal{E}$, this defines the joint conditional weight

$$\widetilde{\gamma}(\sigma, \mathbf{X}, \omega \mid \text{exterior}) \;\propto\; \prod_{e \in \mathcal{E}} \big[p_e \mathbf{1}_{\{\sigma_u = \sigma_v\}}\big]^{\omega_e} \big[(1-p_e) w_e\big]^{1-\omega_e}.$$

An exponent zero means that the corresponding factor is omitted. Counting measure is used also for bonds. The canonical extension $\widetilde{\mu}$ below is defined by these weights. For $r = H, V$, put

$$C_r(\omega) = \{i \in \mathbb{Z}^2 : \omega_{i,r} = 0\}.$$

The center of $(i, i+e_r)$ means its lower-left endpoint i; this identifies horizontal witness centers with horizontal edges.

Summing over ω_e gives $B_{i,r}$. Given the original variables, these auxiliary bonds are conditionally independent with the probabilities prescribed by (2.8)/$B_{i,r}$.

Every original Gibbs state has a canonical extension $\widetilde{\mu}$. For a finite collection of site and bond coordinates, condition on all exterior coordinates. Its conditional density is the product of the factors (2.8) that meet the resampled coordinates, with exterior values fixed. Factors disjoint from them cancel. Indeed, the original DLR density is multiplied by the conditional bond kernels; their denominators B_e cancel the original factors on every affected edge. This is the joint DLR specification. Its site projection is the original specification. Open bonds impose visible equality; closed bonds have no visible-spin dependence.

Symbol	Meaning
q, m, N	CA alphabet size, neighborhood size, neighborhood
$r \in \{H, V\},\ \eta$	orientation; output-reset retention
$h(a,b),\ w_i^r$	collision kernel; positive local collision weight
$c_T,\ z(T)$	collision coefficient; $\varepsilon^{2\lvert T\rvert} c_T$
$C_r(\omega),\ K(\omega)$	closed centers; FK cluster count
$P,\ D(P),\ g$	witness; overlap neighborhood; comparator reward
$\kappa,\ \Delta$	$3^{-(\lvert P\rvert-1)}$; strict comparison displacement
$\lambda^C,\ \mathsf{D}$	conditional hidden measure; Dobrushin matrix
$\rho,\ R_\infty(C,e)$	contraction bound; hidden insertion ratio

3 Collision identities and surjectivity

Let U, U' be independent uniform Bernoulli configurations on $A^{\mathbb{Z}^2}$, and write $Y = F(U)$, $Y' = F(U')$. For finite $T \subset \mathbb{Z}^2$, set

$$c_T = \mathbb{E} \prod_{i \in T} h(Y_i, Y_i'), \qquad c_\varnothing = 1, \qquad z(T) = \varepsilon^{2|T|} c_T. \tag{3.1}$$

Remark (The balance equivalence [29, 36, 32]). A CA on $A^{\mathbb{Z}^2}$ is surjective if and only if it preserves uniform Bernoulli measure.

Proof. Let u be uniform Bernoulli measure and $\nu = F_* u$. Assume first that F is onto. Fix a finite support P and a pattern $a \in A^P$. In a large square Λ, choose translates $P+x_1, \dots, P+x_k$ whose input supports $P+x_j+N$ are pairwise disjoint and contained in $\Lambda+N$, with $P+x_j \subset \Lambda$.

Choosing their centers on a sufficiently widely spaced fixed grid gives $k \geq c|\Lambda| + o(|\Lambda|)$, where $c > 0$ depends only on P, N.

There are $q^{|\Lambda|-k|P|}$ output patterns on Λ that equal the selected translate of a on every $P + x_j$. Surjectivity gives a preimage for every such pattern, and different output patterns require different input patterns on $\Lambda + N$. The number of input patterns satisfying the k prescribed output constraints is exactly

$$q^{|\Lambda+N|}\nu([a])^k, \tag{3.2}$$

because the input supports are disjoint. Hence

$$\left(q^{|P|}\nu([a])\right)^k \geq q^{-(|\Lambda+N|-|\Lambda|)}. \tag{3.3}$$

The exponent on the right is a boundary term. Taking logarithms, dividing by k, and letting Λ grow gives $\nu([a]) \geq q^{-|P|}$. Summing over all patterns a gives equality for each one. Thus $\nu = u$.

Conversely, $F(A^{\mathbb{Z}^2})$ is compact and therefore closed. If $F_*u = u$, this image contains the support of u, which is the whole full shift. Thus F is onto. □

Lemma 3.1 (positivity, factorization, and detection)**.** *For every finite T,*

$$0 \leq c_T \leq (q-1)^{|T|}. \tag{3.4}$$

Every c_T is an effectively computable rational, with denominator dividing $q^{2|T+N|}$. In particular,

$$c_T > 0 \implies c_T \geq q^{-2m|T|}. \tag{3.5}$$

All nonempty c_T vanish if and only if F is surjective.

Let Γ be the graph on output centers in which i, j are adjacent when $(i+N) \cap (j+N) \neq \varnothing$. Its maximum degree plus one is at most m^2. If T_1, T_2 have no Γ-edge between them and are disjoint, then

$$c_{T_1 \cup T_2} = c_{T_1} c_{T_2}. \tag{3.6}$$

Consequently, nonsurjectivity has a finite nonempty connected witness P with $c_P > 0$.

Proof. Choose a real orthonormal basis $1, \chi_1, \ldots, \chi_{q-1}$ for functions on A, with respect to uniform measure. The completeness identity is

$$h(a,b) = \sum_{\alpha=1}^{q-1} \chi_\alpha(a)\chi_\alpha(b). \tag{3.7}$$

Expanding the product and using independence of the two replicas gives

$$c_T = \sum_{\alpha: T \to \{1,\ldots,q-1\}} \left(\mathbb{E} \prod_{i \in T} \chi_{\alpha(i)}(Y_i) \right)^2. \tag{3.8}$$

This proves nonnegativity. The pointwise bound $|h| \leq q - 1$ proves the upper bound. The direct definition averages an integer-valued function over two independent configurations on the finite input set $T + N$; it gives the claimed rationality, computability, and denominator bound.

If F is onto, the balance equivalence makes Y uniform iid, and every nonempty product in (3.8) has mean zero. Conversely, if every nonempty c_T vanishes, every squared Fourier

coefficient in (3.8) vanishes. Every nonconstant tensor Fourier coefficient of every finite-dimensional distribution of Y occurs in one of these expressions, with T its nonconstant-coordinate support. Thus all those distributions are uniform product distributions, and the balance equivalence applies.

The possible differences $j-i$ for overlapping neighborhoods lie in $N-N$, whose cardinality is at most m^2. Independence of disjoint input sets proves (3.6): the random variables used in the two products depend on disjoint coordinates in both replicas. Factorization over connected components and positivity show that a positive nonempty witness has a positive connected component. □

For one orientation, let C be a finite set of closed-edge centers and integrate its hidden variables against the iid prior:

$$W(C) = \mathbb{E} \prod_{i \in C} w_i. \tag{3.9}$$

Expansion gives the fundamental identity

$$\boxed{W(C) = \sum_{T \subseteq C} z(T).} \tag{3.10}$$

In particular, $W(C) \geq 1$, and W is increasing under addition of closed edges.

Summing visible spins in a finite graph gives the cluster weight $2^{K(\omega)}$. With an iid hidden prior on a sufficiently large padded input region, the integrated bond law is

$$\mu(\omega) \;\propto\; 2^{K(\omega)} \prod_e p_e^{\omega_e} (1-p_e)^{1-\omega_e} W(C_H(\omega)) W(C_V(\omega)). \tag{3.11}$$

Free or wired spin counting uses the corresponding cluster count; constants from a wired boundary cluster cancel. Formula (3.11) uses a free hidden boundary: all input variables in the padded region are summed against the iid prior. Fixed hidden boundaries can change insertion ratios near that boundary. Surjectivity also need not make the CA onto on a finite periodic torus: wrap-around constraints can give $W \neq 1$. Section 5 proves the infinite-volume result separately. Every thermodynamic subsequential limit of these finite-volume measures is covered by that result.

4 Witness completion and strict stochastic comparison

Fix a nonempty witness P, let $k = |P|$, and put

$$D(P) = \{j : (j+N) \cap (P+N) \neq \varnothing\}, \qquad |D(P)| \leq m^2 k, \tag{4.1}$$

$$a_P = \frac{\varepsilon^{2k} c_P}{w_{\max}^{|D(P)|}} > 0, \qquad g = \frac{1}{k} \log(1 + a_P), \tag{4.2}$$

$$\kappa = 3^{-(k-1)}, \qquad \Delta = \kappa g > 0. \tag{4.3}$$

Here $D(P)$ denotes an overlap neighborhood; it is not a Dobrushin matrix.

Lemma 4.1 (witness completion)**.** *Let $i \in P$, $i \notin C$, and $P \setminus \{i\} \subseteq C$. Then*

$$\frac{W(C \cup \{i\})}{W(C)} \geq 1 + a_P. \tag{4.4}$$

Without the completion assumption, this ratio is at least one.

Proof. In the positive expansion of $W(C \cup \{i\}) - W(C)$, retain only the subsets $P \cup S$ with $S \subseteq C \setminus D(P)$. Their input supports are disjoint between P and S, so factorization gives

$$W(C \cup \{i\}) - W(C) \geq z(P) W(C \setminus D(P)). \tag{4.5}$$

Pointwise, at most $|D(P)|$ factors were removed from the defining product for $W(C)$, each at most $w_{\max}$. Therefore

$$W(C) \leq w_{\max}^{|D(P)|} W(C \setminus D(P)). \tag{4.6}$$

Dividing proves the claim. The unconditional lower bound follows directly from nonnegative coefficients in (3.10). □

4.1 An attractive comparator

Use witness translates only on horizontal edges. In a finite edge set E, let $\mathcal{T}_E$ contain translates $P + t$ all of whose horizontal edges belong to E. Define

$$V_E(C) = \exp\left(g \sum_{t \in \mathcal{T}_E} \mathbf{1}_{\{P+t \subseteq C\}} \right). \tag{4.7}$$

For a horizontal edge $e \notin C$, write $n_e(C)$ for the number of retained witness translates completed by adding e. Since a translate containing e is determined by its member of P mapped to e, $n_e(C) \leq k$. Lemma 4.1 consequently gives

$$\frac{W(C \cup \{e\})}{W(C)} \geq e^{g n_e(C)}. \tag{4.8}$$

Indeed, if $n_e = 0$ use the unconditional bound; otherwise $e^{g n_e} \leq e^{gk} = 1 + a_P$.

Lemma 4.2 (FKG comparator)**.** *For any free or wired finite-graph FK boundary condition, let*

$$\nu_g(\omega) \propto 2^{K(\omega)} \prod_e p_e^{\omega_e} (1 - p_e)^{1 - \omega_e} V_E(C_H(\omega)). \tag{4.9}$$

This is an FKG measure. The finite law (3.11) satisfies

$$\mu \preceq \nu_g. \tag{4.10}$$

Proof. The function $K(\omega)$ is supermodular in the open-edge set, by submodularity of the graphic-matroid rank. Each all-closed indicator is also supermodular: if I denotes absence of any open edge from a fixed set, then $I(\omega \vee \omega') + I(\omega \wedge \omega') \geq I(\omega) + I(\omega')$. Thus the density in (4.9) is log-supermodular. The FKG inequality [21] applies.

By (4.8), $W(C_H)/V_E(C_H)$ is increasing in closed edges. So is $W(C_V)$. The density of μ relative to ν_g is therefore decreasing in open edges. For increasing f, positive association implies that its covariance with this decreasing density is nonpositive. Normalizing gives $\mu(f) \leq \nu_g(f)$. This is also a special case of Holley's comparison principle [30]; see also [26, Theorems 2.19 and 2.27]. □

4.2 Covariance interpolation

Fix one member $p_* \in P$ as the anchor. Each retained translate $P + t$ has anchor e_t, the horizontal edge centered at $p_* + t$. Distinct translates have distinct anchors. Let $\mathcal{A}_E$ be this anchor set. With log-odds $\theta_e = \log(p_e/(1 - p_e))$, consider

$$\nu_{\theta,s}(\omega) \propto 2^{K(\omega)} \exp\left(\sum_e \theta_e \omega_e + s \sum_{t \in \mathcal{T}_E} I_t(\omega) \right), \qquad I_t = \mathbf{1}_{\{P+t \text{ closed}\}}. \tag{4.11}$$

It is FKG for $s \geq 0$. Assume every horizontal $\theta_e \leq \log 2$.

Lemma 4.3 (the anchor covariance inequality)**.** *For every increasing f and every retained translate,*

$$\mathrm{Cov}_{\nu_{\theta,s}}(f, I_t) \le -3^{-(k-1)} \mathrm{Cov}_{\nu_{\theta,s}}(f, \omega_{e_t}). \tag{4.12}$$

Proof. The event $I_t = 1$ implies $\omega_{e_t} = 0$. Conditional FKG monotonicity gives

$$\nu(f \mid I_t = 1) \le \nu(f \mid \omega_{e_t} = 0). \tag{4.13}$$

Subtracting $\nu(f)$ and multiplying by $\nu(I_t)$ yields

$$\mathrm{Cov}_{\nu}(f, I_t) \le -\frac{\nu(I_t)}{\nu(\omega_{e_t} = 0)} \mathrm{Cov}_{\nu}(f, \omega_{e_t}). \tag{4.14}$$

No pivotal modification is involved.

The conditional open odds of a horizontal edge are at most $e^{\theta_e} \le 2$: its FK cluster factor is at most one, and opening an edge can only remove all-closed rewards. Its conditional closed probability is therefore at least $1/3$. This remains true when only some other edges are conditioned, by averaging. Sequentially closing the other $k-1$ edges gives

$$\frac{\nu(I_t)}{\nu(\omega_{e_t} = 0)} = \nu(I_t \mid \omega_{e_t} = 0) \ge 3^{-(k-1)}. \tag{4.15}$$

The remaining covariance is nonnegative by FKG, proving the displayed inequality. □

Proposition 4.4 (a strictly subcritical ordinary FK upper bound)**.** *The measure ν_g is dominated by the ordinary FK-Ising measure with horizontal log-odds $\log 2 - \Delta$ on $\mathcal{A}_E$, critical log-odds on the other horizontal edges, and vertical log-odds zero, with the same FK boundary condition.*

Proof. Use the path $0 \le s \le g$ with

$$\theta_e(s) = \begin{cases} \log 2 - \kappa g + \kappa s, & e \in \mathcal{A}_E, \\ \log 2, & e \text{ horizontal but not in } \mathcal{A}_E, \\ 0, & e \text{ vertical.} \end{cases} \tag{4.16}$$

Finite-volume differentiation and Lemma 4.3 give

$$\frac{d}{ds}\nu_{\theta(s),s}(f) = \kappa \sum_{e \in \mathcal{A}_E} \mathrm{Cov}(f, \omega_e) + \sum_{t \in \mathcal{T}_E} \mathrm{Cov}(f, I_t) \le 0. \tag{4.17}$$

The endpoint $s = g$ is ν_g, and $s = 0$ is the asserted ordinary FK measure. □

In the infinite-volume interior, this comparison has the homogeneous parameters

$$v'_H = e^{\theta'_H} = 2e^{-\Delta}, \qquad v'_V = 1, \qquad p'_H = \frac{2e^{-\Delta}}{1 + 2e^{-\Delta}}, \quad p'_V = \frac{1}{2}. \tag{4.18}$$

The square-lattice FK-Ising critical curve is $v_H v_V = 2$, equivalently $\sinh(2K_H)\sinh(2K_V) = 1$. Thus (4.18) is strictly subcritical. The precise handling of the unshifted finite boundary strip is included in the all-state proof below.

5 The comparison in every Gibbs state

This section is needed because integrating an iid hidden prior in finite volume does not by itself control all Gibbs states.

5.1 Uniform conditional contraction of the hidden layer

For one orientation, regard (X_i, X_i') as a single hidden variable with alphabet A^2. For a possibly infinite set C of closed centers, the hidden interaction consists of $-\log w_i$ for $i \in C$. Let

$$\ell_w = \log(w_{\max}/w_{\min}). \tag{5.1}$$

For (2.1),

$$\ell_w \le 2q\varepsilon^2. \tag{5.2}$$

Lemma 5.1 (hidden uniqueness uniformly in the bonds)**.** *For every $C \subseteq \mathbb{Z}^2$, the hidden specification has a unique Gibbs measure λ^C. The Dobrushin row sums are bounded, uniformly in C, by*

$$\rho \le \tfrac{1}{2}\ell_w\, m(m-1) < 1. \tag{5.3}$$

Its local distributions are insensitive, exponentially in the separation, to hidden boundary conditions and to changes of closed centers outside a distant region. The corresponding statement also holds conditional on any visible configuration.

Proof. For distinct hidden sites i, j, let n_{ij} count factors whose input supports contain both. Changing the variable at j changes the single-site log weight at i by a function of oscillation at most $2n_{ij}\ell_w$. The likelihood-ratio estimate [24, Proposition 8.8] gives

$$\mathsf{D}_{ij} \le \tanh(n_{ij}\ell_w/2) \le n_{ij}\ell_w/2.$$

Each site belongs to at most m factors, each involving at most $m-1$ other sites. Thus $\sum_j n_{ij} \le m(m-1)$, proving (5.3).

The estimate applies separately to both independent orientation layers. Given visible spins, the factors are $p_r\delta_{\text{same}} + (1-p_r)w_i^r$. Their logarithmic oscillations are no larger than ℓ_w, so the same estimate holds.

The Dobrushin uniqueness and comparison theorem [14, 39] now applies. One may see exponential locality directly from its influence matrix D: it has row sum at most $\rho < 1$, and each nonzero entry joins sites at separation at most $\operatorname{diam} N$. The resolvent $\sum_{n\ge 0}\mathsf{D}^n$ therefore has exponentially small tails in physical distance, with constants allowed to depend on N. It also controls inhomogeneous changes of factors confined outside a distant region. □

If $e \notin C$, define the infinite-volume insertion ratio

$$R_\infty(C, e) = \lambda^C(w_e). \tag{5.4}$$

For a finite exhaustion E_n, the measure obtained by retaining only factors centered in $C \cap E_n$ has product reference outside their input support. Lemma 5.1 gives, uniformly in the ambient C,

$$\frac{W((C\cap E_n)\cup\{e\})}{W(C\cap E_n)} \longrightarrow R_\infty(C, e). \tag{5.5}$$

Lemma 5.2 (insertion bounds and local reweighting)**.** *For every infinite pattern C, with $e \notin C$,*

$$R_\infty(C, e) \ge 1. \tag{5.6}$$

For the horizontal witness rewards of Section 4,

$$R_\infty(C, e) \ge e^{g n_e(C)} \tag{5.7}$$

when $n_e(C)$ counts any collection of completed witness translates. Moreover,

$$\lambda^{C\cup\{e\}} = \frac{w_e}{R_\infty(C, e)}\,\lambda^C, \qquad \lambda^{C\cup\{e\}}(w_e^{-1}) = R_\infty(C, e)^{-1}. \tag{5.8}$$

If F is onto, $R_\infty(C, e) = 1$ for every C, e.

Proof. The finite insertion inequalities pass to the uniform limit (5.5). Only finitely many translates contain a given edge, so their completions are all seen in a sufficiently large truncation. Multiplication of a Gibbs measure by the positive local factor w_e, followed by normalization, changes its specification by precisely that factor. Uniqueness in Lemma 5.1 identifies the result with $\lambda^{C\cup\{e\}}$, proving (5.8). In the onto case all finite W's equal one, by Lemma 3.1, and so do their insertion ratios. □

5.2 Conditional signs and marginal bond odds

Let $\widetilde{\mu}$ be the canonical bond extension of an arbitrary $\mu \in \mathcal{G}(\Phi_F)$.

Lemma 5.3 (conditional independence and fair finite clusters)**.** *Given (σ, ω), the hidden variables have the unique conditional law $\lambda^{C_H(\omega)} \otimes \lambda^{C_V(\omega)}$. Consequently hidden variables and visible cluster signs are independent conditional on ω. Conditional on the bonds, every finite open cluster has an independent fair sign, also independent of all infinite-cluster signs. No absence-of-infinite-clusters assumption is needed.*

Proof. There are countably many finite-volume DLR equations (test cylinder indicators). A regular conditional distribution therefore satisfies all of their hidden-coordinate versions simultaneously outside one null set. Disintegrating these joint DLR equations gives $\mathcal{L}(\mathbf{X} \mid \sigma, \omega) \in \mathcal{G}(\gamma^{C(\omega)})$ almost surely, where $\gamma^{C(\omega)}$ is the product of the two hidden closed-factor specifications. It is independent of σ; Lemma 5.1 gives its unique law. This proves the conditional independence.

For a fixed finite open cluster, flip every visible spin in that cluster. All open equality constraints are preserved, and closed factors do not depend on visible spins. This finite transformation preserves the joint conditional Gibbs weights. Its two cluster signs are therefore equiprobable, conditional on all remaining information. Applying the argument to any finite collection of finite clusters gives independent fair signs. Conditioning on the event that a specified finite vertex set is a cluster makes this a direct finite-support DLR argument; countably many such events cover all finite clusters. □

Lemma 5.4 (exact marginal opening odds)**.** *Fix an edge $e = \{u, v\}$, condition on all other bonds, and let C denote the other closed edges in its orientation,* excluding *e. Put*

$$\alpha_\mu = \widetilde{\mu}(\sigma_u = \sigma_v \mid \omega_{e^c},\, \omega_e = 0). \tag{5.9}$$

Then, almost surely,

$$\frac{\widetilde{\mu}(\omega_e = 1 \mid \omega_{e^c})}{\widetilde{\mu}(\omega_e = 0 \mid \omega_{e^c})} = \frac{p_e}{1 - p_e} \frac{\alpha_\mu}{R_\infty(C, e)}. \tag{5.10}$$

The closed conditional probability is bounded away from zero.

Proof. Changing a closed bond to open multiplies the local weight by

$$\frac{p_e}{1 - p_e} \frac{\mathbf{1}_{\{\sigma_u = \sigma_v\}}}{w_e}. \tag{5.11}$$

On the closed branch, the hidden law is $\lambda^{C\cup\{e\}}$, not λ^C. Lemmas 5.3 and 5.2 factor its conditional expectation into $\alpha_\mu / R_\infty(C, e)$. This proves the odds formula. Its right-hand side is at most $p_e/(1 - p_e)$, because $\alpha_\mu \le 1$ and $R_\infty \ge 1$. Thus the closed probability is at least $1 - p_e$. □

5.3 Wired cores and arbitrary exterior conditions

Proposition 5.5 (all-state FK domination)**.** *For every $\mu \in \mathcal{G}(\Phi_F)$, its bond marginal is stochastically dominated by the wired critical anisotropic FK-Ising measure ϕ^1_c. If $c_P > 0$, it is dominated by the homogeneous strictly subcritical FK-Ising state with parameters* (4.18)*.*

Proof. Let V_n be a square core, E_n its internal nearest-neighbor edges, and identify all boundary vertices of V_n in the reference FK measure. Condition the target bond law on its exterior bonds.

Fix an internal edge e. If its endpoints are disconnected after deleting e in this *fully wired* core, at most one of their components can meet the core boundary. At least one component is therefore contained in the interior of V_n, and is a finite cluster in the full bond configuration. On the closed branch, Lemma 5.3 implies $\alpha_\mu = 1/2$. If the endpoints are wired-connected, use simply $\alpha_\mu \le 1$. Thus, with connectivity in the fully wired core with e deleted,

$$\alpha_\mu \;\le\; \text{wired cluster factor} = \begin{cases} 1, & u \longleftrightarrow_{\mathrm{w}} v, \\ 1/2, & u \not\longleftrightarrow_{\mathrm{w}} v. \end{cases}$$

Retain only witness translates wholly inside E_n, and let $n_e^{V_n}(C)$ count just those completed by closing e. Name the wired-core comparator $\nu_g^{V_n,\mathrm{w}}$: it is (4.9) on E_n, with complete wiring and only these retained translates. By (5.7), the target's horizontal opening odds are at most the ordinary fully wired FK opening odds multiplied by $e^{-g n_e^{V_n}(C)}$. On vertical edges use $R_\infty \ge 1$. Thus every target single-edge opening probability is at most that of the fully wired comparator $\nu_g^{V_n,\mathrm{w}}$ at the same internal configuration.

For clarity, this comparison does not require target FKG. The comparator's single-edge probabilities are increasing in the other edges, by its lattice condition. Couple single-edge heat-bath updates with the same uniform random variable, starting both chains in the all-closed configuration. The pointwise probability comparison and comparator monotonicity preserve the order. The target has downward-closed support because closing any edge has positive conditional probability. Its finite heat-bath chain is irreducible on that support: one can close to the all-closed state and reverse a closing path to any supported configuration. Passing to stationary distributions proves target domination by $\nu_g^{V_n,\mathrm{w}}$. This is the single-edge Holley coupling criterion with a monotone upper law [26, Theorems 2.3 and 2.27]; the support argument covers the target's possible zeros. It holds for almost every conditioned exterior, so also after averaging it.

Apply Proposition 4.4. It reduces the horizontal odds on the anchors of complete witness translates; it leaves an unshifted strip near the boundary. This strip is not ignored. Since P is fixed, there is a fixed width such that every horizontal edge of a smaller concentric core is an eligible anchor. Condition the resulting ordinary FK law on all edges outside this smaller core. By the domain Markov property its induced boundary partition is dominated by complete wiring [26, Lemmas 4.13–4.14]. Therefore its marginal on any fixed still smaller region is dominated by the ordinary homogeneous subcritical FK law in the smaller core with wired boundary. Let $n \to \infty$. The wired measures decrease to their infinite-volume limit [26, Theorem 4.19]; here that is the unique subcritical FK state. This proves the comparison for local increasing events. Finally $\{x \longleftrightarrow y\}$ is the increasing union of events that an open path inside a finite box joins x to y, so continuity from below gives the same connection bound.

For the critical comparison omit witness rewards entirely and use only $R_\infty \ge 1$ in the same wired-core argument. The upper limit is critical wired FK-Ising. □

6 The two branches and completion of the magnetic proof

We use the following external Ising facts, with their exact roles stated.

(I1) **Critical curve.** Rectangular nearest-neighbor ferromagnetic Ising is critical when

$$\sinh(2K_H)\sinh(2K_V) = 1. \tag{6.1}$$

Equivalently $v_H v_V = 2$, for FK odds $v_r = e^{2K_r} - 1$. The square-lattice specialization of the weighted-planar equation [11] is

$$\tanh K_H + \tanh K_V + \tanh K_H \tanh K_V = 1.$$

For $(e^{2K_H}, e^{2K_V}) = (3, 2)$ this reads $1/2 + 1/3 + 1/6 = 1$, explaining the chosen rational critical pair.

(I2) **Critical continuity.** At a critical nondegenerate biperiodic planar Ising model, the plus-state magnetization is zero [15, Theorem 1.5]. Ising monotonicity implies uniqueness: in the monotone coupling of plus and minus states, every single-site disagreement has probability zero, and any Gibbs state lies between them. The Edwards–Sokal identification gives

$$\phi^1_{\mathrm{c}}(0 \longleftrightarrow \infty) = \langle\sigma_0\rangle^+ = 0$$

[26, Theorems 4.91 and 5.17]. Thus critical wired FK has no infinite cluster; its free and wired limits coincide. Indeed, the free Ising state lies between the equal plus and minus states. Sampling bonds independently on equal-spin edges gives the free FK limit from the free Ising state and the wired FK limit from the plus state, so these bond laws agree [26, Theorem 4.91 and its following free-boundary remark]. The FK boundary-condition sandwich [26, Lemma 4.14, Theorem 4.19] gives uniqueness of the FK state.

(I3) **Sharpness.** Strictly subcritical finite-range Ising two-point functions decay exponentially, and hence so do FK-Ising connectivities [16, Theorem 2.1 and Section 2.2, arXiv v3 numbering]. At criticality the susceptibility is infinite [16, Proposition 2.2]; see also the original correlation-inequality argument [40]. These results allow translation-invariant anisotropic couplings. At the displaced point (4.18), fix its couplings and vary β along their ray. The critical-curve theorem gives $\beta_c > 1$; sharpness applies at $\beta = 1$. The original sharpness result is due to Aizenman, Barsky and Fernández [1].

Proposition 6.1 (nonsurjective branch)**.** *If F is not onto, there are $C_F < \infty$, $a_F > 0$, independent of the Gibbs state, such that every $\mu \in \mathcal{G}(\Phi_F)$ satisfies*

$$\mu(\sigma_x) = 0, \qquad 0 \le \mu(\sigma_x\sigma_y) \le C_F e^{-a_F\|x-y\|_1}. \tag{6.2}$$

Proof. Lemma 3.1 supplies a positive connected witness. Proposition 5.5 dominates every bond marginal by the same strictly subcritical FK-Ising state. It follows that all its clusters are finite, and (I3) bounds its connection probabilities exponentially. Fair independent cluster signs give

$$\mu(\sigma_x) = 0, \qquad \mu(\sigma_x\sigma_y) = \widetilde{\mu}(x \longleftrightarrow y). \tag{6.3}$$

Stochastic domination proves the result with the comparator's constants. □

Corollary 6.2 (the entire visible odd sector)**.** *Let f, g be bounded local functions of visible and hidden variables, each odd under simultaneous reversal of all visible spins. If their supports are S, T, respectively, then in the nonsurjective branch*

$$\begin{aligned} &\mu(f) = \mu(g) = 0, \\ &|\operatorname{Cov}_\mu(f,g)| \le \|f\|_\infty\|g\|_\infty\,\widetilde{\mu}(S \longleftrightarrow T) \\ &\qquad\qquad \le C_F|S||T|\|f\|_\infty\|g\|_\infty e^{-a_F\operatorname{dist}_1(S,T)}. \end{aligned} \tag{6.4}$$

Proof. Condition on the bonds and hidden variables. All clusters have independent fair signs. Flipping the finite union of clusters that meets S reverses f, giving conditional mean zero. If no cluster meets both S and T, the same flip leaves g unchanged, so the conditional product has mean zero. On the remaining event bound it by the product of the norms. A union bound and the two-point FK estimate give the last inequality. □

Proposition 6.3 (surjective branch)**.** *If F is onto, the full Gibbs state is unique. Its visible marginal is exactly critical anisotropic Ising with $e^{2K_H} = 3$ and $e^{2K_V} = 2$. Its magnetic two-point function has no exponential envelope.*

Proof. All insertion ratios equal one. First apply the critical comparison in Proposition 5.5. By (I2), critical wired FK has no infinite cluster, so neither does the target bond law.

With all clusters finite, Lemma 5.3 makes the factor α_μ in (5.10) exactly one when the endpoints are already connected without the tested edge, and exactly $1/2$ otherwise. These are precisely the ordinary FK-Ising conditional probabilities. Equivalently, conditioning on any finite core induces the standard outside-connectivity boundary partition. All finite conditional configurations have positive probabilities, so the single-edge conditionals identify the finite-core FK specification. For every exterior partition ξ and core V_n,

$$\phi^0_{V_n} \preceq \phi^\xi_{V_n} \preceq \phi^1_{V_n}.$$

The middle law is the conditional bond law given the exterior (domain Markov property). Averaging preserves this sandwich. On every fixed local increasing event the outer laws tend to ϕ^0_{c} and ϕ^1_{c} [26, Lemmas 4.13–4.14, Theorem 4.19]. They agree by (I2), so the bond marginal is their common, unique critical FK state.

Given the bonds, the cluster signs are uniquely determined in law by fair independent coloring, and the hidden layer is uniquely $\lambda^{C_H} \otimes \lambda^{C_V}$. This identifies the full extended state and hence the original Gibbs state uniquely. Coloring the critical FK measure gives the ordinary critical Ising marginal. The base odds are 2 and 1, so $e^{2K_H} = 1 + 2 = 3$, $e^{2K_V} = 1 + 1 = 2$. Directly,

$$\sinh(\log 3)\sinh(\log 2) = \frac{4}{3} \cdot \frac{3}{4} = 1. \tag{6.5}$$

Finally, (I3) says that the critical susceptibility $\sum_x \mu(\sigma_0\sigma_x)$ is infinite. An exponential envelope would make this sum finite. Thus the required envelope does not exist. Alternatively, exponential decay would make the Duminil-Copin–Tassion finite-set quantity less than one in a large square; continuity would extend that strict inequality past criticality, contradicting their critical-point characterization. □

These propositions establish the variable-alphabet dichotomy, with zero magnetization in every state. On nonsurjective inputs the constants are uniform over *all* Gibbs states and locations. On surjective inputs the unique state is translation invariant and fails exponential decay. Thus either choice of the state quantifier in Theorem A has the same answer on this family.

7 Fixing the cellular-automaton alphabet

Kari's original surjectivity construction [31, Theorem 2, p. 179] places a tile background and a bit at each site. The background supplies local matching constraints and directed arrows. A local error freezes the affected bit; away from errors the update XORs it with the bit at the neighbor named by the directed tile's arrow (the successor). More precisely, Kari adds a symbol B to the directed-tile coordinate. The bit stays unchanged if either tile-component

matching test fails at the cell, or if any site in its 3×3 Moore neighborhood has first component B. Thus one can view the rule as bits propagating along locally checked directed paths.

The precise full-shift hypothesis used below is

$$\boxed{G(q, s) = (q, g_q(s)) \quad \text{for every } q \in Q^{\mathbb{Z}^2},\ s \in \{0,1\}^{\mathbb{Z}^2}.}$$

Here Q is the finite set of allowed tile pairs; invalid global tilings are included. Kari states: "The tile components of the states, both d and t, do not change. Only the bits may be altered." [31, p. 179] This is the full-shift property required by the reduction.

Lemma 7.1 (fixed-background alphabet reduction)**.** *For a fixed finite alphabet B, let $G(q, s) = (q, g_q(s))$ be a CA on $(Q \times B)^{\mathbb{Z}^2}$. There is a computable CA $\widetilde{G}$ on $(\{0,1\} \times B)^{\mathbb{Z}^2}$, commuting with every lattice translation, such that*

$$\widetilde{G} \text{ is onto} \quad \Longleftrightarrow \quad G \text{ is onto.} \tag{7.1}$$

Its neighborhood is allowed to grow.

Proof. Let $\ell = \max(1, \lceil \log_2 |Q| \rceil)$. Choose a computable surjection $d : \{0,1\}^\ell \to Q$ and a right inverse $c : Q \to \{0,1\}^\ell$. Write $D_\ell(n_1, n_2) = (\ell n_1, n_2)$. For a binary background u, decode at every site x by the sliding window

$$\mathcal{D}(u, x) = d(u_x, u_{x+e_1}, \ldots, u_{x+(\ell-1)e_1}). \tag{7.2}$$

Suppose the evolving rule of G is

$$[g_q(s)]_x = f((q_{x+n}, s_{x+n})_{n \in N}). \tag{7.3}$$

Define

$$[\widetilde{G}(u, v)]_x = \left(u_x,\ f\big((\mathcal{D}(u, x + D_\ell n), v_{x+D_\ell n})_{n \in N}\big)\right). \tag{7.4}$$

This is one finite local rule applied at every site. Thus it commutes with *all* shifts; it does not select a privileged residue class.

For the proof only, decompose the signal coordinates into the ℓ classes $(r, 0) + D_\ell \mathbb{Z}^2$, $0 \le r < \ell$, and define

$$q_n^{(r)} = \mathcal{D}(u, (r, 0) + D_\ell n), \qquad v_n^{(r)} = v_{(r,0)+D_\ell n}. \tag{7.5}$$

For fixed u, equation (7.4) gives

$$[\widetilde{G}_u(v)]^{(r)} = g_{q^{(r)}}(v^{(r)}). \tag{7.6}$$

The decoded backgrounds for different residues can overlap in their use of u; the signal coordinates in different residues do not interact.

Because G freezes its background, it is onto exactly when g_q is onto for every q. If this holds, a target (u, w) for $\widetilde{G}$ has a preimage obtained by choosing, separately for each of the finitely many residues, a preimage of $w^{(r)}$ under $g_{q^{(r)}}$.

Conversely, suppose $\widetilde{G}$ is onto. For any target (q, s) for G, place $c(q_{(n,m)})$ in the disjoint horizontal block $\{(\ell n + j, m) : 0 \le j < \ell\}$ of a binary background u. Then $q^{(0)} = q$. Choose a signal target w whose residue-zero component equals s, with arbitrary other components. A preimage under $\widetilde{G}$ has the same background u; its residue-zero signal is a preimage of s under g_q. Thus G is onto. □

Corollary 7.2 (four CA symbols and 512 physical states)**.** *Surjectivity is undecidable for two-dimensional CAs on the fixed four-symbol alphabet $A_* = \{0,1\}^2$, with varying finite neighborhoods. Theorem A follows with $|S_*| = 512$.*

Proof. Apply Lemma 7.1 with $B = \{0, 1\}$ to Kari's hard family. The collision construction applies directly to the resulting four-symbol full-shift CA. There is no need for the decoder to preserve uniform measure: the balance equivalence applies to the resulting CA itself. Propositions 6.1 and 6.3 turn an algorithm deciding (1.1) into an algorithm deciding the complement of surjectivity, and hence surjectivity, a contradiction. □

For explicit bookkeeping, a Moore-neighborhood presentation of Kari's rule can be padded to $N = \{-1, 0, 1\}^2$. The encoded rule can use

$$\widetilde{N} = \{-\ell, -\ell + 1, \dots, 2\ell - 1\} \times \{-1, 0, 1\}, \qquad |\widetilde{N}| = 9\ell. \tag{7.7}$$

The choice (2.1) is then

$$\varepsilon = \frac{1}{400(9\ell + 1)^2}. \tag{7.8}$$

Every operation, including writing the finite local weight table, is effective. The four-symbol surjectivity corollary is a routine encoding of Kari's frozen-background rule; we make no separate novelty claim for that CA statement. Its role here is the physical dimension bound.

8 Exact diagonal quantum transfer

Let $d = |S|$ and let

$$\mathcal{A} = \overline{\bigcup_{\Lambda \Subset \mathbb{Z}^2} \bigotimes_{x \in \Lambda} M_d(\mathbb{C})} \tag{8.1}$$

be the full quasilocal matrix algebra. Diagonalize each classical interaction term in the product basis indexed by S, obtaining $\widehat{\Phi}$. Let E_{diag} be the local dephasing conditional expectation, extended consistently to $\mathcal{A}$.

Remark (The standard diagonal correspondence [7, 41, 19]). For every $\beta > 0$, there is an affine bijection

$$\mathcal{G}_\beta(\Phi) \longleftrightarrow \mathrm{KMS}_\beta(\widehat{\Phi}), \qquad \mu \longmapsto \omega_\mu = \mu \circ E_{\mathrm{diag}}. \tag{8.2}$$

For disjointly supported local quantum observables A, B,

$$\mathrm{Cov}_{\omega_\mu}(A, B) = \mathrm{Cov}_\mu(E_{\mathrm{diag}} A, E_{\mathrm{diag}} B). \tag{8.3}$$

In particular, $Z_x = \sigma_x^z \otimes I$ has exactly the classical magnetic correlations, in every KMS state.

Proof. Every local diagonal observable D is fixed by the dynamics. The KMS identity implies $\omega(AD) = \omega(DA)$ for all local A. Hence ω is invariant under conjugation by every local diagonal unitary. Averaging those unitaries dephases each finite matrix algebra, so $\omega = \omega \circ E_{\mathrm{diag}}$. This argument also removes coherences inside energy-degenerate subspaces.

Let Λ contain the support of a local matrix unit $|a\rangle\langle b|$, and write $E(a, \eta)$ for the sum of interaction energies meeting Λ, with interior configuration a and exterior configuration η. Diagonality and finite range give

$$\alpha_{i\beta}(|a\rangle\langle b|) = e^{-\beta(E(a,\eta) - E(b,\eta))} |a\rangle\langle b|. \tag{8.4}$$

Only finitely many interaction terms meeting the support enter; all other terms commute with it. Thus every local observable is entire analytic.

Use the KMS identity with opposite matrix units, multiplied by arbitrary exterior cylinder functions. It gives the conditional ratio

$$\mu(a \mid \eta) = e^{-\beta(E(a,\eta) - E(b,\eta))} \mu(b \mid \eta). \tag{8.5}$$

These strictly positive ratios, together with normalization, are exactly the DLR conditional probabilities.

Conversely, if μ is DLR, condition on the classical exterior of a finite region containing the supports of two local operators. The conditional interior density matrix is the finite-volume diagonal Gibbs matrix. Cyclicity of its trace gives the KMS identity for the two operators and their local analytic dynamics. Integrating the exterior proves KMS on the local algebra, and hence the full algebra. This establishes the bijection. The same local commuting-Hamiltonian KMS reduction is discussed in [19, Section 3.5].

For disjoint supports, dephasing factorizes: $E_{\mathrm{diag}}(AB) = (E_{\mathrm{diag}}A)(E_{\mathrm{diag}}B)$. Applying μ and subtracting the means proves (8.3). □

Taking the diagonal algebra alone would be incorrect for this assertion: its dynamics is trivial and every probability measure is KMS. The full matrix algebra is essential. The input energies remain log-rational; the theorem makes no claim for the narrower class of rational or algebraic Hamiltonian matrix entries at a fixed temperature.

Proof of Theorem B. Apply the correspondence to Theorem A. The diagonal observable Z_x has precisely the visible classical spin values. The affine bijection preserves translation invariance and all stated spin expectations. It therefore transfers both branches and each state quantifier. □

9 Noise and effective impossibility corollaries

9.1 A prescribed output-reset noise family

Proposition 9.1 (structured reset noise)**.** *Fix a rational retention parameter $\eta \in (0, 1]$. Independently reset each CA output, in each replica and at each site, to a uniform symbol with probability $1 - \eta$, before applying the collision kernel and averaging the resets. The resulting local weight is*

$$w_{i,\eta} = 1 + (\eta\varepsilon)^2 h(F(X)_i, F(X')_i). \tag{9.1}$$

Theorem A, including its every-Gibbs-state conclusion and diagonal quantum transfer, holds for this family at every such fixed $\eta > 0$. At $\eta = 0$ all inputs have the same critical visible marginal.

Proof. The reset kernel is $T_\eta(a, b) = \eta \mathbf{1}_{\{a=b\}} + (1-\eta)/q$. It multiplies each mean-zero function on A by η. By the kernel expansion used in (3.8), averaging two independent resets multiplies h by η^2. The collision activities become $\eta^{2|T|} z(T)$, so their zero/nonzero pattern is unchanged for $\eta > 0$. All completion and covariance estimates apply with ε replaced by $\eta\varepsilon$; the hidden contraction improves. Rational η preserves rationality of the weights. If $\eta = 0$, every w equals one. □

This proves Theorem C.

Proposition 9.2 (fragility under detuning and persistence under readout noise)**.** *The surjective branch is not robust to arbitrary sufficiently small changes of the weights. In contrast, independent sign noise on the visible readout preserves the correlation-decay dichotomy for every nonzero retention.*

Proof. Decrease the base horizontal log-odds from $\log 2$ to $\log 2 - t$, $t > 0$, leaving the collision factors unchanged. The proof using $R_\infty \geq 1$ now dominates every input model by the strictly subcritical ordinary FK model with odds product $2e^{-t} < 2$. Thus all inputs have exponential magnetic decay, however small t is. Rational changes of p_H can be chosen arbitrarily close to $2/3$.

For independent visible flips with mean multiplier $s \neq 0$, distinct-site connected correlations become

$$\operatorname{Cov}(\widetilde{\sigma}_x, \widetilde{\sigma}_y) = s^2 \operatorname{Cov}(\sigma_x, \sigma_y), \qquad x \neq y. \tag{9.2}$$

The same identity holds for a product quantum channel satisfying $Z \mapsto sZ$. Multiplication by a fixed positive constant preserves existence or nonexistence of an exponential envelope. The postprocessed state is not asserted to be a Gibbs state of another finite-range interaction. □

9.2 Effective compactness and no total positive decay modulus

Remark (Finite-volume extrema; standard compactness [41, 19]). For any finite-range interaction with positive rational local weights, and any rational-valued local observable O, let Λ_n exhaust the lattice and contain $\operatorname{supp} O$. Define

$$m_n = \min_{\eta} \gamma^{\eta}_{\Lambda_n}(O), \qquad M_n = \max_{\eta} \gamma^{\eta}_{\Lambda_n}(O), \tag{9.3}$$

where only the finite interaction collar of Λ_n must be enumerated. These numbers are exactly computable rationals, and

$$m_n \longrightarrow \inf_{\mu \in \mathcal{G}(\Phi)} \mu(O), \qquad M_n \longrightarrow \sup_{\mu \in \mathcal{G}(\Phi)} \mu(O). \tag{9.4}$$

No monotonicity or uniqueness assumption is required.

Proof. Finite-volume weights are products of positive rationals. Enumerating finitely many interior configurations and exterior collars computes every conditional expectation and both extrema exactly. The DLR equations place every Gibbs expectation between m_n and M_n.

For any subsequence of maximizing boundary conditions, extend each finite-volume measure by any configuration matching its chosen collar outside. Compactness gives a further weakly convergent subsequence. Every fixed local DLR equation holds for all sufficiently large volumes, and continuity passes it to the limit. Thus the limit is Gibbs. Its O-expectation is the corresponding limit of the M_n's, so no subsequential upper limit can exceed the supremum over Gibbs states. Together with the outward bound this proves convergence. The argument for minima is identical. □

Proposition 9.3 (no total algorithm for positive decay moduli)**.** *There is no total algorithm which, on every input in the constructed family, returns strictly positive rational numbers C_F, a_F, with the guarantee that on every nonsurjective input*

$$\forall \mu \in \mathcal{G}(\Phi_F)\ \forall x: \quad \mu(\sigma_0 \sigma_x) \leq C_F e^{-a_F \|x\|_1}. \tag{9.5}$$

The same impossibility holds if the outputs are computable positive real numbers with effective descriptions.

Proof. Suppose such an algorithm exists. The fixed critical Ising state has a computable expectation for each local observable: apply the finite-volume extrema remark to the rational-weight critical Ising specification, and use its uniqueness to make the lower and upper extrema converge to the same value.

Run the proposed algorithm on F. Since critical magnetic correlations have no exponential envelope, an effective search over x, rational b, and approximation accuracies terminates with

$$C_F e^{-a_F \|x\|_1} < b < \langle \sigma_0 \sigma_x \rangle_{\mathrm{c}}. \tag{9.6}$$

The exponential is a computable real, so the strict inequalities can be certified once true.

For the input interaction Φ_F, compute the finite-collar extrema in the finite-volume extrema remark for $O = \sigma_0\sigma_x$. If F is nonsurjective, all Gibbs expectations lie strictly below b, and eventually $M_n < b$. If F is onto, its unique Gibbs expectation is the critical value strictly above b, and eventually $m_n > b$. Running these tests decides surjectivity, contradicting Kari's theorem and Section 7. Computable-real outputs can either be used in the same strict-inequality search or relaxed to rational C above and a below them. □

This proves Theorem D. Positivity and totality on *every* input are essential. It does not rule out an algorithm which first searches for a collision witness and returns moduli only after finding one; nor does it rule out a total rate description allowed to be zero on surjective inputs.

Corollary 9.4 (semidecidability and explicit witnesses)**.** *The positive answer to* (1.1)*, on this constructed family, is computably enumerable. Given a witness, a strictly positive comparison displacement is explicitly computable.*

Proof. Enumerate finite sets T and compute the rational c_T exactly. The search finds a positive one precisely in the nonsurjective branch. Equations (4.2)–(4.3) give a computable positive Δ. If decay constants are desired, enumerate finite-set sharpness certificates for the resulting ordinary subcritical Ising model. Their strict inequalities are semidecidable for computable couplings, one exists by [16], and its iterative proof gives exponential constants. No computable a priori bound on the witness search is asserted. □

9.3 A general arithmetic upper bound

Here a positive rational finite-range model has a finite spin alphabet and a finite translation-invariant list of strictly positive rational weight tables; the specified on-site observable f has rational values. Put $f_x = f \circ \tau_x$. The following is an upper bound, not a completeness statement.

Proposition 9.5 (arithmetic upper bounds)**.** *The raw-decay property*

$$\forall \mu \in \mathcal{G}(\Phi)\ \exists C_\mu, a_\mu > 0\ \forall x:\ |\mu(f_0 f_x)| \le C_\mu e^{-a_\mu \|x\|_1}$$

belongs to Σ^0_3*. Under the promise that* $\mathcal{G}(\Phi)$ *is a singleton, it belongs to* Σ^0_2*.*

Proof. First the statewise bounds can be made uniform. The nonempty compact metrizable convex set $K = \mathcal{G}(\Phi)$ is covered by the closed sets

$$A_{C,m} = \{\mu:\ |\mu(f_0 f_x)| \le C\, 2^{-\lceil \|x\|_1/m \rceil} \text{ for every } x\}, \qquad C, m \in \mathbb{N}_{>0}.$$

This is equivalent to exponential decay after enlarging C, m. Baire's theorem gives one $A_{C,m}$ with nonempty relative interior, containing a neighborhood of some μ_0. There is a single $\lambda > 0$ so small that $(1-\lambda)\mu_0 + \lambda\nu$ lies in this neighborhood for every $\nu \in K$. Indeed a basic weak neighborhood imposes finitely many expectation constraints on bounded local functions; their changes under this mixture are uniformly $O(\lambda)$. Since also $\mu_0 \in A_{C,m}$, affine linearity gives, for every ν and x,

$$|\nu(f_0 f_x)| \le \lambda^{-1}(2-\lambda) C\, 2^{-\lceil \|x\|_1/m \rceil}.$$

Rounding the new constant up proves uniformity, without requiring that Gibbs states themselves be translation invariant.

Choose nested boxes $\Lambda_L \uparrow \mathbb{Z}^d$ and, once they contain $0, x$, define the computable rational

$$U_L(x) = \max_\tau |\gamma^\tau_{\Lambda_L}(f_0 f_x)|.$$

The tower identity for specifications gives $U_{L+1}(x) \le U_L(x)$: each larger-volume expectation is an average of smaller-volume conditional expectations. Every Gibbs expectation is bounded by $U_L(x)$. Conversely maximizing collars have a subsequential Gibbs limit by the compactness argument above, so

$$U_L(x) \downarrow \sup_{\mu \in K} |\mu(f_0 f_x)|.$$

With slack obtained by enlarging C, m, the property is exactly

$$\exists C, m \in \mathbb{N}_{>0}\ \forall x \in \mathbb{Z}^d\ \exists L: \quad \{0, x\} \subseteq \Lambda_L, \qquad U_L(x) < C\, 2^{-\lceil \|x\|_1/m \rceil}.$$

The matrix predicate is decidable rational arithmetic; this is Σ_3^0.

Under uniqueness, the finite-collar minimum and maximum converge to the same expectation. Search until their gap is at most 2^{-k} to compute a rational $v_{x,k}$ with $|v_{x,k} - \mu(f_0 f_x)| \le 2^{-k}$. This search is uniform and total on the uniqueness promise. The decay property is then equivalent to

$$\exists C, m\ \forall x, k: \quad |v_{x,k}| \le C\, 2^{-\lceil \|x\|_1/m \rceil} + 2^{-k}.$$

For a literal arithmetic formula on input codes, let the universal variable also encode any finite transcript of this deterministic approximation algorithm. Check the displayed inequality only if the transcript is a completed run for (x, k), otherwise accept it. This is a decidable predicate; totality on the promise ensures a completed transcript for each (x, k). Letting $k \to \infty$ proves the converse. Thus the promise restriction has a Σ_2^0 definition. □

Remark (Decidable special cases). In one dimension, positive finite-range weights yield a primitive block transfer matrix [22, Corollary 6.41]. All connected correlations decay exponentially; raw decay of a real on-site f is equivalent to zero Gibbs mean. With rational weights this mean is algebraic, so equality to zero is decidable by real algebraic arithmetic. For square-lattice nearest-neighbor zero-field ferromagnetic Ising and the spin observable, with rational $u = e^{2K_H} > 1$, $v = e^{2K_V} > 1$, raw all-state exponential decay holds exactly when $(u-1)(v-1) < 2$, by (I1)–(I3) and positive magnetization above criticality [11, 16]. Equality is critical and fails. This is a rational comparison. These remarks do not assert decidability for general higher-dimensional models.

10 Potts order with a fixed gap and a computable pressure

The collision machinery also detects a standard order parameter with a fixed gap. We distinguish the CA alphabet size $s = |A|$ from the number Q of visible colours. In this section $Q = 9$ in the theorem; the intermediate comparison statements hold for every integer $Q \ge 2$.

10.1 The local model and the unchanged ingredients

At each site use $S_Q = \{1, \dots, Q\} \times A^4$ with the same two independent replica pairs as in Section 2. Use ε from (2.1), evaluated with $q = s$, and define

$$b_{i,r} = w_i^r + \sqrt{Q}\, \mathbf{1}_{\{\sigma_i = \sigma_{i+e_r}\}}, \qquad w_i^r = 1 + \varepsilon^2 \big(s \mathbf{1}_{\{F(X^r)_i = F(X'^r)_i\}} - 1\big). \tag{10.1}$$

Write Ψ_F for this interaction. Its Hamiltonian is the sum of $-\log b_{i,r}$ over the affected factors, exactly as in Section 2, with counting a-priori measure and no hidden one-site terms. For $Q = 9$ the odds $\sqrt{Q} = 3$ and all entries are strictly positive rationals. Translation invariance, finite range and effective description follow by the proof of Lemma 2.1.

The multiplier in the collision kernel is *exactly* $s = |A|$. Indeed an onto CA with a different multiplier s' would give

$$W(C) = \big(1 + \varepsilon^2 (s'/|A| - 1)\big)^{|C|},$$

which moves the onto branch away from the coexistence point. No such detuning is allowed here.

Use the same c_T, $D(P)$ and hidden insertion ratios as before. Lemmas 3.1, 4.1, 5.1 and 5.2 transfer verbatim, with witness translates allowed in either orientation. The proof of Lemma 5.4 also transfers verbatim, with Lemma 10.4 in place of Lemma 5.3: only its base odds $p_e/(1-p_e)$ become $\sqrt{Q}$. These arguments use the CA collision layer, positivity and conditional hidden uniqueness, not the visible alphabet size. In particular, given visible colours, the oscillation of $\log(\sqrt{Q}\delta_{\text{same}}+w)$ is at most $\log(w_{\max}/w_{\min})$, so the hidden Dobrushin bound is unchanged. We do not repeat these proofs.

With a free padded hidden input region and free or monochromatic visible boundary, the integrated bond density is

$$\widetilde{\mu}_F(\omega) \;\propto\; Q^{K(\omega)}(\sqrt{Q})^{|\omega|}W(C_H(\omega))W(C_V(\omega)). \tag{10.2}$$

Here K counts free clusters; with monochromatic boundary the wired cluster has weight one. This follows by expanding (10.1): open edges contribute $\sqrt{Q}\delta_{\text{same}}$ and closed edges contribute w_e. Conditional on bonds in finite volume, visible cluster colours and hidden variables are independent.

10.2 The changed comparison statements

Let $c_P > 0$, $k = |P|$, $w_{\max} = 1 + \varepsilon^2(s-1)$, and

$$a_P = \frac{\varepsilon^{2k}c_P}{w_{\max}^{|D(P)|}}, \qquad g = \frac{1}{k}\log(1+a_P), \qquad \kappa_Q = (1+\sqrt{Q})^{-(k-1)}, \qquad \Delta_Q = \kappa_Q g. \tag{10.3}$$

For each $r = H, V$, retain every translate of the r-oriented edge pattern with centers P that lies in a finite edge set E. Let $\mathcal{T}_E$ be the union of these *two* collections. Put $I_t = \mathbf{1}_{\{P+t \text{ closed in its orientation}\}}$ and $N_E = \sum_{t\in\mathcal{T}_E} I_t$. Fix $p_* \in P$ and anchor a translate at its edge centered at $p_* + t$, in the same orientation. The anchors are distinct, also between the two layers. Every edge sufficiently far from the boundary is an anchor.

Lemma 10.1 (Q-version of Lemma 4.2)**.** *For $Q \geq 2$, any free or wired FK boundary condition and real edge log-odds θ_e, the law*

$$\nu_{\theta,g}(\omega) \propto Q^{K(\omega)} \exp\Big(\sum_e \theta_e\omega_e + gN_E(\omega)\Big)$$

is FKG for $g \geq 0$. At $\theta_e = \log\sqrt{Q}$, the finite law (10.2) is stochastically dominated by $\nu_{\theta,g}$.

Proof. The proof of Lemma 4.2 changes 2^K to Q^K; K and each I_t remain supermodular. For an edge in either orientation, at most k retained translates are completed by closing it. The completion bound gives an insertion ratio at least e^{gn_e}, since $e^{gk} = 1 + a_P$. Thus $W(C_H)W(C_V)e^{-gN_E}$ is increasing in closed edges. Its density relative to the FKG comparator is decreasing in open edges, proving domination by the same covariance argument. The factor $1/k$ in (10.3) is essential when several translates are completed at once. □

Lemma 10.2 (Q-version of Lemma 4.3)**.** *For $\nu_{\theta,u} \propto Q^K \exp(\sum_e \theta_e\omega_e + uN_E)$ with $u \geq 0$ and every $\theta_e \leq \log\sqrt{Q}$, every increasing f and retained translate t satisfy*

$$\operatorname{Cov}(f, I_t) \leq -\kappa_Q \operatorname{Cov}(f, \omega_{e_t}). \tag{10.4}$$

Proof. The FK cluster factor in the conditional opening odds is in $\{1, 1/Q\}$. Opening can only remove rewards, so the odds are at most $\sqrt{Q}$. The conditional closure probability is at

least $(1+\sqrt{Q})^{-1}$, also under partial conditioning by averaging. Sequentially closing the other $k-1$ edges gives $\nu(I_t \mid \omega_{e_t}=0) \geq \kappa_Q$. Conditional FKG, as in Lemma 4.3, gives

$$\operatorname{Cov}(f, I_t) \leq -\frac{\nu(I_t)}{\nu(\omega_{e_t}=0)} \operatorname{Cov}(f, \omega_{e_t}),$$

and the last covariance is nonnegative. This proves (10.4). At $Q=9$, the odds bound is 3 and $\kappa_9 = 4^{-(k-1)}$. □

Proposition 10.3 (Q-version of Proposition 4.4)**.** *The comparator at critical odds $\sqrt{Q}$ and reward g is dominated by ordinary FK with odds $\sqrt{Q}e^{-\Delta_Q}$ on all retained anchors and $\sqrt{Q}$ on the remaining edges, with the same boundary condition.*

Proof. Use $0 \leq u \leq g$, with $\theta_e(u) = \log\sqrt{Q} - \kappa_Q g + \kappa_Q u$ on anchors and $\theta_e(u) = \log\sqrt{Q}$ otherwise. All odds stay at most $\sqrt{Q}$. Differentiation gives the sum of $\kappa_Q \operatorname{Cov}(f, \omega_e)$ over anchors and $\operatorname{Cov}(f, I_t)$ over translates; distinct anchors and (10.4) make it nonpositive. The two endpoints are exactly the asserted laws. This is the proof of Proposition 4.4, now shifting both orientations. □

Lemma 10.4 (Q-version of Lemma 5.3)**.** *For the canonical bond extension of every $\mu \in \mathcal{G}(\Psi_F)$, the hidden conditional law given (σ, ω) is $\lambda^{C_H(\omega)} \otimes \lambda^{C_V(\omega)}$ and is independent of the colours. Conditional on bonds, all finite clusters have independent uniform colours in $\{1, \dots, Q\}$, independent also of the colours on infinite clusters.*

Proof. The common-null-set disintegration argument in Lemma 5.3 is unchanged. Replace its finite-cluster spin flip by a permutation of the colours of a finite cluster. This preserves open equality constraints and all closed factors. Conditional on everything outside that cluster, the Q colours are equiprobable. Apply this finite-support DLR argument to any finite collection of clusters. Countably many finite-cluster events give the stated conditional law. □

Proposition 10.5 (Q-version of Proposition 5.5)**.** *Every bond marginal of every $\mu \in \mathcal{G}(\Psi_F)$ is dominated by critical wired FK, $\phi^1_{p_c(Q),Q}$, where $p_c(Q) = \sqrt{Q}/(1+\sqrt{Q})$. If $c_P > 0$, it is dominated by the homogeneous wired FK limit at*

$$p_F = \frac{\sqrt{Q}e^{-\Delta_Q}}{1+\sqrt{Q}e^{-\Delta_Q}} < p_c(Q). \tag{10.5}$$

Proof. The exact odds identity of Lemma 5.4 is $\sqrt{Q}\,\alpha_\mu / R_\infty(C, e)$. In a fully wired finite core, if the endpoints are disconnected after deleting e, at least one component is a finite interior cluster. Lemma 10.4 gives $\alpha_\mu = 1/Q$. If the endpoints are wired-connected, $\alpha_\mu \leq 1$. Thus α_μ is bounded by the wired cluster factor, respectively $1/Q$ or 1. The insertion bounds apply in both orientations. Consequently the target's opening probabilities are at most those of the fully wired comparator at the same internal configuration.

Use the monotone-upper-law heat-bath coupling in Proposition 5.5; its downward-support and irreducibility argument applies without target FKG. Apply Proposition 10.3. Keep the fixed-width unshifted boundary strip: every edge in a smaller concentric core is an eligible anchor in its own orientation. Conditional on the strip and exterior edges, the induced ordinary FK boundary partition is bounded by full wiring. The smaller core therefore has a homogeneous upper law with parameter (10.5). Let the cores increase. Wired FK limits and increasing finite-path unions give the comparison, including infinite-cluster events, exactly as in Proposition 5.5. Omitting rewards proves the critical upper bound. Below $p_c(Q)$ the wired limit is unique and has exponential connectivity decay [5]. □

10.3 The two Potts branches

The external inputs are distinct. Beffara–Duminil-Copin [5] prove $p_c(Q) = \sqrt{Q}/(1+\sqrt{Q})$ and exponential decay below this value. Discontinuity at $Q > 4$ is due to Duminil-Copin, Gagnebin, Harel, Manolescu and Tassion [17]. Their Theorem 1.2(2) states

$$\phi^1_{\mathbb{Z}^2,p_c(Q),Q}[\text{there exists an infinite cluster}] = 1,$$

and their Theorem 1.1(2) states, for each $i \in \{1,\dots,Q\}$,

$$\mu^i_c[\sigma_0 = i] > 1/Q.$$

These are the statements used here; criticality alone does not imply them. In particular, with $Q = 9$,

$$m_9 = \phi^1_{3/4,9}(0 \longleftrightarrow \infty) > 0. \tag{10.6}$$

Positivity also follows directly from the first statement and translation invariance: if the probability at each site were zero, a countable union could contain no infinite-cluster vertex.

Proposition 10.6 (Potts branch expectations)**.** *If F is nonsurjective, every Gibbs state of Ψ_F has only finite visible bond clusters, is invariant under S_9 acting on all visible colours, and satisfies*

$$\begin{gathered}\mu(X_x) = 0, \qquad X_x = \frac{9\mathbf{1}_{\{\sigma_x=1\}} - 1}{8}, \\ 0 \le \mu(\mathbf{1}_{\{\sigma_x=\sigma_y\}}) - \tfrac{1}{9} = \tfrac{8}{9}\widetilde{\mu}(x \longleftrightarrow y) \le C_F e^{-a_F\|x-y\|_1},\end{gathered} \tag{10.7}$$

with $C_F < \infty, a_F > 0$ uniform over states and positions. If F is surjective, there are translation-invariant Gibbs states μ^i_F, $1 \le i \le 9$, with

$$\mu^1_F(X_0) = m_9, \qquad \mu^i_F(X_0) = -m_9/8 \quad (i \ne 1).$$

Over all Gibbs states, in the surjective branch,

$$\sup_\mu \mu(X_0) = m_9, \qquad \inf_\mu \mu(X_0) = -m_9/8. \tag{10.8}$$

Proof. Nonsurjectivity supplies a witness by Lemma 3.1. Proposition 10.5 and subcritical sharpness make every cluster finite, with a uniform exponential connection bound. Independent uniform colours, also independent of the hidden layer given bonds, give S_9 invariance of the full state, zero magnetization and (10.7).

If F is onto, $W(C) = 1$ for every finite C. In padded finite volumes with colour-i visible boundary, (10.2) is exactly wired critical FK. The visible-bond marginal is its Edwards–Sokal coupling: the wired cluster has colour i, other clusters have independent uniform colours. The hidden buffer is free and includes all input variables for retained factors. Every fixed interior specification is consequently exact for sufficiently large volumes, so limits are Gibbs for Ψ_F.

The wired bond limit and its colour-i coupling are translation invariant. The conditional hidden law is unique by Lemma 5.1 and covariant under shifts; hence this marginal has a unique full hidden extension and the resulting μ^i_F is translation invariant. The usual wired Edwards–Sokal limit colours the infinite cluster i and finite clusters uniformly [26, Theorem 4.91 and Chapter 5]. In finite volume already,

$$\mu^1_F(X_0) = \phi^1(0 \longleftrightarrow \partial\Lambda),$$

with the appropriate finite-volume notation. Passing to the wired limit gives m_9. For $i \ne 1$ the contribution of the wired infinite cluster is $-1/8$, giving $-m_9/8$.

For an arbitrary Gibbs state let $\vartheta = \widetilde{\mu}(0 \longleftrightarrow \infty)$ and $b = \widetilde{\mu}(0 \longleftrightarrow \infty, \sigma_0 = 1)$. Finite-cluster colours give

$$\mu(X_0) = \frac{9b - \vartheta}{8}, \qquad 0 \le b \le \vartheta \le m_9,$$

where the last bound is the critical wired domination. Thus $-m_9/8 \le \mu(X_0) \le m_9$. The ordered states attain both endpoints. This argument requires no classification of all critical Gibbs states. □

Proof of Theorem E. Apply the CA-only encoding of Corollary 7.2 to the hard inputs and use (10.1) with $Q = 9$. The CA alphabet is four, independently of the visible alphabet. Therefore the local dimension is $9 \cdot 4^4 = 2304$. Proposition 10.6 gives the exact alternative $\sup_\mu \mu(X_0) \in \{0, m_9\}$, also when the supremum is restricted to translation-invariant states. Every fixed rational $m_* \in (0, m_9]$ supplies a promise gap; such a rational exists and can be fixed in the decision problem. No explicit numerical lower bound for m_9 is needed or asserted. A decision algorithm under this promise would decide CA surjectivity, contradicting Kari's theorem. □

10.4 Computable pressure, undecidable differentiability

Add the uniform field $h \sum_x X_x$ in the log weight, so the energy acquires $-h \sum_x X_x$. Use counting a-priori measure and write

$$P_F(h) = \lim_{L \to \infty} |\Lambda_L|^{-1} \log Z_{F,\Lambda_L}(h), \qquad \Lambda_L = \{0, \dots, L-1\}^2.$$

The boundary convention does not affect this limit. Rational weights are asserted at $h = 0$ only; the factors e^{hX_x} need not be rational.

Corollary 10.7 (computable pressure and an undecidable kink)**.** *The real function $P_F(h)$ is uniformly computable from F and a computable-real name for h. Nevertheless differentiability at $h = 0$ is undecidable on this family. Exactly,*

	$P'_F(0+)$	$P'_F(0-)$
F surjective	m_9	$-m_9/8$
F nonsurjective	0	0

(10.9)

Thus the derivative jump is either 0 *or the fixed positive constant* $9m_9/8$.

Proof. Pressure computability follows from Gayral–Sablik–Taati [23, Proposition 14]; their proof is uniform in a computable potential. Here there is also a direct modulus. Choose an integer R containing each factor support in $[-R, R]^2$, and a computable rational upper bound M for $\sum_{r=H,V} \|\log b_{0,r}\|_\infty$. Tile a large square by L-squares and omit the factors crossing tile boundaries. At most $4R/L$ of the centers per orientation are affected in the density estimate, so the safe bound $K_F = 16(R+1)M$ gives

$$\left| P_F(h) - L^{-2} \log Z^{\mathrm{free}}_{F,\Lambda_L}(h) \right| \le K_F/L \qquad (L > 2R).$$

The field has no crossing terms. Finite sums of computable exponentials and logarithms compute the middle term. Also $|P_F(h) - P_F(h')| \le |h - h'|$ since $|X_x| \le 1$, making this uniform for real input names as well as rational fields.

For completeness, the derivative identification is a variational argument. Let $\mathcal{I}$ be the compact set of translation-invariant probability measures on $S_{\mathrm{P}}^{\mathbb{Z}^2}$, $\mathfrak{s}(\nu)$ their entropy density, and $u_F(\nu) = \sum_{r=H,V} \nu(-\log b_{0,r})$ the energy density. The Gibbs variational principle gives

$$P_F(h) = \max_{\nu \in \mathcal{I}} \{A_F(\nu) + h\nu(X_0)\}, \qquad A_F(\nu) = \mathfrak{s}(\nu) - u_F(\nu). \tag{10.10}$$

The maximizers are exactly the translation-invariant DLR states for the field interaction. These facts are the Lanford–Ruelle principle: see Georgii [24, Theorems 15.30 and 15.39, with the converse in Theorem 15.37]. Upper semicontinuity of entropy on a finite-alphabet full shift follows from its finite-window infimum representation [24, Theorem 15.12 and Proposition 15.14]; the energy is continuous. Thus A_F is upper semicontinuous and maximizers exist.

Let $\mathcal{E}_0$ be the maximizers at zero. Substituting any $\nu \in \mathcal{E}_0$ at $h > 0$ gives $(P_F(h) - P_F(0))/h \geq \nu(X_0)$. Conversely choose a maximizer ν_h at h. Since $A_F(\nu_h) \leq P_F(0)$,

$$\frac{P_F(h) - P_F(0)}{h} \leq \nu_h(X_0).$$

As $h \downarrow 0$, compactness supplies limit points. The Lipschitz pressure bound and bounded X_0 give $A_F(\nu_h) \to P_F(0)$; upper semicontinuity puts every such limit point in $\mathcal{E}_0$. Therefore

$$P_F'(0+) = \max_{\nu \in \mathcal{E}_0} \nu(X_0).$$

Apply the same argument to $-X_0$ for negative fields to obtain $P_F'(0-) = \min_{\nu \in \mathcal{E}_0} \nu(X_0)$. Lanford–Ruelle is needed here to identify the limiting variational states with Gibbs states; finite-volume formal differentiation alone would not suffice. Proposition 10.6 bounds all Gibbs states and attains the extremes in translation-invariant ones, yielding (10.9). The hard family from Corollary 7.2 proves undecidability, even under any fixed rational positive lower promise smaller than the nonzero jump. □

The diagonal Gibbs–KMS correspondence of Section 8 transfers Theorem E to on-site dimension 2304, with the diagonal observable X_0 and the same equilibrium pressure. The output-reset calculation of Section 9, first subsection, transfers verbatim: $\varepsilon^2 \mapsto \eta^2 \varepsilon^2$ for fixed rational $\eta \in (0, 1]$ preserves the witness zero set and hence both Potts branches, the fixed gap and the pressure kink.

11 Auxiliary uniqueness routes

Theorems A–E do not use this section. As a heuristic, factorization over the overlap graph rewrites W as a hard-core polymer partition function. At sufficiently small activity its logarithm has the form

$$\log W(C) = L(C) + R(C), \qquad L(C) = \sum_{\substack{P \neq \varnothing \\ P \ \Gamma\text{-connected}}} z(P) \mathbf{1}_{\{P \subseteq C\}}. \tag{11.1}$$

Cluster expansions [33, 20] suggest a quadratic norm bound for R. Combining such a bound with a uniform near-critical susceptibility estimate is a possible uniqueness strategy. We give no proof here of the required lattice-uniform spin estimate, the limiting odd-sector gap, or a boundary-compatible response bound; scaling-limit intuition alone does not supply them. No susceptibility exponent or claimed quadratic room is used to establish uniqueness of the actual $L + R$ model.

11.1 What attractiveness does prove for the L-model

Write $n_e = 1 - \omega_e$, and $I_P = \prod_{e \in P} n_e$. The L-model means the FK-Ising law multiplied by

$$\exp\left(s \sum_{r=H,V} \sum_{P} z_t(P) I_{P,r} \right), \qquad z_t(P) = t^{|P|} c_P, \quad t = \varepsilon^2, \quad s \geq 0.$$

The sum is over nonempty connected center sets and both orientations, with every translate counted once. In particular t is the microscopic parameter, whereas s scales the entire positive interaction. The factor $t^{|P|}$ must be retained: unnormalized powers would destroy the surjective cancellation. At the paper's ε, the activity bound and the connected-set count give

$$\sup_e \sum_{P,r:\, e\in P_r} |P| z_t(P) < \infty$$

with an exponential tail in $|P|$ and diameter. Indeed there are at most $(m^2)^{2(j-1)}$ connected j-sets through a center, and $z_t(P) \le [(q-1)t]^j$; their product has ratio less than one. Thus all boundary energy densities below vanish in growing squares.

Proposition 11.1 (L-model uniqueness at almost every scale)**.** *For each fixed input and $t = \varepsilon^2$, the L-model has a unique bond Gibbs state for every $s \ge 0$ outside a countable set. If all collision coefficients vanish, it is the unique critical FK state at every s. This proposition concerns the L-model only.*

Proof. We first prove the needed finite-energy separation lemma. Let $\nu^- \preceq \nu^+$ be Gibbs measures for an attractive binary specification on coordinates n_e, with every single-coordinate conditional probability of $n_e = 1$ at least $a > 0$. For a finite P and $e \in P$, take a monotone coupling and resample the coordinates of $P \setminus \{e\}$ once each by the common-uniform heat bath. Both marginals remain Gibbs and ordered; n_e is unchanged. At each update the lower coordinate becomes one with conditional probability at least a. Conditional on the original discrepancy at e, the probability that all these lower coordinates finish at one is at least $a^{|P|-1}$. On that event I_P differs by one between the upper and lower configurations. Since I_P is increasing, its difference is never negative. Therefore

$$\nu^+(I_P) - \nu^-(I_P) \ge a^{|P|-1}\big[\nu^+(n_e) - \nu^-(n_e)\big]. \tag{11.2}$$

The FK lattice condition and all positive all-closed rewards give attractiveness in the closed variables. Their monotone extremal limits ν_s^-, ν_s^+ sandwich every bond Gibbs state, by the same finite-volume comparison as for FK. The absolutely summable reward tail preserves the sandwich on passage to infinity. Every finite-volume L law is also dominated in open edges by the ordinary wired critical FK law, since closing gains only positive rewards. Its limits have no infinite open cluster by (I2); exterior connections between a fixed core and increasingly distant boundaries therefore have vanishing probability. The usual FK core specification consequently passes to these limits, together with the summable reward tail. Thus the extremal limits are themselves Gibbs states, not just ordered limiting laws. Opening an edge has FK odds at most 2, and loses rewards; hence its closed conditional probability is at least $a = 1/3$, uniformly in boundary partitions.

Let $\mathcal{P}(s)$ be the limiting pressure per vertex. It exists and is boundary independent: tiling a large square by smaller squares changes the FK cluster count by at most the number of boundary vertices; the summable interaction tail gives a further $o(\text{volume})$ error after first truncating its range. Finite-volume pressures are convex in s, hence so is $\mathcal{P}$. At a differentiability point, their derivatives converge to $\mathcal{P}'(s)$. To verify this without exchanging derivatives and limits, bound each finite derivative between its two difference quotients at $s \pm h$, pass to the pressure limit, and then send $h \downarrow 0$ (at $s = 0$ use right derivatives or omit this single point).

For either extremal boundary, that finite derivative is the mean positive pattern density. Translation invariance of the extremal limits, removal of the vanishing boundary density, and the uniformly summable tail identify its limit as

$$\mathcal{P}'(s) = \sum_{r=H,V} \sum_{P \ni 0} \frac{z_t(P)}{|P|} \nu_s^{\pm}(I_{P,r}).$$

All differences between the upper and lower terms are nonnegative. If a positive witness exists, equality of these sums forces equality of its pattern expectation in each orientation. Translates of that witness cover every edge of that orientation. Equation (11.2) then forces equality of every one-edge marginal of ν_s^+ and ν_s^-. In a monotone coupling every coordinate disagreement has probability zero; countably many coordinates give equality of the measures. The sandwich proves uniqueness. A convex function on an interval has at most countably many nondifferentiability points (the disjoint derivative-jump intervals contain distinct rationals). Add $s = 0$ to the possible exceptional set. If no witness exists, $L = 0$ and (I2) applies directly. □

Remark (Why this does not prove uniqueness in Theorem A). The actual collision marginal has $\log W = L + R$, and need not be attractive. For a binary iid input let $F(X)_i = \mathbf{1}_{\{X_i=1, X_{i+e_1}=0\}}$. Its sign output has mean $1/2$, but the sign product at adjacent centers has mean zero: the two output-one events are disjoint. Thus $c_{\{i\}} = 1/4$ and $c_{\{i,i+e_1\}} = 0$. At $t = \varepsilon^2$,

$$W(\{i, i+e_1\}) - W(\{i\})W(\{i+e_1\}) = -\frac{t^2}{16} < 0.$$

For example embed those two centers as adjacent horizontal edges on a path. Ordinary FK weights there are products, so they do not repair this failed lattice inequality. Pressure differentiability alone cannot identify Gibbs states without the monotone separation argument. Uniqueness of the $L + R$ model used in Theorem A remains open.

12 Open problems

Question 12.1 (Uniqueness). Does every nonsurjective member of the collision family have a unique Gibbs state?

The magnetic estimate controls cluster signs. It does not control the response of closed-pattern energies or hidden observables to an exterior condition. A closed FK cut does not isolate the hidden factors that cross it.

Question 12.2 (All local observables). Can a collision-specific boundary comparison prove exponential clustering for arbitrary bounded local observables in the nonsurjective branch?

For finite supports S, T, the usual target bounds $|\operatorname{Cov}_\mu(f, g)|$ by a constant times $|S||T|\|f\|_\infty\|g\|_\infty e^{-a\,\mathrm{dist}(S,T)}$. Odd observables are covered by Corollary 6.2. Even and hidden observables require another argument.

Question 12.3 (Dimension and range). Can the local dimension and interaction range both be fixed while retaining positive rational weight tables?

Fixing a CA alphabet and its neighborhood leaves only finitely many rules. Rational physical couplings can still carry unbounded input data, so that observation does not rule out a different construction.

Question 12.4 (Perturbations). Can an undecidability construction for connected correlations, or for the raw observable $D_x = \sigma_x - \sigma_{x+e_1}$, be stable under sufficiently small translation-invariant perturbations of its local weights?

The output-reset theorem concerns its prescribed noise family. Lowering the horizontal base odds makes every member here magnetically subcritical, as Section 9 shows.

Remark (Conditional contraction and bounded magnetic clusters do not force uniqueness). Place hidden spins $\tau_x = \pm 1$ on $\mathbb{Z}^2$. For each nearest-neighbor edge $e = \{x, y\}$ introduce 45 disjoint visible dimers $(s_{e,j}, t_{e,j})$ with factor

$$B_{e,j} = \frac{100}{101}\mathbf{1}_{\{s_{e,j}=t_{e,j}\}} + \frac{1}{101}2^{\tau_x+\tau_y}, \qquad e^{h\tau_x} = (67/72)^{45\tau_x}.$$

These are strictly positive rational weights. Given bonds, the hidden law is product. Given visible spins, the maximum hidden coupling contributed by an equal dimer is

$$j = \tfrac{1}{4}\log\frac{10426}{10201}, \qquad 4\tanh(45j) < 45\log\left(1+\frac{225}{10201}\right) < \frac{10125}{10201} < 1.$$

Thus hidden contraction holds under both conditionings, and every visible open cluster has at most two vertices. Summing one dimer gives a factor proportional to $50 + 2^{\tau_x+\tau_y}$. The chosen field cancels its accumulated field, leaving zero-field hidden Ising coupling

$$J = \frac{45}{4}\log\frac{603}{578} > K_c.$$

The strict inequality follows from $603^{45} > 6\,578^{45}$ and $6 > 3 + 2\sqrt{2} = e^{4K_c}$. Hence there are distinct hidden phases. They also change visible dimer energies, since $\mathbb{E}(st \mid \tau) = 50/(50 + 2^{\tau_x+\tau_y})$ is strictly decreasing in each endpoint. The closed-factor integral is log-supermodular: it is a product over sites of $\cosh(h + (\log 2)n_x(C))/\cosh h$. This example invalidates the generic uniqueness inference, not the small collision construction.

Remark (An FKG counterexample with no signed remainder). Take isolated visible dimers indexed by $i \in \mathbb{Z}^2$, with closed indicators $c_i = 1 - \omega_i$. Add the positive reward

$$L(c) = \lambda \sum_{\langle i,j\rangle} c_i c_j, \qquad \lambda = \log 16.$$

Choose ordinary dimer FK parameter $p = 512/513$. Its effective open odds are $p/[2(1-p)] = 256 = e^{2\lambda}$. The bond density is proportional to $\exp(2\lambda\sum_i \omega_i + L(c))$. Substituting $\tau_i = 2c_i - 1$ cancels the field and gives hidden-index Ising coupling $\lambda/4 = \log 2 > K_c$. There are two bond Gibbs states. The density is FKG and is dominated by the underlying product dimer law, whose clusters are bounded. All added weights $16^{c_i c_j}$ are rational. Thus FKG, bounded open clusters and $R = 0$ still do not establish uniqueness.

Remark (Energy-sector reduction; the remaining estimate is open). For an edge put $\tau_e = \sigma_u\sigma_v$. Conditional on visible spins, summing the ordinary ES bond gives a closed-pattern energy

$$v_P(\sigma) = \prod_{e\in P}\left(1 - \frac{p_e}{2} - \frac{p_e}{2}\tau_e\right).$$

Each factor is the conditional closed probability; the sum of absolute polynomial coefficients of this product is at most one. Singleton rewards can be absorbed into the corresponding pair coupling. Since the all-state odd sector already clusters, only even visible information can distinguish states. On a connected lattice every even spin monomial is a product of edge variables τ_e (join its even number of vertices by finite paths). Odd monomials have expectation zero. Thus all finite products of edge energies determine the entire visible law. Given that law, conditional hidden uniqueness from Lemma 5.1 determines the full law as well.

Consequently a sufficient route to uniqueness is a signed massive *annealed* energy-response estimate for these edge products, uniform over the original visible and hidden boundaries. The desired scale has mass at least $c\Delta$, a kernel bounded by

$$K_\Delta(e,f) \le \frac{Ce^{-c\Delta\,\mathrm{dist}(e,f)}}{(1+\mathrm{dist}(e,f))^2}, \qquad \sup_e \sum_f K_\Delta(e,f) \le C[1+\log(1/\Delta)],$$

and a boundary error tending to zero, including rewards induced by fixed hidden boundary variables. Such rewards can be $O(\varepsilon^2)$ even when all bulk activities vanish (a surjective shift with pinned equal replicas gives $w_e = 1 + (q-1)\varepsilon^2$). A bulk quadratic remainder estimate cannot simply absorb that boundary layer. A weaker row bound of order $\Delta^{-7/4}$ could suffice after a suitably small parameter choice and certified constants, but neither bound is established here. This is a constructive-renormalization-scale open estimate, not a theorem supplied by the odd-sector comparison or Proposition 11.1.

13 Discussion

The reduction keeps the temperature fixed. Computation is encoded in finite rational local tables, and the infinite-volume spin correlation detects whether a finite collision witness exists. Nonsurjectivity gives a strictly positive displacement from criticality. Surjectivity cancels every collision correction and leaves the same critical Ising marginal.

Theorem E adds a fixed order-parameter gap: the maximal magnetization is exactly 0 or m_9, and the pressure jump is exactly 0 or $9m_9/8$. Theorem A, in contrast, has no total computable positive decay-rate bound, as Theorem D makes precise. The Potts comparison uses the same witness, rewarded in both orientations; the rational critical odds are 3. Nine is the smallest perfect square greater than four, giving both rational odds and a rigorously discontinuous transition.

Both constructions have the same fine-tuning limitation. The onto branch lies on a critical or coexistence surface of codimension one in the visible coupling parameters. Decreasing both Potts odds below 3 makes every input subcritical, by the reward-free comparison. Thus arbitrarily small rational detuning can erase the dichotomy. The prescribed output-reset family preserves it because it changes only the collision retention and leaves that surface fixed. Uniform-replacement readout noise with retention $\eta > 0$ multiplies $\mu(X_0)$ by η, and the distinct-site excess equal-colour correlation by η^2; this preserves the corresponding distinctions for the postprocessed observables.

Full-state uniqueness in the nonsurjective Potts branch remains open. Proving it for every nonsurjective input would give undecidability of Gibbs uniqueness on this family, since the surjective branch already has nine distinct ordered Gibbs states. No such uniqueness theorem or full-clustering conclusion is claimed here.

The all-state argument is separate from the free-hidden finite-volume expansion. Uniform conditional uniqueness of the replicas provides local insertion ratios in arbitrary infinite bond configurations. Wired-core comparison then controls every Gibbs state without imposing full-state uniqueness in the nonsurjective branch.

The difficult remaining issue is the response of even and hidden observables. The counterexamples show why short magnetic clusters and conditional hidden contraction alone are insufficient. A successful completion must use the collision identities and their quantitative smallness.

Remark (Noetherian obstruction for a fixed collision template). Fix a finite template with finitely many rational bias parameters $z = (z_1, \dots, z_b)$, restricted to an explicitly decidable rational domain where all denominators below are nonzero. Suppose a branch is characterized by exact identities $R_j(z) = 0$, $j \geq 1$, with each R_j a rational function over $\mathbb{Q}$. Clearing denominators gives polynomials p_j. The Hilbert basis theorem [12, Chapter 2] makes the ideal $(p_1, p_2, \dots) \subset \mathbb{Q}[z]$ finitely generated. Moreover finitely many of the original p_j suffice: each of finitely many ideal generators is a finite polynomial combination of original members; collect those members. Consequently all identities hold exactly when this finite sublist vanishes. Exact rational (or algebraic) evaluation decides the branch. This is an existence-of-a-decider statement for each fixed template; it supplies no uniform procedure for finding that finite sublist from an infinite presentation.

For one fixed CA F, with iid input biases in finitely many parameters, every c_T is a rational polynomial in those biases: use the direct finite sum of integer-valued collision products. Apply the preceding argument to all nonempty T. A finite sublist fits in some box $[-N(F), N(F)]^d$ (the bias parametrization is fixed with F). Hence collision-nullity on all T is equivalent to nullity for all nonempty subsets of that finite box. No fixed CA with rational biases can encode an undecidable predicate solely in this exact null locus.

With fixed alphabet, dimension and range there are only finitely many deterministic

local rules, so the same obstruction covers that finite collection of collision templates. In a collision-nullity route with these restrictions, range must therefore grow with the input, as it does here. This conclusion concerns fixed finite templates and exact identities; it is not an obstruction to limit predicates, infinite inequalities, varying alphabets, or every possible fixed-range construction.

Remark (Uniform fields and raw decay; robustness sketches). For translation-invariant positive finite-range Ising-spin models, raw all-state spin decay has empty interior in any real parameter space allowing a uniform field $H \sum_x \sigma_x$. Here is a short proof. A translation-invariant Gibbs state with summable raw spin correlations has zero magnetization: its squared spatial-average mean is bounded by the second moment of that average, which tends to zero by summability. On any compact interval of fields, finite energy bounds single-spin conditional variances below by a constant $c > 0$. Choose a sublattice of positive density whose distinct sites occur in no common interaction term. Conditional on its complement these spins are independent. For $M_\Lambda = \sum_{x\in\Lambda} \sigma_x$, total variance gives $\mathrm{Var}(M_\Lambda) \geq c'|\Lambda|$ up to boundary terms. Therefore the finite-volume pressures have second derivative at least $c' > 0$, and their limit is strongly convex in H. Magnetizations of translation-invariant Gibbs states are its subgradients (the Gibbs variational principle [24, Chapter 15]). Strong convexity allows subgradient zero at at most one field. Since raw all-state decay forces that zero subgradient, it can hold at at most one point of each uniform-field line, proving empty interior. Fields correspond to arbitrary positive real weights; the interior assertion uses this real topology.

For contrast, the following robust-region statements are sketches based on standard mixing and low-temperature phase theory [22, Chapters 6–7]. Dobrushin's strict contraction condition gives an open region of exponential decay for connected correlations. It also gives raw decay for $D_x = \sigma_x - \sigma_{x+e_1}$, since uniqueness and translation invariance imply $\mu(D_x) = 0$. At low temperature the square-lattice checkerboard antiferromagnet has two translation-related phases throughout an open interval of small uniform fields. Their symmetric mixture has persistent staggered correlations and nonzero long-distance D correlations along even displacements, giving open regions of failure for the all-state connected and D-raw properties. Extension to small translation-invariant finite-range perturbations uses the Peierls/contour hypotheses of Pirogov–Sinai theory; that contour argument is not proved here. These sketches make no robustness claim about Theorem A.

Remark (Why computable-real weights make full clustering easy to encode). Enumerate finite Garden-of-Eden witnesses for a CA F by a fixed effective search, at integer times $T \geq 0$, and set

$$\delta_F = \begin{cases} 2^{-(T+2)}, & \text{the first witness appears at time } T, \\ 0, & \text{no witness appears.} \end{cases} \qquad e^{2K_H} = 3 - \delta_F, \quad e^{2K_V} = 2.$$

This is a uniformly computable real input: simulate through time n; if no witness has appeared, output zero with error at most $2^{-(n+3)}$. If one has appeared, its dyadic value is exact. Logarithms of these positive computable reals also give computable energy entries.

A witness exists exactly for nonsurjective F: a nononto continuous image of the compact full shift is closed and omits a cylinder; testing whether a finite pattern has a preimage is a finite local enumeration. If $\delta_F > 0$, the odds product is $2 - \delta_F < 2$, so this nearest-neighbor ferromagnet is strictly subcritical and has a unique Gibbs state with exponential clustering of all bounded local observables. To recall the full-clustering step, (I3) and the Edwards–Sokal boundary coupling bound influence on an interior set by the probability of an FK connection to the boundary; exponential connectivity makes this influence exponentially small. Conditioning a local observable on a distant exterior and integrating then gives exponential covariance bounds for arbitrary finite supports [22, Chapters 3 and 6]. If $\delta_F = 0$, the fixed

critical spin correlation already fails exponential decay. Thus deciding full clustering for these computable-real descriptions would decide CA surjectivity.

The numerical value $3-\delta_F$ is rational in each case, but its presentation here is a computable-real name: no algorithm outputs its exact rational numerator and denominator without deciding whether a witness will ever appear. The content of Theorem A is its finite *rational weight-table* input, not merely that the values happen to be rational.

How this work was done

The authors selected the questions and proof standards, directed the research, and take responsibility for every claim. AI systems from more than one developer assisted with exploration, derivation, programming and drafting.

The mathematical evidence consists of the proofs in this paper and the explicitly cited external theorems. The ancillary programs provide reproducible finite checks. For this paper, the construction and magnetic comparison were independently checked by a different model family. The accompanying suites include symbolic and exact finite computations as well as floating-point enumerations. A further exact-arithmetic checker tests collision identities, the finite covariance comparison and rational counterexample constants. The Potts extension also received an independent refutation-first check, whose six repairs are incorporated in the stated proofs and citations. Its finite checkers are preserved alongside the added exact Potts checks. The ancillary README gives the commands, outputs, dependencies and limitations of each suite.

These finite checks do not establish infinite-volume comparison, uniqueness, the KMS correspondence or undecidability by sampling. Those conclusions rest on the written proofs. The auxiliary arithmetic upper bound, fixed-template obstruction and L-model proposition were written out and checked during this revision; no additional independent-model audit of those new statements is documented. The robustness and energy-response passages retain their stated sketch or open status. This preprint has not been peer reviewed.

A Reproducibility

The accompanying `verification/` directory contains the two review suites, their recorded numerical outputs, fresh execution logs and an additional exact-arithmetic checker. Its README gives relative-path commands, dependency versions, provenance and limitations. `SHA256SUMS` identifies every public bundled file apart from itself. No public file depends on a machine-local path.

The first suite checks collision identities, reset coefficients, finite surjectivity controls, failure of the FKG lattice condition for the collision weight and conditional Dobrushin bounds. The second checks witness completion, the attractive comparator, covariance interpolation, wired boundary strips, insertion-ratio convergence and finite-cylinder Ising correlations. Its network-flow and linear-programming implementations enumerate finite state spaces but use floating-point arithmetic. The larger integer-scaled flow tests are screening calculations with an explicit rounding bound.

The additional checker uses rational arithmetic. It tests the sum-of-squares collision identity on a nonsurjective selector rule and an XOR control, witness-completion inequalities, finite FKG and stochastic comparisons, independent-reset scaling, and both rational counterexample certificates. The checker also computes exact nonzero single-site Dobrushin influences and their row sums over every boundary assignment for a two-factor binary example, and tests the comparator, both dominations and the covariance inequality on all 168 increasing events of a four-edge cycle (the 2×2 vertex box). Its finite Gibbs trace identity is an algebra sanity

check, not verification of the infinite-volume KMS correspondence. The runner writes new replay records to a fresh directory and leaves shipped hashes intact. These checks support finite algebraic steps. They do not replace the analytic arguments.

Build the supplied PDF and source from `paper/` with `python build.py`; the build rejects final TeX/BibTeX warnings and overfull or underfull boxes.

The Potts subdirectory adds the independent review's collision checker and its floating-point nine-colour FK checker, ported byte for byte. The additional `exact_q9_checks.py` compares direct nine-colour spin sums with wired bond sums using integers and fractions, obtaining the finite-volume magnetization identity exactly. It checks a real two-site collision witness in both orientations, and certifies both dominations for every increasing event of a seven-edge graph by integer-capacity monotone-coupling flow. Rational points along a nonconstant-odds interpolation verify the closure and anchor bounds. The exact endpoint test uses a slightly smaller admissible reward to keep all odds rational; the analytic proof above uses the full reward in (10.3). The full replay now contains 17 standalone jobs. These finite checks do not establish the infinite-volume Potts transition or pressure variational principle; those rely on the proofs and cited theorems.

References

[1] Michael Aizenman, David J. Barsky, and Roberto Fernández. The phase transition in a general class of Ising-type models is sharp. *Journal of Statistical Physics*, 47:343–374, 1987.

[2] Michael Aizenman and Geoffrey Grimmett. Strict monotonicity for critical points in percolation and ferromagnetic models. *Journal of Statistical Physics*, 63:817–835, 1991. doi:10.1007/BF01029985.

[3] Paul Balister, Béla Bollobás, and Oliver Riordan. Essential enhancements revisited, 2014. https://arxiv.org/abs/1402.0834.

[4] Johannes Bausch, Toby S. Cubitt, and James D. Watson. Uncomputability of phase diagrams. *Nature Communications*, 12:452, 2021. https://doi.org/10.1038/s41467-020-20504-6.

[5] Vincent Beffara and Hugo Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probability Theory and Related Fields*, 153:511–542, 2012. https://arxiv.org/abs/1006.5073.

[6] Carol E. Bezuidenhout, Geoffrey R. Grimmett, and Harry Kesten. Strict inequality for critical values of Potts models and random-cluster processes. *Communications in Mathematical Physics*, 158:1–16, 1993. doi:10.1007/BF02097229.

[7] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics 2*. Springer, second edition, 1997.

[8] Ana Bušić, Jean Mairesse, and Irène Marcovici. Probabilistic cellular automata, invariant measures, and perfect sampling. *Advances in Applied Probability*, 45(4):960–980, 2013.

[9] Jean-René Chazottes and Michael Hochman. On the zero-temperature limit of Gibbs states. *Communications in Mathematical Physics*, 297:265–281, 2010. https://arxiv.org/abs/0907.0081.

[10] Jean-René Chazottes and Mao Shinoda. On the absence of zero-temperature limit of equilibrium states for finite-range interactions on the lattice F^2, 2020. https://arxiv.org/abs/2010.08998.

[11] D. Cimasoni and H. Duminil-Copin. The critical temperature for the Ising model on planar doubly periodic graphs. *Electronic Journal of Probability*, 18:1–18, 2013. https://arxiv.org/abs/1209.0951; paper 44.

[12] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms*. Springer, fourth edition, 2015. Chapter 2: the Hilbert basis theorem.

[13] Toby S. Cubitt, David Pérez-García, and Michael M. Wolf. Undecidability of the spectral gap. *Nature*, 528:207–211, 2015. https://doi.org/10.1038/nature16059.

[14] R. L. Dobrushin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability and its Applications*, 13:197–224, 1968. https://doi.org/10.1137/1113026.

[15] H. Duminil-Copin and M. Lis. On the double random current nesting field. *Probability Theory and Related Fields*, 175:937–955, 2019. https://doi.org/10.1007/s00440-019-00899-0.

[16] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for bernoulli percolation and the Ising model. *Communications in Mathematical Physics*, 343:725–745, 2016. https://arxiv.org/abs/1502.03050; arXiv v3 numbering: Theorem 2.1 and Proposition 2.2.

[17] Hugo Duminil-Copin, Maxime Gagnebin, Matan Harel, Ioan Manolescu, and Vincent Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *Annales scientifiques de l'École normale supérieure*, 54:1363–1413, 2021. https://arxiv.org/abs/1611.09877; Theorems 1.1(2) and 1.2(2).

[18] R. G. Edwards and A. D. Sokal. Generalization of the fortuin–kasteleyn–swendsen–wang representation and monte carlo algorithm. *Physical Review D*, 38:2009–2012, 1988. https://doi.org/10.1103/PhysRevD.38.2009.

[19] H. Fawzi, O. Fawzi, and S. O. Scalet. Certified algorithms for equilibrium states of local quantum hamiltonians. *Nature Communications*, 15:7394, 2024. https://arxiv.org/abs/2311.18706.

[20] R. Fernández and A. Procacci. Cluster expansion for abstract polymer models. new bounds from an old approach. *Communications in Mathematical Physics*, 274:123–140, 2007. https://arxiv.org/abs/math-ph/0605041.

[21] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, 22:89–103, 1971.

[22] Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, 2017. https://doi.org/10.1017/9781316882603.

[23] L. Gayral, Mathieu Sablik, and Siamak Taati. Characterisation of the set of ground states of uniformly chaotic finite-range lattice models, 2023. https://arxiv.org/abs/2302.07326.

[24] Hans-Otto Georgii. *Gibbs Measures and Phase Transitions*. De Gruyter, second edition, 2011. Proposition 8.8; Theorems 15.12, 15.30, 15.37, 15.39; Proposition 15.14, https://doi.org/10.1515/9783110250329.

[25] Sheldon Goldstein, Roelof Kuik, Joel L. Lebowitz, and Christian Maes. From PCA's to equilibrium systems and back. *Communications in Mathematical Physics*, 125:71–79, 1989. https://doi.org/10.1007/BF01217769.

[26] Geoffrey Grimmett. *The Random-Cluster Model*, volume 333 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2006. Corrected reprint 2009; https://www.statslab.cam.ac.uk/~grg1000/books/rcm1-1.pdf.

[27] Mile Gu, Christian Weedbrook, Alvaro Perales, and Michael A. Nielsen. More really is different. *Physica D*, 238:835–839, 2009. https://arxiv.org/abs/0809.0151.

[28] Theodore E. Harris. A lower bound for the critical probability in a certain percolation process. *Proceedings of the Cambridge Philosophical Society*, 56:13–20, 1960.

[29] Gustav A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Mathematical Systems Theory*, 3:320–375, 1969. doi:10.1007/BF01691062.

[30] R. Holley. Remarks on the FKG inequalities. *Communications in Mathematical Physics*, 36:227–231, 1974.

[31] J. Kari. Reversibility and surjectivity problems of cellular automata. *Journal of Computer and System Sciences*, 48:149–182, 1994. https://doi.org/10.1016/S0022-0000(05)80025-X.

[32] Jarkko Kari and Siamak Taati. Statistical mechanics of surjective cellular automata. *Journal of Statistical Physics*, 160:1198–1243, 2015. https://doi.org/10.1007/s10955-015-1281-2.

[33] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103:491–498, 1986. https://doi.org/10.1007/BF01211762.

[34] Joel L. Lebowitz, Christian Maes, and Eugene R. Speer. Statistical mechanics of probabilistic cellular automata. *Journal of Statistical Physics*, 59:117–170, 1990. https://doi.org/10.1007/BF01015566.

[35] Irène Marcovici and Siamak Taati. Positive-rate PCA and IPS with stationary bernoulli measures are rapidly forgetful, 2026. https://arxiv.org/abs/2605.16904.

[36] Akira Maruoka and Masayuki Kimura. Condition for injectivity of global maps for tessellation automata. *Information and Control*, 32:158–162, 1976. doi:10.1016/S0019-9958(76)90195-9.

[37] A. Perales-Eceiza, T. Cubitt, M. Gu, D. Pérez-García, and M. M. Wolf. Undecidability in physics: a review. *Physics Reports*, 1138:1–29, 2025. doi:10.1016/j.physrep.2025.06.004; https://arxiv.org/abs/2410.16532, Section 5.1.4, arXiv v2.

[38] James Purcell, Zhi Li, and Toby Cubitt. Chaitin phase transition, 2024. https://arxiv.org/abs/2410.02600.

[39] P. Rebeschini and R. van Handel. Comparison theorems for Gibbs measures. *Journal of Statistical Physics*, 157:234–281, 2014. https://arxiv.org/abs/1308.4117.

[40] Barry Simon. Correlation inequalities and the decay of correlations in ferromagnets. *Communications in Mathematical Physics*, 77:111–126, 1980. https://doi.org/10.1007/BF01982711.

[41] Barry Simon. *The Statistical Mechanics of Lattice Gases*, volume I. Princeton University Press, 1993. Section III.5.

[42] Andrei Toom. Algorithmical unsolvability of the ergodicity problem for binary cellular automata. *Markov Processes and Related Fields*, 6(4):569–577, 2000.