Forgetting under noise is undecidable

Seth Douglas Nidhal Mghirbi

Version 3.0.0, November 2026

Abstract

Can ongoing positive noise make the long-term behaviour of a finite local rule decidable? A natural noisy Wang rule is ergodic exactly when its raw tile set does not tile the plane, at every replacement rate up to 10^{-24} on $\mathbb{Z}^6$ and 10^{-30} on $\mathbb{Z}^4$. A related construction on $\mathbb{Z}^3$ gives the same equivalence for north/east deterministic tile sets. On non-tiling inputs the lower-dimensional all-input nonconstancy probability on a finite window F is at most $|F|C_R \exp(-t/24000)$, with an effectively findable input-dependent prefactor uniform in the rate; the exponent $1/4$ is retained for rates up to 10^{-4535}. On those inputs their normalized space–time interactions have uniqueness among all Gibbs measures and exponential weak spatial mixing. Ergodicity, Gibbs uniqueness and exponential weak spatial mixing are therefore undecidable at one fixed rational rate. A different family on $\mathbb{Z}^6$, of fixed range 11, makes uniqueness among translation-invariant stationary laws Π_2^0-complete at the single rational rate 10^{-22}, under a stated published-source condition C_T; the corresponding translation-invariant Gibbs problem is Π_2^0-complete on $\mathbb{Z}^7$. The enlarged intervals are paper proofs with exact arithmetic. Lean covers only the version 2.2.0 intervals of the retained natural-rule cores; the machine-family core is unchanged. Finite forbidden patterns give forgetting bounds, while infinite tilings protect distinct stationary laws.

AI-assisted derivation and checking; the authors' verification is described in "How this work was done".

Contents

1 Introduction **4**
1.1 Context and the quantitative distinction 6
1.2 Consequences and the retained second route 6

2 The natural Wang rule and its proof **9**
2.1 Proof architecture 9
2.2 Tiles, configurations and noise 10
2.3 Five fixed provider boxes 10
2.4 The geometry behind the two budgets 11
2.5 Three elementary inputs 12
2.6 The all-input coefficient and the theorem 13
2.7 A marked field determined only by flags 14
2.8 Five balanced charges 15
2.9 Counting flag positions 18
2.10 Packing contours along a marked path 20
2.11 Successful and unaffected providers 21
2.12 Same-site coherence at finite age 21
2.13 Cardinal compatibility at finite age 22

2.14 Finite certificates of non-tiling . . . 23
2.15 Deterministic mortality outside the marked field . . . 23
2.16 Age, span and planar growth in one picture . . . 24
2.17 Tracing all possible discrepancies at once . . . 25
2.18 From common maps to ergodicity . . . 26
2.19 A tiling protects two reference phases . . . 27
2.20 Fully translation-invariant stationary laws . . . 28
2.21 Complete histories: the qualitative route . . . 28
2.22 Completion and dependence of the estimates . . . 29

3 Consequences for the natural rule **30**
3.1 The domino reduction and nilpotency . . . 30
3.2 All space–time Gibbs measures . . . 31
3.3 Robustness on the same neighbourhood . . . 32
3.4 Exact complexity on this family . . . 34
3.5 Other replacement laws and the forgetting time . . . 34
3.6 Canonical phases and the conditional continuum . . . 35
3.7 The stack on $\mathbb{Z}^7$ . . . 37
3.8 Why this mechanism uses four transverse dimensions . . . 38

4 The model and the token game **40**
4.1 Wang sets, blocks and the checker . . . 40
4.2 The quorum rule and replacement noise . . . 41
4.3 Ergodicity, forgetting and memory . . . 42
4.4 Token tables . . . 44
4.5 The token game . . . 45
4.6 One token, and finite token width . . . 48

5 Bounded-width refutations give forgetting **49**
5.1 Setting and statement . . . 50
5.2 Instant rules and transport . . . 51
5.3 A covering field computed from the noise . . . 52
5.4 Contours . . . 53
5.5 Packing contours along a discrepancy path . . . 57
5.6 Counting, and the proof of the theorem . . . 57
5.7 All starts at once . . . 58

6 Tilings give memory **59**
6.1 The rule and the cloned tiling . . . 60
6.2 Two bad parents . . . 61
6.3 Domination by the binary field . . . 61
6.4 Counting, and the proof of the theorem . . . 62

7 A hierarchical scaffold read by a fixed number of tokens **64**
7.1 Ollinger's substitution . . . 64
7.2 The ruler decoration . . . 67
7.3 Reading a block from a fixed number of sites . . . 69
7.4 Anchors and certificates . . . 70
7.5 Transfer to pebble families, and anchors of every depth . . . 74

8 Routing a machine onto the scaffold **76**
8.1 Towers and red paths . . . 77
8.2 Quarters, walls, flows and roles . . . 79
8.3 Frames and free coordinates . . . 82
8.4 A routing model on the plane . . . 83
8.5 The initialized board in every word . . . 84
8.6 The pointed Horn layer . . . 85
8.7 The family . . . 88

9 Every halting member has a bounded-width refutation **89**
9.1 Inputs from the setting, scaffold and family . . . 89
9.2 One adjusted quarter phase . . . 91
9.3 The index family . . . 92
9.4 Pins and the forced word . . . 92
9.5 The forced role chart . . . 93
9.6 The selected component . . . 93
9.7 The projected meet . . . 94
9.8 Compression and the contradiction . . . 94
9.9 The token budget . . . 95

10 The machine-family theorem and its corollaries **95**
10.1 Corollaries . . . 96
10.2 Perturbed kernels . . . 102
10.3 Stacked automata: one phase or a continuum . . . 107

11 Three tokens and dimension eight **111**
11.1 Lists read by one moving token . . . 112
11.2 Arc consistency on the scaffold . . . 113
11.3 Corner pairs at every scale . . . 118
11.4 The forced word and the contradiction . . . 120
11.5 The three tokens . . . 122
11.6 Dimension eight . . . 123

12 Lower dimensions **124**
12.1 Why the budget-only estimate stops at six . . . 124
12.2 Three boxes and an all-path density lemma . . . 126
12.3 The natural rule on four spatial dimensions . . . 129
12.4 Three dimensions via Kari's two-input generator . . . 131
12.5 Exponential all-input coalescence for non-tiling inputs . . . 134
12.6 All space–time Gibbs measures in lower dimensions . . . 137

13 Open problems **140**

14 The degree of translation-invariant uniqueness **142**
14.1 The source and its precise dependency . . . 142
14.2 A finite checker and five providers . . . 143
14.3 The two noise estimates . . . 143
14.4 Clean coherence and exhaustion on totality . . . 145
14.5 Two stationary laws on nontotality . . . 146

14.6 Arithmetic upper bounds and the Gibbs consequence 147

How this work was done **148**

Verification status **151**

A Tables for the scaffold **153**

B The one-sided machine **156**

C Formal proofs in Lean 4 **158**

D Proof of the totality source **165**

References **167**

1 Introduction

A finite set of Wang tiles determines a natural noisy rule: check the four edges around each tile, then retain a tile only when a fixed quorum of nearby copies agrees on it. Our main equivalence is

$$R \text{ tiles } \mathbb{Z}^2 \quad \Longleftrightarrow \quad \text{the noisy rule } G_R \text{ remembers its start.}$$

The input letters are the tiles of R themselves, together with one blank. The same local formula and the same positive interval of noise rates work for every finite R. No machine, hierarchy, check bit or block alphabet is inserted between the input and the dynamics.

Here "ergodic" means that there is exactly one invariant probability measure and that the law from every initial probability measure converges weakly to it. This includes spatially inhomogeneous starts. Noise replaces the whole output letter independently at each site and time, uniformly over the finite alphabet. Every output letter then has positive probability. A fixed rate means a fixed replacement probability, not a fixed lower bound on all table entries as the alphabet grows.

Theorem D (Natural Wang rule; Theorem 2.1). Put $K = 16^4$, $q = 5K$, $h = 4K$ and

$$\varepsilon_* = 10^{-24}.$$

For every finite Wang set R, a cardinal checker followed by an h-of-q quorum in five fixed transverse boxes of K offsets each defines a rule G_R on $\mathbb{Z}^6$ with alphabet $R \sqcup \{\perp\}$. At every $0 < \varepsilon \le \varepsilon_*$, uniform whole-symbol replacement after G_R is ergodic if and only if R does not tile the plane. If R tiles, there are two distinct translation-invariant stationary laws.

This answers Problem 10.3 of version 1.4.0 of this paper. The rule is the clustered data checker of our noise-diagram paper [12, Section 5], simplified to raw tiles and cardinal checks. The proof below is self-contained: it gives the finite charge-arm estimate, coherence, finite forbidden-square delay, all-input coupling and memory. Its antecedents are [12, Lemmas 5.1–5.3 and 8.5, Theorem 9.5 and Appendix A]; the new point is the classification of the bare rule on arbitrary raw Wang sets. Theorem 9.5 and Appendix A of that paper belong to its version 1.1.0 package directed by Nidhal Mghirbi.

Theorem A (Fixed-rate undecidability on $\mathbb{Z}^6$; Corollary 3.1). There is a total computable map $M \mapsto W_M$ from Turing machines to finite Wang sets such that, for every $\varepsilon \in (0, \varepsilon_*]$, the positive-rate automaton obtained from G_{W_M} is ergodic if and only if M halts.

Theorem E (Natural Wang rule on $\mathbb{Z}^4$; Theorem 12.5). Put $K_4 = 16^2$, $q_4 = 3K_4 = 768$, $h_4 = 2K_4 = 512$ and

$$e_0 = 10^{-30}.$$

For every finite raw Wang set R, the cardinal checker followed by an h_4-of-q_4 quorum in the three transverse boxes $\{0, \dots, 15\}^2$ at $(0,0)$, $(160,0)$ and $(0,160)$ defines $G_R^{(4)}$ on $\mathbb{Z}^2 \times \mathbb{Z}^2 = \mathbb{Z}^4$ with alphabet $R \sqcup \{\perp\}$. At every $0 < \varepsilon \le e_0$, uniform whole-symbol replacement after this rule is ergodic if and only if R does not tile the plane. Tiling gives two distinct invariant laws; non-tiling gives finite-window all-start coupling from the past and attraction from every initial law. The neighbourhood $\Gamma \times S$ and the interval are common to every R.

Theorem F (Fixed-rate undecidability on $\mathbb{Z}^3$; Theorem 12.11). There is a total computable family indexed by all finite north/east deterministic Wang sets T, with alphabet $T \sqcup \{\perp\}$ and the common neighbourhood $\{0,1\} \times S \subset \mathbb{Z}^3$, where S is the same three-box set as in Theorem E. Each provider applies the two-input matching function f_T of Section 12, followed by the same 512-of-768 quorum. At every $0 < \varepsilon \le e_0 = 10^{-30}$, uniform whole-symbol replacement gives ergodicity if and only if T does not tile the plane. Tiling gives two distinct invariant laws; non-tiling gives finite-window all-start coupling from the past and attraction from every initial law. Ergodicity is undecidable across this family at the single fixed rational rate e_0.

Theorem F is a noise-robust transverse-quorum lift of Kari's 1992 nilpotency reduction [26, 51]. Its input restriction is essential to the stated construction; Theorem E applies to every raw Wang set. Their common argument uses an all-path marked-parent density lemma and amortized coherence. For non-tiling inputs, both lower-dimensional families now have an exponential all-input coalescence rate with universal exponent $1/24000$, and $1/4$ for $\varepsilon \le 10^{-4535}$, and an effectively findable input-dependent prefactor (Theorem 12.12). Corollary 12.14 applies the space–time transfer of [9, Corollary 13.7] to these common maps: the normalized interactions on $\mathbb{Z}^5$ and $\mathbb{Z}^4$ have all-Gibbs uniqueness and arbitrary-volume exponential weak spatial mixing exactly on non-tiling inputs. Tiling inputs have at least two fully translation-invariant Gibbs measures. The qualitative $\mathbb{Z}^4$ classification of Theorem 12.5 (Theorem E) and its undecidability corollaries have a Lean core (Section C) on the version 2.2.0 interval $0 < \varepsilon \le 382205952^{-33720}$. The enlarged interval is a paper proof with exact arithmetic. The $\mathbb{Z}^3$ classification, the lower-dimensional rates and Corollary 12.14 are not formalized; the robustness theorem and the scope of the $\mathbb{Z}^6$ Lean core are not transferred to these families.

The only undecidability input in the $\mathbb{Z}^6$ route is the classical domino reduction of Berger and Robinson [3, 43]. Undecidability then follows by the identity reduction on raw tile sets. The classification's enlarged interval uses explicit rational inequalities certified by exact arithmetic in `verification-thresholds/`; it does not use the corner-configuration computation of Theorem C, which belongs to the retained machine route. Its explicit interval is now much larger than the retained machine route's: $\log_{10} \varepsilon_* = -24$, compared with $\log_{10}(490000^{-99}) \approx -563.3$. Both sentences of Theorem D and Theorem A on $\mathbb{Z}^6$ are verified in Lean, without a named hypothesis, on the version 2.2.0 interval $0 < \varepsilon \le (2(18q)^2)^{-200}$; the present interval $0 < \varepsilon \le 10^{-24}$ is proved on paper with exact arithmetic. The domino reduction itself is proved there through our family R_M (Corollary 3.1 and Section C). The existing $\mathbb{Z}^8$ core is retained.

1.1 Context and the quantitative distinction

Kari's undecidability of nilpotency [26], following the tiling-to-nilpotency reduction of Čulik, Pachl and Yu [7], is the noise-free analogue. Indeed G_R itself is nilpotent exactly when R does not tile (Corollary 3.2). Bušić, Mairesse and Marcovici identify ergodicity and nilpotency for deterministic one-dimensional cellular automata [5, Theorem 3.3]. Marcovici, Sablik and Taati study noisy nilpotent automata, with a noise threshold depending on the automaton [34]. The uniform threshold over an undecidable raw-input family is the distinction here. Our nilpotent-memory example [11] also warns that nilpotency alone does not imply forgetting at every positive rate.

Kurdyumov [29] and Toom [48] proved undecidability of ergodicity with forbidden transitions. Toom's stable rule [47] and Gács's reliable automata [18] show that positive noise can preserve memory. The contour proof used here follows Toom's method in the charge-arm form of Berman–Simon and Gács [4, 19]; see also [46]. Marsan, Sablik and Törmä exhibit two ergodicity phase transitions for a perturbed automaton [36]. Marcovici and Taati study positive-rate rules with stationary Bernoulli measures [35].

For non-tiling R, let n_R witness that no legal $n_R \times n_R$ pattern exists. With $\kappa = 1/8$, our estimate is

$$a_t \le \frac{\exp\{c(762n_R + 1 - t)\}}{1-\kappa}, \qquad c = \frac{\log(1/\kappa)}{16} = \frac{3}{16}\log 2.$$

Here a_t tests one realization of the random maps against *all* initial configurations at once. The slope c is universal and computable; the forbidden square controls only the delay. No total computable function of R bounds the forgetting time. This is the same separation between a refutation-dependent delay and member-independent decay already visible after Corollary 10.7, now without a bounded-width hypothesis.

1.2 Consequences and the retained second route

On the $\mathbb{Z}^6$ natural family, non-tiling, nilpotency, noisy ergodicity, exponential all-input coalescence, all-DLR uniqueness and arbitrary-volume exponential weak spatial mixing coincide (Section 3). At every fixed permitted rational rate the three decision properties ergodicity, Gibbs uniqueness and exponential weak mixing are Σ_1^0-complete; their complements are Π_1^0-complete. A total-variation ball of radius ε_* on the same alphabet and neighbourhood preserves both dynamical labels. Any configuration-independent replacement law works dynamically; the Gibbs conclusions require full support. An extra stacking coordinate gives an unconditional one-measure-or-continuum alternative on $\mathbb{Z}^7$. For G_R itself the continuum conclusion is proved when its tiling space is uncountable; countable tiling spaces remain open.

The inspector theory explains why the earlier thin-provider rule required more work. Its two halves remain unchanged.

Theorem B (Theorems 5.1 and 6.2). Fix $j \ge 1$, $k \ge \max(2, 2j)$ and $0 < \varepsilon \le \varepsilon_k$, with the explicit rate of Remark 5.2. For every finite Wang set R, the thin-provider quorum rule remembers a bit forever and is not ergodic if R tiles; if R is j-token refutable, it forgets its initial configuration exponentially fast after a delay set by the refutation, and is ergodic.

A token is a local inspector retaining a legal 3×3 block. Token refutability is a bounded-width local-consistency condition related to existential pebble games (Theorem 4.25). Thick provider boxes synchronize any finite transverse span by induction; single providers cannot pay that overlap budget. The resulting budget-only dimension obstruction is Toom's eroder

criterion with Radon's theorem for this mechanism (Remark 3.12); Theorem 12.1 and Corollary 12.2 strengthen it to actual saturation for the stated cover class. The new density argument evades that exclusion certificate.

Theorem C (Theorems 9.8 and 11.1). There is a total computable family $M \mapsto R_M$ such that R_M tiles if and only if M does not halt, and every halting member is 3-token refutable.

Together with Theorem B this retains the $\mathbb{Z}^8$ machine-family theorem at every $0 < \varepsilon \leq 490000^{-99}$, including its formally verified core. The hierarchical scaffold, Horn computation and three-token argument are all retained, as are the old robustness, Gibbs, timing and $\mathbb{Z}^9$ stacking results. A two-token family remains an open problem about that second route.

Changes from version 1.4.0. The natural rule and its complete proof now lead, followed by its consequences. The inspector theory and machine construction move behind them; every old proved statement is retained with its content. New results are the direct raw-Wang classification on $\mathbb{Z}^6$, its quantitative forbidden-square proof, nilpotency, the transfer consequences of Section 3 and the $\mathbb{Z}^7$ stack. Problems 10.1 and 10.2 of version 1.4.0 are restated for the new bounds, Problem 10.3 is answered and replaced by lower-dimensional instances, and a problem on countable tiling spaces is added. The old Theorem A is now called the machine-family theorem (Theorem 10.2 and Corollary 11.21); the recoded map of Corollary 10.5 is renamed H_R, because G_R now denotes the natural rule, and the sentence after Corollary 10.5 is updated. No proved statement of version 1.4.0 is corrected. The scopes of the new results are explicit: robustness fixes the neighbourhood (Remark 3.6), the $\{0,1\}$-energy form needs an inverse temperature depending on R (Remark 3.4), nonergodicity is Π_1^0-complete (Corollary 3.7), and only the delay, not the exponential slope, is uncomputable (Corollary 3.9). The added lower-dimensional results are the natural classification on $\mathbb{Z}^4$ and fixed-rate undecidability on $\mathbb{Z}^3$, with complete qualitative proofs in Section 12. They use amortized coherence instead of the budget-only estimate. The dimension problems are restated accordingly; stronger forms and token width two remain open. The existing $\mathbb{Z}^8$ Lean core is retained. On the version 2.2.0 interval $0 < \varepsilon \leq (2(18q)^2)^{-200}$, a new $\mathbb{Z}^6$ core verifies both sentences of Theorem D, Theorem A, all of Corollary 3.1, the domino reduction and the undecidability of the domino problem, without named hypotheses; the coalescence bound, nilpotency and the other consequences of Section 3 remain outside its scope (Section C).

Changes from version 2.0.0. Theorem 12.12, appended at the end of Section 12, adds exponential all-input coalescence for every non-tiling input in either lower-dimensional family, uniformly for $0 < \varepsilon \leq 382205952^{-33720}$. The exponent is $1/4$; only an effectively findable prefactor depends on the input. Example 12.13 shows why non-tiling is essential. Theorems E and F, the abstract and verification status now include this rate, and the corresponding rate question is removed from the open problems. Every existing numbered result and equation keeps its number. The Lean appendix and existing formalizations are unchanged; the new rate is not formalized. No lower-dimensional all-Gibbs uniqueness or robustness claim is added.

Changes from version 2.1.0. Corollary 12.14, added after Example 12.13, applies [9, Corollary 13.7] to the lower-dimensional all-input bound. It adds all-Gibbs uniqueness and arbitrary-volume exponential weak spatial mixing on $\mathbb{Z}^5$ and $\mathbb{Z}^4$ for non-tiling inputs, two fully translation-invariant Gibbs measures on tiling inputs, and fixed-rate undecidability. The qualitative $\mathbb{Z}^4$ classification and its undecidability corollaries now have the `lean-lower/` core

(Section C). Existing statements and equation numbers are preserved; scope disclaimers are updated. The version 2.1.0 change note above is historical. The $\mathbb{Z}^3$ classification, the lower-dimensional rates and the Gibbs corollary are not formalized; lower-dimensional robustness is not claimed.

Changes in 3.0.0. The natural-rule noise thresholds become 10^{-24} on $\mathbb{Z}^6$ and 10^{-30} on $\mathbb{Z}^4$ and $\mathbb{Z}^3$. A count of distinct flag positions using polynomial volume growth, together with improved arm accounting, gives these intervals with the same rules and provider geometry. The lower-dimensional exponent is $1/24000$, with $1/4$ retained for $\varepsilon \le 10^{-4535}$; the $\mathbb{Z}^6$ memory error bound becomes $p_* = 10^{-4}$ instead of 2ε (the bound 2ε is recovered for $\varepsilon \le 10^{-40}$), and the $\mathbb{Z}^6$ coalescence slope becomes $(3/16)\log 2$ on the whole interval. The three-box setup of Section 12 now requires logical offsets in $\{-1, 0, 1\}^r$, which all its rules satisfy. The formal developments cover only the version 2.2.0 intervals. Building on the same $\mathbb{Z}^6$ family and totality dichotomy at 160000^{-101} in [52, Lemma 6.1], whose Theorem C gives the $\mathbb{Z}^7$ stationary degree, Section 14 extends the rate interval to $(0, 10^{-22}]$, writes the proof in full, and explicitly proves Π^0_2-completeness of translation-invariant stationary uniqueness on $\mathbb{Z}^6$ at fixed rational rates through 10^{-22}, and of the associated translation-invariant Gibbs uniqueness on $\mathbb{Z}^7$, under its stated source condition. It answers the translation-invariant variant of Open problem 13.2. The all-law ergodicity question remains open. Sandhu's independent fixed-temperature Gibbs result and the effective-equilibrium work of Gayral, Sablik and Taati are credited below.

Translation-invariant uniqueness. The raw-Wang families above have first-level arithmetic degrees. Under a stated published-source condition C_T (Lemma 14.2), a different family on $\mathbb{Z}^6$, of fixed range 11, makes uniqueness among spatially translation-invariant stationary laws Π^0_2-complete at each fixed rational rate in $(0, 10^{-22}]$ (Theorem 14.1); the corresponding translation-invariant Gibbs problem is Π^0_2-complete on $\mathbb{Z}^7$. At the rate 160000^{-101} the same family and dichotomy appear in our Rice paper [52, Lemma 6.1 and Theorem C]. Törmä's deterministic unique-ergodicity degree [55, Proposition 48] and Sandhu's fixed-temperature family [54] are the closest independent comparisons. The proof controls every translation-invariant stationary law, not uniqueness among all stationary laws.

Reading guide. Section 2 proves Theorem D without using any later section; Section 3 gives its consequences. Sections 4 to 6 contain the inspector theory. Sections 7 to 11 give the formally checked second route and its further results. Section 12 proves the saturation theorem and the lower-dimensional classifications. Open problems are in Section 13; Section 14 gives the translation-invariant degree and can be read with the source lemma as its only self-simulation input; Section D gives that lemma's proof. The tables, one-sided machine and precise Lean scope remain in Sections A to C.

2 The natural Wang rule and its proof

This section proves the dynamical equivalence on arbitrary raw Wang inputs. No result about token tables, machine simulation or a computed finite configuration is used. We first specify the rule, then prove one flag-only contour estimate. Two deterministic uses of that estimate finish the argument: coherence turns an old unmarked live region into a forbidden finite tiling, while a genuine tiling protects a live phase against the same marked field. All probabilities below concern independent update randomness; no averaging over tile sets is involved.

2.1 Proof architecture

There are three objects in the argument and two different uses of compactness. The deterministic rule reads tile letters. The random maps couple arbitrary starts. The marked field forgets the letters entirely and records only where a two-cluster explanation can end in flags. We prove estimates for that field before asking whether the tiles tile. This order makes the noise threshold independent of the input and prevents a hidden union over the input alphabet.

For a non-tiling input the logical certificate is finite: one square has no legal filling. Coherence first turns transverse comparisons into ordinary matching equations. Backward growth then produces a legal filling of that square from any sufficiently old unmarked live output. This deterministic contradiction places every late live output inside the marked field. Only at the last step do we compare two arbitrary starts: any discrepancy must contain a live letter in one of them and therefore follows a marked path. The path-tail estimate controls the event that *some* pair could differ.

For a tiling input no forbidden square is available. Instead the tiling supplies a fixed reference configuration. A wrong output requires more failed providers than a single cluster can hold. Its error set is thus covered by the very same marked field. The blank reference has its own error cover. Sparse one-point marks, rather than a path-tail estimate, are enough to keep their one-site densities apart forever.

The first compactness use is logical: legal finite squares of all sizes would yield a plane tiling, so non-tiling has a finite witness. The second is probabilistic: Cesàro limits give stationary laws on a compact configuration space. Neither compactness argument asserts a computable bound on when a limit or certificate appears. The quantitative information comes separately from finite-age coherence and the finite-contour counting.

Ingredient	Non-tiling direction	Tiling direction
Finite certificate or reference	forbidden n_R-square	a plane tiling θ
Deterministic use of the quorum	coherence and backward growth	two-cluster error cover
Required flag estimate	tail of marked paths	probability of a marked root
Stationary conclusion	one law attracting all starts	two laws with a fixed density gap
Input-dependent quantity	delay $762n_R + 1$	reference letters only

Table 1: The two uses of one marked field. The contour constants and the permitted noise interval are the same in both columns.

2.2 Tiles, configurations and noise

A Wang tile is an oriented unit square with four named edge colours. A finite Wang set R tiles the plane if there is $\theta : \mathbb{Z}^2 \to R$ whose shared horizontal and vertical edge colours match. Translations are allowed, rotations and reflections are not. The tiling space $\mathcal{T}(R)$ is the closed subset of $R^{\mathbb{Z}^2}$ given by these matching equations. A finite legal pattern need only match its internal edges; no extendibility beyond its domain is assumed. Keeping this distinction is essential: a non-tiling set can have very large legal finite patterns.

The alphabet is

$$A_R = R \sqcup \{\perp\}, \qquad \Omega_R = A_R^{\mathbb{Z}^6}.$$

The new symbol $\perp$ is blank; every other letter is *live*. We write a spatial site as $v = (u, P)$, where $u \in \mathbb{Z}^2$ is planar and $P \in \mathbb{Z}^4$ is transverse. The product topology of discrete alphabets makes Ω_R compact and metrizable. A cylinder event reads finitely many sites; it is both open and closed. A probability measure is stationary if applying one update leaves it unchanged. Spatial translation invariance always means invariance under every translation of $\mathbb{Z}^6$, unless a smaller group is explicitly named.

Let

$$\Gamma = \{0, e_1, -e_1, e_2, -e_2\}, \qquad \Lambda = \{-1, 0, 1\}^2.$$

The cardinal checker C_R applied in a planar layer retains the central tile precisely when all five letters at $u + \Gamma$ are live and the four shared edges at u match. Otherwise it outputs $\perp$. A successful check therefore retains an *original* tile of R. It introduces no bit, block presentation or auxiliary symbol.

Given a deterministic finite-range rule G_R, use independent variables

$$\eta_t(v) \sim \text{Bernoulli}(\varepsilon), \qquad \zeta_t(v) \sim \text{Unif}(A_R)$$

at every site and update time. The two fields are independent. From a start x at time zero,

$$X_t^x(v) = \begin{cases} \zeta_t(v), & \eta_t(v) = 1, \\ G_R(X_{t-1}^x)(v), & \eta_t(v) = 0. \end{cases} \tag{1}$$

Every start uses the same fields. An indicated replacement is charged as a possible error even when its value happens to be correct. For $m = |A_R|$ the stochastic table is

$$p_{R,\varepsilon}(a \mid y) = (1 - \varepsilon)\mathbf{1}_{\{a = G_R(y)\}} + \varepsilon/m. \tag{2}$$

Thus every entry is at least $\varepsilon/m > 0$. The operation replaces the whole symbol after the deterministic update.

2.3 Five fixed provider boxes

In the transverse lattice put

$$K = 16^4, \qquad C_0 = 0, \qquad C_i = 160e_i \quad (1 \le i \le 4),$$
$$S_i = C_i + \{0, \dots, 15\}^4, \qquad S = \bigcup_{i=0}^{4} S_i, \qquad q = 5K, \quad h = 4K. \tag{3}$$

The boxes are disjoint. Each contains K offsets, and S contains q offsets. We call the S_i clusters. Set

$$y(u, Q) = C_R(x(\cdot, Q))(u).$$

The clustered quorum is

$$G_R(x)(u,P) = \begin{cases} s, & s \in R, \quad |\{a \in S : y(u,P+a) = s\}| \geq h, \\ \perp, & \text{if no such live } s \text{ exists.} \end{cases} \tag{4}$$

Two distinct live letters cannot both qualify, since $2h = 8K > 5K = q$. Blank readings do not vote. All boundary equalities in this definition are intentional: exactly h agreeing readings suffice.

The actual neighbourhood is $\Gamma \times S$. We also use the fixed enclosing neighbourhood

$$\overline{N} = \Lambda \times S. \tag{5}$$

The latter contains zero; redundant inputs are ignored in (4). The geometry is fixed before R: only the alphabet and its four matching tests vary. Exhaustive enumeration of parent words computes the finite table from R. Its size is very large but finite, which is all that computability requires.

For estimates put

$$\varepsilon_* = 10^{-24}, \qquad \kappa = \frac{1}{8}, \qquad c = \frac{\log(1/\kappa)}{16} = \frac{3}{16}\log 2. \tag{6}$$

These are effective constants independent of R. They come from the provider geometry alone, through the arm coefficient 16 of Lemma 2.3 and the volume growth of the space–time graph in Lemma 2.3A; the rational inequality behind κ is displayed after (16). Numerical sizes are orientation only; every estimate follows from displayed rational inequalities. (Version 2.2.0 used $\varepsilon_* = (2(18q)^2)^{-200}$, $\kappa = 4B^{-99}$ and $c = \log(1/\kappa)/50$ with $B = 2(18q)^2$. The new interval contains the old one, so every statement proved here for $0 < \varepsilon \leq \varepsilon_*$ holds in particular on the old interval. The Lean statements of Section C are stated with the old value and cover only the old interval.)

2.4 The geometry behind the two budgets

The boxes serve two jobs. Their separation supplies affine charges with positive drift away from one bad cluster per charge. Their thickness supplies many offset pairs one lattice step apart. Only the second job is used for coherence; only the first is used to count contours. A large separation cannot compensate for an overlap deficit, and a thick box alone cannot supply the balanced charge condition. The choices 16 and 160 meet both requirements with explicit slack.

The number q counts provider offsets, not checker input sites. One provider reads five cardinal letters; nearby providers in the same transverse box lie in different planes. During the proof the enclosing Moore graph allows nine planar offsets per provider and hence $9q$ backward neighbours. Its polynomial volume growth (Lemma 2.3A), rather than its degree, enters the flag-position count. This enclosing graph also allows the robustness statement to use every input of $\overline{N}$.

A live output guarantees at least h successful providers. Failure is therefore permitted at $q - h = K$ offsets. Independently of their letters, an unmarked root has affected inputs in at most one cluster, also K offsets. Successful and unaffected providers can thus be found by discarding at most $2K$ offsets. When two roots are compared, injectivity of the offset pairing turns their two exclusion budgets into at most $4K$ lost pairs. This is why the within-box unit-shift count must exceed $4K$, rather than merely exceed the number of failed votes.

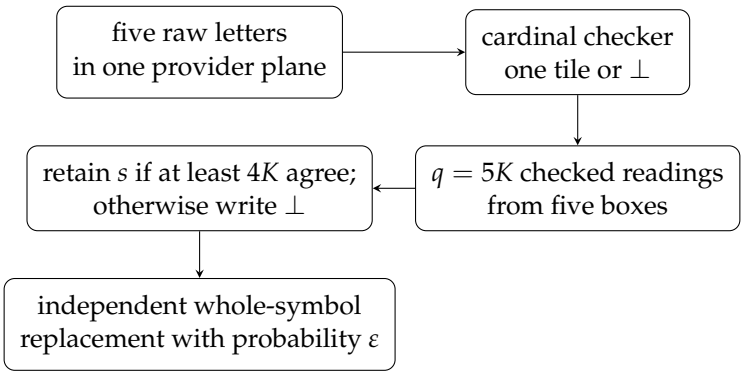

Figure 1: One local update. The checker is repeated once per provider offset. The noise is applied after the quorum and replaces the entire output letter.

For memory the inequality is read in the opposite direction. To erase a correct reference tile, there must be $q - h + 1 = K + 1$ failed providers. The extra one matters: K failures would leave exactly the quorum of $4K$ correct votes. Since no one box has $K + 1$ providers, those failures occupy two clusters and seed the mark recurrence. This explains the use of a nonstrict quorum threshold together with a strict failure count.

2.5 Three elementary inputs

The empty input needs no convention about a uniform law on an empty set: the actual alphabet is $A_\varnothing = \{\perp\}$, and replacement is uniform on this one-point alphabet. The checker, quorum and noisy process all have one possible output. This input is non-tiling and its dynamics is immediately ergodic. Separating it explicitly also avoids applying a later positive-table theorem whose standing assumptions require at least two letters.

Next let $R = \{r\}$ with its opposite edge colours matching. There is exactly one planar tiling, the constant r tiling. Writing $b(u, P) = \mathbf{1}_{\{x(u,P)=r\}}$, the deterministic rule has the completely explicit binary form

$$G_R(b)(u, P) = \mathbf{1}\left\{ \sum_{a \in S} \prod_{d \in \Gamma} b(u + d, P + a) \geq 4K \right\}.$$

A checked vote is one exactly when all five cardinal inputs are one. Uniform replacement writes zero and one with equal conditional probability at a flag. The two deterministic constant configurations are references for the two-cluster error cover. The theorem proves two distinct stationary laws for this noisy binary rule throughout the common interval. It makes no claim here that these exhaust the stationary laws or that the original unstacked rule has a continuum of extremes. That is precisely the countable-tiling case left open later.

Finally let $R = \{r\}$ but give r different north and south colours. A live centre cannot pass both vertical matching tests, whatever its neighbours, since the only live neighbour letter is again r. Thus C_R and G_R are identically blank on every input configuration. Under uniform replacement, the next row is independent of the old row and has independent live indicators of probability $\varepsilon/2$. It has a unique product stationary law, reached after one step. The general forbidden-square estimate is valid but far from sharp in this example: the first forbidden square only provides a convenient certificate for a proof that must also handle arbitrarily complicated inputs.

These examples illustrate what is held fixed in the family. The two singleton inputs have the same binary alphabet operation, the same boxes, the same quorum and the same noise

rate. Only the tile's edge-matching test changes. On a general input the labels r themselves are retained by a successful check, so one is still applying the same formula to raw tiles. There is no preliminary recognition of whether the input is one of these elementary examples, nor any conditional choice of a machine simulation. The finite table is generated uniformly by the checking prescription.

The mismatched singleton is also useful for the scope of robustness. Its base dynamics has no dependence on the past at all. Nevertheless a nearby table with an arbitrarily enlarged neighbourhood can read enough sites to recover a stable binary majority signal. The counterexample in Remark 3.6 therefore cannot be dismissed as a perturbation of a finely tuned complicated input. What must be bounded in the robustness theorem is the specified input neighbourhood, not the logical difficulty of the tile set.

2.6 The all-input coefficient and the theorem

Each innovation $(\eta_t(v), \zeta_t(v))$ specifies a local map from parent words to output letters: either G_R or a constant map. Composing these maps through a depth-t cone at v gives a random function of the initial row. Define

$$a_t = \mathbb{P}\{x \longmapsto X_t^x(v) \text{ is not constant on } \Omega_R\}. \tag{7}$$

Translation covariance makes this probability independent of v. At a fixed finite time only finitely many initial sites can affect v. Thus constancy is a finite test for each innovation realization, and the event in (7) is measurable. The pair of starts witnessing nonconstancy may be chosen *after* the innovations are known.

This is stronger than a bound for every fixed pair of starts. If a single noise event covers all possible discrepancies, it bounds a_t without a union over starts. For a finite observation set O, a union over the output sites gives

$$\mathbb{P}\{x \longmapsto X_t^x|_O \text{ is not constant}\} \le |O| a_t. \tag{8}$$

The same estimate bounds the total-variation distance between the two O-marginals from any two initial laws, by common-map coupling. Independence of the chosen initial states from future innovations is sufficient.

Theorem 2.1 (Natural Wang classification)**.** *For every finite Wang set R and every $0 < \varepsilon \le \varepsilon_*$, the rule (4) with uniform whole-symbol replacement is ergodic if and only if R does not tile $\mathbb{Z}^2$. If R tiles, there are two distinct spatially translation-invariant stationary laws. If R has no legal $n \times n$ pattern, $n \ge 1$, then for every $t \ge 0$,*

$$a_t \le \frac{1}{1-\kappa} \exp\{c(762n + 1 - t)\}. \tag{9}$$

The constants ε_*, κ, c are the same for all R.

The proof occupies the rest of this section. Its contour, path and coherence arguments are the raw-letter forms of [12, Lemmas 5.1–5.3]; the complete-history compactness argument below corresponds to that paper's Lemma 8.5. Lemma 5.4 there concerns deep wrong readings of an auxiliary forecast; that extra device is absent from the bare rule and is not required here. The quantitative finite-history method follows its Theorem 9.5 and Appendix A, specifically Lemmas A.1–A.2. Those results belong to the version 1.1.0 package directed by Nidhal Mghirbi. The finite certificate here is simply a forbidden square, which removes the statistical eraser and grouped rare-birth estimates needed there. The charge-arm method itself is due to Toom, Berman–Simon and Gács [47, 4, 19].

2.7 A marked field determined only by flags

For now sample the Bernoulli flags at all integer times, including auxiliary negative times. An ordinary backward edge is

$$(u, P, t) \longrightarrow (u + d, P + a, t - 1), \qquad d \in \Lambda,\ a \in S. \tag{10}$$

The associated undirected graph has degree at most $18q$; its volume growth in Lemma 2.3A enters the estimates. Each edge changes time by exactly one. The larger Moore neighbourhood is convenient for estimates and robustness; the deterministic rule uses only its cardinal subset.

For an integer t_0, start a marked set empty at time t_0, and thereafter mark every flagged vertex and every vertex having marked parents in two distinct transverse clusters. The union as $t_0 \to -\infty$ is denoted $\mathcal{M}$. Equivalently it is the least set containing all flags and closed under the rule

$$\begin{aligned} &(u + d, P + a, t - 1) \in \mathcal{M}, \quad (u + d', P + a', t - 1) \in \mathcal{M}, \\ &a \in S_i,\ a' \in S_j,\ i \neq j \quad \Longrightarrow \quad (u, P, t) \in \mathcal{M}. \end{aligned} \tag{11}$$

The closure is finitary: every marked point has a finite derivation from flags. Taking derivation depth as rank and selecting two lower-rank supports at each nonflagged mark yields a finite root-reachable acyclic support graph. Branches may merge. Its terminals are distinct physical flagged vertices, not independent copies of repeated appearances of one flag.

Least closure and finite supports. One can define the field without reference to a start time. Let E be the set of flags, let $\Phi(J)$ add to J every vertex with two distinct-cluster parents in J, and put $J_0 = E$, $J_{r+1} = \Phi(J_r)$. Then $\bigcup_{r \geq 0} J_r$ is closed: a supporting pair belongs to two finite ranks and therefore to their maximum rank. It is contained in every closed superset of E, so is the least one. This agrees with the increasing finite-start construction. A finite-rank derivation uses finitely many parents recursively and reaches a finite earliest time; starting empty before that time realizes the same mark. Conversely every finite-start mark has a finite-rank derivation by induction on elapsed time.

This equivalence distinguishes the least closure from any larger solution of the recurrence. For example, marking all vertices would satisfy the propagation implication but would not have finite flag witnesses. Every argument below concerns the least closure. The supports selected for a nonflagged marked vertex can be chosen at strictly smaller derivation ranks. Terminating at flags gives a finite graph, even though the full lattice and its past are infinite. If the same ancestor is reached by two branches it is retained once as a physical vertex. The graph can have undirected cycles, but no directed cycle, because each edge decreases time. These are exactly the properties used by the frontier construction.

Lemma 2.2 (The unmarked exclusion budget)**.** *An unmarked vertex is unflagged. Among its provider offsets $a \in S$, at most K have any marked Moore parent $(u + d, P + a, t - 1)$, $d \in \Lambda$.*

Proof. If marked parents occurred in two different clusters, (11) would mark the vertex. They therefore occur in at most one cluster, containing K offsets. A flag is a mark by definition. □

This assertion depends on the flag field alone. It holds simultaneously for every trajectory and every choice of letters written at flags. In particular it remains valid for starts selected with knowledge of the noise; no causality assertion about such selected starts is needed for this deterministic step.

2.8 Five balanced charges

Ignore u and assign the potentials

$$\psi_0(P,t) = 4\sum_{i=1}^{4} P_i + 560t, \qquad \psi_i(P,t) = -4P_i - 140t \quad (1 \le i \le 4). \tag{12}$$

Their sum is identically zero. A backward edge with transverse displacement a has gains $4\sum_i a_i - 560$ for charge zero and $140 - 4a_i$ for charge i. If its cluster is nonzero, the first gain is at least $4 \cdot 160 - 560 = 80$. If its cluster differs from i, the second gain is at least $140 - 4 \cdot 15 = 80$. All gains, including the other choices, are at least -560.

For any two selected supports in distinct clusters, every charge can choose one of them with gain at least 80: charge zero avoids cluster zero, and charge $i \ge 1$ avoids cluster i. Call that support good for the charge. If neither support lies in cluster c, both are good for charge c. The choices for different charges can differ. A support graph is not a tree of five independent paths; the same physical edge or terminal may be used by several arms. The next construction keeps that multiplicity for the potential accounting and removes it only when counting physical witnesses.

Checking every gain analytically. For charge zero, offsets in S_0 have coordinate sum between zero and 60, so the gain lies between -560 and -320. In any nonzero cluster that sum is between 160 and 220, giving gains between 80 and 320. For charge $i \ge 1$, a provider in cluster i has coordinate a_i between 160 and 175, giving gains between -560 and -500. In any other cluster a_i is between zero and 15, giving gains between 80 and 140. These interval computations check all offsets at once.

Charge	its excluded cluster	every other cluster
0	$[-560, -320]$ in S_0	$[80, 320]$
$i \in \{1,2,3,4\}$	$[-560, -500]$ in S_i	$[80, 140]$

Table 2: Backward gains. A support pair in distinct clusters always includes a good choice for every charge, and both supports are good for each of the three charges whose cluster contains neither. The sum of the five potentials is zero at every source and every terminal.

The balance is algebraic, not a probabilistic cancellation. Each transverse coefficient $4P_i$ of ψ_0 cancels the coefficient $-4P_i$ of ψ_i, and $560t$ cancels four copies of $-140t$. The possible negative gain pays only for first edges introduced when a frontier component splits, and at one split at most two of the five first edges can be bad. That is why the edge bound is proportional to the number of new sources, which is the number of distinct selected flags minus one.

Lemma 2.3 (Finite charge-arm witness)**.** *For a marked root there is a connected physical ordinary-edge witness containing the root and $f \ge 1$ distinct flags. It has an abstract copied-arm realization with at most*

$$e_{\mathrm{arm}} \le 16(f-1) \tag{13}$$

edge occurrences. All its vertices lie between its earliest terminal time and the root time, and its physical edge count and its time span are at most e_{arm}. A flagged root may use the singleton witness $f = 1$, $e_{\mathrm{arm}} = 0$; an unflagged root has $f \ge 2$.

Proof. Stop the finite support graph at flags. Name a terminal by its physical space–time coordinate, and for each vertex v let $D(v)$ be its nonempty set of descendant terminal names.

At a nonterminal with selected supports a, b,

$$D(v) = D(a) \cup D(b).$$

We first construct arms so that distinct terminal names, rather than unfolded tree leaves, pay for their length.

Expose the graph from root to terminals in a topological order: only nonterminal vertices are exposed, each after every vertex from which it is reached, and terminals stay in the frontier until the end. A frontier is the set of unexposed vertices currently reached. Join two frontier vertices when their descendant terminal sets overlap. Initially the frontier is just the root. Exposing v replaces it by its supports, retaining one physical copy of an already present support. Because $D(a)$ and $D(b)$ are subsets of $D(v)$, the replacement cannot meet an outside overlap component. It splits its own component into at most two. To verify the latter assertion, follow an old overlap path from a remaining vertex toward v. Its last common terminal with v belongs to $D(a)$ or $D(b)$, so its remaining part attaches to one of the supports. Every resulting component therefore contains a or b. In a genuine split the two daughter components have disjoint sets of descendant terminal names.

Fix a good-parent choice for every charge at every nonterminal. Starting at a given support and repeatedly following good parents reaches a terminal, since the support graph is finite and acyclic. Begin with one source occurrence at the root and five good arms, one for each charge. Track each arm by its eventual terminal name. Maintain the invariant that in each frontier overlap component there is exactly one endpoint name for each charge, counting arm occurrences.

When exposure causes no split, the invariant persists: a terminal name in the old component still lies in its one successor component. At a split at v, the existing endpoint for each charge belongs to exactly one daughter. Add one new source occurrence at v. For that charge choose the support in the daughter missing its old endpoint, take this first edge, and then follow good parents to a terminal there. The first edge need not be good; every later edge is good. Do this for the five charges. The new source restores one endpoint of each charge in each daughter. If the split is at the original root, keep the new source occurrence distinct from the original one even though their physical coordinates coincide.

At the final frontier each component is one terminal name. For each charge, source occurrences and terminal names are in bijection via their arms. Now take separate abstract copies of all arm interiors. Identify the five starts belonging to a source occurrence, and identify endpoints with the same terminal name. Keep different source occurrences distinct even when their coordinates agree. Retain the connected component containing the original root source. Restricting the charge bijections to it is valid, since every arm connects its source and endpoint within a single component. If f terminal names remain, exactly f source occurrences remain, with one incidence of each charge at each source and terminal.

The accounting uses only the following facts, each established above. (A1) Every non-flagged mark has its two selected supports in two distinct clusters, by (11); so they are distinct vertices, and for every charge at least one of them is good. (A2) The bad set of charge c is exactly one cluster (S_0 for $c = 0$, S_i for $c = i$; see the gain table above), so whether an edge is good for c depends only on the cluster of its offset. (A3) The support graph is finite, acyclic and reachable from the root, every edge lowers time by exactly one, and flags are terminals even when they also have two marked parents. (A4) Terminal names are physical space–time coordinates; a vertex reached twice is one physical vertex. (A5) Only nonterminals are exposed, in a topological order, and terminals stay in the frontier. (A6) The potentials ignore u, so a gain depends only on the transverse offset. (A7) The first edge of a new arm

goes from the split vertex to one of its two selected supports. (A8) The five arms of the original root source start with good-parent edges. (A9) Edge occurrences are counted with multiplicity on both sides of the telescoping sum.

Only the first edge from a nonoriginal source can be bad for its charge; every edge of the five arms of the original root source, and every later edge of every arm, is a good-parent edge and gains at least 80. Consider a nonoriginal source occurrence at a split vertex w, with selected supports a in cluster S_i and b in cluster S_j, $i \neq j$. By construction the first edge of each charge goes from w to a or to b (A7). By (A2) the edge to a is bad only for charge i and the edge to b only for charge j. For each of the three charges outside $\{i, j\}$ both possibilities are good, so that first edge gains at least 80. Each of the two remaining first edges gains at least -560. Hence the five first edges at one nonoriginal source gain at least $3 \cdot 80 - 2 \cdot 560 = -880$ in total.

Sum potential increments along all abstract arms, with their multiplicities. At each source the sum of the five starting potentials is zero; at each terminal the sum of the five ending potentials is zero. Thus the total increment equals zero. There are exactly $f - 1$ nonoriginal sources, each with five first edges, and every other edge occurrence gains at least 80. Consequently

$$0 \;\geq\; 80\big(e_{\mathrm{arm}} - 5(f-1)\big) - 880(f-1) \;=\; 80\, e_{\mathrm{arm}} - 1280(f-1),$$

which gives (13).

An unflagged original source has outgoing good edges, so $e_{\mathrm{arm}} > 0$ and $f \geq 2$.

Projection of the retained abstract component to physical edges is connected, contains the root and the f distinct flags, and can only reduce the edge count. Every projected vertex is at or before the root time. A vertex of minimum time in the abstract graph must be a terminal: a source or arm interior has an outgoing edge to an earlier time. A simple undirected path in this connected graph from the root to a minimum-time terminal uses at most e_{arm} edge occurrences. Each changes time by exactly one, so its time span is at most e_{arm}. Every vertex lies between those two times. No unary edge or identification changes the clock, because there are no unary links and terminal names are actual space–time coordinates. □

The separation of abstract arms from physical edges is load-bearing. Potential telescoping uses one arm of each charge per terminal, whereas probability uses one Bernoulli factor per *distinct* flag. Unfolding a merged support graph into a tree and counting its leaves as independent would not justify the bound.

What a split preserves. The frontier is a partition of the remaining terminal names into overlap components, rather than a partition of the original support vertices into independent branches. After exposing a vertex, all terminal names previously reachable are still reachable through one of its supports. If the component splits, its two terminal-name unions are disjoint. Each old arm endpoint therefore has one definite daughter, even if several abstract arms pass through the same physical intermediate vertex. Adding a new source supplies only the missing daughter for each charge; it does not duplicate an already counted flag.

At the end a source-to-terminal arm graph may have several connected components. Discarding those away from the original root is legitimate for two reasons. First, every retained arm connects one retained source to one retained terminal, so each charge's bijection restricts to the retained component. Second, the potential sum is balanced at each individual source and terminal, so it telescopes within that component without using discarded arms. This is why the proof can obtain a connected physical witness and an exact count of distinct

terminal flags at the same time. Neither connectedness nor distinctness is inferred from the support graph being a tree.

2.9 Counting flag positions

We bound the probability of a bad event by a union over the ordered positions of the distinct flags that it forces, weighted against the polynomial volume growth of the space–time graph. Witness shapes and paths are never counted: two witnesses with the same flags give the same event. Hence neither the degree of the graph nor the number of possible witnesses enters the estimates. For a fixed set of f distinct vertices, the probability that all are flags is exactly ε^f. The remaining support conditions are discarded. Thus adaptive selection of a support graph never creates a conditional independence claim: we take a countable union over deterministic finite sequences of flag positions.

Let $\mathcal{G}$ be the undirected graph on $\mathbb{Z}^2 \times \mathbb{Z}^4 \times \mathbb{Z}$ whose edges are the ordinary edges (10), and let d be its graph distance. For $0 < y < 1$ put

$$Z(y) = \sum_x y^{d(x_0,x)},$$

which does not depend on x_0 by translation invariance. Let

$$A_7(y) = 1 + 120y + 1191y^2 + 2416y^3 + 1191y^4 + 120y^5 + y^6,$$
$$A_5(y) = 1 + 26y + 66y^2 + 26y^3 + y^4.$$

$$\widehat{Z}_6(y) = \frac{4 \cdot 175^4\, A_7(y)}{(1-y)^8}, \qquad \widehat{Z}_4(y) = \frac{4 \cdot 175^2\, A_5(y)}{(1-y)^6}.$$

Lemma 2.3A (Volume growth)**.** *Let $m \in \{1,2\}$, $k \geq 1$, let $J \subseteq \{-1,0,1\}^m$ and let $S' \subseteq \{0,\dots,175\}^k$ be finite and nonempty. Consider the undirected graph on $\mathbb{Z}^m \times \mathbb{Z}^k \times \mathbb{Z}$ with an edge between (u,P,t) and $(u+j,P+a,t-1)$ for every $j \in J$ and $a \in S'$. For every vertex x_0 and every integer $r \geq 0$, the number of vertices at graph distance exactly r from x_0 is at most $(r+1)(2r+1)^m(175r+1)^k$.*

- *(a) For $\mathcal{G}$ ($m = 2$, $k = 4$, $J = \Lambda$, $S' = S$): this number is at most $4 \cdot 175^4(r+1)^7$, and $Z(y) \leq \widehat{Z}_6(y)$ for $0 < y < 1$.*
- *(b) For the ordinary graphs of Section 12 ($k = 2$, $S' = S$ of (45), and $m = 2$ with any $J \subseteq \{-1,0,1\}^2$, or $m = 1$ with any $J \subseteq \{-1,0,1\}$; the checker uses $J = \{-1,0,1\}^2$ and the generator $J = \{0,1\}$): this number is at most $4 \cdot 175^2(r+1)^5$, and the corresponding generating function is at most $\widehat{Z}_4(y)$ for $0 < y < 1$.*

Proof. A vertex at distance r is the endpoint of a walk of exactly r edges from x_0. Suppose j of its edges are traversed backward, changing (u,P,t) by $(d,a,-1)$ with $d \in J$, $a \in S'$, and $r-j$ forward, changing it by $(-d,-a,+1)$. The time coordinate of the endpoint is $t_0 + r - 2j$, which takes at most $r+1$ values. Each logical coordinate changes by at most one per edge, so takes at most $2r+1$ values. For a fixed value of j, each transverse coordinate of the endpoint lies in the interval $[P_{0,i} - 175(r-j),\ P_{0,i} + 175j]$, because $S' \subseteq [0,175]^k$; this interval contains at most $175r+1$ integers. Multiplying over the $r+1$ values of j (which fix the time), the m logical and the k transverse coordinates gives the first bound. For (a) use $(2r+1)^2 \leq 4(r+1)^2$ and $(175r+1)^4 \leq 175^4(r+1)^4$. For (b) use $2r+1 \leq (2r+1)^2 \leq 4(r+1)^2$ and $(175r+1)^2 \leq 175^2(r+1)^2$.

For the generating functions, $Z(y) = \sum_{r\geq 0} y^r \,\#\{x : d(x_0, x) = r\}$, and it remains to show $\sum_{r\geq 0}(r+1)^n y^r = A_n(y)/(1-y)^{n+1}$ for $n = 7, 5$. Multiply the left side by $(1-y)^{n+1} = \sum_i (-1)^i \binom{n+1}{i} y^i$. The coefficient of y^r in the product is $\sum_{i=0}^{n+1}(-1)^i\binom{n+1}{i}(r+1-i)^n$, where terms with $r < i$ are omitted. For $r \geq n$ every omitted term vanishes (for $r = n$ the only one has $i = n+1$ and equals $0^n = 0$), so the coefficient is an $(n+1)$-st finite difference of the polynomial $s \mapsto s^n$ of degree n, hence zero. For $r < n$ the coefficients are those of A_7 and A_5 displayed above, as one checks directly. Since all terms are nonnegative, the bounds on spheres give the bounds on Z. □

Lemma 2.3B (Counting flag positions)**.** *Let G be a graph on a countable vertex set with graph distance d, let $0 < y < 1$, and suppose $Z_G(y) := \sup_{x_0} \sum_x y^{d(x_0,x)} < \infty$. Let the vertices carry independent Bernoulli(e) flags. Let x_0 be a vertex, Φ a set of positive integers and $L : \Phi \to [0, \infty)$. Let E be an event on which there exist $f \in \Phi$ and distinct flagged vertices $T_1, \ldots, T_f$ with $\sum_{j=1}^{f} d(T_{j-1}, T_j) \leq L(f)$, where $T_0 = x_0$. Then*

$$\mathbb{P}(E) \leq \sum_{f\in\Phi} e^f \, y^{-L(f)} \, Z_G(y)^f. \tag{14}$$

Proof. E is contained in the countable union, over $f \in \Phi$ and over sequences $(x_1, \ldots, x_f)$ of distinct vertices with $\sum_j d(x_{j-1}, x_j) \leq L(f)$ (where x_0 is the given vertex), of the events that $x_1, \ldots, x_f$ are all flagged. Each such event has probability e^f, since the vertices are distinct and the flags independent. For fixed f, since $\mathbf{1}[s \leq L] \leq y^{s-L}$ for $0 < y < 1$, the number of admissible sequences is at most

$$\sum_{x_1,\ldots,x_f} y^{\sum_{j=1}^{f} d(x_{j-1},x_j) - L(f)} \leq y^{-L(f)} Z_G(y)^f,$$

by summing first over x_f, then over x_{f-1}, and so on down to x_1, each sum being at most $Z_G(y)$. □

Lemma 2.3C (Walk ordering)**.** *Let w be a walk of length ℓ starting at x_0 in a graph with distance d, and let $T_1, \ldots, T_f$ be distinct vertices visited by w, listed in the order of their first visits. Then $\sum_{j=1}^{f} d(T_{j-1}, T_j) \leq \ell$, where $T_0 = x_0$. If H is a finite connected subgraph with at most k edges and $x_0 \in H$, there is a closed walk from x_0 of length at most $2k$ that visits every vertex of H.*

Proof. Let $0 \leq \tau_1 < \cdots < \tau_f \leq \ell$ be the first-visit times, and $\tau_0 = 0$. The piece of w between times τ_{j-1} and τ_j is a walk from T_{j-1} to T_j of length $\tau_j - \tau_{j-1} \geq d(T_{j-1}, T_j)$. Summing gives at most $\tau_f \leq \ell$. For the second claim, H has a spanning tree with $|V(H)| - 1 \leq k$ edges, and the depth-first tour of that tree from x_0 traverses each tree edge twice. □

Lemma 2.4 (Sparse roots)**.** *Let the marked set $\mathcal{M}$ be built from independent flags of density e. For every vertex v and every $0 < y < 1$ with $\lambda := e\,\widehat{Z}_6(y)\, y^{-32} < 1$,*

$$\mathbb{P}(v \in \mathcal{M}) \leq e + e\,\widehat{Z}_6(y)\frac{\lambda}{1-\lambda} \leq e + \frac{\lambda^2}{1-\lambda}. \tag{15}$$

For every $0 < e \leq 2\varepsilon_*$, uniformly in R, this gives $\mathbb{P}(v \in \mathcal{M}) \leq p_* := 10^{-4}$. In particular $\mathbb{P}(v \in \mathcal{M}) \leq p_*$ at every $0 < \varepsilon \leq \varepsilon_*$.

Proof. A flagged root costs e. If v is marked and unflagged, Lemma 2.3 gives a connected physical witness H containing v and $f \geq 2$ distinct flags, with at most $16(f-1)$ edges. By

Lemma 2.3C there is a closed walk from v of length at most $32(f-1)$ visiting all of them; listing the flags in the order of first visits gives distinct flagged $T_1, \dots, T_f$ with $\sum_j d(T_{j-1}, T_j) \le 32(f-1)$. Lemma 2.3B with $\Phi = \{2, 3, \dots\}$, $L(f) = 32(f-1)$ and Lemma 2.3A(a) bound this case by

$$\sum_{f \ge 2} e^f y^{-32(f-1)} \widehat{Z}_6(y)^f = e\, \widehat{Z}_6(y) \frac{\lambda}{1-\lambda}.$$

Finally $e\, \widehat{Z}_6(y) = \lambda y^{32} \le \lambda$. For the numerical claim take $y = 785/1000$. Then $\widehat{Z}_6(785/1000) \le 2.043 \cdot 10^{18}$ and $(1000/785)^{32} \le 2313$, so for $e \le 2\varepsilon_* = 2 \cdot 10^{-24}$ we get $\lambda \le 95/10^4$ and $e + \lambda^2/(1-\lambda) \le 2 \cdot 10^{-24} + (95/10^4)^2/(1 - 95/10^4) < 10^{-4}$. The bound is increasing in e, so it holds throughout $0 < e \le 2\varepsilon_*$. □

Remark. For $0 < e \le 10^{-40}$ the first bound in (15) with $y = 5/8$ gives $\widehat{Z}_6(5/8)\, \lambda/(1-\lambda) < 1$, hence the sharper form $\mathbb{P}(v \in \mathcal{M}) < 2e$ of version 2.2.0, in particular on its interval $0 < e \le (2(18q)^2)^{-200}$. For $e \le 2\varepsilon_*$ the first bound at $y = 785/1000$ gives $\mathbb{P}(v \in \mathcal{M}) \le 4 \cdot 10^{-8}$. Only $\mathbb{P}(v \in \mathcal{M}) \le p_*$ is used below.

The estimate does not mention the alphabet size, the replacement values, or a property of R. It controls one binary field for all starts and all tile sets at once. A sparse one-point field alone does not exclude an infinite path; the following packing argument supplies that additional fact and its finite-prefix tail.

2.10 Packing contours along a marked path

Lemma 2.5 (Finite marked-path tail)**.** *Let $\mathcal{M}$ be built from independent flags of density e. For a specified root, $N \ge 1$ and every $0 < y < 1$ with $\kappa_y(e) := e\, \widehat{Z}_6(y)\, y^{-48} < 1$, the probability of an ordinary backward path of N edges all of whose vertices belong to $\mathcal{M}$ is at most*

$$\sum_{f \ge \lceil N/16 \rceil} \kappa_y(e)^f \;\le\; \frac{\kappa_y(e)^{N/16}}{1 - \kappa_y(e)}. \tag{16}$$

With $y = 849/1000$, $\kappa_y(e) \le \kappa = 1/8$ for $0 < e \le \varepsilon_*$ and $\kappa_y(e) \le 1/4$ for $0 < e \le 2\varepsilon_*$. Almost surely there is no infinite such path, simultaneously at every root.

Proof. Fix a finite marked path whenever one exists, with vertices $p_1 = v$ at the root and so on. At p_1 choose the full witness W_1 of Lemma 2.3, including its f_1 selected flags. If its earliest terminal has time s_1, restart at the path vertex p_2 at time $s_1 - 1$, provided the path has not ended. Repeat, obtaining witnesses $W_1, \dots, W_m$ with $f_1, \dots, f_m$ flags. Every vertex of W_i lies between its earliest terminal time s_i and the time of p_i, and p_{i+1} is earlier than s_i, so the witnesses lie in disjoint closed time slabs and all $f = \sum_i f_i$ selected flags are distinct. This uses the span assertion of Lemma 2.3, not just an edge count.

The span of W_i is at most $16(f_i - 1)$. The path piece from p_i to p_{i+1}, including the restart edge, has at most $16(f_i - 1) + 1 \le 16 f_i$ edges. For the final witness retain the whole contour, even if its earliest terminal is earlier than the path endpoint, and truncate only the connecting path piece. Hence $N \le 16f$, that is, $f \ge \lceil N/16 \rceil$. A flagged path vertex is allowed: its singleton witness has no edge and covers one restart step.

Now form a walk from v: for $i = 1, \dots, m$, perform the closed tour of a spanning tree of W_i from p_i, of length at most $32(f_i - 1)$ by Lemma 2.3C, and then follow the path piece from p_i to p_{i+1} (or to the path end). Its length is at most $\sum_i \big(48(f_i - 1) + 1\big) \le 48f$, and it visits all f distinct flags. Lemma 2.3C, Lemma 2.3B with $\Phi = \{f : f \ge \lceil N/16 \rceil\}$, $L(f) = 48f$,

and Lemma 2.3A(a) give the bound $\sum_{f \geq \lceil N/16 \rceil} e^f y^{-48f} \widehat{Z}_6(y)^f$, which is (16). No path and no witness shape is counted; only the positions of the distinct flags are.

The numerical claims follow from $\widehat{Z}_6(849/1000) \leq 4.33 \cdot 10^{19}$ and $(1000/849)^{48} \leq 2586$: then $\kappa_y(10^{-24}) \leq 4.33 \cdot 10^{19} \cdot 2586 \cdot 10^{-24} < 1/8$ and $\kappa_y(2 \cdot 10^{-24}) < 1/4$; $\kappa_y(e)$ is linear in e. The tail tends to zero as $N \to \infty$. An infinite marked path gives a finite marked path of every length from its root; these events decrease, and their probabilities tend to zero. Countably many roots give the simultaneous almost-sure statement. □

A ledger for the path witness. It helps to separate three counts. The number f_i is the number of distinct flags selected by one contour, e_i its abstract arm length and n_i its number of physical edges after projection. Their relations are $n_i \leq e_i \leq 16(f_i - 1)$ and span at most e_i. The connecting path pieces are walked separately; they may cross edges of a contour already visited, which only lengthens the walk and so only weakens the bound. Successive contours can have geometric projections that meet, but their selected physical time slabs are disjoint, so their terminal flags remain distinct. Spatial overlap never changes their independent Bernoulli cost. The last contour is kept whole because cut vertices at the path endpoint need not be flags. The walk budget is 48 per flag: 32 for the closed tour of a contour and 16 for its connector. At the maximum permitted source density for the unperturbed rule the base is at most $1/8$; for robustness the source density doubles and the base is at most $1/4$. None of these margins depends on how late the first forbidden square appears.

2.11 Successful and unaffected providers

We turn from probability to deterministic geometry. Fix any flag realization, any replacement values and one arbitrary trajectory started at time zero. All coherence statements refer to that one trajectory; starts will be compared only after the deterministic statements are established.

An unmarked live output of age $t \geq 1$ is an ordinary G_R output, say s. Its successful-provider set contains at least h offsets returning s. At most $q - h = K$ offsets fail. By Lemma 2.2, at most K offsets have any marked Moore input. Call a provider *good* if it is successful and unaffected by marked inputs. At least $h - K = 3K$ providers are good. A good provider supplies five unmarked live cardinal inputs one time earlier, with central letter s and matched central edges.

There are two different budgets here. For a single root we exclude at most $2K$ offsets to find good providers. To compare two roots through an injective pairing, we exclude at most $4K$ pairs. A provider can be affected even when its own central input is unmarked: all Moore inputs are charged so the exclusion is valid for every checker input. This conservative choice will also cover arbitrary maps on the enclosing neighbourhood in Corollary 3.5.

2.12 Same-site coherence at finite age

Lemma 2.6 (Equality at one planar site)**.** *In one trajectory started at time zero, two unmarked live readings at (u, P, t) and (u, Q, t) have the same tile whenever*

$$d = \|Q - P\|_1 \leq t.$$

Proof. Induct on d, uniformly over the time and planar site. If $d = 0$, both readings are at the same physical vertex. For $d > 0$, choose a coordinate j on which $Q - P$ is nonzero, and let σ be its sign. Pair offsets inside each box by

$$a \longleftrightarrow a' = a - \sigma e_j, \qquad a, a' \in S_i.$$

For each cluster there are $15 \cdot 16^3$ pairs. Both coordinate projections of the pairing are injective. The total count is

$$5 \cdot 15 \cdot 16^3 = \frac{75}{16}K.$$

Each root excludes at most K failed and at most K affected offsets. Injectivity means these exclusions remove at most $4K$ pairs altogether, even if all exclusion sets are disjoint. There remain at least

$$\left(\frac{75}{16} - 4\right) K = \frac{11}{16}K > 0 \tag{17}$$

pairs. Select one. Its central parents at $(u, P + a, t - 1)$ and $(u, Q + a', t - 1)$ are unmarked and live and retain the two respective root letters. Their transverse distance is exactly $d - 1$, because the displacement on coordinate j has moved one step toward zero and the other coordinates are unchanged. Also $t - 1 \geq d - 1$. The induction hypothesis identifies the parent letters, hence the root letters. □

The induction is on transverse span, not on the number of initial configurations. It never assumes that different planes agree at age zero. Finite age is the resource that pays for finite span. In a complete trajectory the same finite induction works at every span because an arbitrarily long finite past is available.

Why the reduction of span is exact. Write $Q - P = (d_1, \ldots, d_4)$ and choose j with $d_j \neq 0$. With $\sigma = \operatorname{sign}(d_j)$ and $a' = a - \sigma e_j$, the parent displacement is $(Q + a') - (P + a) = (Q - P) - \sigma e_j$. Its jth absolute coordinate drops by one and its other coordinates do not change. Thus its ℓ^1 norm is exactly $\|Q - P\|_1 - 1$ for either sign, including $|d_j| = 1$. Pairing is not restricted by the size of $Q - P$: only the relative unit shift of the offsets must remain inside the box. Hence the same overlap surplus works for arbitrarily distant root planes, one induction step at a time.

Within one box, fifteen choices remain in the shifted coordinate and sixteen in each other coordinate. If $\sigma = 1$, omit the face with local coordinate zero on the first side; if $\sigma = -1$, omit the face with local coordinate fifteen. The pairing is one-to-one in both directions. A discarded offset on either root side can therefore destroy at most one candidate pair. No independence or favourable correlation of the two roots' bad-provider sets is assumed. The inequality survives their worst-case placement.

This pairing also explains why comparing roots by directly intersecting their absolute provider locations is unnecessary. Distant translates of S can be disjoint. The proof compares the *relative* offsets a, a' and uses the previous-age coherence assertion at the resulting parent span. What persists across the induction is the root letter: a successful check retains its central parent's raw tile, so the two letters being compared are carried unchanged into the induction hypothesis.

2.13 Cardinal compatibility at finite age

Lemma 2.7 (Neighbouring planar readings)**.** *Let $\gamma \in \{\pm e_1, \pm e_2\}$. Unmarked live readings at (u, P, t) and $(u + \gamma, Q, t)$ have matching shared edges if*

$$t \geq \|Q - P\|_1 + 1.$$

Proof. Again induct on $d = \|Q - P\|_1$. When $d = 0$, both ages are positive. Their successful-provider sets each have size at least h, so they overlap in at least

$$2h - q = 3K > 0$$

offsets. At a common offset a, the check at u retains the first root tile s, and the check at $u + \gamma$ retains the second tile s'. Their old central sites are adjacent in the same provider plane. The first check certifies the matching edge between those two actual old letters; the second identifies the neighbour letter as s'. Thus the two retained tiles match. Unmarkedness of those old inputs is not needed in this base case.

For $d > 0$ use the same injective unit-shift pairing as in Lemma 2.6. Its surplus (17) supplies a good provider for each root. Their central parents are unmarked, live and cardinal-adjacent in the planar coordinate, carry the root letters, and have transverse span $d - 1$. Their age satisfies $t - 1 \geq d = (d - 1) + 1$. The induction hypothesis certifies their shared edge and hence the root edge. □

The extra unit of age in this lemma is necessary for the proof: at equal transverse coordinates, arbitrary initial adjacent tiles need not match. Equality at the same planar site has a tautological age-zero base, whereas compatibility first uses an actual checker update.

2.14 Finite certificates of non-tiling

Lemma 2.8 (Forbidden square)**.** *If a finite Wang set R does not tile the plane, there is $n_R \geq 1$ for which no legal $n_R \times n_R$ pattern exists. A finite search for such an n_R terminates exactly on non-tiling inputs.*

Proof. For each n enumerate all maps from $\{0, \ldots, n-1\}^2$ to R and test the internal shared edges. This is finite. If the test has no survivor, a full tiling is impossible. Conversely, suppose patterns exist for every n. Taking patterns on centered squares of arbitrarily large odd side, extend each outside its square arbitrarily to a configuration in $R^{\mathbb{Z}^2}$; when R is nonempty this is possible. A subsequence stabilizes successively on each centered square by finite pigeonhole and diagonal selection. Its limiting configuration matches every edge, since that edge lies in all sufficiently large squares. This is a tiling. For $R = \varnothing$ the first test is already empty. The contrapositive proves termination on every non-tiler. □

The search is a semidecision. On a tiling input it never finds an empty square. It supplies no total computable upper bound for the first empty square as a function of the finite input. The proof of forgetting may use that finite certificate without computing it uniformly on all inputs.

The relevant transverse diameter is

$$D = \operatorname{diam}_{\ell^1}(S) = 380. \tag{18}$$

For two distinct nonzero cluster centres the distance is 320; for cluster zero and a nonzero cluster it is 160. Box corners contribute at most $4 \cdot 15 = 60$. Equality 380 is attained, for example by taking the two centres $160e_1, 160e_2$ with opposite allowable corners in every coordinate. A sum of r provider offsets therefore has pairwise ℓ^1 spread at most Dr: pair the corresponding summands in two sums and apply the triangle inequality. This statement concerns the r-fold sumset $S + \cdots + S$, not the dilation $\{ra : a \in S\}$.

2.15 Deterministic mortality outside the marked field

Lemma 2.9 (Forbidden-square delay)**.** *Suppose R has no legal $n \times n$ pattern. Put*

$$r = 2n, \qquad T_R = (D + 1)r + 1 = 762n + 1. \tag{19}$$

For every realization of flags and replacement values, for every initial configuration and every $t \geq T_R$, every live output at age t lies in $\mathcal{M}$.

Proof. Assume instead that an unmarked live reading occurs at (u_0, P_0, t). At every unmarked live positive-time output, a good provider exists because $h - K = 3K > 0$. It supplies unmarked live cardinal inputs at all five sites one step earlier. Repeat this construction for r steps. For each $w \in \mathbb{Z}^2$ with $\|w\|_1 \leq r$, choose a cardinal path of exactly r planar steps from u_0 to $u_0 + w$, padding a shorter path by zero steps. Following good providers along it produces an unmarked live reading at

$$(u_0 + w, Q_w, t - r), \qquad Q_w \in P_0 + \underbrace{S + \cdots + S}_{r \text{ summands}}.$$

Different paths need not use the same providers or produce the same Q_w. Choose one reading for each w in the ball. Any pair of these transverse coordinates is within Dr. Their common age obeys

$$t - r \geq Dr + 1.$$

The same-site lemma makes any alternative chosen reading at a planar site agree, and Lemma 2.7 makes every internal cardinal edge of the ball match. Hence the selected raw letters form a legal pattern on that whole planar ball. The ball of radius $2n$ contains $[-n, n]^2$, and thus contains an $n \times n$ square. Restricting gives the forbidden pattern, a contradiction.

All uses of the update occur at positive ages, since $t - r \geq 1$. No stationary law, statistical erasure event, or probability estimate has entered. Thus the contradiction holds for every noise realization and every start simultaneously. □

The factor 762 is only a convenient explicit choice: $2n$ planar growth steps times transverse diameter plus age cost, followed by one checker step. It is not a sharp extinction horizon. Its significance is that the certificate changes the delay, while all contour constants were fixed before the certificate and before R.

2.16 Age, span and planar growth in one picture

The coherence induction spends one unit of age to remove one unit of transverse ℓ^1 span. The backward-growth argument spends r units of age to cover a planar ball, but can spread the transverse coordinate by as much as Dr. These are different costs. After growth, the readings to be compared all lie on the same old time slice. That slice must still have enough age for the neighbour lemma. The budget is therefore $t - r \geq Dr + 1$, exactly the condition $t \geq (D + 1)r + 1$ used above.

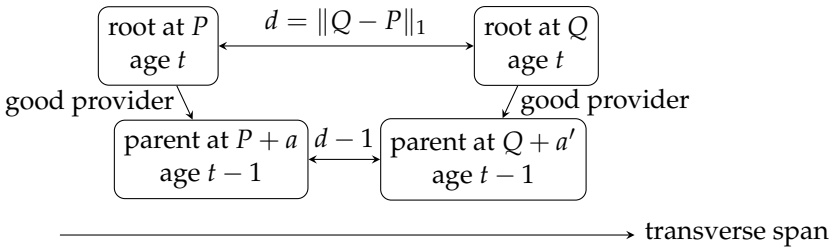

Figure 2: One coherence step, drawn schematically. The raw labels are retained along both arrows. Paired offsets $a' = a - \sigma e_j$ reduce the span while both ages drop by one. Absolute parent locations need not lie between the roots; only the relative span is depicted.

For the growth step, every planar displacement in the radius-r ball is a sum of exactly r members of Γ because zero is allowed. This permits all supplied readings to be placed at the same past time. If one used only shortest paths, their ages would differ, and the stated coherence lemmas would not directly compare them. Padding by zero steps is the small but essential device that puts them on one common time slice.

Transverse readings at that slice lie in a sumset, because at every step a possibly different provider is chosen. The choices may depend on the trajectory and on the branch. Pairwise separation is nevertheless at most Dr, by matching the r summands in any two selected paths. There is no requirement that one provider be good for all branches or that the whole grown patch lie in one transverse plane. Finite-age coherence supplies exactly the matching needed after the branchwise choices have been made.

2.17 Tracing all possible discrepancies at once

Lemma 2.10 (All-input coalescence)**.** *Under the hypothesis of Lemma 2.9,* (9) *holds for all $t \geq 0$. The bound is uniform over $0 < \varepsilon \leq \varepsilon_*$ and over all replacement-value laws that are independent of the configuration and shared across starts.*

Proof. Fix the entire innovation realization. If the output map at (v, t) is nonconstant, choose starts x, x' for which the outputs differ. A discrepant positive-time vertex is not flagged, since a common replacement is constant in the parent word and writes the same value in both trajectories. Different deterministic outputs require a difference at some actual parent in $\Gamma \times S \subset \overline{N}$. Select such a parent and repeat.

At every discrepant vertex of age at least T_R, at least one of its two letters is live: two different letters cannot both equal $\perp$. By Lemma 2.9 that vertex belongs to the single field $\mathcal{M}$. Consequently, for $t > T_R$, existence of *any* discrepant pair forces an ordinary marked backward path of $N = t - T_R$ edges from (v, t) down to age T_R. This is a deterministic inclusion of events. Although the pair was chosen after seeing the innovations, its existence is covered by the same flag-only event; no sum over pairs or initial configurations is taken.

By Lemma 2.5 and (6),

$$a_t \leq \sum_{f \geq \lceil (t-T_R)/16 \rceil} \kappa^f \leq \frac{1}{1-\kappa} \exp\Big(-\frac{\log(1/\kappa)}{16}(t - T_R)\Big).$$

For $t \leq T_R$ the same bound is at least one, so $a_t \leq 1$ completes the estimate for all ages. The auxiliary negative-time flags caused no change to the positive-time process: the deterministic inclusion holds for every extension of its positive-time flags, so averaging over an independent extension is legitimate. Replacement values never enter the tail calculation. □

An event concerning only the number of bad letters for a chosen start would be insufficient here. The mortality statement is already universal over starts, and the marked-path event contains every discrepancy that could occur. This order of quantifiers is what later allows the all-DLR Gibbs theorem to be applied. The alphabet size creates neither a count of candidate trajectories nor a multiplicative factor in the rate.

Finite dependence and the order of quantifiers. For a fixed output vertex and a fixed time t, every initial site read by the composite lies in the t-fold sumset of the finite neighbourhood. There are finitely many such sites and finitely many initial words on them. Thus a pair witnessing nonconstancy can be selected from a finite set, without any measurable-selection

issue. The proof does not sum the probabilities for those pairs. Instead it shows pointwise in the innovation realization that each possible discrepancy, whichever pair witnesses it, produces a path in the same pre-existing marked field.

The resulting probability statement is therefore $\mathbb{P}(\exists x, x' : X_t^x(v) \neq X_t^{x'}(v))$, rather than merely $\sup_{x,x'} \mathbb{P}(X_t^x(v) \neq X_t^{x'}(v))$. The former bounds the latter but need not equal it for a general random-map representation. The stronger event gives a finite set of common maps on which every input row has the same output. This is the hypothesis used by the space–time transfer, and it is also the natural analogue of the retained all-input statement Corollary 5.12, the strengthened form of Theorem B(ii).

No initial word contributes a factor to the contour count. Even a start depending on the entire noise realization is covered by the pointwise inclusion, although attraction of an initial probability law is subsequently proved with an ordinary independent initial row. This distinction prevents any assumption that an adversarial choice of a witness pair remains independent of the marks. Only the fixed terminal subsets in the witness union require independence, and they consist of distinct Bernoulli flag variables.

2.18 From common maps to ergodicity

Lemma 2.11 (Stationary existence and all-law attraction)**.** *A finite-range finite-alphabet PCA is Feller and has a stationary probability measure. If its all-input coefficient tends to zero, its stationary measure is unique and attracts every initial law. In the present non-tiling case, for its stationary law π_R and every finite O,*

$$\sup_{\lambda} d_{\mathrm{TV}}((\lambda P_{R,\varepsilon}^t)_O, (\pi_R)_O) \leq |O| \frac{e^{c(T_R - t)}}{1 - \kappa}. \tag{20}$$

Proof. The expectation of a cylinder function after one step depends on finitely many old coordinates and is continuous. Cylinder functions uniformly approximate continuous functions on the compact product space, so the transition operator preserves continuity. Thus it is Feller.

For any initial law λ, its Cesàro means

$$\nu_N = \frac{1}{N} \sum_{s=0}^{N-1} \lambda P^s$$

have a weakly convergent subsequence by compactness of probability measures on a compact metrizable space. For a bounded continuous f,

$$(\nu_N P - \nu_N)(f) = \frac{(\lambda P^N)(f) - \lambda(f)}{N} \longrightarrow 0.$$

The Feller property passes this identity to a weak limit, which is stationary.

Choose a row of that stationary law π_R and a row of an arbitrary law λ, independently of future innovations, and evolve both by the common maps. On any realization where the time-t map is constant on O, the two output words agree. Their total-variation distance is bounded by $|O| a_t$ through (8). This tends to zero and gives (20). Replacing λ by another stationary law shows equality of its every finite marginal with π_R, hence equality of the laws. Convergence of all finite marginals implies weak convergence: uniformly approximate a bounded continuous function by a cylinder function and then use finite-marginal convergence. No spatial invariance of λ is assumed. □

If $R = \varnothing$, A_R is a singleton and there is only one configuration. All assertions in the non-tiling direction are then immediate. It is harmless to apply the elementary lemmas above

to that case, but later theorems stated for alphabets of at least two symbols will instead treat it separately.

2.19 A tiling protects two reference phases

Let $\theta \in \mathcal{T}(R)$ and clone it transversely:

$$\xi_\theta(u, P) = \theta(u).$$

Every checker retains its central tile and all q votes agree, so $G_R(\xi_\theta) = \xi_\theta$. Also $G_R(\perp) = \perp$ for the all-blank configuration. These are deterministic fixed configurations; their noisy trajectories are controlled relative to the same references.

Lemma 2.12 (Two-cluster memory cover)**.** *Start one trajectory X from ξ_θ and another Y from all blank. For every time, every site and every realization, the errors $X_t(u, P) \neq \theta(u)$ and the live readings $Y_t(u, P) \neq \perp$ are contained in the marked field begun empty at time zero, hence in $\mathcal{M}$. This inclusion holds for arbitrary, even adaptive, letters written at the indicated flags.*

Proof. Both error sets are empty at time zero. Every flagged output can be charged as an error regardless of its value. Consider an unflagged wrong X output at the next time. If at most K checked providers were different from $\theta(u)$, the other $q - K = h$ would force the correct output. Therefore at least $K + 1$ providers are wrong. A checked provider can be wrong only if at least one of its five cardinal inputs differs from the corresponding cloned tile, because the five correct letters form a legal cardinal neighbourhood and the checker would retain the centre. Those $K + 1$ provider offsets cannot all lie in one K-element cluster. Thus there are old errors at parents in two distinct clusters.

An unflagged live Y output requires at least $h = 4K$ live checked providers. Their old central letters are live. Since $h > K$, at least two clusters contain such live old centres. These old errors provide the two supports in (11). Induction on time proves both inclusions. The argument uses no distributional information about the values at flags and remains valid if those values are chosen adaptively. □

Lemma 2.13 (Persistent separation)**.** *For every $v = (u, P)$ and $t \geq 0$,*

$$\mathbb{P}_{\xi_\theta}\{X_t(v) \neq \theta(u)\} \leq p_*, \qquad \mathbb{P}_{\perp}\{X_t(v) \neq \perp\} \leq p_*, \qquad p_* = 10^{-4}. \tag{21}$$

At positive times the two laws have a fixed one-site cylinder gap at least $1 - 2p_* > 1/2$, uniformly in time. There are two distinct stationary laws.

Proof. Apply Lemma 2.4 to the two inclusions of Lemma 2.12. The clopen cylinder $E = \{x : x(v) = \theta(u)\}$ has probability at least $1 - p_*$ from the clone and at most p_* from blank. Its gap persists under time averages. The Feller compactness argument of Lemma 2.11 produces subsequential stationary limits for each average. The event is clopen, so its probability converges under these weak limits. They are distinct by the same gap. At time zero separation is exact. Since $p_* = 10^{-4}$, the displayed positive gap is in fact greater than $1/2$ throughout our interval. □

This already proves nonergodicity without a periodic tiling or a spatially invariant reference. It also proves that a_t cannot tend to zero on tiling inputs: the all-input discrepancy probability dominates the total-variation distance of this fixed pair of starts, which is at least $1 - 2p_*$.

2.20 Fully translation-invariant stationary laws

The stationary laws just selected need not be spatially invariant, because a chosen tiling need not be. To obtain the stronger statement of Theorem 2.1, first average the Dirac law at θ over larger and larger planar squares. A weak subsequential limit ν exists on the compact tiling space. For each fixed planar translation z, the fraction of the averaging square changed by z tends to zero; therefore ν is shift invariant and remains supported on tilings.

Clone a tiling sampled from ν in all transverse coordinates. Its row law is invariant under all translations of $\mathbb{Z}^6$. The kernel commutes with those translations. All its time marginals and their Cesàro limits are therefore spatially invariant. Integrating the first bound of (21) over ν shows that the live probability at every site is at least $1 - p_*$. From the all-blank start it is at most p_*. The corresponding invariant limits are distinguished by the fixed event $\{x : x(v) \neq \perp\}$. This proves the two fully translation-invariant stationary laws, even for an aperiodic tile set.

These laws are not asserted to be extreme in the *full* stationary simplex merely because they are translation invariant. Extremality in that full simplex is a separate issue treated for the canonical phases in Theorem 3.10. The present argument needs only stationarity, invariance and the live-density separation.

2.21 Complete histories: the qualitative route

For clarity we also give the complete-history version of the non-tiling proof, corresponding to [12, Lemmas 5.3 and 8.5]. It is not needed to improve the quantitative estimate, but isolates why no bounded-width refutation hypothesis is required.

Fix all two-sided flags and values. A complete trajectory is a sequence $(X_t)_{t\in\mathbb{Z}}$ satisfying the prescribed local update at every integer time. For any fixed finite transverse span, the inductions of Lemmas 2.6 and 2.7 use only finitely many past steps. With an unlimited past their conclusions hold at every span. Thus all unmarked live readings at a planar site and time have one letter, and neighbouring planar readings match.

Let $V_t \subset \mathbb{Z}^2$ be the planar sites admitting an unmarked live reading at time t. A good provider exists at every such output, and supplies five unmarked live old inputs. Consequently

$$V_t + \Gamma \subset V_{t-1}.$$

If $u \in V_t$, iteration gives $u + [-n, n]^2 \subset V_{t-2n}$ for every n. The coherent letters there form legal square patterns of arbitrarily large size. Compactness as in Lemma 2.8 yields a tiling. Therefore on a non-tiling input every live reading of every complete trajectory lies in $\mathcal{M}$.

If two complete trajectories with the same innovations differ at a vertex, that vertex is unflagged, is live in at least one trajectory, and hence is marked. It has a discrepant actual parent. Iteration would give an infinite ordinary marked path, excluded almost surely by Lemma 2.5. This exclusion is simultaneous at every root and the deterministic implication is simultaneous for every complete trajectory. Thus there is almost surely at most one complete history.

There is at least one. To see this without a stationarity assumption, take trajectories from successively earlier arbitrary starts. On any fixed finite set of sites and times there are finitely many possible histories. A diagonal subsequence in the countable product of finite alphabets stabilizes on all such finite sets. Every local update equation is a closed finite-coordinate condition and holds eventually, so it holds in the limit. Extend forward by the specified maps. This constructs a complete trajectory for the fixed innovation realization.

For a more explicit coalescence formulation, let K_j be the image at time zero of all configurations started at time $-j$. The sets K_j are nonempty, compact and nested. Their intersection is exactly the time-zero states of complete histories, by the same diagonal extension. Uniqueness makes the intersection a singleton. For a finite O, projection commutes with this nested compact intersection: any pattern surviving in every projection has witnesses in a nested family of nonempty closed subsets and therefore a common witness. The finite sets of possible O-patterns must eventually be singletons. This gives almost-sure finite pullback coalescence on O.

The failure events decrease and their probabilities tend to zero. Time stationarity of the independent innovations identifies those probabilities with the forward nonconstancy probabilities. No summability argument is necessary. This qualitative route proves all-law ergodicity by Lemma 2.11; the forbidden-square argument earlier supplies the explicit exponential bound and its universal slope.

2.22 Completion and dependence of the estimates

Proof of Theorem 2.1. If R does not tile, Lemma 2.8 supplies a finite n_R. The deterministic mortality lemma and marked-path estimate give (9), and Lemma 2.11 gives uniqueness and attraction from every initial law. If R tiles, Lemmas 2.12 and 2.13 and planar averaging give two distinct fully translation-invariant stationary laws. The empty input is the one-point case. These exhaust all finite inputs. □

The dependencies can be read directly from the proof. The flag field, graph volume growth, potentials, arm count and path-tail base are fixed by the provider geometry. The raw alphabet has no role in their numerical constants. The non-tiling certificate enters only through $T_R = 762n_R + 1$. For the memory side a tiling supplies the reference, but no period or invariant tiling is required until spatial averaging is requested. The replacement values enter only through equality of the common constant maps at flags; their probabilities never enter the contour estimate.

This is also why the new route is shorter than the old inspector route. Thick boxes give the strict pairing surplus (17), which transports equality through any finite transverse distance one unit at a time. A finite forbidden square then kills every unmarked live output deterministically after a delay. The thin-provider rule in Section 4 has no comparable unit-shift surplus, so its proof transports a fixed finite collection of inspectors and assumes token refutability. Neither argument asserts that every non-tiling Wang set has bounded token width.

3 Consequences for the natural rule

Throughout this section G_R, $\overline{N}$, ε_*, κ, c and p_* have the meanings of Section 2. Unless otherwise stated $0 < \varepsilon \le \varepsilon_*$ and replacement is uniform. The old machine-family corollaries remain separately in Section 10; this section applies directly to the raw input R.

3.1 The domino reduction and nilpotency

Corollary 3.1 (Theorem A on $\mathbb{Z}^6$)**.** *There is a total computable map $M \mapsto W_M$ to finite Wang sets such that W_M tiles if and only if M does not halt. For every $\varepsilon \in (0, \varepsilon_*]$, G_{W_M} with uniform replacement is a positive-rate PCA on $\mathbb{Z}^6$ that is ergodic if and only if M halts. In particular ergodicity is undecidable at each fixed rational rate in this interval.*

Proof. Use the classical effective unseeded domino reduction of Berger and Robinson [3, 43], with translations of edge-coloured squares and no distinguished origin tile. Its output tiles precisely on the nonhalting branch. Apply Theorem 2.1 to that same output. The rule table is an explicit finite construction from its raw tiles. A decision procedure for ergodicity would decide halting. □

The reduction through our machine family. The printed proof uses the classical Berger–Robinson reduction [3, 43]. Our total computable family R_M (Definition 8.17) also supplies it: R_M tiles if and only if M does not halt. The nonhalting direction is Proposition 8.18. If M halts, Theorem 11.1 makes R_M 3-token refutable, so Theorem B (Theorem 5.1) makes its quorum rule on $\mathbb{Z}^8$ ergodic at $0 < \varepsilon \le 490000^{-99}$. A tiling set would instead give memory by Theorem B (Theorem 6.2), hence nonergodicity by Lemma 4.13, a contradiction. This is the route taken by the Lean development: `FNR.domino_reduction` and `FNR.domino_undecidable` prove the reduction and the undecidability of the domino problem, and `FNR.theorem_A` and `FNR.machine_corollary` verify Theorem A on $\mathbb{Z}^6$ and all of Corollary 3.1 on the version 2.2.0 interval $0 < \varepsilon \le (2(18q)^2)^{-200}$, without a named hypothesis (Section C). Berger and Robinson are the classical sources of the domino reduction and its undecidability consequence.

The classical proof above uses no theorem about bounded token width. The detailed machine family gives the interval $0 < \varepsilon \le 490000^{-99}$ for its own quorum rule on $\mathbb{Z}^8$. The natural-rule interval $0 < \varepsilon \le 10^{-24}$ is now larger. The machine-route constant is unchanged.

Corollary 3.2 (Nilpotency)**.** *At zero noise G_R is nilpotent if and only if R does not tile. More precisely, if no legal $n \times n$ pattern exists, then*

$$G_R^{762n+1}(x) \equiv \perp \quad \textit{for every } x \in \Omega_R.$$

Proof. With no flags the least marked field is empty. The deterministic mortality lemma therefore makes every output blank after the stated time. Conversely a tiling clone is a nonblank fixed point, while blank is also fixed, so no iterate can be constant on all inputs. □

This is the thick-quorum form of the tiling-to-nilpotency construction of Čulik–Pachl–Yu [7], with Kari's one-dimensional nilpotency theorem [26] as the classical undecidability counterpart. Generic stability of a particular nilpotent rule under sufficiently small noise [34] does not by itself provide a threshold independent of its extinction time. Here the marked-path estimate provides that uniform threshold.

3.2 All space–time Gibbs measures

For $z = (v, t) \in \mathbb{Z}^6 \times \mathbb{Z}$, let

$$P_z = \{(v, t)\} \cup \{(v + a, t - 1) : a \in \overline{N}\}.$$

Define the normalized finite-range interaction

$$\Psi_{R,\varepsilon,z}(\sigma) = -\log p_{R,\varepsilon}(\sigma(v, t) \mid (\sigma(v + a, t - 1))_{a \in \overline{N}}). \tag{22}$$

Its finite-volume specification γ_V^τ multiplies all factors with $P_z \cap V \neq \varnothing$, including factors of exterior children, and normalizes with exterior configuration τ. A Gibbs measure means a DLR state for this specification, with no invariance imposed. This is the space–time convention of [9, Section 2] and [21].

Corollary 3.3 (Gibbs uniqueness and weak spatial mixing)**.** *If R does not tile, (22) on $\mathbb{Z}^7$ has exactly one Gibbs measure $\bar{\pi}_R$, with no invariance assumed. There are $C_*, c_* > 0$ such that for every finite $V \subset \mathbb{Z}^7$, every $E \subset V$ and all exteriors τ, τ',*

$$d_{\mathrm{TV}}((\gamma_V^\tau)_E, (\bar{\pi}_R)_E) \leq C_* \sum_{z \in E} e^{-c_* \operatorname{dist}_\infty(z, V^c)}, \tag{23}$$

$$d_{\mathrm{TV}}((\gamma_V^\tau)_E, (\gamma_V^{\tau'})_E) \leq 2C_* \sum_{z \in E} e^{-c_* \operatorname{dist}_\infty(z, V^c)}. \tag{24}$$

If R tiles, there are at least two fully space–time translation-invariant Gibbs measures. Thus uniqueness, uniqueness among fully translation-invariant Gibbs measures, and exponential weak spatial mixing each hold exactly on non-tiling inputs.

Proof. For $R \neq \varnothing$, the alphabet is finite with at least two symbols, the fixed finite neighbourhood $\overline{N}$ contains zero, and the table is bounded below by $\varepsilon/|A_R| > 0$. Equation (1) is an independent local-map representation. Its common-map coefficient is exactly (7); Lemma 2.10 gives $a_t \leq C_R e^{-ct}$ for every $t \geq 0$. These check every hypothesis of [9, Corollary 13.7], in spatial dimension six. That corollary, through its prefix-cap Theorem 13.3, gives all-DLR uniqueness and (23)–(24). Pairwise convergence would not have supplied its all-input hypothesis. The one-symbol case $R = \varnothing$ has a single space–time configuration and is immediate.

For tiling R, take the two spatially translation-invariant stationary laws constructed in Section 2. Their two-sided stationary Markov trajectories are DLR states by [9, Lemma 2.4], which does not assume spatial invariance. They are fully translation invariant by spatial invariance and stationarity. Their time-zero live densities differ, so the Gibbs measures are distinct. Finally weak spatial mixing implies uniqueness: condition any DLR state outside a volume containing E at distance tending to infinity and integrate (24); the right side tends to zero. Thus weak mixing is impossible on the tiling branch. □

Distance in these estimates is the sup norm, as in [9, Definition 2.3]. Volumes can have holes and need not be connected. The boundary distance is to the entire complement, not just the set on which two exteriors disagree; strong spatial mixing is not claimed. The constants may depend on R and ε. In particular no uniformity of the Gibbs constants as $\varepsilon \downarrow 0$ follows from the universal dynamical slope. The correspondence has its classical source in Goldstein–Kuik–Lebowitz–Maes [21]; the transfer template is the retained Corollary 10.8, and the translation-invariant converse used for stacking is [9, Corollary 10.4].

Remark 3.4 (Binary energies and their temperature). Let $m = |A_R|$, $a = 1 - \varepsilon + \varepsilon/m$ and $b = \varepsilon/m$. Then

$$p_{R,\varepsilon}(s \mid y) = a \exp\{-\beta(R,\varepsilon)\mathbf{1}_{\{s \neq G_R(y)\}}\}, \qquad \beta(R,\varepsilon) = \log\left(1 + \frac{m(1-\varepsilon)}{\varepsilon}\right).$$

The constant a cancels from each finite-volume normalization. Thus the same specification is given by the $\{0,1\}$ energies $\mathbf{1}_{\{\sigma_z \neq G_R(\sigma_{\text{parents}})\}}$ at this inverse temperature, which depends on R through m. This is the conversion of [12, Corollary B], with inverse $\varepsilon(\beta) = m/(e^\beta + m - 1)$.

No common finite β gives the tiling classification for this direct binary-energy family. At fixed β, the depth-t basic composite can be nonconstant only along a backward path of nonreplacement updates, so

$$a_t \leq [|\overline{N}|(1 - \varepsilon(\beta))]^t, \qquad 1 - \varepsilon(\beta) = \frac{e^\beta - 1}{e^\beta + m - 1} \longrightarrow 0.$$

For all large m this gives exponential coalescence and hence all-DLR uniqueness, even though arbitrarily large tiling sets exist (take many isolated self-matching tiles). This high-noise mechanism is consistent with the Dobrushin and Dobrushin–Shlosman regimes [15, 16]; see [9, Remark 13.9]. At each fixed β, uniqueness for this direct family is decidable nonuniformly by a finite table of the remaining bounded-size colour-equality patterns. We use the normalized fixed-ε convention throughout.

3.3 Robustness on the same neighbourhood

For tables on a specified alphabet and neighbourhood write

$$d(P,Q) = \max_y d_{\mathrm{TV}}(P(\cdot \mid y), Q(\cdot \mid y)).$$

The independence of the stochastic updates at distinct space–time sites is part of the definition of a PCA table; table distance does not allow correlated update randomness.

Corollary 3.5 (A uniform radius)**.** *For every finite R, every $0 < \varepsilon \leq \varepsilon_*$, and every local PCA kernel Q on A_R with neighbourhood $\overline{N}$ satisfying*

$$d(Q, p_{R,\varepsilon}) \leq \varepsilon_*,$$

Q is ergodic if and only if R does not tile. The radius is the same for every R and every permitted ε. The dynamical assertion allows zero entries in Q.

Proof. Put $e_0 = 2\varepsilon_* = 2 \cdot 10^{-24}$. At every parent word,

$$Q(G_R(y) \mid y) \geq 1 - \varepsilon - \varepsilon_* \geq 1 - e_0.$$

Entrywise subtraction therefore gives a stochastic table Q' with

$$Q = (1 - e_0)\delta_{G_R} + e_0 Q'.$$

At each exceptional flag sample a local map by drawing one $Q'(\cdot \mid y)$ output for each parent word y; samples at different update sites are independent. This finite random-map representation evaluates identical parent words identically. It is the exceptional-map device of Corollary 10.12 and Remark 10.13 and [12, Theorem B.1], which in turn cites [9, Proposition 4.5].

Use the same marked closure with source density e_0. Away from marks the update is still G_R, so the exclusion budget, coherence and forbidden-square mortality are unchanged. At an exceptional flag two different output values need not coincide. They still require a discrepant parent in $\overline{N}$, because the shared sampled map is a function. At every late discrepant vertex a live symbol marks it by mortality. Thus the whole backward discrepancy path is marked, even when it passes through exceptional flags. By Lemma 2.5, at source density e_0 the tail base is at most $\kappa_1 = 1/4$ on the same enclosing graph. The all-input estimate has slope $\log(1/\kappa_1)/16 = (\log 2)/8$ and the same delay T_R, hence is at most $(4/3)\exp\{((\log 2)/8)(T_R - t)\}$.

It gives exponential coalescence on every non-tiler. Feller compactness and coupling give ergodicity.

On a tiler, charge every exceptional flag as a source. Off flags the two-cluster memory cover is unchanged. The sparse-root estimate at density e_0 bounds both errors by $p_* = 10^{-4}$, leaving a live-density gap $1 - 2p_* > 0$. Planar tiling averaging and temporal Cesàro limits give two fully translation-invariant stationary laws. □

If a Gibbs statement for *every* nearby table is desired, also impose $d(Q, p_{R,\varepsilon}) \le \varepsilon/(2|A_R|)$. Then every entry of Q is at least $\varepsilon/(2|A_R|) > 0$; the same all-input proof and [9, Corollary 13.7] give the corresponding normalized Gibbs dichotomy. This smaller positivity radius depends on the alphabet size. The dynamical radius ε_* does not. Both statements fix the alphabet and the neighbourhood, as do Corollary 10.12 and Remark 10.13 and [12, Theorem B.1].

Remark 3.6 (Why a neighbourhood restriction is essential). Take R consisting of a single tile with different north and south colours. Its checker, and hence G_R, is identically blank. Nevertheless every total-variation ball about $p_{R,\varepsilon}$ contains a strictly positive finite-range kernel on the same binary alphabet with two translation-invariant stationary laws, if its neighbourhood may grow. Here is the construction and its finite estimate.

Code the tile by one and blank by zero, so the base kernel writes one with probability $p = \varepsilon/2$. Given a target radius choose $0 < a < \min\{\text{radius}, 1 - p\}$ and $\vartheta = p + a/2$. Use three coordinates (z_1, z_2, w) of $\mathbb{Z}^6$. For $A = \{0, e_1, e_2\}$ set

$$T_i(x; z, w) = \mathbf{1}\left\{\frac{1}{L}\sum_{k=0}^{L-1} x(z + i, w + k) \ge \vartheta\right\}, \qquad F_L = \text{majority}(T_0, T_{e_1}, T_{e_2}),$$

and $Q_L(1 \mid x) = p + aF_L(x)$. It is strictly positive and $d(P, Q_L) = a$. Write $B_t = F_L(X_t)$ and realize $X_{t+1} = \mathbf{1}_{\{U_{t+1} \le p + aB_t\}}$ with independent uniforms.

Partition the w-axis into blocks $Lj + \{0, \dots, L - 1\}$. A source at $(z, j, t + 1)$ is any failure of the L possible tests in that block, at any of the three offsets and either hypothetical Bernoulli level p or $p + a$. Hoeffding's inequality [24] gives source probability at most $s_L = 6Le^{-La^2/2}$. Its uniform footprint is $(z + A) \times [Lj, Lj + 2L - 2]$ at one time. Colour (z_1, z_2, j) by coordinate residues modulo two. The eight colours have disjoint same-colour footprints, and different times are independent. Hence a fixed set of f sources has probability at most $s_L^{f/8} = e_L^f$.

Off a source a wrong majority requires wrong tests at two different offsets. Each wrong test must contain a wrong old B value, since otherwise all its bits equal the hypothetical correct Bernoulli level. The coarse parents lie in $A \times \{0, 1\}$, with two distinct planar-offset clusters. This is Toom's north-east-centre majority mechanism [47]. The balanced potentials

$$3(z_1 + z_2) + 2t, \quad -3z_1 - t, \quad -3z_2 - t$$

have good gains at least one and arbitrary gains at least -2. The finite-arm argument with three charges gives at most $9(f - 1)$ edges. The undirected degree is at most 12. The doubled-edge Euler encoding in the proof of Lemma 5.11, which uses only a degree bound, shows that

at most 12^{2^n} connected n-edge sets contain a given root; with at most 2^{n+1} terminal subsets this gives at most $2B_0^n$ encodings, $B_0 = 2 \cdot 12^2 = 288$. Together with the fixed-source bound it gives error probability

$$\rho \le e_L \left(1 + \frac{2\lambda_L}{1-\lambda_L}\right), \qquad \lambda_L = 288 e_L^{1/9}.$$

Choose a finite L so that $e_L < 1/8$ and $\lambda_L < 1/3$. Then $\rho < 1/4$ in both homogeneous phases, uniformly in time. Their next-row one-site densities are at most $p + a\rho$ and at least $p + a(1-\rho)$, respectively. The gap exceeds $a/2$. The homogeneous starts and kernel are translation invariant, so temporal Cesàro limits give two distinct translation-invariant stationary laws. This disproves unrestricted-neighbourhood robustness without changing the alphabet. The construction uses eight footprint colours; twenty-seven would also suffice.

3.4 Exact complexity on this family

Corollary 3.7 (Completeness at a fixed rational rate)**.** *At each fixed rational $0 < \varepsilon \le \varepsilon_*$, the following sets of raw finite Wang inputs coincide and are Σ_1^0-complete:*

(i) *the non-tiling inputs;*

(ii) *the inputs for which the noisy G_R is ergodic;*

(iii) *the inputs for which (22) has exactly one Gibbs measure;*

(iv) *the inputs whose specification has arbitrary-volume exponential weak spatial mixing.*

Their complements are Π_1^0-complete. Existence of two distinct fully translation-invariant stationary laws, and existence of two distinct fully translation-invariant space–time Gibbs measures, are also Π_1^0-complete on this family.

Proof. Non-tiling is computably enumerable by the forbidden-square search. The Berger–Robinson reduction sends halting to non-tiling and nonhalting to tiling, so non-tiling is many-one complete among c.e. sets. Theorem 2.1 and Corollary 3.3 identify all four predicates with it. The memory construction gives two fully translation-invariant stationary laws and two fully translation-invariant Gibbs measures on every tiler; uniqueness on non-tilers gives the converses. Taking complements proves the Π_1^0 statements. □

The directions matter: nonergodicity is Π_1^0-complete here, not Σ_1^0-complete. These are exact degrees on the specified natural family, not a claim about the class of all positive-rate tables. The latter still has the Π_2^0 upper bound recalled in Remark 10.9 from [9, Remark 10.9]. The precursors for the restricted-family placement are Corollary 10.5 and Remark 10.9 and [12, Corollary 9.12].

3.5 Other replacement laws and the forgetting time

Corollary 3.8 (Replacement values)**.** *For every fixed configuration-independent law ν_R on A_R, use independent replacements with law ν_R at independent Bernoulli flags of rate ε. The dynamical classification, two-translation-invariant-law conclusion and exponential all-input estimate of Theorem 2.1 still hold, with the same constants. No support assumption is needed dynamically. If ν_R has full support, the normalized positive table also has the Gibbs dichotomy and weak mixing of Corollary 3.3. The memory bounds (21) hold even when the values written at the flags are arbitrary and adaptive.*

Proof. All contour estimates concern only the flags. Coherence and mortality use only updates outside them. At a flag the shared configuration-independent value makes the local map constant, whatever its law, so all-input discrepancy tracing is unchanged. The two-cluster memory cover is deterministic for every choice of the values and already proves its adaptive form. With full support the table is bounded below by $\varepsilon \min_a \nu_R(a) > 0$, and [9, Corollary 13.7] applies. Without full support positivity can fail, so that Gibbs import is unavailable. □

In particular $\nu_R = \delta_\perp$ is allowed dynamically: the blank configuration then absorbs, but a distinct tiling phase still persists. This is the raw-rule analogue of Corollary 10.6 and Remark 10.14(a). The adaptive statement is an error bound; a general adaptive choice need not define a time-homogeneous PCA or a stationary process.

For the timing statement write

$$\delta_R(t) = \max_{x,x'} d_{\mathrm{TV}}\big(\mathrm{Law}_x\, X_t(0), \mathrm{Law}_{x'}\, X_t(0)\big).$$

The maximum is over finitely many words on the depth-t initial cone, so it exists.

Corollary 3.9 (No total forgetting-time bound)**.** *Fix rational $0 < \varepsilon \le \varepsilon_*$. No total computable function of R bounds, on every non-tiling input, the first time $a_t < 1/2$. The same is true for the first time $\delta_R(t) < 1/2$. There is a partial algorithm terminating on every non-tiling input that produces a valid delay and prefactor for* (9)*, hence a forgetting modulus for every prescribed accuracy.*

Proof. At rational rate, finite-cone enumeration computes both a_t and $\delta_R(t)$ exactly as rationals. Enumerate all initial words on the finite cone and all flags and replacement letters in it; the weights are rational. For a_t test constancy for each fixed realization, and for $\delta_R(t)$ compute the finite family of output distributions and its largest total-variation distance.

Both sequences are nonincreasing. For a_t, use nested pullback images under common maps and time stationarity. For $\delta_R(t)$, any time-$(t+1)$ output law is a mixture of time-t output laws, and taking their convex hull cannot increase total-variation diameter. On non-tilers they tend to zero. On tilers the fixed clone-versus-blank cylinder gap gives $\delta_R(t) \ge 1 - 2p_* > 1/2$ at every t, and $a_t \ge \delta_R(t)$.

If a total computable bound $b(R)$ worked on all non-tilers, evaluate the corresponding rational coefficient at $b(R)$. Monotonicity would make it less than $1/2$ on every non-tiler and greater than $1/2$ on every tiler. This would decide tiling, a contradiction. For the partial algorithm search for a forbidden square, compute $T_R = 762 n_R + 1$, then use the explicit geometric bound for the requested accuracy. □

This specializes the timing argument of Corollary 10.7; compare [12, Corollaries 9.10–9.11]. The slope c is universal and computable. The impossible object is a *total* bound on the delay, prefactor or forgetting modulus over all finite inputs, even though it is required to be valid only on non-tilers. A partial computation of these quantities on the non-tiling domain is fully consistent with undecidability.

3.6 Canonical phases and the conditional continuum

Theorem 3.10 (An extreme phase for each tiling)**.** *For $\xi = \perp$ and for each clone $\xi = \xi_\theta$, $\theta \in \mathcal{T}(R)$, there is a canonical stationary law μ_ξ, extreme among* all *stationary laws, whose stationary trajectory is mixing under time translations. The map $\theta \mapsto \mu_\theta$ is continuous and injective. If $\mathcal{T}(R)$ is uncountable, the noisy G_R has exactly continuum many extreme stationary laws and* (22) *has exactly continuum many extremal Gibbs measures. No spatial invariance is asserted for the individual tiling-labelled extremes.*

Proof. Use the two-sided innovations and start the reference ξ at times $-s$ and $-t$, where $s > t$. By Lemma 2.12, errors of both runs relative to the same reference lie in the single global $\mathcal{M}$. A discrepancy at time zero traces a backward discrepant path through the shared time interval of length t. Every vertex of this path is an error in at least one of the two runs and is therefore marked. This inclusion holds for existence of *any* $s > t$, without a union over s. Hence at a fixed root

$$\mathbb{P}\{\exists s > t : X_0^{(-s,\xi)}(v) \neq X_0^{(-t,\xi)}(v)\} \leq \frac{e^{-ct}}{1-\kappa}. \tag{25}$$

These events decrease: once all longer pullbacks agree with the depth-t value, they agree with one another. Their probabilities tend to zero. Thus the reference pullbacks eventually stabilize almost surely at each root. Countability gives simultaneous limits $H_\xi(v,t)$ at all space–time roots. Finite parent sets stabilize together, so the limiting field satisfies the update equations. It depends only on past innovations, and fresh innovations retain the required conditional kernel. Time translation just reindexes the constant-in-time reference starts. The history is therefore a time-equivariant measurable factor of independent whole-noise rows, so is stationary and mixing in time.

The stationary-history map from stationary row laws is affine and injective, with inverse the time-zero marginal. If μ_ξ admitted a nontrivial stationary decomposition, its stationary Markov history would admit the corresponding nontrivial decomposition into time-invariant histories. This contradicts its time ergodicity. Thus μ_ξ is extreme among all stationary laws.

The error bounds pass to the limits:

$$\mu_\theta\{x(u,P) = \theta(u)\} \geq 1 - p_*, \qquad \mu_\perp\{x(u,P) \neq \perp\} \leq p_*.$$

Distinct tilings differ at some planar site; their laws are separated there by at least $1 - 2p_*$. In particular the map is injective. For continuity, approximate any finite marginal uniformly in θ by a finite-depth reference pullback using (25). At fixed depth it reads only a finite planar part of θ and is locally constant as a function of that tiling. Uniform approximation proves continuity. Compactness of $\mathcal{T}(R)$ makes this a homeomorphism onto its image.

An uncountable compact metrizable tiling space has cardinal continuum. The embedding gives that many stationary extremes, and the space of all probability laws on a countable finite-alphabet product has cardinal at most continuum. For Gibbs extremes, pass the tiling-labelled row laws to their stationary trajectories, which are DLR by [9, Lemma 2.4]. Their time-zero cylinders give an uncountable family uniformly separated in total variation.

Apply the cardinality argument of [12, Theorem 9.6]. The DLR set is compact, convex and metrizable. If it had only countably many extreme points, the metrizable Choquet theorem [6] would represent every state as a countable mixture of them. Truncating the mixture and rationalizing its weights would give a countable total-variation-dense subset, contradicting the separated family. Its extreme set is G_δ: the nonextreme points form the union, over $j \geq 1$, of compact sets of midpoints of pairs at weak distance at least $1/j$. Every nonextreme point is such a midpoint after shortening a segment through it. The perfect-set theorem [27] then gives continuum many Gibbs extremes. This argument does not identify each stationary extreme with a Gibbs extreme; that implication is not used. □

The reference-pullback construction is related to coupling from the past [41, 49]; eventual agreement of these particular reference runs does not certify constancy against arbitrary older starts. No finitary driving-map claim on tiling inputs follows from it. The conditional continuum is already applicable to strongly aperiodic sets: a nonempty minimal tiling subsystem

with an isolated point would be finite, whereas aperiodicity excludes a finite orbit. Such a subsystem is therefore perfect and uncountable. Berger and Robinson supply the classical aperiodic context [3, 43].

For arbitrary tiling R the countable case remains genuinely open here. A single self-matching tile has one tiling, and a countably infinite example can be made from a single interface position. Tileability alone does not give the uncountable hypothesis. The next construction supplies the unconditional continuum after adding one coordinate, without asserting that the gap for the original rule has been closed.

3.7 The stack on $\mathbb{Z}^7$

Define $\widehat{G}_R$ on $\mathbb{Z}^6 \times \mathbb{Z}$ by updating slice h from slice $h+1$ with G_R, using independent replacement innovations at every site of every slice. Thus its parent offsets are $(a, 1)$, $a \in \overline{N}$. Along each characteristic $h + t = j$ the process is an independent copy of the original PCA. This is exactly the stacking device of Definition 10.15, Lemma 10.16, and Corollary 10.17, now with the natural rule as its base.

Corollary 3.11 (One invariant measure or an arc)**.** *For every $0 < \varepsilon \le \varepsilon_*$:*

(i) *If R does not tile, $\widehat{G}_R$ has exactly one invariant measure, attracts every initial law exponentially on every finite window, and has the same one-site all-input coefficient and one-site forgetting diameter as its base rule.*

(ii) *If R tiles, $\widehat{G}_R$ has an arc $\{\mu_p : 0 \le p \le 1\}$ of distinct invariant measures, each fully spatially translation invariant, extreme among all invariant measures, and mixing in time. The parameterization is a homeomorphism onto its image. For $0 < p < 1$ these measures need not be ergodic under translations within a slice; in this construction they are not.*

(iii) *Its normalized space–time interaction on $\mathbb{Z}^8$ has exactly one Gibbs measure on the non-tiling branch. On the tiling branch the stationary trajectories of the arc are fully translation-invariant Gibbs measures, extreme in the translation-invariant Gibbs class. For $0 < p < 1$ they are not extreme among all Gibbs measures.*

(iv) *At each fixed rational permitted rate, no algorithm decides which branch holds. The one-site forgetting time on non-tilers has no total computable bound.*

Proof. For each j set $Z^j_s(v) = X_s(v, j - s)$. Substituting the update equation shows that Z^j follows the base rule from initial slice j. The innovation families for different j are disjoint. As the entire stack start varies, its slice $h + t$ ranges over all base starts. Thus the one-site common-map coefficient and the supremum one-site output-law diameter equal their base counterparts. This proves the exponential and timing claims in (i) and (iv), by Lemmas 2.10 and 2.11 and Corollary 3.9.

On a tiler choose the two fully translation-invariant stationary base laws ν^-, ν^+ with live densities $a_- \le p_*$ and $a_+ \ge 1 - p_*$. Put $\rho_p = (1-p)\nu^- + p\nu^+$ and let μ_p have independent slices, each of law ρ_p. Updating from slice $h + 1$ preserves this product, since ρ_p is stationary and characteristic noises are independent. The product is fully spatially invariant. Finite cylinder probabilities are polynomials in p, and the one-site live probability $(1-p)a_- + pa_+$ is strictly increasing. Hence the parameterization is continuous and injective, and compactness gives the homeomorphism claim.

For extremality, suppose a stationary law α satisfies $\lambda\alpha \le \mu_p$ for some $\lambda > 0$. Its bounded density can be approximated in $L^1(\mu_p)$ by a function f of finitely many slices. For a cylinder

observable g, the function $P^t g$ depends only on its original slice indices translated by t. At large t those slices are disjoint from the ones read by f, and product independence gives

$$\alpha(g) = \alpha(P^t g) \approx \mu_p(f)\mu_p(P^t g) = \mu_p(f)\mu_p(g).$$

The error is at most $\|g\|_\infty$ times the L^1 approximation error, while $\mu_p(f)$ tends to one. Thus $\alpha(g) = \mu_p(g)$ for every cylinder, so $\alpha = \mu_p$. This proves extremality in the full stationary set. Functions of two finite time windows read disjoint finite sets of characteristic indices at a sufficiently large time separation, giving exact independence and hence time mixing. For $0 < p < 1$, the single-slice law ρ_p is a proper mixture of distinct slice-translation-invariant laws, so it has a nontrivial slice-translation-invariant event. Its lift to slice zero proves the last assertion of (ii).

For (iii), add the ignored origin parent if necessary; factors introduced only through that unused parent cancel from finite-volume normalizations. On a non-tiler the unchanged all-input coefficient checks [9, Corollary 13.7] in spatial dimension seven. On a tiler, [9, Lemma 2.4 and Corollary 10.4] identify the fully translation-invariant Gibbs class affinely with fully translation-invariant stationary trajectories. Taking time-zero marginals and using (ii) proves extremality in that class and continuity of the arc.

Finally every interaction factor lies within one characteristic hyperplane $h + t = j$. The specification factors over these hyperplanes. Under the trajectory of μ_p, their restrictions are independent, each with law $(1-p)\bar{\nu}^- + p\bar{\nu}^+$. For $0 < p < 1$, decomposing just the hyperplane $j = 0$ into its two Gibbs laws, while leaving the others unchanged, gives a proper convex decomposition into distinct Gibbs states. Thus the trajectory is not extreme in the full DLR set. This is precisely the distinction in Corollary 10.17(iii); no stronger Gibbs extremality is inferred. □

3.8 Why this mechanism uses four transverse dimensions

Remark 3.12 (The counting certificate and Toom's criterion). Consider c_0 disjoint equal boxes of side m_0, each with K_0 providers, and quorum h_0 among $q_0 = c_0 K_0$ providers. Suppose the proof uses the same two-cluster mark rule and the same independent two-root exclusion budget. Unit-shift pairing supplies $c_0(1 - 1/m_0)K_0$ pairs. Each root can discard $(q_0 - h_0) + K_0$ offsets. Strict guaranteed surplus requires

$$2h_0 > (c_0 + 2 + c_0/m_0)K_0.$$

The two-cluster false-erasure cover requires $q_0 - h_0 \geq K_0$, hence $h_0 \leq (c_0 - 1)K_0$. Together these imply

$$c_0 > 4 + c_0/m_0, \qquad c_0 \geq 5.$$

This is necessity for the *pair-overlap counting certificate*, not for every possible coherence argument. The five-cluster count was first isolated in the cross-family check of the natural-rule proof (see "How this work was done").

Now suppose $d_\perp \leq 3$ and choose five cluster representatives $a_1, \ldots, a_5 \in \mathbb{R}^{d_\perp}$. The vectors $(a_i, 1)$ are linearly dependent. Splitting a nonzero affine dependence into its positive and negative parts and normalizing gives a Radon point z in the convex hulls of two disjoint nonempty groups [42]. At least one group avoids any given index, so

$$z \in \bigcap_j \operatorname{conv}\{a_i : i \neq j\}. \tag{26}$$

A balanced strict affine-charge certificate asks every charge to have a positive-gain step on one parent of every distinct-cluster support pair. Choose planar displacement zero. Each charge then has an affine gain function on the representatives and at most one nonpositive representative gain. Evaluate it at the deletion hull omitting that representative; it is positive at z. Every charge is therefore positive at z, contradicting the identically zero balanced sum. Such a certificate needs at least four transverse dimensions.

This is Toom's eroder criterion, not a new general obstruction to low-dimensional memory. For the two-cluster binary mark rule the minimal zero-sets are $Z_j = \Lambda \times (S \setminus S_j)$. Its erosion criterion is $\bigcap_j \operatorname{conv}(Z_j) = \varnothing$ [47]; see the precise formulation of the criterion in [46, Proposition 3]. The representatives in (26), with zero planar coordinate, lie in every corresponding deletion hull and make that intersection nonempty. Thus this particular mark rule fails the eroder criterion in $d_\perp \leq 3$. Nothing here excludes a different natural rule, a different support restriction, or another proof on $\mathbb{Z}^5$ or $\mathbb{Z}^4$.

Helly's theorem [23] explains why the same certificate obstruction persists for positive weighted quorums and higher mark thresholds. If marks require $k_0 \geq 2$ clusters, let W be the sum of the largest $k_0 - 1$ cluster masses. The same budgets require total mass $q_0 > 4W$, hence $c_0 > 4(k_0 - 1)$. In transverse dimension at most three, any four hulls formed by deleting $k_0 - 1$ representatives share a surviving representative. Finite Helly gives a point common to all such hulls. Each charge has at most $k_0 - 1$ nonpositive representatives; deleting those makes its value positive at the common point, again contradicting balance. This remains an obstruction to this architecture and its certificate inequalities only.

Property at $0 < \varepsilon \leq \varepsilon_*$	R does not tile	R tiles
G_R nilpotent at zero noise	yes	no
Noisy G_R ergodic	yes	no
Exponential all-input coalescence	universal slope	no
Normalized space–time Gibbs measures	exactly one	at least two TI
Arbitrary-volume exponential weak mixing	yes	no
Invariant measures of the stack	exactly one	an arc of TI extremes

Table 3: The raw Wang equivalence and its consequences. The coalescence delay is input dependent; the continuum for the unstacked rule requires uncountably many tilings. TI means fully translation invariant in the relevant lattice.

4 The model and the token game

The next three sections retain the inspector theory of version 1.4.0. Their rule $F_R^{[k]}$ uses thin provider planes and a block alphabet; it is distinct from the raw clustered rule G_R of Section 2. All definitions and proved statements of this theory are retained.

This section fixes the cellular automaton, its noise, and the notions of ergodicity, forgetting and memory used in the paper. It then defines token tables, a finite procedure that may refute a Wang set. Theorem 4.25 shows two things. The procedure never refutes a Wang set that tiles. When it fails to refute a set, it leaves behind an existential pebble family, and the proof of Theorem 9.8 works with that family.

For $r \geq 0$ write $V_r = \{-r, \dots, r\}^2 \subseteq \mathbb{Z}^2$, and put $\Lambda = V_1 = \{-1, 0, 1\}^2$. For $u \in \mathbb{Z}^2$, $|u|_\infty$ is the largest absolute value of a coordinate of u. A *window* is a finite set of sites.

4.1 Wang sets, blocks and the checker

Definition 4.1 (Wang sets)**.** A *Wang set* is a finite set R of tiles. Each tile has a north, east, south and west edge colour, written with the subscripts $\mathrm{N}, \mathrm{E}, \mathrm{S}, \mathrm{W}$. A *pattern* on a set $U \subseteq \mathbb{Z}^2$ is a map $p\colon U \to R$. It is *legal* if $p(u)_\mathrm{E} = p(u + (1,0))_\mathrm{W}$ whenever $u, u + (1,0) \in U$, and $p(u)_\mathrm{N} = p(u + (0,1))_\mathrm{S}$ whenever $u, u + (0,1) \in U$. A *tiling* is a legal pattern on $\mathbb{Z}^2$, and R *tiles* if it has a tiling.

Restrictions and translates of legal patterns are legal.

Lemma 4.2 (Compactness)**.** *If R does not tile, then there is $n \geq 1$ such that no legal pattern on V_n exists. For such n, no square of side at least $2n + 1$ carries a legal pattern.*

Proof. Take the legal patterns on the boxes $V_0, V_1, \dots$ as the nodes of a tree. The parent of a pattern on V_{n+1} is its restriction to V_n, and a root is joined to every pattern on V_0. Each box carries finitely many patterns, so the tree is finitely branching. If every box carried a legal pattern, the tree would be infinite, and König's lemma would give a branch $p_0, p_1, \dots$ in which each p_n is the restriction of p_{n+1}. The union of the branch would be a tiling. So some box V_{n_0} carries no legal pattern, and neither does any larger box; take $n = \max(n_0, 1)$. A legal pattern on a square of side at least $2n + 1$ would restrict and translate to a legal pattern on V_n. □

Definition 4.3 (Blocks, live symbols and compatibility)**.** Let R be a Wang set. The set T_R of *legal* 3×3 *blocks* is the set of legal patterns on Λ. A *live symbol* is a pair $s = (\lambda, b)$ of a block $\lambda \in T_R$ and a *checker bit* $b \in \{0, 1\}$, and $L_R = T_R \times \{0, 1\}$ is the set of live symbols. The alphabet is $\mathcal{A}_R = L_R \cup \{\perp\}$, where $\perp$ (bottom) is one further symbol. The *Moore directions* are $D = \Lambda \setminus \{(0,0)\}$. For live symbols $s = (\lambda, b)$ and $s' = (\lambda', b')$ and for $d \in D$, the symbol s' is *compatible with s at displacement d* if $b = b'$ and $\lambda(w) = \lambda'(w - d)$ for every $w \in \Lambda \cap (d + \Lambda)$. The *centre* of $s = (\lambda, b)$ is the tile $\mathrm{ctr}(s) = \lambda(0,0)$.

Place s at a site u and s' at $u + d$. Their blocks cover $u + \Lambda$ and $u + d + \Lambda$. Compatibility says that the two bits are equal and that the two blocks agree on every common cell. The relation is symmetric: s' is compatible with s at displacement d exactly when s is compatible with s' at displacement $-d$, as the substitution $w \mapsto w - d$ shows. The alphabet has $2|T_R| + 1$ symbols; if T_R is empty, $\perp$ is the only symbol.

Lemma 4.4 (Centres match)**.** *Let s, s' be live. If s' is compatible with s at displacement $(1,0)$, then $\mathrm{ctr}(s)_\mathrm{E} = \mathrm{ctr}(s')_\mathrm{W}$. If s' is compatible with s at displacement $(0,1)$, then $\mathrm{ctr}(s)_\mathrm{N} = \mathrm{ctr}(s')_\mathrm{S}$.*

Proof. Write $s = (\lambda, b)$ and $s' = (\lambda', b)$; the bits agree by compatibility. Let $d = (1, 0)$. The cell $w = (1, 0)$ lies in $\Lambda \cap (d + \Lambda)$, so compatibility gives $\lambda(1, 0) = \lambda'(0, 0) = \mathrm{ctr}(s')$. The block λ is legal, so $\lambda(0, 0)_{\mathrm{E}} = \lambda(1, 0)_{\mathrm{W}}$. Hence $\mathrm{ctr}(s)_{\mathrm{E}} = \mathrm{ctr}(s')_{\mathrm{W}}$. The vertical case is the same, with $d = (0, 1)$ and the colours N and S. □

For a tiling z and a site u, the *block of z at u* is $z^{[u]} \in T_R$, given by $z^{[u]}(w) = z(u + w)$ for $w \in \Lambda$. The *block configuration* of z is $\hat{z} \in L_R^{\mathbb{Z}^2}$, $\hat{z}(u) = (z^{[u]}, 0)$. For every u and every $d \in D$, the symbol $\hat{z}(u + d)$ is compatible with $\hat{z}(u)$ at displacement d: both bits are 0, and $z^{[u]}(w) = z(u + w) = z^{[u+d]}(w - d)$ for $w \in \Lambda \cap (d + \Lambda)$.

Definition 4.5 (Planar checker)**.** The *checker* $C \colon \mathcal{A}_R^{\mathbb{Z}^2} \to \mathcal{A}_R^{\mathbb{Z}^2}$ is given by $C(x)(u) = x(u)$ if $x(u)$ is live and, for every $d \in D$, $x(u + d)$ is live and compatible with $x(u)$ at displacement d. Otherwise $C(x)(u) = \perp$.

Thus $C(x)(u) \in \{x(u), \perp\}$ for every x and u, and $C(\hat{z}) = \hat{z}$ for every tiling z.

Remark 4.6 (The checker reduction)*.* The checker is the planar form of the classical reduction from tilings to the nilpotency of cellular automata [7, 26]. If R tiles, C has the live fixed point $\hat{z}$. If R does not tile, take $n \geq 1$ from Lemma 4.2; then $C^n(x) = \perp$ for every $x \in \mathcal{A}_R^{\mathbb{Z}^2}$.

To see this, let $m \geq 1$, $x \in \mathcal{A}_R^{\mathbb{Z}^2}$ and $u \in \mathbb{Z}^2$ with $C^m(x)(u)$ live. We show that R has a legal pattern on $u + V_m$. A live output of C at a site needs live inputs at the nine sites of its window, so induction on m shows that $C(x)(v)$ is live for every $v \in u + V_{m-1}$. Hence, for $v \in u + V_{m-1}$, the symbol $x(v)$ is live, say with block λ_v, and $x(v + d)$ is compatible with $x(v)$ at displacement d for every $d \in D$. For $w \in u + V_m$ choose $v \in u + V_{m-1}$ with $|w - v|_\infty \leq 1$ and put $p(w) = \lambda_v(w - v)$. The value does not depend on the choice. For two choices v, v' with $|v - v'|_\infty \leq 1$ this is compatibility. Otherwise define v'' coordinatewise by $v''_c = v_c$ if $|v_c - v'_c| \leq 1$ and $v''_c = w_c$ if $|v_c - v'_c| = 2$. Then $v'' \in u + V_{m-1}$, and v'' lies within distance 1 of v, of v' and of w, so two applications of compatibility link the values at v and v'. Finally, let $\{w, w + \Delta\}$ with $\Delta \in \{(1, 0), (0, 1)\}$ be an edge inside $u + V_m$, and let v be a point of $u + V_{m-1}$ nearest to $w + \Delta/2$ in each coordinate. Both w and $w + \Delta$ lie in $v + \Lambda$, and λ_v is legal, so the edge matches. So p is legal. With $m = n$ this contradicts the choice of n.

4.2 The quorum rule and replacement noise

Definition 4.7 (Quorum rule)**.** Let $k \geq 2$. A *site* is a pair $(u, P) \in \mathbb{Z}^2 \times \mathbb{Z}^k$, and the sites with a fixed transverse coordinate P form the *plane* P. The configuration space is $\Omega = \mathcal{A}_R^{\mathbb{Z}^2 \times \mathbb{Z}^k}$. For $x \in \Omega$, $x(\cdot, P) \in \mathcal{A}_R^{\mathbb{Z}^2}$ is its restriction to the plane P. The *providers* are $c_0 = 0$ and $c_i = e_i$ for $i = 1, \ldots, k$, where e_i is the i-th unit vector of $\mathbb{Z}^k$. Put $y(u, P) = C(x(\cdot, P))(u)$. The *quorum rule* is

$$F_R^{[k]}(x)(u, P) = \begin{cases} s & \text{if } s \text{ is live and } \#\{i \in \{0, \ldots, k\} : y(u, P + c_i) = s\} \geq k, \\ \perp & \text{otherwise.} \end{cases}$$

Each plane runs the checker. The rule then reads the checked symbol at the site itself and at its k transverse neighbours $(u, P + e_i)$, and keeps a live symbol only when at least k of these $k + 1$ readings agree on it. Two different live symbols cannot both be read k times, since $2k > k + 1$; so $F_R^{[k]}$ is well defined. The value $F_R^{[k]}(x)(v)$ depends only on x restricted to $v + \mathcal{N}$, where $\mathcal{N} = \Lambda \times \{c_0, \ldots, c_k\}$. So $F_R^{[k]}$ is a cellular automaton on $\mathbb{Z}^2 \times \mathbb{Z}^k = \mathbb{Z}^{2+k}$, and its local rule is a finite table computable from R and k. If z is a tiling, the *cloned tiling configuration* $\xi_z \in \Omega$, $\xi_z(u, P) = \hat{z}(u)$, satisfies $F_R^{[k]}(\xi_z) = \xi_z$, because all $k + 1$ providers read $\hat{z}(u)$.

Definition 4.8 (Replacement noise)**.** Fix $\varepsilon \in (0,1)$. Let $\eta_t(v) \in \{0,1\}$ and $\zeta_t(v) \in \mathcal{A}_R$, for $t \geq 1$ and $v \in \mathbb{Z}^2 \times \mathbb{Z}^k$, be independent random variables with $\mathbb{P}(\eta_t(v) = 1) = \varepsilon$ and $\zeta_t(v)$ uniform on $\mathcal{A}_R$. The *process from* $x \in \Omega$ is given by $X_0 = x$ and, for $t \geq 0$,

$$X_{t+1}(v) = \begin{cases} \zeta_{t+1}(v) & \text{if } \eta_{t+1}(v) = 1, \\ F_R^{[k]}(X_t)(v) & \text{if } \eta_{t+1}(v) = 0. \end{cases}$$

A site v at a time $t \geq 1$ with $\eta_t(v) = 1$ is a *replacement event*, whatever value $\zeta_t(v)$ takes.

So after each application of $F_R^{[k]}$, each site independently is replaced, with probability ε, by a uniformly random symbol of $\mathcal{A}_R$. Given X_t, the values $X_{t+1}(v)$ are independent, and for every symbol s

$$\mathbb{P}\big(X_{t+1}(v) = s \mid X_t\big) = (1-\varepsilon)\,\mathbf{1}\big[F_R^{[k]}(X_t)(v) = s\big] + \frac{\varepsilon}{|\mathcal{A}_R|} \;\geq\; \frac{\varepsilon}{|\mathcal{A}_R|} > 0.$$

The process is therefore a probabilistic cellular automaton with positive rates. We write Law_x for the law of $(X_t)_{t\geq 0}$ from x, and $\mathbb{P}_x$, $\mathbb{E}_x$ for probabilities and expectations under it. The symbol $\perp$ also denotes the all-bottom configuration. For a probability measure μ on Ω, Law_μ is the law of the process when X_0 has law μ and is independent of the variables $\eta_t(v)$, $\zeta_t(v)$.

By induction on t, $X_t(v)$ is a function of X_0 restricted to the finite set $v + \mathcal{N}_t$, where $\mathcal{N}_0 = \{0\}$ and $\mathcal{N}_{t+1} = \mathcal{N}_t + \mathcal{N}$, and of finitely many noise variables. Hence, for a window O, the law of $X_t|_O$ under Law_x depends on only finitely many coordinates of x, and

$$\mathbb{P}_\mu(X_t|_O \in \mathcal{E}) = \int \mathbb{P}_x(X_t|_O \in \mathcal{E})\,\mu(dx) \qquad \text{for every } \mathcal{E} \subseteq \mathcal{A}_R^O.$$

Running all starts with the same variables $\eta_t(v), \zeta_t(v)$ couples them. This is the common-noise coupling used in Theorem 5.1.

Write $\mu\Phi^t$ for the law of X_t under Law_μ, and $\Phi f(x) = \mathbb{E}_x f(X_1)$ for bounded measurable $f\colon \Omega \to \mathbb{R}$, so that $(\mu\Phi)(f) = \mu(\Phi f)$. The noise after time t is independent of X_t, so given X_t the process $(X_{t+t'})_{t'\geq 0}$ is the process from X_t. In particular $\mu\Phi^{t+1} = (\mu\Phi^t)\Phi$. A probability measure μ on Ω is *invariant* if $\mu\Phi = \mu$.

4.3 Ergodicity, forgetting and memory

The space Ω carries the product topology; it is compact and metrizable.

Definition 4.9 (Ergodicity)**.** The process is *ergodic* if it has exactly one invariant probability measure π and $\mu\Phi^t \to \pi$ weakly as $t \to \infty$ for every probability measure μ on Ω.

For a window O and $\mathcal{E} \subseteq \mathcal{A}_R^O$, the cylinder $\{\omega \in \Omega : \omega|_O \in \mathcal{E}\}$ is open and closed. Functions that depend on finitely many coordinates are finite combinations of cylinder indicators, and they are dense in the continuous functions on Ω. Hence $\mu_t \to \mu$ weakly if and only if $d_{\mathrm{TV}}(\mu_t|_O, \mu|_O) \to 0$ for every window O. Here $\mu|_O$ is the marginal law on $\mathcal{A}_R^O$, and $d_{\mathrm{TV}}(\nu,\nu') = \max_{\mathcal{E}\subseteq\mathcal{A}_R^O} |\nu(\mathcal{E}) - \nu'(\mathcal{E})|$.

Lemma 4.10 (Invariant measures exist)**.** *For every continuous $f\colon \Omega \to \mathbb{R}$, the function Φf is continuous. For every probability measure μ on Ω and every sequence $n_m \to \infty$, the averages $\bar\mu_n = \frac{1}{n}\sum_{t=0}^{n-1} \mu\Phi^t$ have a weak limit along a subsequence of (n_m), and every such limit is invariant.*

Proof. If f depends only on the coordinates in a window O, then $\Phi f(x)$ depends only on x restricted to the window $O + \mathcal{N}$, so Φf is continuous. A general continuous f is uniformly continuous on the compact space Ω, so it is a uniform limit of functions f_ℓ that depend on finitely many coordinates. Since $\|\Phi g\|_\infty \le \|g\|_\infty$, the function Φf is the uniform limit of the continuous functions Φf_ℓ, hence continuous. The probability measures on the compact metrizable space Ω form a weakly sequentially compact set, which gives the convergent subsequence. Let $\bar{\mu}_{n_m} \to \pi$ weakly. For continuous f, using that Φf is continuous,

$$|\pi(\Phi f) - \pi(f)| = \lim_{m\to\infty} \left|\bar{\mu}_{n_m}(\Phi f) - \bar{\mu}_{n_m}(f)\right| = \lim_{m\to\infty} \frac{|\mu\Phi^{n_m}(f) - \mu(f)|}{n_m} \le \lim_{m\to\infty} \frac{2\|f\|_\infty}{n_m} = 0.$$

So $\pi\Phi = \pi$. □

Definition 4.11 (Forgetting and memory)**.** For a window O and $t \ge 0$ put

$$\delta_O(t) = \sup_{x\in\Omega} d_{\mathrm{TV}}\big(\mathrm{Law}_x(X_t|_O),\, \mathrm{Law}_\perp(X_t|_O)\big).$$

The process *forgets* if $\delta_O(t) \to 0$ as $t \to \infty$ for every window O. It has *memory* if there are configurations $x, x' \in \Omega$, a window O, a set $\mathcal{E} \subseteq \mathcal{A}_R^O$ and $\alpha > 0$ such that

$$\mathbb{P}_x(X_t|_O \in \mathcal{E}) \;\ge\; \mathbb{P}_{x'}(X_t|_O \in \mathcal{E}) + \alpha \qquad \text{for every } t \ge 0.$$

Under memory the laws of $X_t|_O$ from x and from x' stay at total variation distance at least α for all t, and one fixed event sees the distance. In Theorem 6.2 the two starts are a cloned tiling configuration and the all-bottom configuration, and the event is that one site carries its tiling symbol. For forgetting, any coupling of Law_x and $\mathrm{Law}_\perp$ bounds the distance on O by the probability that the two processes differ somewhere on O; Theorem 5.1 bounds $\delta_O(t)$ in this way.

Lemma 4.12 (Forgetting implies ergodicity)**.** *If the process forgets, it is ergodic. Its invariant measure π satisfies $d_{\mathrm{TV}}(\mu\Phi^t|_O, \pi|_O) \le 2\delta_O(t)$ for every probability measure μ, every window O and every $t \ge 0$.*

Proof. Let μ, ν be probability measures on Ω, let O be a window, $t \ge 0$ and $\mathcal{E} \subseteq \mathcal{A}_R^O$. By the mixture formula,

$$\mathbb{P}_\mu(X_t|_O \in \mathcal{E}) - \mathbb{P}_\nu(X_t|_O \in \mathcal{E}) = \iint \big(\mathbb{P}_x(X_t|_O \in \mathcal{E}) - \mathbb{P}_{x'}(X_t|_O \in \mathcal{E})\big)\, \mu(dx)\, \nu(dx').$$

By the triangle inequality through $\mathrm{Law}_\perp$, the integrand has absolute value at most $2\delta_O(t)$. Hence $d_{\mathrm{TV}}(\mu\Phi^t|_O, \nu\Phi^t|_O) \le 2\delta_O(t)$. By Lemma 4.10 there is an invariant measure π. Taking $\nu = \pi$, for which $\pi\Phi^t = \pi$, gives the stated bound, and so $\mu\Phi^t \to \pi$ weakly for every μ. If π' is invariant, the bound with $\mu = \pi'$ gives $d_{\mathrm{TV}}(\pi'|_O, \pi|_O) \le 2\delta_O(t)$ for every t. So π' and π agree on every cylinder, and $\pi' = \pi$. □

Lemma 4.13 (Memory excludes ergodicity)**.** *If the process has memory, it has two different invariant measures. In particular it is not ergodic.*

Proof. Let $\bar{\mu}_n$ and $\bar{\mu}'_n$ be the averages of Lemma 4.10 for the point masses at x and at x'. Choose $n_m \to \infty$ with $\bar{\mu}_{n_m} \to \pi$ weakly, and then pass to a subsequence along which also $\bar{\mu}'_{n_m} \to \pi'$. Both limits are invariant by Lemma 4.10. The indicator of the cylinder $\Gamma = \{\omega : \omega|_O \in \mathcal{E}\}$ is continuous, so

$$\pi(\Gamma) - \pi'(\Gamma) = \lim_{m\to\infty} \frac{1}{n_m} \sum_{t=0}^{n_m-1} \big(\mathbb{P}_x(X_t|_O \in \mathcal{E}) - \mathbb{P}_{x'}(X_t|_O \in \mathcal{E})\big) \;\ge\; \alpha.$$

Hence $\pi \ne \pi'$. □

A weaker hypothesis already excludes ergodicity: if $d_{\mathrm{TV}}(\mathrm{Law}_x(X_t|_O), \mathrm{Law}_{x'}(X_t|_O)) \geq \alpha$ for every t, then the process is not ergodic, since under ergodicity both marginals converge to $\pi|_O$. That hypothesis alone does not produce two invariant measures through averages, because the separating event may change with t; memory fixes the event for this reason.

4.4 Token tables

Fix a Wang set R, a number $j \geq 1$ of tokens and a radius $r \geq 0$, and put $V = V_r$. A position tuple $\mathbf{u} = (u_1, \ldots, u_j) \in V^j$ and a symbol tuple $\mathbf{s} = (s_1, \ldots, s_j) \in L_R^j$ describe j tokens: token i sits at u_i and carries s_i. Write $\mathbf{s}[i \leftarrow s']$ for $\mathbf{s}$ with entry i replaced by s', and $\mathbf{u}[i \leftarrow w]$ likewise. A *table* S assigns to each $\mathbf{u} \in V^j$ a set $S(\mathbf{u}) \subseteq L_R^j$. Tables are compared by inclusion at every $\mathbf{u}$.

Definition 4.14 (Token tables)**.** The deletion rules act on a table S as follows.

(D1) Coincidence rule. Delete $\mathbf{s}$ from $S(\mathbf{u})$ if $u_i = u_{i'}$ and $s_i \neq s_{i'}$ for some i, i'.

(D2) Move rule. Delete $\mathbf{s}$ from $S(\mathbf{u})$ if there are an index i and a direction $d \in D$ with $u_i + d \in V$ such that no live symbol s' compatible with s_i at displacement d has $\mathbf{s}[i \leftarrow s'] \in S(\mathbf{u}[i \leftarrow u_i + d])$.

(D3) Clone rule. Delete $\mathbf{s}$ from $S(\mathbf{u})$ if $\mathbf{s}[i' \leftarrow s_i] \notin S(\mathbf{u}[i' \leftarrow u_i])$ for some $i \neq i'$.

$\mathrm{Inst}(S)$ is obtained from S by applying (D1) and (D3) until neither applies. $\mathrm{Timed}(S)$ is obtained from S by applying (D2) once, simultaneously, to every tuple of S, with the condition evaluated in S. The *token tables* are

$$S_0 = \mathrm{Inst}(\text{all tuples}), \qquad S_{n+1} = \mathrm{Inst}\big(\mathrm{Timed}(S_n)\big) \quad (n \geq 0),$$

where the table of all tuples has $S(\mathbf{u}) = L_R^j$ for every $\mathbf{u} \in V^j$. We write $S_n^{(r)}$ when the radius matters.

The clone rule makes token i' jump onto token i and take its symbol. The move rule moves token i by one Moore step inside V and asks for a compatible new symbol at the new site. The rules commute with translations of $\mathbb{Z}^2$, so tables on a translate $u + V_r$ are defined by translating positions.

Lemma 4.15 (Tables are well defined)**.** *Fix j and r. Call a table* closed *if neither (D1) nor (D3) deletes anything from it.*

(i) $\mathrm{Inst}(S)$ *is the largest closed table contained in S. In particular it does not depend on the order of the deletions.*

(ii) If $S \subseteq S'$, then $\mathrm{Inst}(S) \subseteq \mathrm{Inst}(S')$ *and* $\mathrm{Timed}(S) \subseteq \mathrm{Timed}(S')$*. Moreover $S_{n+1} \subseteq S_n$ for every $n \geq 0$.*

(iii) There is n_0 with $S_n = S_{n_0}$ for every $n \geq n_0$. We write $S_\infty = S_\infty^{(r)}$ for this final table.

Proof. (i) Closedness asks, of each tuple separately, that coincident tokens carry equal symbols and that its clones lie in the table. So a union of closed tables is closed, and S contains a largest closed table S^*. Run the deletions in any order. By induction on the number of deletions, the current table contains S^*. A tuple deleted by (D1) violates the coincidence condition, so it is not in S^*. A tuple deleted by (D3) has a clone missing from the current table, hence missing from S^*; since S^* is closed, the tuple is not in S^*. The run stops, because the table is finite, and it stops at a closed table contained in S, hence contained in S^*. So it stops at S^*.

(ii) If $S \subseteq S'$, then $\mathrm{Inst}(S)$ is a closed table contained in S', so it is contained in $\mathrm{Inst}(S')$ by (i). A tuple of S that survives (D2) evaluated in S has, for each allowed move, a witness in $S \subseteq S'$, so it survives (D2) evaluated in S'. Hence $\mathrm{Timed}(S) \subseteq \mathrm{Timed}(S')$. Now $S_1 \subseteq \mathrm{Timed}(S_0) \subseteq S_0$, and if $S_n \subseteq S_{n-1}$, then $S_{n+1} = \mathrm{Inst}(\mathrm{Timed}(S_n)) \subseteq \mathrm{Inst}(\mathrm{Timed}(S_{n-1})) = S_n$.

(iii) The tables are subsets of the finite set $V^j \times L_R^j$, and they decrease by (ii). □

Let $\mathbf{0} = (0, \ldots, 0) \in V^j$. By (D1), every tuple in $S_n(\mathbf{0})$ is constant, $(s, \ldots, s)$ with s live.

Definition 4.16 (Refutability)**.** The Wang set R is *j-token refutable* if $S_\tau^{(r)}(\mathbf{0}) = \emptyset$ for some radius $r \geq 0$ and some time $\tau \geq 0$. For a j-token refutable R, the *refutation time* $\tau(R)$ is the least τ for which some r works. The *token width* of R is the least j for which R is j-token refutable, and ∞ if there is none.

So a refutation deletes every live symbol at the origin. Each table $S_n^{(r)}$ is a finite computation from R, j, r and n. Hence, for each j, the j-token refutable Wang sets form a computably enumerable set.

4.5 The token game

The tables have a game reading. Two players act on j tokens. Spoiler either makes one token jump onto another, so that it takes over that token's position and symbol, or moves one token by a Moore step. After a move, Duplicator chooses the new symbol of the moved token, which must be compatible with its old symbol at that displacement. Duplicator loses when two tokens on one site carry different symbols, or when she has no reply. In a box, Spoiler may move only inside the box. We work with the following sets of game states, from which Duplicator can play forever by keeping the play inside the set.

Definition 4.17 (States and safe sets)**.** A *j-token state* is a pair $(\mathbf{u}, \mathbf{s}) \in (\mathbb{Z}^2)^j \times L_R^j$. For $i \neq i'$, the *clone* of token i onto token i' is the state $(\mathbf{u}[i' \leftarrow u_i], \mathbf{s}[i' \leftarrow s_i])$. For an index i and $d \in D$, a *move* of token i by d is a state $(\mathbf{u}[i \leftarrow u_i + d], \mathbf{s}[i \leftarrow s'])$ with s' live and compatible with s_i at displacement d. Let V be a box V_r or the plane $\mathbb{Z}^2$. A set $\mathcal{W}$ of states with positions in V^j is *V-safe* if every $(\mathbf{u}, \mathbf{s}) \in \mathcal{W}$ satisfies:

(W1) if $u_i = u_{i'}$, then $s_i = s_{i'}$;

(W2) for all $i \neq i'$, the clone of token i onto token i' lies in $\mathcal{W}$;

(W3) for every i and every $d \in D$ with $u_i + d \in V$, some move of token i by d lies in $\mathcal{W}$.

A *plane-safe* set is a $\mathbb{Z}^2$-safe set. For a state satisfying (W1), $h_{\mathbf{u},\mathbf{s}}$ denotes the map $\{u_1, \ldots, u_j\} \to L_R$ with $h_{\mathbf{u},\mathbf{s}}(u_i) = s_i$.

We identify a table S on V_r with the set of states $(\mathbf{u}, \mathbf{s})$ with $\mathbf{s} \in S(\mathbf{u})$. Then (W1) says that (D1) deletes nothing, (W2) says that (D3) deletes nothing, and (W3) with $V = V_r$ says that (D2) deletes nothing.

Lemma 4.18 (Walking to the origin)**.** *Let V be a box V_r or the plane $\mathbb{Z}^2$. Every nonempty V-safe set contains a state $(\mathbf{0}, \mathbf{s})$.*

Proof. Take a state $(\mathbf{u}, \mathbf{s})$ in the set. Cloning token 1 onto each other token gives, by (W2), a state of the set with all tokens at u_1. While token 1 is not at the origin, move it by the step $d \in D$ whose coordinates are the signs of the coordinates of its negated position, with the sign of 0 equal to 0. The new position lies in V, since no coordinate grows in absolute value, so (W3) gives a move in the set. After finitely many steps token 1 is at the origin. Cloning it onto the other tokens gives a state $(\mathbf{0}, \mathbf{s}')$ of the set. □

Lemma 4.19 (Finite time and finite boxes)**.** *Fix $j \geq 1$.*

(i) For every r, $S_\infty^{(r)}$ is the largest V_r-safe set.

(ii) If $r \leq r'$, the states of $S_\infty^{(r')}$ with positions in V_r^j lie in $S_\infty^{(r)}$.

(iii) R is j-token refutable if and only if $S_\infty^{(r)} = \emptyset$ for some r. In that case $S_\infty^{(r')} = \emptyset$ for every $r' \geq r$.

Proof. (i) By Lemma 4.15(iii), $S_\infty = S_{n_0} = S_{n_0+1} = \mathrm{Inst}(\mathrm{Timed}(S_\infty)) \subseteq \mathrm{Timed}(S_\infty) \subseteq S_\infty$. So $\mathrm{Timed}(S_\infty) = S_\infty$, which is (W3), and $\mathrm{Inst}(S_\infty) = S_\infty$, which is (W1) and (W2). Conversely, let $\mathcal{W}$ be V_r-safe. It is a closed table, so $\mathcal{W} \subseteq S_0$ by Lemma 4.15(i). If $\mathcal{W} \subseteq S_n$, then every state of $\mathcal{W}$ survives (D2) evaluated in S_n, because (W3) supplies its witnesses in $\mathcal{W} \subseteq S_n$. So $\mathcal{W} \subseteq \mathrm{Timed}(S_n)$, and since $\mathcal{W}$ is closed, $\mathcal{W} \subseteq \mathrm{Inst}(\mathrm{Timed}(S_n)) = S_{n+1}$. Hence $\mathcal{W} \subseteq S_\infty$.

(ii) The states of $S_\infty^{(r')}$ with positions in V_r^j form a V_r-safe set. Indeed (W1) is inherited, clones do not create new positions, and a move with $u_i + d \in V_r$ is allowed in $V_{r'}$ and its witness has positions in V_r^j. Apply (i).

(iii) If $S_\tau^{(r)}(\mathbf{0}) = \emptyset$, then $S_\infty^{(r)}(\mathbf{0}) = \emptyset$ by Lemma 4.15(ii). By (i) and Lemma 4.18, a nonempty $S_\infty^{(r)}$ would contain a state at $\mathbf{0}$, so $S_\infty^{(r)} = \emptyset$. Conversely, if $S_\infty^{(r)} = \emptyset$, then $S_{n_0}^{(r)}(\mathbf{0}) = \emptyset$. If $r' \geq r$, then (ii) gives $S_\infty^{(r')}(\mathbf{0}) \subseteq S_\infty^{(r)}(\mathbf{0}) = \emptyset$, and the same argument gives $S_\infty^{(r')} = \emptyset$. □

Lemma 4.20 (Boxes and the plane)**.** *R is j-token refutable if and only if there is no nonempty plane-safe set of j-token states.*

Proof. Let $\mathcal{W}$ be a nonempty plane-safe set. By Lemma 4.18 it contains a state $(\mathbf{0}, \mathbf{s})$. For every r, the states of $\mathcal{W}$ with positions in V_r^j form a V_r-safe set: (W1) is inherited, clones keep positions in V_r^j, and for a move with $u_i + d \in V_r$ the move supplied by (W3) in the plane has positions in V_r^j. By Lemma 4.19(i) this set lies in $S_\infty^{(r)}$, so $(\mathbf{0}, \mathbf{s}) \in S_\infty^{(r)} \subseteq S_n^{(r)}$ for every n. Hence $S_n^{(r)}(\mathbf{0}) \neq \emptyset$ for all r and n, and R is not j-token refutable.

Conversely, suppose that R is not j-token refutable. Then $S_\infty^{(r)}(\mathbf{0}) \neq \emptyset$ for every r. Let $\mathcal{W}$ be the set of states that lie in $S_\infty^{(r)}$ for infinitely many r. Since L_R^j is finite, some $\mathbf{s}$ lies in $S_\infty^{(r)}(\mathbf{0})$ for infinitely many r, so $\mathcal{W}$ is nonempty. Each $S_\infty^{(r)}$ is V_r-safe, and we check (W1)–(W3) for $\mathcal{W}$. Property (W1) holds because it holds in each $S_\infty^{(r)}$. If $(\mathbf{u}, \mathbf{s}) \in \mathcal{W}$, each of its clones lies in $S_\infty^{(r)}$ whenever $(\mathbf{u}, \mathbf{s})$ does, by (W2) in $S_\infty^{(r)}$, so the clone lies in $\mathcal{W}$. For (W3), fix i and $d \in D$. For each of the infinitely many $r \geq |u_i|_\infty + 1$ with $(\mathbf{u}, \mathbf{s}) \in S_\infty^{(r)}$, the site $u_i + d$ lies in V_r, so (W3) in $S_\infty^{(r)}$ gives a live s'_r with $(\mathbf{u}[i \leftarrow u_i + d], \mathbf{s}[i \leftarrow s'_r]) \in S_\infty^{(r)}$ and s'_r compatible with s_i at displacement d. There are finitely many live symbols, so one value s' occurs for infinitely many r, and the corresponding move lies in $\mathcal{W}$. So $\mathcal{W}$ is a nonempty plane-safe set. □

The next lemma is where the move rule and the coincidence rule combine.

Lemma 4.21 (Neighbouring tokens are compatible)**.** *Let $\mathcal{W}$ be plane-safe, $(\mathbf{u}, \mathbf{s}) \in \mathcal{W}$, and $u_{i'} = u_i + d$ with $d \in D$. Then $s_{i'}$ is compatible with s_i at displacement d.*

Proof. By (W3) some move $(\mathbf{u}[i \leftarrow u_i + d], \mathbf{s}[i \leftarrow s'])$ lies in $\mathcal{W}$, with s' compatible with s_i at displacement d. In this state tokens i and i' both sit at $u_{i'}$; note $i \neq i'$ because $d \neq 0$. By (W1), $s' = s_{i'}$. □

Definition 4.22 (Existential pebble families)**.** A *partial homomorphism into the live target* is a map $h\colon U \to L_R$ on a finite set $U \subseteq \mathbb{Z}^2$ such that $h(u + d)$ is compatible with $h(u)$ at displacement d whenever $d \in D$ and $u, u + d \in U$. A *partial homomorphism into R* is a legal pattern on a finite

set $U \subseteq \mathbb{Z}^2$. Fix one of the two targets. An *existential j-pebble family*, or *E_j family*, is a nonempty set H of partial homomorphisms into the target such that:

(Size) every $h \in H$ has at most j sites in its domain;

(Restriction) if $h \in H$ and $U \subseteq \operatorname{dom} h$, then $h|_U \in H$;

(Extension) if $h \in H$, $|\operatorname{dom} h| < j$ and $w \in \mathbb{Z}^2$, then some $h' \in H$ extends h and has $w \in \operatorname{dom} h'$.

These families are the winning strategies of Duplicator in the existential j-pebble game of Kolaitis and Vardi [28], played from the grid $\mathbb{Z}^2$ to the target; see also [1]. In that game Spoiler places and removes up to j pebbles on sites, Duplicator answers each placement with a target symbol, and Duplicator must keep the pebbled correspondence a partial homomorphism. For the live target the grid carries the eight Moore relations; for R it carries the horizontal and vertical edge relations.

Remark 4.23. Let H be an E_j family. By (Restriction) it contains the empty map. For every set $U \subseteq \mathbb{Z}^2$ with $|U| \leq j$, some $h \in H$ has domain exactly U: add the sites of U one at a time with (Extension), and after each step restrict with (Restriction) to the sites added so far. For $1 \leq K \leq j$, the members of H with at most K sites form an E_K family, since an extension supplied by (Extension) can be restricted to the old domain plus the new site.

Lemma 4.24 (Pebble families from safe sets)**.** *Let $\mathcal{W}$ be a nonempty plane-safe set of j-token states. Let $H(\mathcal{W})$ be the set of restrictions $h_{\mathbf{u},\mathbf{s}}|_U$ with $(\mathbf{u},\mathbf{s}) \in \mathcal{W}$ and $U \subseteq \{u_1, \dots, u_j\}$. Then $H(\mathcal{W})$ is an E_j family for the live target, and $\operatorname{ctr} \circ H(\mathcal{W}) = \{\operatorname{ctr} \circ h : h \in H(\mathcal{W})\}$ is an E_j family for R.*

Proof. Each $h_{\mathbf{u},\mathbf{s}}$ is well defined by (W1). It is a partial homomorphism into the live target: if u and $u + d$ are sites of $\mathbf{u}$ with $d \in D$, pick i, i' with $u_i = u$ and $u_{i'} = u + d$ and apply Lemma 4.21. Restrictions of partial homomorphisms are partial homomorphisms. The family is nonempty, has at most j sites per map, and is closed under restriction by construction.

For (Extension), let $h = h_{\mathbf{u},\mathbf{s}}|_U$ with $(\mathbf{u},\mathbf{s}) \in \mathcal{W}$ and $|U| < j$, and let $w \in \mathbb{Z}^2$. If $w \in U$ there is nothing to do. Otherwise, for each $u \in U$ choose one index $i(u)$ with $u_{i(u)} = u$. These are fewer than j indices, so some index i_0 is not among them. Choose finitely many Moore steps leading from u_{i_0} to w, and apply (W3) to token i_0 once for each step. This gives states of $\mathcal{W}$ in which every token other than i_0 keeps its position and symbol, ending with a state $(\mathbf{u}^*, \mathbf{s}^*) \in \mathcal{W}$ with $u^*_{i_0} = w$. The map $h' = h_{\mathbf{u}^*,\mathbf{s}^*}|_{U \cup \{w\}}$ lies in $H(\mathcal{W})$, contains w in its domain, and extends h, since $h'(u) = s^*_{i(u)} = s_{i(u)} = h(u)$ for $u \in U$.

For the second family, Lemma 4.4 shows that $\operatorname{ctr} \circ h$ is a legal pattern whenever h is a partial homomorphism into the live target, because $(1,0)$ and $(0,1)$ are Moore directions. Restriction commutes with composition by ctr. For (Extension), write a member as $\operatorname{ctr} \circ h$ with $h \in H(\mathcal{W})$, extend h in $H(\mathcal{W})$, and compose the extension with ctr. □

Theorem 4.25 (Tokens and pebble families)**.** *Let R be a finite Wang set and $j \geq 1$.*

(i) R is j-token refutable if and only if there is no nonempty plane-safe set of j-token states.

(ii) For every nonempty plane-safe set $\mathcal{W}$ of j-token states, the restrictions of the maps $h_{\mathbf{u},\mathbf{s}}$, $(\mathbf{u},\mathbf{s}) \in \mathcal{W}$, form an E_j family for the live target, and their compositions with ctr *form an E_j family for R.*

(iii) If R is not j-token refutable, there is an E_j family for R. Equivalently, if there is no E_j family for R, then R is j-token refutable.

(iv) If R tiles the plane, then R is not j-token refutable.

(v) If there is an E_{j+1} family for the live target, then R is not j-token refutable.

Proof. Part (i) is Lemma 4.20, and part (ii) is Lemma 4.24. Part (iii) follows: if R is not j-token refutable, (i) gives a nonempty plane-safe set, and (ii) turns it into an E_j family for R.

For (iv), let z be a tiling and let $\mathcal{W}_z$ be the set of states $(\mathbf{u}, (\hat{z}(u_1), \dots, \hat{z}(u_j)))$ with $\mathbf{u} \in (\mathbb{Z}^2)^j$. It is nonempty. Property (W1) holds because equal positions carry equal symbols. The clone of token i onto token i' is the state of $\mathcal{W}_z$ at the positions $\mathbf{u}[i' \leftarrow u_i]$, so (W2) holds. For (W3), the symbol $\hat{z}(u_i + d)$ is compatible with $\hat{z}(u_i)$ at displacement d, and the resulting move is the state of $\mathcal{W}_z$ at the positions $\mathbf{u}[i \leftarrow u_i + d]$. So $\mathcal{W}_z$ is plane-safe, and (i) shows that R is not j-token refutable. Concretely, for every r and n the tuple $(\hat{z}(0), \dots, \hat{z}(0))$ lies in $S_n^{(r)}(\mathbf{0})$.

For (v), let H be an E_{j+1} family for the live target, and let $\mathcal{W}$ be the set of j-token states $(\mathbf{u}, \mathbf{s})$ that satisfy (W1) and have $h_{\mathbf{u},\mathbf{s}} \in H$. Let $g \in H$, and let $(\mathbf{u}', \mathbf{s}')$ be a j-token state with $u'_i \in \operatorname{dom} g$ and $s'_i = g(u'_i)$ for every i. Then $(\mathbf{u}', \mathbf{s}')$ satisfies (W1), and $h_{\mathbf{u}',\mathbf{s}'}$ is the restriction of g to $\{u'_1, \dots, u'_j\}$, so it lies in H by (Restriction); hence $(\mathbf{u}', \mathbf{s}') \in \mathcal{W}$. By Remark 4.23, H contains the empty map and, for every $w \in \mathbb{Z}^2$, a map with domain $\{w\}$. The state with all j tokens at w, each carrying the value of that map, therefore lies in $\mathcal{W}$, and $\mathcal{W}$ is nonempty. Now let $(\mathbf{u}, \mathbf{s}) \in \mathcal{W}$ and $h = h_{\mathbf{u},\mathbf{s}}$. Every token of a clone of $(\mathbf{u}, \mathbf{s})$ sits at a site of $\operatorname{dom} h$ and carries the value of h there. So the map of the clone is a restriction of h, the clone lies in $\mathcal{W}$, and (W2) holds. For (W3), let i be an index and $d \in D$. The map h has at most j sites, fewer than $j + 1$, so (Extension) gives $h' \in H$ that extends h and has $u_i + d \in \operatorname{dom} h'$. Put $s' = h'(u_i + d)$. Since h' is a partial homomorphism into the live target and $h'(u_i) = s_i$, the live symbol s' is compatible with s_i at displacement d, so $(\mathbf{u}[i \leftarrow u_i + d], \mathbf{s}[i \leftarrow s'])$ is a move of token i by d. Every token of this move sits at a site of $\operatorname{dom} h'$ and carries the value of h' there. So the map of the moved state is a restriction of h', and the move lies in $\mathcal{W}$. Hence $\mathcal{W}$ is a nonempty plane-safe set, and (i) shows that R is not j-token refutable. □

Part (ii) applies to every nonempty plane-safe set. An argument may therefore first pass to a smaller nonempty plane-safe set with extra properties, and then take its family.

4.6 One token, and finite token width

Proposition 4.26 (One token is arc consistency)**.** *Let $L^{(0)} = L_R$, and for $n \geq 0$ let $L^{(n+1)}$ be the set of $s \in L^{(n)}$ such that for every $d \in D$ some $s' \in L^{(n)}$ is compatible with s at displacement d. The sets $L^{(n)}$ decrease, and $L^{(n)} = L^{(|L_R|)}$ for every $n \geq |L_R|$; call this set $L^{(\infty)}$. Then R is 1-token refutable if and only if $L^{(\infty)} = \emptyset$. In particular, 1-token refutability is decidable.*

Proof. By definition $L^{(n+1)} \subseteq L^{(n)}$. Once $L^{(n+1)} = L^{(n)}$, the sequence is constant from n on, and each strict decrease removes at least one of the $|L_R|$ symbols; this gives the second claim. For $j = 1$ there is no pair $i \neq i'$, so (W1) and (W2) hold for every set of 1-token states, and such a set is plane-safe exactly when (W3) holds.

If $L^{(\infty)} \neq \emptyset$, the set $\mathbb{Z}^2 \times L^{(\infty)}$ of 1-token states is plane-safe. Indeed $L^{(\infty)} = L^{(|L_R|+1)}$, so every $s \in L^{(\infty)}$ has, for every $d \in D$, a partner $s' \in L^{(\infty)}$ compatible with s at displacement d. By Theorem 4.25(i), R is not 1-token refutable.

Conversely, let $\mathcal{W}$ be a nonempty plane-safe set of 1-token states, and let Q be the set of symbols carried by its states. Then $Q \neq \emptyset$ and $Q \subseteq L^{(0)}$. If $Q \subseteq L^{(n)}$ and $(u, s) \in \mathcal{W}$, then for each $d \in D$, (W3) gives $(u + d, s') \in \mathcal{W}$ with s' compatible with s at displacement d, and $s' \in Q \subseteq L^{(n)}$; so $s \in L^{(n+1)}$. Hence $\emptyset \neq Q \subseteq L^{(\infty)}$. By Theorem 4.25(i), if R is not 1-token refutable then $L^{(\infty)} \neq \emptyset$.

The set $L^{(\infty)}$ is computed from R by a finite procedure, which gives decidability. □

The proof also shows that $L^{(\infty)}$ is the largest set of live symbols in which every symbol has a compatible partner in every Moore direction. This is arc consistency for the network on $\mathbb{Z}^2$ in which every site has domain L_R and Moore neighbours must carry compatible symbols.

Corollary 4.27 (One token is not enough)**.** *Let $\ell \mapsto R_\ell$ be a total computable map from $\mathbb{N}$ to finite Wang sets such that every R_ℓ that does not tile is 1-token refutable. Then $\{\ell : R_\ell \text{ tiles}\}$ is decidable.*

Proof. Given ℓ, compute $L^{(\infty)}$ for R_ℓ. If $L^{(\infty)} = \emptyset$, then R_ℓ is 1-token refutable by Proposition 4.26, so it does not tile by Theorem 4.25(iv). If $L^{(\infty)} \neq \emptyset$, then R_ℓ is not 1-token refutable, so it tiles by the hypothesis on the family. □

So the fixed number of tokens in Theorem 9.8 is at least 2. For $j \geq 2$ the tokens of a state can be arbitrarily far apart, the states up to translation form an infinite set, and the argument of Proposition 4.26 does not apply.

Proposition 4.28 (Finite token width)**.** *If R is j-token refutable, then it is $(j+1)$-token refutable. If no legal pattern exists on a square of side m, then R is m^2-token refutable. Hence the token width of R is finite if and only if R does not tile, and R is j-token refutable exactly when j is at least its token width.*

Proof. Let $\mathcal{W}$ be a nonempty plane-safe set of $(j+1)$-token states. Deleting the last token from every state gives a nonempty set of j-token states, and this set is plane-safe. Indeed (W1) is inherited, a clone among the first j tokens is the image of a clone in $\mathcal{W}$, and a move of one of the first j tokens is the image of a move in $\mathcal{W}$. By Theorem 4.25(i), if R is not $(j+1)$-token refutable, it is not j-token refutable. This is the first claim.

Suppose that no legal pattern exists on the square $U = \{0, \dots, m-1\}^2$, and that R is not m^2-token refutable. By Theorem 4.25(iii) there is an E_{m^2} family H for R. Since $|U| = m^2$, Remark 4.23 gives a member of H with domain U. It is a legal pattern on U, a contradiction. This is the second claim.

If R does not tile, Lemma 4.2 gives such a square, so the token width is finite. If R tiles, Theorem 4.25(iv) shows that it is not j-token refutable for any j. The last claim follows from the first. □

5 Bounded-width refutations give forgetting

This section proves Theorem 5.1. If R is j-token refutable and $k \geq \max(2, 2j)$, then the noisy rule $F_R^{[k]}$ forgets its initial configuration at every noise rate below a threshold that depends only on k. The refutation time enters the bound only through a constant factor.

The proof has four steps. First, a violated tuple of tokens that carries no replacement has violated parent tuples in at least $k + 1 - j$ provider planes, all on one common set of at most j planar sites (Lemma 5.4). Second, a pigeonhole step turns this into a binary field E. The field is computed from the replacement indicators alone, and after the refutation time it covers every live site of every run (Lemma 5.6 and Corollary 5.7). Third, Toom's contour argument bounds the size of an explanation of a marked point by the number of replacements in it (Lemma 5.8). Fourth, contours placed along a discrepancy path in disjoint time slabs force linearly many distinct replacements (Lemma 5.10), and a Peierls count ends the proof (Section 5.6). The field E (Definition 5.5) and the contour lemma (Lemma 5.8) are stated in the generality in which the proof of Theorem 6.2 uses them again. Section 5.7 shows that the bound holds for all starts at once (Corollary 5.12); Corollary 10.8 uses this form.

5.1 Setting and statement

Throughout this section $j \geq 1$ and $k \geq \max(2, 2j)$ are integers, $q = k + 1$, and R is a Wang set with live symbols L_R. We write $F = F_R^{[k]}$. A site of $\mathbb{Z}^{2+k} = \mathbb{Z}^2 \times \mathbb{Z}^k$ is written $v = (u, P)$, with planar part $u \in \mathbb{Z}^2$ and transverse part $P = (P_1, \dots, P_k) \in \mathbb{Z}^k$. The providers are $c_0 = 0$ and $c_i = e_i$ for $1 \leq i \leq k$, and $\bar{D} = D \cup \{0\} = \{-1, 0, 1\}^2$. The value $C(x(\cdot, P))(u)$ depends only on the symbols $x(u + d, P)$ with $d \in \bar{D}$. Hence $F(x)(v)$ depends only on the restriction of x to

$$\mathcal{N}(v) = \big\{(u + d,\ P + c_i) :\ d \in \bar{D},\ 0 \leq i \leq k\big\}. \tag{27}$$

Noise and coupling. Let $\eta_t(v) \in \{0, 1\}$ and $\zeta_t(v) \in \mathcal{A}_R$, for $t \geq 1$ and $v \in \mathbb{Z}^{2+k}$, be independent random variables with $\mathbb{P}(\eta_t(v) = 1) = \varepsilon$ and $\zeta_t(v)$ uniform on $\mathcal{A}_R$. The run from a configuration x is given by $X_0 = x$ and, for $t \geq 1$,

$$X_t(v) = \begin{cases} \zeta_t(v) & \text{if } \eta_t(v) = 1, \\ F(X_{t-1})(v) & \text{if } \eta_t(v) = 0. \end{cases} \tag{28}$$

Its law is Law_x. A space-time point (v, t) with $t \geq 1$ and $\eta_t(v) = 1$ is an *event*. Being an event does not depend on the value $\zeta_t(v)$: a replacement whose value equals $F(X_{t-1})(v)$ is still an event. The *common-noise coupling* runs X from x and Y from $\perp$ with the same variables η and ζ. Then Y has law $\mathrm{Law}_\perp$.

Token tables. We recall the token tables. Fix $r \geq 0$ and the box $V = \{-r, \dots, r\}^2$. A *table* S assigns to each position tuple $\mathbf{z} = (z_1, \dots, z_j) \in V^j$ a set $S(\mathbf{z})$ of symbol tuples $\mathbf{b} = (b_1, \dots, b_j) \in L_R^j$. Token l sits at z_l and carries b_l. We write $\mathbf{b}[l \leftarrow b]$ and $\mathbf{z}[l \leftarrow z]$ for the tuples with entry l replaced. We write $S \subseteq S'$ if $S(\mathbf{z}) \subseteq S'(\mathbf{z})$ for every $\mathbf{z}$. Three rules delete $\mathbf{b}$ from $S(\mathbf{z})$, with S itself as the reference table:

- the *coincidence rule*, if $z_l = z_{l'}$ and $b_l \neq b_{l'}$ for some l, l';
- the *clone rule*, if $\mathbf{b}[l' \leftarrow b_l] \notin S(\mathbf{z}[l' \leftarrow z_l])$ for some $l \neq l'$;
- the *move rule*, if for some l and some $d \in D$ with $z_l + d \in V$ there is no live symbol b such that b_l at z_l and b at $z_l + d$ are compatible and $\mathbf{b}[l \leftarrow b] \in S(\mathbf{z}[l \leftarrow z_l + d])$.

$\mathrm{Inst}(S)$ is obtained from S by applying the coincidence and clone rules, one deletion at a time, until neither applies. $\mathrm{Timed}(S)$ is obtained from S by deleting at once every tuple that the move rule deletes with reference S. Let $S_{\mathrm{all}}(\mathbf{z}) = L_R^j$ for every $\mathbf{z}$, and put

$$S_0 = \mathrm{Inst}(S_{\mathrm{all}}), \qquad S_{n+1} = \mathrm{Inst}(\mathrm{Timed}(S_n)) \quad (n \geq 0).$$

The set R is *j-token refutable with time τ* if $S_\tau(0, \dots, 0) = \emptyset$ for some r.

For $w \in \mathbb{Z}^2$ and $\mathbf{z} \in (w + V)^j$ put $S_n^w(\mathbf{z}) = S_n(\mathbf{z} - w)$, where $\mathbf{z} - w = (z_1 - w, \dots, z_j - w)$. Fix a run X. A pair $(\mathbf{z}, \mathbf{b})$ with $\mathbf{z} \in (w + V)^j$ and $\mathbf{b} \in L_R^j$ is *present at* (P, t) if $X_t(z_l, P) = b_l$ for every l. It is *clean* if none of the points $((z_l, P), t)$ is an event. It *violates* S_n^w if $\mathbf{b} \notin S_n^w(\mathbf{z})$.

Theorem 5.1 (forgetting). *Let $j \geq 1$ and $k \geq \max(2, 2j)$, and put $q = k + 1$, $a = 2q^2 + 1$ and $D_0 = 50q$. Let R be a Wang set that is j-token refutable with time τ. Let $0 < \varepsilon < 1$ be such that $\beta = 2D_0^2 \varepsilon^{1/a} < 1$. Under the common-noise coupling, for every configuration x, every site v and every time $t \geq 0$,*

$$\mathbb{P}\big(X_t(v) \neq Y_t(v)\big) \leq 2\varepsilon^{-\tau/a}\, \frac{\beta^t}{1 - \beta}. \tag{29}$$

Consequently, for every finite set $O \subset \mathbb{Z}^{2+k}$ and every $t \geq 0$,

$$\sup_x \, d_{\mathrm{TV}}\big(\mathrm{Law}_x(X_t|_O),\, \mathrm{Law}_\perp(X_t|_O)\big) \leq 2|O|\,\varepsilon^{-\tau/a}\,\frac{\beta^t}{1-\beta}, \tag{30}$$

which tends to 0 as $t \to \infty$. So the noisy rule forgets its initial configuration, and it is ergodic.

Remark 5.2. The rate β depends only on k and ε. The set R, the box and the time τ enter only through the factor $\varepsilon^{-\tau/a}$. If $0 < \varepsilon \leq \varepsilon_k = (4D_0^2)^{-a}$, then $\beta \leq 2D_0^2(4D_0^2)^{-1} = 1/2$. The hypothesis $k \geq 2j$ is used once, in the pigeonhole step of Lemma 5.6.

5.2 Instant rules and transport

A table S' is *closed* if neither the coincidence rule nor the clone rule, with reference S', deletes a tuple of S', as in Lemma 4.15.

Lemma 5.3. *Let S be a table.*

(a) $\mathrm{Inst}(S)$ *is the largest closed table contained in S. In particular it does not depend on the order of the deletions.*

(b) Let $\mathcal{Q}$ be a set of pairs $(\mathbf{z}, \mathbf{b}) \in V^j \times L_R^j$ with three properties: if $z_l = z_{l'}$ then $b_l = b_{l'}$; if $(\mathbf{z}, \mathbf{b}) \in \mathcal{Q}$ and $l \neq l'$, then $(\mathbf{z}[l' \leftarrow z_l], \mathbf{b}[l' \leftarrow b_l]) \in \mathcal{Q}$; and $\mathbf{b} \in S(\mathbf{z})$ for every $(\mathbf{z}, \mathbf{b}) \in \mathcal{Q}$. Then $\mathbf{b} \in \mathrm{Inst}(S)(\mathbf{z})$ for every $(\mathbf{z}, \mathbf{b}) \in \mathcal{Q}$.

Proof. (a) This is Lemma 4.15(i).

(b) Let $S^* = \mathrm{Inst}(S)$, which is closed by (a), and let $S'(\mathbf{z}) = S^*(\mathbf{z}) \cup \{\mathbf{b} : (\mathbf{z}, \mathbf{b}) \in \mathcal{Q}\}$. Then $S' \subseteq S$. The table S' is closed. For a pair of $\mathcal{Q}$, the coincidence rule does not apply by the first property, and the clone rule consults a member of $\mathcal{Q}$ by the second. A tuple of S^* is kept because S^* is closed and $S^* \subseteq S'$. By (a), $S' \subseteq S^* = \mathrm{Inst}(S)$. □

Present pairs give such families. Fix a run, w, P, t and a set $Z \subseteq w + V$. Let $\mathcal{Q}$ consist of the pairs $(\mathbf{z} - w, \mathbf{b})$ such that $(\mathbf{z}, \mathbf{b})$ is present at (P, t) and all positions of $\mathbf{z}$ lie in Z. Two tokens on one site read the one symbol there, and cloning a token of a present pair gives a present pair. So $\mathcal{Q}$ has the first two properties of Lemma 5.3(b).

The rule F commutes with planar translations. Replacing $X_t(u, P)$ by $X_t(u + w, P)$, and the noise variables likewise, therefore turns a run into a run, present pairs on $w + V$ into present pairs on V, and S_n^w into S_n. So it suffices to state the next lemma for every w and to prove it for $w = 0$.

Lemma 5.4 (transport)**.** *Fix a configuration x, a realisation of the noise, and the run X from x. Let $w \in \mathbb{Z}^2$, $P \in \mathbb{Z}^k$, $t \geq 1$ and $n \geq 1$. Let $(\mathbf{z}, \mathbf{b})$, with $\mathbf{z} \in (w + V)^j$, be a clean pair present at (P, t) that violates S_n^w. Then there exist*

- *a set $\Gamma \subseteq w + V$ with $|\Gamma| \leq j$,*
- *a map $\pi : \Gamma \to \{z_1, \ldots, z_j\}$ with $z - \pi(z) \in \bar{D}$ for every $z \in \Gamma$, and*
- *a set $I \subseteq \{0, \ldots, k\}$ with $|I| \geq k + 1 - j$,*

such that the following holds for every $i \in I$: there is a pair $(\mathbf{z}', \mathbf{b}')$ with $\mathbf{z}' \in \Gamma^j$ that is present at $(P + c_i, t - 1)$ and violates S_{n-1}^w.

In the statement, Γ and π are chosen before the index i. The same set Γ and the same map π serve every $i \in I$; only the pair $(\mathbf{z}', \mathbf{b}')$ depends on i.

Proof. We take $w = 0$ and write S_n for S_n^0. For $t \geq 1$ let y_{t-1} be the configuration obtained by applying C to X_{t-1} in every plane, so that $y_{t-1}(u, P') = C(X_{t-1}(\cdot, P'))(u)$.

Step 1: a tuple that fails the move round. Let $\mathcal{Q}$ be the set of pairs present at (P, t) whose positions lie in $\{z_1, \dots, z_j\}$. It has the first two properties of Lemma 5.3(b). If every member $(\mathbf{z}', \mathbf{b}')$ of $\mathcal{Q}$ had $\mathbf{b}' \in \mathrm{Timed}(S_{n-1})(\mathbf{z}')$, then Lemma 5.3(b) with $S = \mathrm{Timed}(S_{n-1})$ would give $\mathcal{Q} \subseteq \mathrm{Inst}(\mathrm{Timed}(S_{n-1})) = S_n$. This is false, since $(\mathbf{z}, \mathbf{b}) \in \mathcal{Q}$ and $\mathbf{b} \notin S_n(\mathbf{z})$. So some $(\mathbf{z}^*, \mathbf{b}^*) \in \mathcal{Q}$ has $\mathbf{b}^* \notin \mathrm{Timed}(S_{n-1})(\mathbf{z}^*)$. Then either $\mathbf{b}^* \notin S_{n-1}(\mathbf{z}^*)$, or $\mathbf{b}^* \in S_{n-1}(\mathbf{z}^*)$ and the move rule with reference S_{n-1} deletes it.

Step 2: common providers. The positions of $\mathbf{z}^*$ form a set of at most j sites among $z_1, \dots, z_j$. Let z be one of them. Say that the index i *succeeds* at z if $y_{t-1}(z, P + c_i) = X_t(z, P)$. The point $((z, P), t)$ is not an event, because $(\mathbf{z}, \mathbf{b})$ is clean, so $X_t(z, P) = F(X_{t-1})(z, P)$, and this value is live. By the definition of F, at least k of the $k + 1$ indices succeed at z, so at most one fails. Let I be the set of indices that succeed at every position of $\mathbf{z}^*$. Then $|I| \geq k + 1 - j$. Let $i \in I$ and $1 \leq l \leq j$. Then $y_{t-1}(z_l^*, P + c_i) = X_t(z_l^*, P) = b_l^*$, which is live. By the definition of C, $X_{t-1}(z_l^*, P + c_i) = b_l^*$, and for every $d \in D$ the symbol $X_{t-1}(z_l^* + d, P + c_i)$ is live and compatible with b_l^* (the latter at z_l^*, the former at $z_l^* + d$).

Step 3: the parents. Suppose first that $\mathbf{b}^* \notin S_{n-1}(\mathbf{z}^*)$. Let Γ be the set of positions of $\mathbf{z}^*$ and let π be the identity. By Step 2, $(\mathbf{z}^*, \mathbf{b}^*)$ is present at $(P + c_i, t - 1)$ for every $i \in I$, and it violates S_{n-1}.

Otherwise the move rule deletes $\mathbf{b}^*$: there are l and $d \in D$ with $z_l^* + d \in V$ such that no live symbol b compatible with b_l^* (with b_l^* at z_l^* and b at $z_l^* + d$) has $\mathbf{b}^*[l \leftarrow b] \in S_{n-1}(\mathbf{z}')$, where $\mathbf{z}' = \mathbf{z}^*[l \leftarrow z_l^* + d]$. Let Γ be the set of positions of $\mathbf{z}'$. Put $\pi(z_l^* + d) = z_l^*$ and $\pi(z) = z$ for the other points z of Γ; then $z - \pi(z) \in \bar{D}$ for every $z \in \Gamma$. For $i \in I$ let $b_i = X_{t-1}(z_l^* + d, P + c_i)$. By Step 2, b_i is live and compatible with b_l^*, so $\mathbf{b}^*[l \leftarrow b_i] \notin S_{n-1}(\mathbf{z}')$. The pair $(\mathbf{z}', \mathbf{b}^*[l \leftarrow b_i])$ is present at $(P + c_i, t - 1)$: the position $z_{l'}^*$, $l' \neq l$, carries $b_{l'}^*$ by Step 2, and the position $z_l^* + d$ carries b_i. If $z_l^* + d$ equals some $z_{l'}^*$, both statements describe the one symbol there.

In both cases $\Gamma \subseteq V$ and $|\Gamma| \leq j$. The positions of $\mathbf{z}^*$ lie in $\{z_1, \dots, z_j\}$, so π maps into this set. The set I was fixed in Step 2, and Γ and π are determined by $\mathbf{z}^*$, $\mathbf{b}^*$ and, in the second case, l and d. All of these were chosen before any $i \in I$ was considered. □

The lemma is the only place where the quorum enters. Every provider that succeeds at all positions of the tuple carries a violating parent, and all these parents sit on the same set Γ.

5.3 A covering field computed from the noise

As in Toom's stability proofs [47, 19], we dominate the errors by a binary field that is driven by the noise alone.

Definition 5.5 (domination field)**.** Let $\eta = (\eta_t(v))_{t \geq 1, v \in \mathbb{Z}^{2+k}}$ be the replacement indicators. The *domination field* $E_t(v) \in \{0, 1\}$, for $t \geq 0$ and $v \in \mathbb{Z}^{2+k}$, is defined recursively by $E_0 \equiv 0$ and, for $t \geq 1$ and $v = (u, P)$, as follows. If $\eta_t(v) = 1$, then $E_t(v) = 1$. If $\eta_t(v) = 0$, then $E_t(v) = 1$ if and only if there are indices $i \neq i'$ in $\{0, \dots, k\}$ and offsets $d, d' \in \bar{D}$ with

$$E_{t-1}(u + d,\, P + c_i) = E_{t-1}(u + d',\, P + c_{i'}) = 1. \tag{31}$$

The field E is a function of the indicators η alone. It does not depend on R, on the initial configuration or on the replacement values ζ. In (31) the two points with value 1 lie in different planes, their planar offsets may differ, and the offset 0 is allowed.

Lemma 5.6 (cover)**.** *Let $k \geq 2j$. Fix a configuration x, a realisation of the noise and the run X from x. Let $0 \leq n \leq t$, $w \in \mathbb{Z}^2$ and $P \in \mathbb{Z}^k$. If a pair $(\mathbf{z}, \mathbf{b})$ with $\mathbf{z} \in (w+V)^j$ is present at (P,t) and violates S_n^w, then $E_t(z_l, P) = 1$ for some l.*

Proof. We prove by induction on $n \geq 0$ the statement

(C_n) for every $t \geq n$, every $w \in \mathbb{Z}^2$, every $P \in \mathbb{Z}^k$ and every pair $(\mathbf{z}, \mathbf{b})$ with $\mathbf{z} \in (w+V)^j$ that is present at (P,t) and violates S_n^w, there is an index l with $E_t(z_l, P) = 1$.

The configuration x and the realisation of the noise stay fixed throughout.

The case $n = 0$. Fix t, w and P. The family $\mathcal{Q}$ of all pairs $(\mathbf{z} - w, \mathbf{b})$ with $(\mathbf{z}, \mathbf{b})$ present at (P,t) has the first two properties of Lemma 5.3(b), and it lies in S_{all}. So every present pair lies in S_0^w. No pair satisfies the hypothesis of (C_0), and (C_0) holds.

From $n-1$ to n. Let $n \geq 1$ and assume (C_{n-1}). Fix $t \geq n$, w, P and a pair $(\mathbf{z}, \mathbf{b})$ with $\mathbf{z} \in (w+V)^j$ that is present at (P,t) and violates S_n^w.

If some point $((z_l, P), t)$ is an event, then $E_t(z_l, P) = 1$ by Definition 5.5, since $t \geq 1$. We are done in this case.

Otherwise the pair is clean. Since $t \geq 1$ and $n \geq 1$, Lemma 5.4 applies. It gives a set Γ, a map π and a set I as there. We fix them now; nothing below changes them. For each $i \in I$ the lemma gives a pair $(\mathbf{z}^{(i)}, \mathbf{b}^{(i)})$ with $\mathbf{z}^{(i)} \in \Gamma^j$ that is present at $(P + c_i, t-1)$ and violates S_{n-1}^w; fix one such pair for each i. Since $0 \leq n-1 \leq t-1$, the statement (C_{n-1}), applied with time $t-1$, translate w, plane $P + c_i$ and the pair $(\mathbf{z}^{(i)}, \mathbf{b}^{(i)})$, gives an index l_i with

$$E_{t-1}\big(z_i',\, P + c_i\big) = 1, \qquad \text{where } z_i' = z_{l_i}^{(i)} \in \Gamma.$$

Now consider the map $i \mapsto \pi(z_i')$ from I to $\{z_1, \dots, z_j\}$. Its domain has $|I| \geq k + 1 - j \geq j + 1$ elements, because $k \geq 2j$, and its target has at most j elements. So there are indices $i \neq i'$ in I with $\pi(z_i') = \pi(z_{i'}')$; call this point z_l. This comparison is possible because π is one map, fixed before the indices i were considered. Put $d = z_i' - z_l$ and $d' = z_{i'}' - z_l$. Since $z_l = \pi(z_i')$, the defining property of π gives $d \in \bar{D}$, and likewise $d' \in \bar{D}$. So

$$E_{t-1}(z_l + d,\, P + c_i) = E_{t-1}(z_l + d',\, P + c_{i'}) = 1, \qquad i \neq i'.$$

The point $((z_l, P), t)$ is not an event, because the pair is clean. By Definition 5.5, $E_t(z_l, P) = 1$. This proves (C_n). □

Corollary 5.7. *Let $k \geq 2j$, and let $S_\tau(0, \dots, 0) = \emptyset$ for the box V. Fix a configuration x, a realisation of the noise and the run X from x. For every $t \geq \tau$ and every site v, if $X_t(v)$ is live then $E_t(v) = 1$.*

Proof. Let $v = (u, P)$ and $X_t(v) = b$ live. The diagonal pair $((u, \dots, u), (b, \dots, b))$ is present at (P,t) on $u + V$. It violates S_τ^u, because $S_\tau^u(u, \dots, u) = S_\tau(0, \dots, 0) = \emptyset$. Lemma 5.6 with $n = \tau \leq t$ and $w = u$ gives $E_t(u, P) = 1$, since all positions equal u. □

The corollary holds for every initial configuration and every realisation of η and ζ, with the same field E. In particular it holds for both runs of the common-noise coupling.

5.4 Contours

Let $\mathcal{G}$ be the graph with vertex set $\mathbb{Z}^2 \times \mathbb{Z}^k \times \{0, 1, 2, \dots\}$ in which each vertex (u, P, t) with $t \geq 1$ is joined to the points $(u + d, P + c_i, t-1)$, $d \in \bar{D}$, $0 \leq i \leq k$. We call $(u + d, P + c_i, t-1)$ a *lower neighbour of (u, P, t) through provider i*. We identify (u, P, t) with the space-time point (v, t), $v = (u, P)$. Every edge of $\mathcal{G}$ joins two consecutive times. A vertex has at most $9q$ lower

and $9q$ upper neighbours, so every degree is at most $18q \le D_0$. By (27), $(v', t-1)$ is a lower neighbour of (v,t) whenever $v' \in \mathcal{N}(v)$.

The next lemma is Toom's contour argument [47]; see also [19]. The poles, the charges, the linear functions that sum to zero, and the splitting of frontier components in Step 3 below follow [46, Theorem 33, Definitions 35 and 37, Lemma 36]. We state it for any set of vertices in which every non-fault has two lower neighbours through different providers. Only the sinks need to be distinct faults; arms may meet and may pass through the same vertices. We include the proof to keep the section self-contained.

Lemma 5.8 (contours)**.** *Let $\mathcal{U}$ be a set of vertices of $\mathcal{G}$, all of positive time, and let $\mathcal{U}_{\mathrm{f}} \subseteq \mathcal{U}$; we call the points of $\mathcal{U}_{\mathrm{f}}$* faults*. Suppose that every vertex of $\mathcal{U} \setminus \mathcal{U}_{\mathrm{f}}$ has two lower neighbours in $\mathcal{U}$ through different providers. Let $\rho = (v, t_\rho) \in \mathcal{U} \setminus \mathcal{U}_{\mathrm{f}}$. Then there are an integer $m \ge 2$, a set $\Sigma \subseteq \mathcal{U}_{\mathrm{f}}$ of m faults and a finite set K of edges of $\mathcal{G}$ with the following properties.*

(i) The edges of K form a connected graph. Each edge of K joins a vertex of $\mathcal{U} \setminus \mathcal{U}_{\mathrm{f}}$ to one of its lower neighbours in $\mathcal{U}$. The vertices of K lie in $\mathcal{U}$ and include ρ and every point of Σ.

(ii) $1 \le |K| \le q^2(m-1)$. In particular $|\Sigma| = m \ge 1 + |K|/q^2$.

(iii) Let s be the least time of a point of Σ. Every vertex of K has time in $[s, t_\rho]$, and $t_\rho - s \le q^2(m-1)$.

In particular, fix a realisation of the indicators η, let E be the domination field of Definition 5.5, and let $\rho = (v, t_\rho)$ with $t_\rho \ge 1$, $E_{t_\rho}(v) = 1$ and $\eta_{t_\rho}(v) = 0$. Then there is a set K of $n = |K| \ge 1$ edges of $\mathcal{G}$, each joining a vertex to one of its lower neighbours, that forms a connected graph containing ρ, whose vertices include m distinct events, and with $n \le q^2(m-1)$.

A vertex (v,t) is *marked* if $E_t(v) = 1$. The last assertion of Lemma 5.8 is the case in which $\mathcal{U}$ is the set of marked vertices and $\mathcal{U}_{\mathrm{f}}$ is the set of events. These sets satisfy the hypothesis of the lemma. Indeed, marked vertices have positive time because $E_0 \equiv 0$; events are marked by Definition 5.5; and by (31) a marked vertex that is not an event has two marked lower neighbours through different providers, at offsets in $\bar{D}$.

Proof of Lemma 5.8. The last assertion follows from the general one, applied to the sets of the preceding paragraph, which satisfy its hypothesis; we prove the general assertion.

Step 1: the support graph. For every vertex $\theta \in \mathcal{U} \setminus \mathcal{U}_{\mathrm{f}}$ fix once and for all two lower neighbours $\theta^{(1)}, \theta^{(2)} \in \mathcal{U}$ of θ through different providers, the *supports* of θ. They lie in different planes, so they are distinct. Let $\mathcal{B}$ be the set of vertices reached from ρ by following supports, including ρ. It is contained in $\mathcal{U}$. Times decrease by one along each support and stay positive, so $\mathcal{B}$ is finite. The faults in $\mathcal{B}$ are the *terminals*; let $\Phi = \mathcal{B} \cap \mathcal{U}_{\mathrm{f}}$ be the set of terminals. Every other vertex of $\mathcal{B}$, including ρ, has both supports in $\mathcal{B}$. For $\theta \in \mathcal{B}$ let $\Phi(\theta)$ be the set of terminals reached from θ. Following supports from θ ends at a terminal, so $\Phi(\theta) \ne \emptyset$. Moreover $\Phi(\rho) = \Phi$, and

$$\Phi(\varphi) = \{\varphi\} \text{ for } \varphi \in \Phi, \qquad \Phi(\theta) = \Phi(\theta^{(1)}) \cup \Phi(\theta^{(2)}) \text{ for } \theta \in \mathcal{B} \setminus \Phi. \tag{32}$$

Step 2: charges, arms and the potential. The *charges* are the indices $i \in \{0, \dots, k\}$. An edge from a vertex to its lower neighbour through provider i' is *good* for charge i if $i' \ne i$. The two supports of a non-terminal vertex θ come through different providers, so for each charge i at least one of them is reached by a good edge; fix such a support $g_i(\theta)$. A *pole* σ consists of a non-terminal vertex $\theta_\sigma \in \mathcal{B}$, its *location*, and a *first step* $f_i(\sigma) \in \{\theta_\sigma^{(1)}, \theta_\sigma^{(2)}\}$ for each charge i. Its *arm* of charge i is the path

$$\theta_\sigma,\ f_i(\sigma),\ g_i(f_i(\sigma)),\ g_i(g_i(f_i(\sigma))),\ \dots$$

stopped at its first terminal. Let $\mathrm{end}_i(\sigma) \in \Phi(f_i(\sigma))$ be the last vertex of this arm. Every arm has at least one edge, and every edge of an arm except possibly the first is good for its charge. The *root pole* σ_ρ has location ρ and first steps $f_i(\sigma_\rho) = g_i(\rho)$, so all its edges are good.

For $P \in \mathbb{Z}^k$ and $t \geq 0$ put

$$\psi_0(P,t) = q(P_1 + \cdots + P_k) + kt, \qquad \psi_i(P,t) = -qP_i - t \quad (1 \leq i \leq k),$$

and $\psi_i(\theta) = \psi_i(P,t)$ for $\theta = (u,P,t)$. Then

$$\sum_{i=0}^{k} \psi_i(\theta) = 0 \qquad \text{for every vertex } \theta. \tag{33}$$

Let θ' be a lower neighbour of θ through provider i'. The function ψ_0 changes by $-k$ if $i' = 0$ and by $q - k = 1$ otherwise. For $i \geq 1$, the function ψ_i changes by $1 - q = -k$ if $i' = i$ and by 1 otherwise. Hence

$$\psi_i(\theta') - \psi_i(\theta) = \begin{cases} 1 & \text{if } i' \neq i, \\ -k & \text{if } i' = i. \end{cases} \tag{34}$$

So a good edge raises the function of its charge by exactly 1, and any edge lowers it by at most k.

Step 3: the poles. List the non-terminal vertices of $\mathcal{B}$ as $\theta_1 = \rho, \theta_2, \ldots, \theta_N$ in order of non-increasing time. Put $W_0 = \{\rho\}$. For $1 \leq h \leq N$ let W_h be the set of vertices of $\mathcal{B} \setminus \{\theta_1, \ldots, \theta_h\}$ that are supports of at least one of $\theta_1, \ldots, \theta_h$. On W_h join θ and θ' if $\Phi(\theta) \cap \Phi(\theta') \neq \emptyset$. For a connected component $\mathcal{K}$ of this graph put $\Phi(\mathcal{K}) = \bigcup_{\theta \in \mathcal{K}} \Phi(\theta)$. We use four facts.

(a) For $0 \leq h < N$, we have $\theta_{h+1} \in W_h$ and $W_{h+1} = (W_h \setminus \{\theta_{h+1}\}) \cup \{\theta_{h+1}^{(1)}, \theta_{h+1}^{(2)}\}$. For $h = 0$ this is the definition. Let $h \geq 1$. The vertex $\theta_{h+1} \neq \rho$ is a support of some vertex of $\mathcal{B}$. That vertex is non-terminal and has larger time, so it is among $\theta_1, \ldots, \theta_h$. The supports of θ_{h+1} have smaller time than θ_{h+1}, so they are not among $\theta_1, \ldots, \theta_{h+1}$. The formula follows.

(b) For each h, the sets $\Phi(\mathcal{K})$, over the components $\mathcal{K}$ of W_h, partition Φ. They are disjoint by the definition of the graph. Their union $\bigcup_{\theta \in W_h} \Phi(\theta)$ equals $\Phi(\rho) = \Phi$ for $h = 0$, and by (a) and (32) it does not change from h to $h + 1$.

(c) $W_N = \Phi$, and the components of W_N are single terminals. Indeed W_N contains no non-terminal vertex, so $\Phi(\theta) = \{\theta\}$ for $\theta \in W_N$, and (b) gives $W_N = \Phi$. Two distinct terminals are not joined.

(d) Let $\theta = \theta_{h+1}$, let $\mathcal{K}$ be its component in W_h, and let $\mathcal{K}' = (\mathcal{K} \setminus \{\theta\}) \cup \{\theta^{(1)}, \theta^{(2)}\}$. Then the components of W_{h+1} are the components of W_h other than $\mathcal{K}$, together with one or two components whose union is $\mathcal{K}'$. Each of the latter contains $\theta^{(1)}$ or $\theta^{(2)}$, and $\Phi(\mathcal{K}') = \Phi(\mathcal{K})$.

To prove (d), note first that a support $\theta^{(l)}$ lying in W_h lies in $\mathcal{K}$, because $\emptyset \neq \Phi(\theta^{(l)}) \subseteq \Phi(\theta)$. So W_{h+1} is the disjoint union of $W_h \setminus \mathcal{K}$ and $\mathcal{K}'$. By (32), $\Phi(\mathcal{K}') = \Phi(\mathcal{K})$, and this set is disjoint from $\Phi(\theta')$ for every $\theta' \in W_h \setminus \mathcal{K}$. So no vertex of $\mathcal{K}'$ is joined to a vertex outside $\mathcal{K}'$, and the components of W_h other than $\mathcal{K}$ remain components. Now let $\theta' \in \mathcal{K} \setminus \{\theta\}$, and take a shortest path $\theta' = \theta'_0, \theta'_1, \ldots, \theta'_p = \theta$ in $\mathcal{K}$. Its vertices other than θ lie in $\mathcal{K}'$, and $\Phi(\theta'_{p-1})$ meets $\Phi(\theta) = \Phi(\theta^{(1)}) \cup \Phi(\theta^{(2)})$. So θ' is joined inside $\mathcal{K}'$ to $\theta^{(1)}$ or to $\theta^{(2)}$.

We now choose finite sets $\mathcal{P}_0 \subseteq \mathcal{P}_1 \subseteq \cdots \subseteq \mathcal{P}_N$ of poles with the following property.

(I_h) For every component $\mathcal{K}$ of W_h and every charge i, exactly one pole $\sigma \in \mathcal{P}_h$ has $\mathrm{end}_i(\sigma) \in \Phi(\mathcal{K})$.

Put $\mathcal{P}_0 = \{\sigma_\rho\}$. Property ($\mathrm{I}_0$) holds, since W_0 has the single component $\{\rho\}$ and $\Phi(\rho) = \Phi$. Suppose (I_h) holds, and let θ, $\mathcal{K}$ and $\mathcal{K}'$ be as in (d). If $\mathcal{K}'$ is connected, put $\mathcal{P}_{h+1} = \mathcal{P}_h$. Otherwise $\mathcal{K}'$ has two components, $\mathcal{K}_1 \ni \theta^{(1)}$ and $\mathcal{K}_2 \ni \theta^{(2)}$. For each charge i, the unique

pole of $\mathcal{P}_h$ whose arm of charge i ends in $\Phi(\mathcal{K}) = \Phi(\mathcal{K}_1) \cup \Phi(\mathcal{K}_2)$ ends in exactly one of these two sets; let $\iota(i) \in \{1, 2\}$ be the index of the other one. Let $\mathcal{P}_{h+1}$ be $\mathcal{P}_h$ together with the pole σ_θ located at θ with first steps $f_i(\sigma_\theta) = \theta^{(\iota(i))}$. Its arm of charge i ends in $\Phi(\theta^{(\iota(i))}) \subseteq \Phi(\mathcal{K}_{\iota(i)})$. In both cases (I_{h+1}) follows from (d): the other components keep their terminal sets and their poles, and the arms of the new pole end in $\Phi(\mathcal{K})$.

Let $\mathcal{P} = \mathcal{P}_N$. By (c) and (I_N),

$$\text{for every terminal } \varphi \text{ and every charge } i, \text{ exactly one pole } \sigma \in \mathcal{P} \text{ has } \mathrm{end}_i(\sigma) = \varphi. \tag{35}$$

Every pole other than σ_ρ was added at a split; its first steps are arbitrary supports, and all later edges of its arms are good.

Step 4: the contour. Build an abstract graph H as follows. Take a separate copy of every arm of every pole of $\mathcal{P}$, as a path graph whose vertices are labelled by the vertices of $\mathcal{B}$ that the arm visits. Identify the initial vertices of the q arms of each pole σ into one vertex, the *source* of σ. Identify all final vertices labelled by the same terminal φ into one vertex, the *sink* of φ. The other vertices of H are *interior*. Each edge of H joins two vertices whose labels are joined in $\mathcal{G}$.

Let H_ρ be the connected component of H that contains the source of σ_ρ. Let $\mathcal{P}_\rho$ be the set of poles whose sources lie in H_ρ, and Φ_ρ the set of terminals whose sinks lie in H_ρ. The arms of a pole in $\mathcal{P}_\rho$ lie in H_ρ, and every edge of H_ρ lies on such an arm. For each charge i, the map end_i restricts to a bijection from $\mathcal{P}_\rho$ onto Φ_ρ. Indeed, end_i is injective on $\mathcal{P}$ by (35); an arm ends in the component of its source; and by (35) the unique arm of charge i that ends at a sink of H_ρ starts at a source of H_ρ. Let $m = |\mathcal{P}_\rho| = |\Phi_\rho|$, and let e_ρ be the number of edges of H_ρ.

Consider the sum of $\psi_i(\mathrm{end}_i(\sigma)) - \psi_i(\theta_\sigma)$ over all $\sigma \in \mathcal{P}_\rho$ and all charges i. By the bijections, it equals

$$\sum_{\varphi \in \Phi_\rho} \sum_{i=0}^{k} \psi_i(\varphi) - \sum_{\sigma \in \mathcal{P}_\rho} \sum_{i=0}^{k} \psi_i(\theta_\sigma) = 0,$$

by (33). On the other hand, each difference is the sum of the increments (34) along the edges of the arm. A good edge contributes 1 and any other edge at least $-k$. Only the first edges of the poles in $\mathcal{P}_\rho \setminus \{\sigma_\rho\}$ can fail to be good, and there are at most $q(m-1)$ of them. Hence

$$0 \geq \big(e_\rho - q(m-1)\big) - kq(m-1) = e_\rho - q^2(m-1),$$

so $e_\rho \leq q^2(m-1)$. The arms of σ_ρ have at least one edge each, so $e_\rho \geq q \geq 1$, and therefore $m \geq 2$.

Let K be the set of edges of $\mathcal{G}$ that are labels of edges of H_ρ, and let $\Sigma = \Phi_\rho$. Then $|K| \leq e_\rho \leq q^2(m-1)$, and K contains the first edges of the arms of σ_ρ, so $|K| \geq 1$. This is (ii). The labelling maps the connected graph H_ρ onto the graph formed by K, so the latter is connected. Each edge of an arm goes from a non-terminal vertex $\theta \in \mathcal{B}$ to one of its supports, so each edge of K joins a vertex of $\mathcal{U} \setminus \mathcal{U}_\mathrm{f}$ to one of its lower neighbours in $\mathcal{U}$. The vertices of K are labels, so they lie in $\mathcal{B} \subseteq \mathcal{U}$. They include ρ, the label of the source of σ_ρ, and every point of Σ, the labels of the sinks. Distinct sinks carry distinct terminals, so Σ consists of m distinct faults. This proves (i). All labels lie in $\mathcal{B}$, so every vertex of K has time at most t_ρ. Let θ be a vertex of H_ρ whose label has the least time. If θ were a source or an interior vertex, an arm would continue from θ to a vertex of H_ρ whose label has time one less. So θ is a sink, and the least time of a vertex of K is s. Finally, a path in H_ρ from the source of σ_ρ to a sink labelled at time s uses distinct edges of H_ρ, so it has at most e_ρ edges, and the time changes by one along each of them. Hence $t_\rho - s \leq e_\rho \leq q^2(m-1)$, which is (iii). □

5.5 Packing contours along a discrepancy path

Lemma 5.9. *Fix a configuration x and a realisation of the noise, and let X and Y be the runs from x and from $\perp$. If $t \geq 1$ and $X_t(v) \neq Y_t(v)$, then there are sites $v_t = v, v_{t-1}, \dots, v_0$ such that, for $1 \leq s \leq t$, the point (v_s, s) is not an event, $v_{s-1} \in \mathcal{N}(v_s)$, and $X_{s-1}(v_{s-1}) \neq Y_{s-1}(v_{s-1})$. In particular (v_s, s) and $(v_{s-1}, s-1)$ are joined in $\mathcal{G}$.*

Proof. Let $s \geq 1$ and $X_s(v_s) \neq Y_s(v_s)$. Then $\eta_s(v_s) = 0$, since otherwise both values equal $\zeta_s(v_s)$. So $F(X_{s-1})(v_s) \neq F(Y_{s-1})(v_s)$. By (27), X_{s-1} and Y_{s-1} differ at some $v_{s-1} \in \mathcal{N}(v_s)$. Repeat down to $s = 1$. □

Lemma 5.10 (packing)**.** *Let $k \geq 2j$, and let $S_\tau(0, \dots, 0) = \emptyset$ for the box V, with $\tau \geq 1$. Fix a configuration x and a realisation of the noise, and let X and Y be the runs from x and from $\perp$. If $t \geq \tau$ and $X_t(v) \neq Y_t(v)$, then there are an integer $e \geq t$, a set K of e edges of $\mathcal{G}$ and a set Σ of events with the following properties: the edges of K form a connected graph; its vertices include (v, t) and every point of Σ; and $|\Sigma| \geq (e - \tau)/a$.*

Proof. Let $v_t, \dots, v_0$ be the sites of Lemma 5.9. For $\tau \leq s \leq t$ the point (v_s, s) is not an event, because $s \geq 1$. One of $X_s(v_s)$ and $Y_s(v_s)$ is live, because they differ. Corollary 5.7, applied to that run, gives $E_s(v_s) = 1$. So (v_s, s) is a marked vertex that is not an event.

Define times $t_1 > t_2 > \dots$ as follows. Put $t_1 = t$. Given $t_l \geq \tau$, apply Lemma 5.8 with $\mathcal{U}$ the set of marked vertices and $\mathcal{U}_{\mathrm{f}}$ the set of events, at $\rho_l = (v_{t_l}, t_l)$. It gives m_l, a set Σ_l of m_l events, a set K_l of edges, and the least time s_l of a point of Σ_l. If $s_l - 1 \geq \tau$, put $t_{l+1} = s_l - 1$. Otherwise stop, and let $L = l$. Since $s_l \leq t_l$, the times decrease and the construction stops.

By Lemma 5.8(iii), the vertices of K_l have times in $[s_l, t_l]$. These intervals are pairwise disjoint, since $t_{l+1} = s_l - 1$. Hence the sets Σ_l are pairwise disjoint, and $\Sigma = \Sigma_1 \cup \dots \cup \Sigma_L$ has $m = m_1 + \dots + m_L$ elements. One event may lie in the support graphs of many marked vertices; the disjoint slabs ensure that it is counted at most once. The intervals $[s_l, t_l]$ cover $[s_L, t]$, and $s_L \leq \tau$ by the stopping rule. By Lemma 5.8(iii),

$$t - \tau + 1 \leq t - s_L + 1 = \sum_{l=1}^{L} (t_l - s_l + 1) \leq \sum_{l=1}^{L} \bigl(q^2(m_l - 1) + 1\bigr) \leq q^2 m.$$

Let K consist of the t edges joining (v_s, s) and $(v_{s-1}, s-1)$, $1 \leq s \leq t$, together with $K_1, \dots, K_L$. The path is connected, and each K_l is connected and contains the path vertex ρ_l. So K forms a connected graph, and its vertices include (v, t) and Σ. Let $e = |K|$. The t path edges join different pairs of times, so they are distinct and $e \geq t$. By Lemma 5.8(ii),

$$e \leq t + \sum_{l=1}^{L} |K_l| \leq t + q^2 \sum_{l=1}^{L} (m_l - 1) \leq t + q^2 m \leq \tau - 1 + 2q^2 m.$$

So $2q^2 m > e - \tau$. If $e \geq \tau$ this gives $m \geq (e - \tau)/(2q^2) \geq (e - \tau)/a$. If $e < \tau$ the bound $m \geq (e - \tau)/a$ is trivial. □

5.6 Counting, and the proof of the theorem

Lemma 5.11. *Let θ_0 be a vertex of $\mathcal{G}$ and $e \geq 1$. There are at most D_0^{2e} sets K of e edges of $\mathcal{G}$ that form a connected graph with θ_0 among its vertices.*

Proof. Every vertex of $\mathcal{G}$ has degree at most D_0. Fix, for every vertex, a numbering of its neighbours by numbers in $\{1, \dots, D_0\}$. Let K be such a set of edges. Doubling every edge of

K gives a connected multigraph in which every degree is even. By Euler's theorem it has a closed walk $p_0 = \theta_0, p_1, \dots, p_{2e} = \theta_0$ that uses each of its $2e$ edges once. Fix one such walk for each K, and encode it by the word $(\omega_1, \dots, \omega_{2e})$, where ω_s is the number of p_s among the neighbours of p_{s-1}. The word and θ_0 determine the walk, and the walk determines K as the set of edges it uses. So distinct sets K have distinct words, and there are at most D_0^{2e} words. □

Proof of Theorem 5.1. If $L_R = \emptyset$, then $\mathcal{A}_R = \{\perp\}$, so $X_t = Y_t$ and there is nothing to prove. Let $L_R \neq \emptyset$, and fix r with $S_\tau(0, \dots, 0) = \emptyset$ for the box $V = \{-r, \dots, r\}^2$. For a live symbol b, the one-element family $\{((0, \dots, 0), (b, \dots, b))\}$ has the three properties of Lemma 5.3(b) with $S = S_{\text{all}}$. So $(b, \dots, b) \in S_0(0, \dots, 0)$, and therefore $\tau \geq 1$.

Fix x, v and t. If $t < \tau$, the right side of (29) equals $2(2D_0^2)^t \varepsilon^{(t-\tau)/a}/(1-\beta) \geq 2$, because $0 < \varepsilon < 1$; so (29) holds. Let $t \geq \tau$. By Lemma 5.10, on the event $X_t(v) \neq Y_t(v)$ there are an integer $e \geq t$, a set K of e edges of $\mathcal{G}$ that forms a connected graph with (v, t) among its vertices, and a set Σ of vertices of K, all of positive time, such that $|\Sigma| \geq (e-\tau)/a$ and $\eta = 1$ on Σ. The pairs (K, Σ) range over a countable family of fixed sets, so the union bound applies. For a fixed set Σ of distinct points of positive time, independence gives $\mathbb{P}(\eta = 1 \text{ on } \Sigma) = \varepsilon^{|\Sigma|}$, and this is at most $\varepsilon^{(e-\tau)/a}$ when $|\Sigma| \geq (e-\tau)/a$, since $0 < \varepsilon < 1$. By Lemma 5.11 there are at most D_0^{2e} choices of K. A connected graph with e edges has at most $e+1$ vertices, so there are at most 2^{e+1} choices of Σ. Therefore

$$\mathbb{P}\big(X_t(v) \neq Y_t(v)\big) \leq \sum_{e \geq t} D_0^{2e}\, 2^{e+1}\, \varepsilon^{(e-\tau)/a} = 2\varepsilon^{-\tau/a} \sum_{e \geq t} \big(2D_0^2 \varepsilon^{1/a}\big)^e = 2\varepsilon^{-\tau/a}\, \frac{\beta^t}{1-\beta}.$$

The right side does not depend on x or v. This proves (29).

For a finite set O, the pair $(X_t|_O, Y_t|_O)$ is a coupling of $\mathrm{Law}_x(X_t|_O)$ and $\mathrm{Law}_\perp(X_t|_O)$. So

$$d_{\mathrm{TV}}\big(\mathrm{Law}_x(X_t|_O), \mathrm{Law}_\perp(X_t|_O)\big) \leq \mathbb{P}\big(X_t|_O \neq Y_t|_O\big) \leq \sum_{v \in O} \mathbb{P}\big(X_t(v) \neq Y_t(v)\big),$$

and (29) gives (30). Since $\beta < 1$, the right side of (30) tends to 0, so the rule forgets its initial configuration. Lemma 4.12 gives ergodicity. □

5.7 All starts at once

The bound (29) compares two given starts. Its proof gives more: one realisation of the noise makes the value at a space-time point the same for all starts, outside an event that depends on the replacement indicators alone.

Run every configuration $x \in \Omega$ with the same variables $\eta_t(v)$ and $\zeta_t(v)$, and write X^x for the run from x; this is the common-noise coupling of all starts. By Section 4, $X_t^x(v)$ depends on x only through its restriction to the finite set $v + \mathcal{N}_t$. So the event in the next corollary is a finite union of events, and it is measurable.

Corollary 5.12 (Coalescence for all starts)**.** *Under the hypotheses of Theorem 5.1, for every site v and every time $t \geq 0$,*

$$\mathbb{P}\big(\textit{the map } x \mapsto X_t^x(v) \textit{ is not constant on } \Omega\big) \leq 2\varepsilon^{-\tau/a}\, \frac{\beta^t}{1-\beta}. \tag{36}$$

Proof. If $L_R = \emptyset$, then Ω has one point and there is nothing to prove. Let $L_R \neq \emptyset$. As in the proof of Theorem 5.1, $\tau \geq 1$, and we fix a box V with $S_\tau(0, \dots, 0) = \emptyset$. For $t < \tau$ the right side of (36) is at least 2, as shown there. Let $t \geq \tau$.

The witness event. For $e \geq t$ let $\mathcal{K}_e$ be the set of pairs (K, Σ) such that K is a set of e edges of $\mathcal{G}$ that forms a connected graph with (v, t) among its vertices, and Σ is a set of vertices of K, all of positive time, with $|\Sigma| \geq (e - \tau)/a$. The sets $\mathcal{K}_e$ depend on v, t, e, τ and a, and on nothing else; in particular they do not depend on a start. Put

$$\mathcal{B}_t(v) = \bigcup_{e \geq t} \bigcup_{(K,\Sigma) \in \mathcal{K}_e} \{\eta = 1 \text{ on } \Sigma\}.$$

It is a countable union of events, and it is a function of the indicators η alone.

Every discrepancy lies in the witness event. Fix a realisation of the noise and a configuration x with $X_t^x(v) \neq X_t^\perp(v)$. Lemma 5.10, applied to the runs $X = X^x$ and $Y = X^\perp$, gives an integer $e \geq t$, a set K of e edges of $\mathcal{G}$ that forms a connected graph with (v, t) among its vertices, and a set Σ of events, all vertices of K, with $|\Sigma| \geq (e - \tau)/a$. Events have positive time, so $(K, \Sigma) \in \mathcal{K}_e$, and $\eta = 1$ on Σ. Hence the realisation lies in $\mathcal{B}_t(v)$. The pair (K, Σ) depends on x, through the discrepancy path of Lemma 5.9; the contours along that path are contours of the marked set of the one field E, which is computed from η alone and covers the live values of every run (Corollary 5.7). But the conclusion is membership in the one event $\mathcal{B}_t(v)$, which does not depend on x. Therefore

$$\bigcup_{x \in \Omega} \{X_t^x(v) \neq X_t^\perp(v)\} \subseteq \mathcal{B}_t(v).$$

If $x \mapsto X_t^x(v)$ is not constant, some start x gives a value different from the value at the start $\perp$. So the event on the left of (36) is contained in $\mathcal{B}_t(v)$.

Counting. By Lemma 5.11 there are at most D_0^{2e} sets K for each e. A connected graph with e edges has at most $e + 1$ vertices, so each K has at most 2^{e+1} sets Σ. For a fixed set Σ of points of positive time, $\mathbb{P}(\eta = 1 \text{ on } \Sigma) = \varepsilon^{|\Sigma|} \leq \varepsilon^{(e-\tau)/a}$. The union bound gives

$$\mathbb{P}\big(\mathcal{B}_t(v)\big) \leq \sum_{e \geq t} D_0^{2e}\, 2^{e+1}\, \varepsilon^{(e-\tau)/a} = 2\varepsilon^{-\tau/a}\, \frac{\beta^t}{1-\beta},$$

as in the proof of Theorem 5.1. □

In the language of random maps, uniform replacement applies at each space-time point one random map of the neighbourhood $\mathcal{N}(v)$: the local rule of F if $\eta_t(v) = 0$, and the constant map with value $\zeta_t(v)$ if $\eta_t(v) = 1$. Corollary 5.12 bounds the probability that the composite of these maps from time 0 to time t, read at one site, is not constant. This is the *common-map coefficient* of [9, eq. (4)], and Corollary 10.8 uses it in that form.

6 Tilings give memory

This section proves Theorem 6.2, which is part (i) of Theorem B. When R tiles the plane, copy one tiling onto every plane. The copy is a fixed point of $F_R^{[k]}$. Under noise of small rate, each site stays in the copy with high probability at every time. From the all-bottom start, each site stays bottom with high probability. So the process remembers where it started, and it is not ergodic.

The proof is Toom's contour argument [47]; see also [19]. A site that has left the copy, and was not hit by noise, has two such sites one step earlier, in two different planes. So the sites that have left the copy lie inside the binary field used in the proof of Theorem 5.1. The contour lemma for that field, Lemma 5.8, which builds contours from poles in the manner of

[46, Theorem 33, Definitions 35 and 37, Lemma 36], then forces many noise events near such a site, and a Peierls count bounds the probability. Copying a rule onto parallel planes that correct one another by a Toom-type vote goes back to the repeated planes of [20]. Toom-stable clones of Wang-tile models appear in [37].

Throughout, $k \geq 2$, $q = k+1$, $a = 2q^2+1$ and $D_0 = 50q$. A site is $v = (u, P)$ with $u \in \mathbb{Z}^2$ and $P \in \mathbb{Z}^k$. A space-time point is (u, P, t) with $t \in \mathbb{Z}$. The providers are $c_0 = 0$ and $c_1, \dots, c_k$, the unit vectors of $\mathbb{Z}^k$.

Definition 6.1. Let $\theta\colon \mathbb{Z}^2 \to R$ be a tiling. For $u \in \mathbb{Z}^2$ let $\theta[u]$ be the block $w \mapsto \theta(u+w)$ on $\{-1,0,1\}^2$; it lies in T_R because θ matches on every edge. The *cloned tiling* of θ is the configuration $\xi \in \mathcal{A}_R^{\mathbb{Z}^2 \times \mathbb{Z}^k}$ given by $\xi(u,P) = (\theta[u], 0)$ for all u and P.

Theorem 6.2 (Memory)**.** *Let $k \geq 2$, let $0 < \varepsilon \leq (4D_0^2)^{-a}$, and put $\lambda = 2D_0^2\,\varepsilon^{1/q^2}$. Let R be a Wang set that tiles $\mathbb{Z}^2$ and let ξ be a cloned tiling. Let (X_t) be the noisy $F_R^{[k]}$ at rate ε started from ξ, and (Y_t) the same process started from $\perp$. Then for every site v and every time $t \geq 0$,*

$$\mathbb{P}\big(X_t(v) \neq \xi(v)\big) \leq \varepsilon\Big(1 + \frac{2\lambda}{1-\lambda}\Big) < 1.01\,\varepsilon \quad \text{and} \quad \mathbb{P}\big(Y_t(v) \neq \perp\big) \leq \varepsilon\Big(1 + \frac{2\lambda}{1-\lambda}\Big) < 1.01\,\varepsilon. \tag{37}$$

Hence the noisy $F_R^{[k]}$ has two distinct invariant measures, and it is not ergodic.

In particular Theorem 6.2 applies to every $\varepsilon \in (0, \varepsilon_k]$.

6.1 The rule and the cloned tiling

Write $\overline{C}$ for the planar checker applied in every plane, $(\overline{C}x)(u,P) = C(x(\cdot,P))(u)$. Write Q_k for the quorum step: $Q_k(y)(u,P) = s$ if s is live and $y(u, P+c_i) = s$ for at least k of the q indices i, and $Q_k(y)(u,P) = \perp$ otherwise. Thus $F_R^{[k]} = Q_k \circ \overline{C}$. Order $\mathcal{A}_R$ flatly: $\perp \preceq s$ for every symbol s, and distinct live symbols are incomparable. Order configurations site by site.

We realize the noise as follows. For every site v and time $t \geq 1$ there are an indicator $\eta_t(v)$, equal to 1 with probability ε, and a value $\zeta_t(v)$, uniform on $\mathcal{A}_R$; all of these are independent. The run from x is $X_0 = x$ and, for $t \geq 1$, $X_t(v) = \zeta_t(v)$ if $\eta_t(v) = 1$ and $X_t(v) = F_R^{[k]}(X_{t-1})(v)$ otherwise. The space-time point (v,t) is an *event* if $\eta_t(v) = 1$, whatever the value $\zeta_t(v)$. Runs from different starts with the same η and ζ form the *common-noise coupling*.

Lemma 6.3. *Let $k \geq 2$.*

(a) For every configuration y and site (u,P), at most one live symbol s satisfies the quorum condition. So Q_k and $F_R^{[k]}$ are well defined.

(b) C, $\overline{C}$ and Q_k are monotone for $\preceq$, and hence so is $F_R^{[k]}$.

(c) Under the common-noise coupling, $X_0 \preceq X_0'$ implies $X_t \preceq X_t'$ for every t. In particular the run from $\perp$ lies below every run.

Proof. (a) Two distinct symbols, each at k or more of the $k+1$ provider positions, would need $2k \leq k+1$ positions, that is $k \leq 1$.

(b) Let $x \preceq x'$ be planar configurations and let $C(x)(u) = s$ be live. Then $x(u) = s$, the symbols $x(u)$ and $x(u+d)$, $d \in D$, are live, and each $x(u+d)$ is compatible with $x(u)$. A live symbol lies below only itself, so x' agrees with x on these nine sites, and $C(x')(u) = s$. If $C(x)(u) = \perp$ there is nothing to prove. So C is monotone, and so is $\overline{C}$, plane by plane. If

$y \preceq y'$ and $Q_k(y)(u,P) = s$ is live, the k or more provider positions with $y(u,P+c_i) = s$ also carry s in y', so $Q_k(y')(u,P) = s$.

(c) Induction on t. At a non-event the claim follows from (b). At an event both runs take the same value $\zeta_t(v)$. The last sentence holds because $\perp \preceq x$ for every x. □

Lemma 6.4. *Let ξ be a cloned tiling. If a configuration x agrees with ξ at the nine sites $(u+d,P)$, $d \in \{-1,0,1\}^2$, then $(\overline{C}x)(u,P) = \xi(u,P)$. In particular $\overline{C}\xi = \xi$ and $F_R^{[k]}(\xi) = \xi$.*

Proof. The nine symbols $\xi(u+d,P) = (\theta[u+d],0)$ are live. For $d \in D$ the bits of $\xi(u+d,P)$ and $\xi(u,P)$ are both 0, and at a common cell z of $u+\{-1,0,1\}^2$ and $u+d+\{-1,0,1\}^2$ both blocks show $\theta(z)$. So every neighbour is compatible with the centre, and C keeps $\xi(u,P)$. For the last claim, $(\overline{C}\xi)(u,P+c_i) = \xi(u,P+c_i) = \xi(u,P)$ for all q indices i, and $\xi(u,P)$ is live. □

6.2 Two bad parents

Lemma 6.5 (Two bad parents)**.** *Let $k \geq 2$, and fix any realization of the events and of the replacement values.*

(a) Run from ξ, and call (u,P,t) bad *if $X_t(u,P) \neq \xi(u,P)$. No point of time 0 is bad. If (u,P,t) is bad and not an event, there are providers $c \neq c'$ and offsets $d,d' \in \{-1,0,1\}^2$ such that $(u+d,P+c,t-1)$ and $(u+d',P+c',t-1)$ are bad.*

(b) Run from $\perp$, and call (u,P,t) bad *if $Y_t(u,P)$ is live. No point of time 0 is bad. If (u,P,t) is bad and not an event, there are providers $c \neq c'$ such that $(u,P+c,t-1)$ and $(u,P+c',t-1)$ are bad.*

Proof. At time 0 the runs equal ξ and $\perp$, so no point is bad. Let $t \geq 1$.

(a) Put $x = X_{t-1}$, $y = \overline{C}x$ and $s = \xi(u,P)$, which is live. As (u,P,t) is not an event, $F_R^{[k]}(x)(u,P) = X_t(u,P) \neq s$. If $y(u,P+c_i) = s$ held for k or more indices i, the quorum step would output s. Hence $y(u,P+c_i) \neq s$ for at least two indices $i \neq i'$. Fix such an i. Since $\xi(u,P+c_i) = s$, Lemma 6.4 shows that x differs from ξ at some site $(u+d,P+c_i)$ with $d \in \{-1,0,1\}^2$; that is, $(u+d,P+c_i,t-1)$ is bad. The same holds for i', and $c_i \neq c_{i'}$.

(b) Let $s = Y_t(u,P)$, which is live. As (u,P,t) is not an event, $(\overline{C}Y_{t-1})(u,P+c_i) = s$ for at least $k \geq 2$ indices i. The checker outputs a live symbol only by keeping its centre, so $Y_{t-1}(u,P+c_i) = s$ is live for these indices. □

In both cases one bad plane among the providers is outvoted; a bad point that is not an event needs two.

6.3 Domination by the binary field

Define a field E on space-time points from the event indicators alone. Put $E_0 \equiv 0$. For $t \geq 1$, put $E_t(u,P) = 1$ if (u,P,t) is an event; otherwise put $E_t(u,P) = 1$ if and only if there are providers $c \neq c'$ and offsets $d,d' \in \{-1,0,1\}^2$ with $E_{t-1}(u+d,P+c) = E_{t-1}(u+d',P+c') = 1$. This is the binary field used in the proof of Theorem 5.1.

Lemma 6.6. *For every realization of the events and of the replacement values, every bad point of either run in Lemma 6.5 has $E = 1$.*

Proof. Induction on time. At time 0 no point is bad. Let (u,P,t) be bad with $t \geq 1$. If it is an event, $E_t(u,P) = 1$. Otherwise Lemma 6.5 gives two bad points at time $t-1$, at providers $c \neq c'$ and planar offsets in $\{-1,0,1\}^2$. By induction both have $E_{t-1} = 1$, so $E_t(u,P) = 1$. □

6.4 Counting, and the proof of the theorem

Let H_k be the graph on $\mathbb{Z}^2 \times \mathbb{Z}^k \times \mathbb{Z}$ in which each point (u, P, t) is joined to $(u+d, P+c, t-1)$ for every $d \in \{-1, 0, 1\}^2$ and every provider c. Each point has $9q$ neighbours one step lower and $9q$ one step higher, so every vertex of H_k has $18q \le D_0$ neighbours.

Lemma 6.7. *Let H be a graph in which every vertex has at most Δ neighbours, let z be a vertex, and let $n \ge 1$. At most Δ^{2n} sets I of n edges of H form a connected graph having z as an end of one of their edges.*

Proof. Number the edges at each vertex of H injectively by $\{1, \dots, \Delta\}$. Double every edge of such a set I. The resulting multigraph is connected and all its degrees are even, so it has a closed walk from z of length $2n$ that uses each of its edges once. For each step of the walk, record the number of the edge it uses at the vertex where the step starts. Starting at z, the sequence determines the walk, and the walk determines I. So distinct sets I give distinct sequences in $\{1, \dots, \Delta\}^{2n}$. □

Lemma 6.8. *Let $k \ge 2$, $0 < \varepsilon < 1$ and $\lambda = 2D_0^2 \varepsilon^{1/q^2} < 1$. For every site v and time $t \ge 0$,*

$$\mathbb{P}\big(E_t(v) = 1\big) \le \varepsilon\Big(1 + \frac{2\lambda}{1-\lambda}\Big).$$

Proof. For $t = 0$ the probability is 0; let $t \ge 1$. Fix a realization of the events. The field E has $E_0 \equiv 0$, and every point with $E = 1$ that is not an event has two points with $E = 1$ one step earlier, at distinct providers and planar offsets in $\{-1, 0, 1\}^2$; choose two such points for each of them. Suppose $E_t(v) = 1$ and (v, t) is not an event. Lemma 5.8 gives a set I of $n \ge 1$ edges, each joining a point with $E = 1$ to one of its two chosen points, such that I forms a connected graph with (v, t) as an end, and the ends of edges in I include m distinct events with $n \le q^2(m-1)$. So I is a set of edges of H_k, and its ends include a set S of at least $1 + n/q^2$ events. A connected graph with n edges has at most $n+1$ vertices, so S is one of at most 2^{n+1} subsets of the ends of I. Hence

$$\{E_t(v) = 1\} \subseteq \{(v,t) \text{ is an event}\} \cup \bigcup_{n \ge 1} \bigcup_{(I,S)} \{\text{every point of } S \text{ is an event}\},$$

where (I, S) runs over the connected n-edge sets I of H_k having (v, t) as an end and the sets S of at least $1 + n/q^2$ ends of edges of I. By Lemma 6.7 there are at most $D_0^{2n} 2^{n+1}$ such pairs. Events occur only at positive times, independently with probability ε, so each pair contributes at most $\varepsilon^{|S|} \le \varepsilon^{1+n/q^2}$. Therefore

$$\mathbb{P}\big(E_t(v) = 1\big) \le \varepsilon + \sum_{n \ge 1} D_0^{2n}\, 2^{n+1}\, \varepsilon^{1+n/q^2} = \varepsilon + 2\varepsilon \sum_{n \ge 1} \lambda^n = \varepsilon\Big(1 + \frac{2\lambda}{1-\lambda}\Big). \qquad \square$$

Proof of Theorem 6.2. Since $a/q^2 > 2$ and $4D_0^2 > 1$, the hypothesis $\varepsilon \le (4D_0^2)^{-a}$ gives $\varepsilon^{1/q^2} \le (4D_0^2)^{-2}$, so

$$\lambda \le \frac{2D_0^2}{16 D_0^4} = \frac{1}{8D_0^2} \le \frac{1}{180000},$$

because $D_0 = 50q \ge 150$. Then $2\lambda/(1-\lambda) \le 3\lambda < 0.01$. By Lemma 6.6, the sets $\{X_t(v) \ne \xi(v)\}$ and $\{Y_t(v) \ne \perp\}$ are contained in $\{E_t(v) = 1\}$, and Lemma 6.8 gives (37).

Put $\delta = 1 - 2.02\,\varepsilon$, which is positive. Fix a site v and let W be the set of configurations x with $x(v) = \xi(v)$. Since $\xi(v)$ is live, (37) gives, for every t,

$$\mathbb{P}(X_t \in W) > 1 - 1.01\,\varepsilon, \qquad \mathbb{P}(Y_t \in W) \le \mathbb{P}\big(Y_t(v) \ne \perp\big) < 1.01\,\varepsilon. \tag{38}$$

So $d_{\mathrm{TV}}\big(\mathrm{Law}_{\tilde{\zeta}}(X_t(v)), \mathrm{Law}_{\perp}(Y_t(v))\big) > \delta$ for every t: the process has memory.

By (38) the process has memory, and Lemma 4.13 gives two distinct invariant measures. □

Remark 6.9. The field E depends only on the event indicators, Lemma 6.6 holds for every realization of the replacement values, and Lemma 6.8 uses the noise only through the bound $\mathbb{P}(\text{every point of } S \text{ is an event}) \leq \varepsilon^{|S|}$ for finite sets S of points of positive time. So (37) holds for any law of the replacement values, which may depend on the events and on the past, as long as that bound holds.

7 A hierarchical scaffold read by a fixed number of tokens

We now retain the machine family as a second route. Together with the inspector theory it gives Corollary 11.21 on $\mathbb{Z}^8$, the form covered by the unchanged Lean core. The raw-input route on $\mathbb{Z}^6$ in Corollary 3.1 does not use this construction.

This section builds one fixed Wang set, the scaffold, and shows that a fixed number of tokens can read its hierarchy at every depth. Hierarchical tilings go back to Berger [3] and Robinson [43]. Mozes showed that substitution tilings with unique derivation are enforced by local rules [39]; recognizability of primitive aperiodic word substitutions was proved by Mossé [38], and unique composition of nonperiodic self-similar translationally finite tilings by Solomyak [45]. We use Ollinger's 104-tile substitution [40], in the transcription of [31], and prove the facts we need from its tables, which are printed in Section A.

The plan is as follows. Section 7.1 recalls the 104 tiles and proves that every tiling has one parent tiling. Section 7.2 adds two bits to every tile, one constant on columns and one constant on rows. The result is a Wang set $\mathbb{A}$ of 276 letters with a substitution σ of expansion 4 (Proposition 7.11). The bits record the depth of each column and row modulo 2, like a ruler. Section 7.3 reads the full letter of a depth-n block from $\ell = \binom{276}{2} + 1$ fixed sites. Section 7.4 defines anchors and builds the type, edge and point certificates, graphs of bounded tree width that force the readout (Theorem 7.18). Section 7.5 carries their conclusions to every existential pebble family (Lemma 7.19) and shows that every such family contains anchors of every depth (Proposition 7.21).

Throughout, $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$. Intervals are sets of integers: $[a,b] = \{a,\dots,b\}$ and $[0,m) = \{0,\dots,m-1\}$. Wang sets, legal patterns and tilings are as in Definition 4.1. Tiles t, t' *match in direction* $\mathbf{e}_1$ if $t_{\mathrm{E}} = t'_{\mathrm{W}}$, and *in direction* $\mathbf{e}_2$ if $t_{\mathrm{N}} = t'_{\mathrm{S}}$. A pattern on a finite set is *extendable* if it is the restriction of a tiling.

7.1 Ollinger's substitution

A tile of the base set is a triple (θ, κ, ω) of a *thin* type $\theta \in \{\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d}\}$, a *thick* type κ and a *black* type $\omega \in \{\mathsf{A},\dots,\mathsf{E}\}$. A thick type is one of four corners $\angle\mathsf{a}, \angle\mathsf{b}, \angle\mathsf{c}, \angle\mathsf{d}$ or one of sixteen lines $\langle \alpha, \gamma \rangle$, with a vertical atom $\alpha \in \{\mathsf{r}_0, \mathsf{r}_2, \mathsf{g}_0, \mathsf{g}_2\}$ and a horizontal atom $\gamma \in \{\mathsf{r}_1, \mathsf{r}_3, \mathsf{g}_1, \mathsf{g}_3\}$. Each component has an endpoint ∂_X on each side $X \in \{\mathrm{N},\mathrm{E},\mathrm{S},\mathrm{W}\}$, given in Table 4.

Definition 7.1 (The base set)**.** $\mathbb{O}$ is the Wang set of the 104 triples listed in Table 5. The colour of (θ, κ, ω) on side X is the triple $(\partial_X\theta, \partial_X\kappa, \partial_X\omega)$.

This is the 104-tile set of the Lean development [31], a transcription of Ollinger's set [40], in a compact presentation. No agreement with the published figures of [40] is claimed or used: the proofs below use only the tables of Section A. The four exclusions in Table 5 replace a list of 104 cases. Distinct triples give distinct tiles. Indeed, the four thin types have distinct north endpoints. A line $\langle \alpha, \gamma \rangle$ has equal north and south endpoints α and east endpoint γ, while each corner has distinct north and south endpoints and the four corners have distinct north endpoints. The five black types have distinct bit tuples.

Lemma 7.2 (Thin phase)**.** *Let f be a tiling of $\mathbb{O}$. There is a unique coset $c + 2\mathbb{Z}^2$ such that the thin type of $f(c + 2v + \varsigma)$ is $\mathsf{a},\mathsf{c},\mathsf{d},\mathsf{b}$ for $\varsigma = (0,0),(1,0),(0,1),(1,1)$ respectively, for every $v \in \mathbb{Z}^2$.*

Proof. The east endpoints of the thin types are distinct, and so are the west endpoints. By Table 4, the thin type of the east neighbour is therefore determined: $\mathsf{a} \leftrightarrow \mathsf{c}$ and $\mathsf{b} \leftrightarrow \mathsf{d}$. In the same way the north neighbour is determined: $\mathsf{a} \leftrightarrow \mathsf{d}$ and $\mathsf{b} \leftrightarrow \mathsf{c}$. These two involutions

commute; their composite exchanges $\mathsf{a} \leftrightarrow \mathsf{b}$ and $\mathsf{c} \leftrightarrow \mathsf{d}$. Hence the thin type at $u + (x_1, x_2)$ is obtained from the thin type at u by applying the first involution x_1 times and the second x_2 times. The sites of thin type a form one coset $c + 2\mathbb{Z}^2$, and the other three types sit as stated. □

Definition 7.3 (The substitution)**.** For a triple $t = (\theta, \kappa, \omega)$, the pattern $\varphi(t)$ on $[0,2)^2$ is given by Table 6 (left). It uses the maps F, se and nw of Section A. For a pattern w on a set $U \subseteq \mathbb{Z}^2$, the pattern $\varphi(w)$ on $2U + [0,2)^2$ is $\varphi(w)(2v + \varsigma) = \varphi(w(v))(\varsigma)$.

The definition makes sense for every triple of a thin type, a thick type and a black type, whether or not it lies in $\mathbb{O}$. We call such a triple a *formal triple*.

Lemma 7.4 (The substitution is sound)**.** *(a) For every $t \in \mathbb{O}$, $\varphi(t)$ is a legal pattern of tiles of $\mathbb{O}$.*

(b) For every pattern w of tiles of $\mathbb{O}$ on a set $U \subseteq \mathbb{Z}^2$, $\varphi(w)$ is legal if and only if w is legal.

(c) φ is injective on formal triples.

Proof. (a) Let $t = (\theta, \kappa, \omega) \in \mathbb{O}$. Child $(0,0)$ is $(\mathsf{a}, \kappa, \omega)$. It lies in $\mathbb{O}$ because every pair (κ, ω) allowed for some thin type is allowed for a (Table 5). Child $(1,1)$ is $(\mathsf{b}, F(\theta), \mathsf{A})$, a corner with black A, which is allowed. Child $(1,0)$ is the line $\langle \partial_{\mathrm{S}}\theta, \partial_{\mathrm{E}}\kappa \rangle$ with thin c and black $\mathrm{se}(\omega) \in \{\mathsf{B}, \mathsf{C}\}$. It is excluded only if $\mathrm{se}(\omega) = \mathsf{B}$ and $\partial_{\mathrm{S}}\theta = \mathsf{r}_0$, or $\mathrm{se}(\omega) = \mathsf{C}$ and $\partial_{\mathrm{S}}\theta = \mathsf{r}_2$. Now $\partial_{\mathrm{S}}\theta = \mathsf{r}_0$ only for $\theta = \mathsf{d}$, where $\omega \in \{\mathsf{D}, \mathsf{E}\}$ and $\mathrm{se}(\omega) = \mathsf{C}$; and $\partial_{\mathrm{S}}\theta = \mathsf{r}_2$ only for $\theta = \mathsf{b}$, where $\omega = \mathsf{A}$ and $\mathrm{se}(\omega) = \mathsf{B}$. Child $(0,1)$ is the line $\langle \partial_{\mathrm{N}}\kappa, \partial_{\mathrm{W}}\theta \rangle$ with thin d and black $\mathrm{nw}(\omega) \in \{\mathsf{D}, \mathsf{E}\}$. It is excluded only if $\mathrm{nw}(\omega) = \mathsf{D}$ and $\partial_{\mathrm{W}}\theta = \mathsf{r}_3$, which needs $\theta = \mathsf{b}$ and $\omega = \mathsf{A}$, where $\mathrm{nw}(\omega) = \mathsf{E}$; or $\mathrm{nw}(\omega) = \mathsf{E}$ and $\partial_{\mathrm{W}}\theta = \mathsf{r}_1$, which needs $\theta = \mathsf{c}$ and $\omega \in \{\mathsf{B}, \mathsf{C}\}$, where $\mathrm{nw}(\omega) = \mathsf{D}$. So all four children lie in $\mathbb{O}$.

The four internal edges match. The thin pattern $\mathsf{a}, \mathsf{c}, \mathsf{d}, \mathsf{b}$ matches by the thin matching rules listed in Section A. For the thick components, the edge from $(0,0)$ to $(1,0)$ carries $\partial_{\mathrm{E}}\kappa$ on both sides, the edge from $(0,0)$ to $(0,1)$ carries $\partial_{\mathrm{N}}\kappa$, the edge from $(1,0)$ to $(1,1)$ carries $\partial_{\mathrm{S}}\theta = \partial_{\mathrm{S}}F(\theta)$, and the edge from $(0,1)$ to $(1,1)$ carries $\partial_{\mathrm{W}}\theta = \partial_{\mathrm{W}}F(\theta)$. Here we use that $F(\theta)$ has the four endpoints of θ. For the black components, $\partial_{\mathrm{W}}\,\mathrm{se}(\omega) = \partial_{\mathrm{E}}\omega$, $\partial_{\mathrm{S}}\,\mathrm{nw}(\omega) = \partial_{\mathrm{N}}\omega$, and the two edges at child $(1,1)$ carry $\partial_{\mathrm{N}}\,\mathrm{se}(\omega) = 1 = \partial_{\mathrm{S}}\mathsf{A}$ and $\partial_{\mathrm{E}}\,\mathrm{nw}(\omega) = 1 = \partial_{\mathrm{W}}\mathsf{A}$.

(b) By (a) the edges inside each block $\varphi(w(v))$ match. Let $v, v + \mathbf{e}_1 \in U$ with $w(v) = (\theta, \kappa, \omega)$ and $w(v + \mathbf{e}_1) = (\theta', \kappa', \omega')$. Two fine edges cross the seam between their blocks. The lower one joins child $(1,0)$ of the first block to child $(0,0)$ of the second. Its thin endpoints are g_3 and g_3; its thick endpoints are $\partial_{\mathrm{E}}\kappa$ and $\partial_{\mathrm{W}}\kappa'$; its black bits are $\partial_{\mathrm{E}}\,\mathrm{se}(\omega) = \partial_{\mathrm{E}}\omega$ and $\partial_{\mathrm{W}}\omega'$. The upper one joins child $(1,1)$ to child $(0,1)$. Its thin endpoints are g_1 and g_1; its thick endpoints are $\partial_{\mathrm{E}}F(\theta) = \partial_{\mathrm{E}}\theta$ and $\partial_{\mathrm{W}}\theta'$; its black bits are $\partial_{\mathrm{E}}\mathsf{A} = 0$ and $\partial_{\mathrm{W}}\,\mathrm{nw}(\omega') = 0$. So both fine edges match if and only if the three components of the coarse edge match. For $v, v + \mathbf{e}_2 \in U$ the argument is the same: the left fine edge joins child $(0,1)$ to child $(0,0)$ and records $\partial_{\mathrm{N}}\kappa = \partial_{\mathrm{S}}\kappa'$ and $\partial_{\mathrm{N}}\,\mathrm{nw}(\omega) = \partial_{\mathrm{N}}\omega = \partial_{\mathrm{S}}\omega'$, with thin endpoints $\mathsf{g}_2, \mathsf{g}_2$; the right fine edge joins child $(1,1)$ to child $(1,0)$ and records $\partial_{\mathrm{N}}\theta = \partial_{\mathrm{S}}\theta'$, with thin endpoints $\mathsf{g}_0, \mathsf{g}_0$ and black bits $0, 0$.

(c) Child $(0,0)$ shows κ and ω, and child $(1,1)$ shows $F(\theta)$. The map F is a bijection from thin types to corners. □

The proof of (b) compares components, so no list of the 104^2 pairs is needed.

Lemma 7.5 (Parents)**.** *Let w be a legal pattern of tiles of $\mathbb{O}$ on $[0,2)^2$ whose thin types are $\mathsf{a}, \mathsf{c}, \mathsf{d}, \mathsf{b}$ at $(0,0), (1,0), (0,1), (1,1)$. Then $w = \varphi(t)$ for exactly one formal triple t. Exactly 128 formal*

triples arise in this way: the 104 *tiles of* $\mathbb{O}$ *and* 24 *others (Table 7). If w extends to a legal pattern on $[0,2)^2 \cup \{(0,-1),(1,-1),(-1,0),(-1,1)\}$, then $t \in \mathbb{O}$.*

Proof. The tile $w(1,1)$ has thin type b, so it is a corner with black A; let $\theta = F^{-1}$ of that corner. Let κ and ω be the thick and black types of $w(0,0)$, and $t = (\theta, \kappa, \omega)$. The tile $w(1,0)$ has thin type c, so it is a line with black B or C. Its west and north thick endpoints must equal $\partial_E \kappa$ and $\partial_S F(\theta) = \partial_S \theta$; so it is the line $\langle \partial_S \theta, \partial_E \kappa \rangle$. Its black type has west bit $\partial_E \omega$, and the west bits of B and C are 0 and 1; so it is $\mathrm{se}(\omega)$. In the same way $w(0,1)$ is $(\mathsf{d}, \langle \partial_N \kappa, \partial_W \theta \rangle, \mathrm{nw}(\omega))$, because the south bits of D and E are 1 and 0. Hence $w = \varphi(t)$, and t is unique by Lemma 7.4(c).

Which t occur? Child $(0,0)$ must lie in $\mathbb{O}$, which leaves exactly the 52 pairs (κ, ω) of thin type a in Table 5. Child $(1,1)$ is always allowed. Child $(1,0)$ is allowed unless $\mathrm{se}(\omega) = \mathsf{B}$ and $\partial_S \theta = \mathsf{r}_0$, or $\mathrm{se}(\omega) = \mathsf{C}$ and $\partial_S \theta = \mathsf{r}_2$. For $\theta = \mathsf{b}$ ($\partial_S \theta = \mathsf{r}_2$) this excludes $\omega \in \{\mathsf{C}, \mathsf{D}, \mathsf{E}\}$; for $\theta = \mathsf{d}$ ($\partial_S \theta = \mathsf{r}_0$) it excludes $\omega \in \{\mathsf{A}, \mathsf{B}\}$. Child $(0,1)$ is allowed unless $\mathrm{nw}(\omega) = \mathsf{D}$ and $\partial_W \theta = \mathsf{r}_3$, or $\mathrm{nw}(\omega) = \mathsf{E}$ and $\partial_W \theta = \mathsf{r}_1$. For $\theta = \mathsf{b}$ this excludes $\omega \in \{\mathsf{B}, \mathsf{C}, \mathsf{D}\}$; for $\theta = \mathsf{c}$ it excludes $\omega \in \{\mathsf{A}, \mathsf{E}\}$. For $\theta = \mathsf{a}$ the new atoms g_2 and g_3 are green and exclude nothing. This gives the second column of Table 7. Counting the allowed pairs (κ, ω) in each case gives $52 + 4 + 36 + 36 = 128$ formal triples. Comparing with Table 5, the triples outside $\mathbb{O}$ are of two kinds: $\theta = \mathsf{c}$, $\omega = \mathsf{D}$ and κ a line with $\gamma \neq \mathsf{r}_3$; and $\theta = \mathsf{d}$, $\omega = \mathsf{C}$ and κ a line with $\alpha \neq \mathsf{r}_2$. There are 12 of each.

Let t be of the first kind, $t = (\mathsf{c}, \langle \alpha, \gamma \rangle, \mathsf{D})$ with $\gamma \neq \mathsf{r}_3$, and suppose w extends below. Then $w(1,0)$ is the line $\langle \mathsf{g}_0, \gamma \rangle$. By the thin matching rules (Section A), the tile at $(1,-1)$ has thin type b, so it is a corner; its north endpoint is g_0, so it is $\angle\mathsf{d}$, with west endpoint r_3. The tile at $(0,-1)$ has thin type d, so it is a line with black D or E. Its north bit equals the south bit of D, which is 1, so its black is D. Its east endpoint equals the west endpoint r_3 of $\angle\mathsf{d}$, so its horizontal atom is r_3. But thin d with black D forbids $\gamma = \mathsf{r}_3$. So the extension does not exist.

Let t be of the second kind, $t = (\mathsf{d}, \langle \alpha, \gamma \rangle, \mathsf{C})$ with $\alpha \neq \mathsf{r}_2$, and suppose w extends to the left. Then $w(0,1)$ is the line $\langle \alpha, \mathsf{g}_1 \rangle$. The tile at $(-1,1)$ has thin type b, so it is a corner; its east endpoint is g_1, so it is $\angle\mathsf{d}$, with south endpoint r_2. The tile at $(-1,0)$ has thin type c, so it is a line with black B or C. Its east bit equals the west bit of C, which is 1, so its black is C. Its north endpoint equals r_2, so its vertical atom is r_2. But thin c with black C forbids $\alpha = \mathsf{r}_2$. So this extension does not exist either. □

Proposition 7.6 (Every tiling has a parent)**.** *$\mathbb{O}$ tiles. For every tiling f of $\mathbb{O}$ there are a unique coset $c + 2\mathbb{Z}^2$ and, for each representative c, a unique tiling z of $\mathbb{O}$ with $f(c + 2v + \zeta) = \varphi(z(v))(\zeta)$ for all $v \in \mathbb{Z}^2$ and $\zeta \in [0,2)^2$.*

Proof. Let $c + 2\mathbb{Z}^2$ be the coset of Lemma 7.2; no other coset has thin type a at all its sites, so the coset is unique. Fix a representative c. For each v, the block $f|_{c+2v+[0,2)^2}$ has the thin pattern of Lemma 7.5, and the tiling supplies the four cells below and to the left. So it equals $\varphi(z(v))$ for a unique $z(v) \in \mathbb{O}$. By Lemma 7.4(b), z is legal, since $\varphi(z)$ is a translate of f. For existence, $\varphi^m(t)$ is a legal pattern on $[0, 2^m)^2$ for every $t \in \mathbb{O}$ and $m \geq 0$, by Lemma 7.4(a),(b). So every square carries a legal pattern, and $\mathbb{O}$ tiles by König's lemma, as in the proof of Lemma 4.2. □

The proposition holds for every tiling. No assumption is made that f lies in the closure of the substituted patterns, so tilings with faults along lines or quadrants are covered.

7.2 The ruler decoration

We add a column bit and a row bit to every tile. A *red crossing* is a line $\langle \alpha, \gamma \rangle$ with both atoms red, that is $\alpha \in \{\mathsf{r}_0, \mathsf{r}_2\}$ and $\gamma \in \{\mathsf{r}_1, \mathsf{r}_3\}$.

Definition 7.7 (The scaffold)**.** The scaffold $\mathbb{A}$ is the set of tuples $\mathfrak{a} = (\theta, \kappa, \omega, b_1, b_2)$ with $(\theta, \kappa, \omega) \in \mathbb{O}$ and $b_1, b_2 \in \{0, 1\}$ that satisfy:

(R1) if $\theta \in \{\mathsf{b}, \mathsf{c}\}$ then $b_1 = 0$;

(R2) if $\theta \in \{\mathsf{b}, \mathsf{d}\}$ then $b_2 = 0$;

(R3) if κ is a corner then $b_1 = b_2$;

(R4) if κ is a red crossing then $b_1 \neq b_2$.

The *column bit* is $\chi_1(\mathfrak{a}) = b_1$ and the *row bit* is $\chi_2(\mathfrak{a}) = b_2$. The north and south colours of $\mathfrak{a}$ are the colours of (θ, κ, ω) in $\mathbb{O}$ together with b_1; the east and west colours are those colours together with b_2. The *base* of $\mathfrak{a}$ is $\bar{\mathfrak{a}} = (\theta, \kappa, \omega)$.

So $\mathbb{A}$ is a Wang set. Table 8 counts it: $184 + 4 + 44 + 44 = 276$ letters. By construction, two letters that match in direction $\mathbf{e}_2$ have the same column bit, and two letters that match in direction $\mathbf{e}_1$ have the same row bit. We call this the *edgewise bit property*. It holds for every matching pair, not only for pairs that occur in tilings, and the certificates of Section 7.4 use it in that form.

Definition 7.8 (Decorated substitution)**.** For $\mathfrak{a} \in \mathbb{A}$ with bits (b_1, b_2), the pattern $\varphi'(\mathfrak{a})$ on $[0, 2)^2$ has base $\varphi(\bar{\mathfrak{a}})$ and the bits of Table 6 (right). Patterns $\varphi'(w)$ are defined blockwise as in Definition 7.3. Put $\sigma = \varphi' \circ \varphi'$, that is, $\sigma(\mathfrak{a})(2\tau + \varsigma) = \varphi'\big(\varphi'(\mathfrak{a})(\tau)\big)(\varsigma)$ for $\tau, \varsigma \in [0, 2)^2$. Then $\sigma(\mathfrak{a})$ is a pattern on $[0, 4)^2$, and $\sigma^n(\mathfrak{a})$ is a pattern on $[0, 4^n)^2$ with $\sigma^{n+1}(\mathfrak{a})(4^n d + y) = \sigma^n\big(\sigma(\mathfrak{a})(d)\big)(y)$ for $d \in [0, 4)^2$ and $y \in [0, 4^n)^2$.

Lemma 7.9 (Decorated substitution is sound)**.** *(a) For every $\mathfrak{a} \in \mathbb{A}$, $\varphi'(\mathfrak{a})$ is a legal pattern of letters of $\mathbb{A}$.*

(b) For every pattern w of $\mathbb{A}$ on a set $U \subseteq \mathbb{Z}^2$, $\varphi'(w)$ is legal if and only if w is legal.

(c) φ' is injective.

Proof. (a) By Lemma 7.4(a) the four bases lie in $\mathbb{O}$ and match. Let $\mathfrak{a} = (\theta, \kappa, \omega, b_1, b_2)$. Child $(0, 0)$ has thin type a, so (R1) and (R2) are void; it has thick type κ and bits $(1 - b_1, 1 - b_2)$, and (R3), (R4) are unchanged when both bits are complemented. Child $(1, 1)$ has thin type b, a corner and bits $(0, 0)$, which satisfies (R1)–(R3). Child $(1, 0)$ has thin type c and $b_1 = 0$, which gives (R1). Its thick type $\langle \partial_{\mathrm{S}}\theta, \partial_{\mathrm{E}}\kappa \rangle$ is a red crossing only if $\partial_{\mathrm{S}}\theta$ is red, that is $\theta \in \{\mathsf{b}, \mathsf{d}\}$. Then (R2) for $\mathfrak{a}$ gives $b_2 = 0$, so the child has bits $(0, 1)$ and satisfies (R4). Child $(0, 1)$ is symmetric: thin type d and row bit 0; its thick type $\langle \partial_{\mathrm{N}}\kappa, \partial_{\mathrm{W}}\theta \rangle$ is a red crossing only if $\theta \in \{\mathsf{b}, \mathsf{c}\}$, and then $b_1 = 0$ and the child has bits $(1, 0)$. Inside the block, the vertical edges carry column bits $1 - b_1, 1 - b_1$ and $0, 0$, and the horizontal edges carry row bits $1 - b_2, 1 - b_2$ and $0, 0$.

(b) The bases match or fail as in Lemma 7.4(b). Across a vertical seam between the blocks of $\mathfrak{a}$ (left) and $\mathfrak{a}'$ (right), the lower fine edge carries row bits $1 - b_2$ and $1 - b_2'$, and the upper one carries 0 and 0. Across a horizontal seam, the left fine edge carries column bits $1 - b_1$ and $1 - b_1'$, and the right one carries 0 and 0. So the fine bits match if and only if the coarse bits match.

(c) Child $(0, 0)$ shows κ, ω and the complements of both bits, and child $(1, 1)$ shows θ. □

Lemma 7.10 (Every decorated tiling has a parent)**.** *For every tiling f of $\mathbb{A}$ there are a unique coset $c + 2\mathbb{Z}^2$ and, for each representative c, a unique tiling z of $\mathbb{A}$ with $f(c + 2v + \varsigma) = \varphi'(z(v))(\varsigma)$ for all $v \in \mathbb{Z}^2$ and $\varsigma \in [0,2)^2$.*

Proof. The bases of f form a tiling $\bar{f}$ of $\mathbb{O}$. Let c and $\bar{z}$ be given by Proposition 7.6 for $\bar{f}$. By the edgewise bit property, $\chi_1(f(x))$ depends only on the column x_1 and $\chi_2(f(x))$ only on the row x_2; write them $X(x_1)$ and $Y(x_2)$. A column $c_1 + 2v_1 + 1$ has thin types c and b, so $X = 0$ there by (R1). A row $c_2 + 2v_2 + 1$ has thin types d and b, so $Y = 0$ there by (R2). Define

$$z(v) = \big(\bar{z}(v),\ 1 - X(c_1 + 2v_1),\ 1 - Y(c_2 + 2v_2)\big).$$

Comparing with Table 6, $f(c + 2v + \varsigma) = \varphi'(z(v))(\varsigma)$ for all v and ς. The bits of z are constant on columns and on rows and the bases form the tiling $\bar{z}$, so z is legal once each $z(v)$ lies in $\mathbb{A}$.

Conditions (R3) and (R4) for $z(v)$ follow from the same conditions for $f(c + 2v)$, which has the thick type of $\bar{z}(v)$ and both bits complemented. For (R1) and (R2) we go down two levels. Apply Proposition 7.6 to $\bar{z}$: there are c' and a tiling $\bar{w}$ of $\mathbb{O}$ with $\bar{z}(c' + 2m + \tau) = \varphi(\bar{w}(m))(\tau)$. Put $o = c + 2c'$. Then $\bar{f}(o + 4m + y) = \varphi^2(\bar{w}(m))(y)$ for $y \in [0,4)^2$. In this depth-two block, Table 9 shows that the cells $(2,1)$ and $(2,3)$ have thin type d and that exactly one of them is a red crossing, whatever the thin type of $\bar{w}(m)$. The row bits of both cells are 0 by (R2). By (R4) the red crossing has column bit 1, so $X(o_1 + 4m_1 + 2) = 1$. In the same way the cells $(1,2)$ and $(3,2)$ have thin type c, hence column bit 0, and one of them is a red crossing; so $Y(o_2 + 4m_2 + 2) = 1$.

Now let $v = c' + 2m + \tau$. The letter $z(v)$ has thin type b or c exactly when $\tau_1 = 1$. Then $c_1 + 2v_1 = o_1 + 4m_1 + 2$, so $\chi_1(z(v)) = 1 - X(o_1 + 4m_1 + 2) = 0$, which is (R1). In the same way, thin type b or d means $\tau_2 = 1$, and then $\chi_2(z(v)) = 0$, which is (R2).

Uniqueness: the coset is the thin-a coset of $\bar{f}$, the bases of z are unique by Proposition 7.6, and the bits of z are read off from f at child $(0,0)$. □

One level does not suffice for (R1) and (R2). The witnesses for a coarse column of thin type b or c sit in the depth-two block, and the offset $o = c + 2c'$ combines both levels.

Proposition 7.11 (The scaffold)**.** *The scaffold $\mathbb{A}$ is a Wang set of 276 letters, and it tiles. Its substitution $\sigma = (\varphi')^2$ has expansion 4 and the following properties.*

(i) σ is injective, and $\sigma^n(\mathfrak{a})$ is a legal pattern on $[0,4^n)^2$ for every $\mathfrak{a} \in \mathbb{A}$ and $n \geq 0$.

(ii) For every pattern w of $\mathbb{A}$ on a set $U \subseteq \mathbb{Z}^2$, $\sigma(w)$ is legal if and only if w is legal.

(iii) For every tiling f of $\mathbb{A}$ there is a unique class $\pi(f) \in (\mathbb{Z}/4\mathbb{Z})^2$, the phase *of f, such that for every representative p of $\pi(f)$ there is a unique tiling y of $\mathbb{A}$ with $f(p + 4v + i) = \sigma(y(v))(i)$ for all $v \in \mathbb{Z}^2$ and $i \in [0,4)^2$.*

(iv) For every $\mathfrak{a} \in \mathbb{A}$ and $i = (i_1, i_2) \in [0,4)^2$,

$$\chi_1\big(\sigma(\mathfrak{a})(i)\big) = \begin{cases} \chi_1(\mathfrak{a}) & i_1 = 0, \\ 0 & i_1 \in \{1,3\}, \\ 1 & i_1 = 2, \end{cases} \qquad \chi_2\big(\sigma(\mathfrak{a})(i)\big) = \begin{cases} \chi_2(\mathfrak{a}) & i_2 = 0, \\ 0 & i_2 \in \{1,3\}, \\ 1 & i_2 = 2. \end{cases}$$

(v) Letters that match in direction $\mathbf{e}_2$ have equal column bits, and letters that match in direction $\mathbf{e}_1$ have equal row bits.

Proof. (i) and (ii) follow from Lemma 7.9 applied twice, and (v) is the edgewise bit property. By (i), every square carries a legal pattern, so $\mathbb{A}$ tiles by König's lemma.

(iii) Let f be a tiling. Lemma 7.10 gives $f(c+2v+\varsigma) = \varphi'(z(v))(\varsigma)$, and again $z(c'+2m+\tau) = \varphi'(y(m))(\tau)$. Substituting, $f(c+2c'+4m+2\tau+\varsigma) = \sigma(y(m))(2\tau+\varsigma)$, so the class of $p = c+2c'$ works; for another representative $p+4m_0$ use $y(\cdot+m_0)$. Conversely, suppose $f(p+4m+i) = \sigma(y(m))(i)$ for a tiling y. Then $f(p+2v+\varsigma) = \varphi'(z'(v))(\varsigma)$ with $z'(2m+\tau) = \varphi'(y(m))(\tau)$, and z' is a tiling by Lemma 7.9. By the uniqueness in Lemma 7.10, $p \equiv c \pmod 2$. Write $p = c+2d$; uniqueness gives $z' = z(\cdot+d)$. So $z(d+2m+\tau) = \varphi'(y(m))(\tau)$, and uniqueness for z gives $d \equiv c' \pmod 2$. Hence $p \equiv c+2c' \pmod 4$. For a fixed representative, y is unique because σ is injective.

(iv) Write $i_1 = 2\tau_1+\varsigma_1$. By Table 6, the column bit of $\varphi'(\mathfrak{b})(\varsigma)$ is $1-\chi_1(\mathfrak{b})$ if $\varsigma_1 = 0$ and 0 if $\varsigma_1 = 1$. Applying this twice: for $i_1 = 0$ we get $1-(1-\chi_1(\mathfrak{a})) = \chi_1(\mathfrak{a})$; for $i_1 \in \{1,3\}$ we get 0; for $i_1 = 2$ the middle letter has column bit 0, so we get 1. The row bit is the same with τ_2, ς_2. □

Iterating (iv), the letter $\sigma^n(\mathfrak{a})(0,0)$ has the bits of $\mathfrak{a}$ for every n. In a tiling of phase π, the column bit of a column x_1 is 0 if $x_1 - p_1 \equiv 1,3 \pmod 4$ and 1 if $x_1 - p_1 \equiv 2 \pmod 4$, for any representative p of π. Only the columns $x_1 \equiv p_1 \pmod 4$ inherit a bit from the parent. So the column bits form a ruler sequence, like the parity of the 2-adic valuation of a column index. The rows behave the same way.

7.3 Reading a block from a fixed number of sites

A depth-n block $\sigma^n(\mathfrak{a})$ has 16^n cells, but its root letter can be read from at most ℓ of them, at every depth.

Let $\mathcal{P}$ be the set of unordered pairs of distinct letters of $\mathbb{A}$; it has $\binom{276}{2} = 37950$ elements. For each $p = \{\mathfrak{a},\mathfrak{b}\} \in \mathcal{P}$ fix once and for all a child $\delta(p) \in [0,4)^2$ with $\sigma(\mathfrak{a})(\delta(p)) \neq \sigma(\mathfrak{b})(\delta(p))$. It exists because σ is injective. Put $\nu(p) = \{\sigma(\mathfrak{a})(\delta(p)), \sigma(\mathfrak{b})(\delta(p))\}$, again an element of $\mathcal{P}$. Define offsets $y_n(p) \in [0,4^n)^2$ by

$$y_0(p) = (0,0), \qquad y_{n+1}(p) = 4^n\delta(p) + y_n(\nu(p)),$$

and put

$$P_n = \{(0,0)\} \cup \{y_n(p) : p \in \mathcal{P}\} \subseteq [0,4^n)^2, \qquad \mathrm{Port}_n(u) = u + P_n, \qquad \ell = \tbinom{276}{2} + 1 = 37951.$$

So $|\mathrm{Port}_n(u)| \leq \ell$; coincident offsets are counted once. The *readout* of $\mathfrak{a}$ at depth n is the restriction of $\sigma^n(\mathfrak{a})$ to P_n. A map g into $\mathbb{A}$ defined on $\mathrm{Port}_n(v)$ *reads* $\mathfrak{a}$ if $g(v+y) = \sigma^n(\mathfrak{a})(y)$ for every $y \in P_n$.

Lemma 7.12 (Pair ports)**.** *(a) For every $n \geq 0$, the readout at depth n is injective on $\mathbb{A}$.*

(b) Every $y \in P_{n+1}$ has the form $y = 4^n d + y'$ with $d \in [0,4)^2$ and $y' \in P_n$, and then $\sigma^{n+1}(\mathfrak{a})(y) = \sigma^n\big(\sigma(\mathfrak{a})(d)\big)(y')$ for every $\mathfrak{a}$.

(c) $(0,0) \in P_n$, and $\sigma^n(\mathfrak{a})(0,0)$ has the column bit and the row bit of $\mathfrak{a}$.

Proof. (a) We show $\sigma^n(\mathfrak{a})(y_n(p)) \neq \sigma^n(\mathfrak{b})(y_n(p))$ for every $p = \{\mathfrak{a},\mathfrak{b}\}$ and every n. For $n = 0$ the two values are $\mathfrak{a}$ and $\mathfrak{b}$. For $n+1$, Definition 7.8 gives $\sigma^{n+1}(\mathfrak{a})(y_{n+1}(p)) = \sigma^n(\mathfrak{a}')(y_n(\nu(p)))$ with $\mathfrak{a}' = \sigma(\mathfrak{a})(\delta(p))$, and the same for $\mathfrak{b}$ with $\mathfrak{b}' = \sigma(\mathfrak{b})(\delta(p))$. Since $\nu(p) = \{\mathfrak{a}',\mathfrak{b}'\}$, the induction hypothesis applies. So the offset $y_n(\{\mathfrak{a},\mathfrak{b}\})$ separates any two distinct letters.

(b) The offset $(0,0)$ is $4^n \cdot (0,0) + (0,0)$, and $y_{n+1}(p) = 4^n \delta(p) + y_n(\nu(p))$ with $y_n(\nu(p)) \in P_n$. The identity is Definition 7.8.

(c) This is Proposition 7.11(iv) at $i = (0,0)$, applied n times. □

Part (b) says that each port of a depth-$(n+1)$ block is a port of one of its 16 children. So the readouts of all depths are nested and their number of sites does not grow with n. The same construction works for any injective substitution on m letters, with $\binom{m}{2} + 1$ sites.

7.4 Anchors and certificates

For $u \in \mathbb{Z}^2$ and $n \geq 0$ put

$$\mathrm{Anc}_n(u) = \{u,\ u + 4^n \mathbf{e}_1,\ u + 4^n \mathbf{e}_2\}.$$

Definition 7.13 (Anchors)**.** A map g into $\mathbb{A}$ whose domain contains $\mathrm{Anc}_n(u)$ is a *good n-anchor at u* if $\chi_1(g(u)) \neq \chi_1(g(u + 4^n \mathbf{e}_1))$ and $\chi_2(g(u)) \neq \chi_2(g(u + 4^n \mathbf{e}_2))$.

The condition involves three sites and their labels only. It does not assume that the labels come from a block of depth n; the certificates below derive this.

Definition 7.14 (Certificates)**.** A *certificate* Γ consists of a finite vertex set V, a map $\rho \colon V \to \mathbb{Z}^2$ (its *grid map*), a set of *edges*, and a finite set $\partial\Gamma \subseteq \mathbb{Z}^2$ of sites, its *boundary*. Each edge is an ordered pair (v, v') of vertices with a label $e \in \{\mathbf{e}_1, \mathbf{e}_2\}$ such that $\rho(v') = \rho(v) + e$. Each boundary site x has exactly one *boundary vertex* v_x, with $\rho(v_x) = x$. A *homomorphism* of Γ is a map $\psi \colon V \to \mathbb{A}$ such that $\psi(v)$ and $\psi(v')$ match in direction e for every edge (v, v') with label e. For a boundary site x we write $\psi(x)$ for $\psi(v_x)$.

A *rooted tree decomposition* of Γ is a finite rooted tree with a *bag* $\beta_t \subseteq V$ at each node t such that every vertex lies in some bag, the nodes whose bags contain a given vertex form a connected subtree, both ends of every edge lie in a common bag, and the root bag contains every boundary vertex. Its *width* is the largest bag size.

Distinct vertices may have the same image under ρ, and a homomorphism may give them different letters. Only the edges and the boundary tie vertices together. This freedom is what keeps the width bounded: a certificate may use the same site in many independent sub-certificates without identifying the copies. Certificates of bounded tree width are the structures for which the existential pebble game is decisive, by the work of Dalmau, Kolaitis and Vardi [8]; Lemma 7.19 below is the form we need.

A finite recognition chart. The next lemma replaces an explicit table of local phases by compactness. We need only that the constants exist; they are never computed.

Lemma 7.15 (Recognition chart)**.** *(a) For every finite set $D \subseteq \mathbb{Z}^2$ there is $L(D) \geq 0$ such that, for every $u \in \mathbb{Z}^2$, every legal pattern of $\mathbb{A}$ on $u + D + [-L(D), L(D)]^2$ restricts to an extendable pattern on $u + D$.*

(b) There is $L_0 \geq 8$ such that any two tilings of $\mathbb{A}$ that agree on $[-L_0, L_0]^2$ have the same phase.

Proof. (a) Legality and extendability are invariant under translation, so take $u = 0$. Let w be a pattern on D that is not extendable. Suppose that for every L it extends to a legal pattern on $D + [-L, L]^2$. These extensions form a finitely branching tree under restriction, so König's lemma gives a legal pattern on $\mathbb{Z}^2$ that extends w, a contradiction. So w has a radius L_w

beyond which it has no legal extension. There are finitely many patterns on D; let $L(D)$ be the largest L_w.

(b) For a class $\pi \in (\mathbb{Z}/4\mathbb{Z})^2$, the set of tilings of phase π is closed in $\mathbb{A}^{\mathbb{Z}^2}$. Indeed, let $f_k \to f$ be tilings of phase π and fix a representative p. Write $f_k(p + 4v + i) = \sigma(y_k(v))(i)$. Since σ is injective, $y_k(v)$ is determined by f_k on the block $p + 4v + [0,4)^2$, so y_k converges to a map y with $f(p + 4v + i) = \sigma(y(v))(i)$. The pattern $\sigma(y)$ is a translate of f, hence legal, so y is a tiling by Proposition 7.11(ii). Now suppose (b) fails. Then for every L there are tilings f_L, f'_L that agree on $[-L, L]^2$ and have different phases. Passing to a subsequence, the phases are constant classes $\pi \neq \pi'$, and $f_L \to f$, $f'_L \to f'$ by compactness of $\mathbb{A}^{\mathbb{Z}^2}$. The limits agree on every box, so $f = f'$, and by closedness f has both phases π and π'. This contradicts Proposition 7.11(iii). Enlarging L_0 keeps (b) true, so we may take $L_0 \geq 8$. □

The phase of a translate changes by the translation: if $f(p + 4v + i) = \sigma(y(v))(i)$ then $f(\cdot + u)$ has phase $\pi(f) - u$. So by (b), two tilings that agree on $u + [-L_0, L_0]^2$ have the same phase, for every u.

Definition 7.16 (Constants)**.** Fix L_0 as in Lemma 7.15(b). Put $\mathsf{D}_0 = [-L_0, L_0]^2$, $Q = [-L_0 - 1, L_0 + 1]^2$, $L_1 = L(Q)$ and $\varrho = L_0 + 1 + L_1$. Put $L_2 = L([0,5)^2)$ and $L_3 = L([0,32)^2)$. Define

$$B = \max\Big\{\ell(2\varrho + 3)^2 + 2\ell + 26,\ \ (2L_2 + 5)^2,\ \ 3 + \ell\,(2L_3 + 32)^2\Big\}.$$

The integer B depends only on $\mathbb{A}$, σ and the fixed choice of δ. It does not depend on any depth, site, distance or machine. Its first term bounds the bags of the certificates; the other two are the sizes of the fixed boxes used in Proposition 7.21. In particular $B \geq \ell + 4$. We only use that B exists.

Corridors. A *corridor walk* is a sequence $w_0, \dots, w_m \in \mathbb{Z}^2$ in which each step $w_{k+1} - w_k$ is one of $\pm\mathbf{e}_1, \pm\mathbf{e}_2$ and each coordinate of w_k is monotone in k. Its *corridor* is $S = \bigcup_k \big(w_k + [-\varrho, \varrho]^2\big)$.

Lemma 7.17 (Phase transport)**.** *Let $w_0, \dots, w_m$ be a corridor walk with corridor S, and let $\mathfrak{c}$ be a legal pattern of $\mathbb{A}$ on S. Suppose that $\chi_1(\mathfrak{c}(w_0)) \neq \chi_1(\mathfrak{c}(w_0 + 4\mathbf{e}_1))$ and $\chi_2(\mathfrak{c}(w_0)) \neq \chi_2(\mathfrak{c}(w_0 + 4\mathbf{e}_2))$. Let $v \equiv w_0 \pmod 4$ with $v + [0,8)^2 \subseteq w_k + \mathsf{D}_0$ for some k. Then there are letters $\mathfrak{a}, \mathfrak{b}_1, \mathfrak{b}_2$ such that, for $i \in [0,4)^2$ and $e \in \{\mathbf{e}_1, \mathbf{e}_2\}$,*

$$\mathfrak{c}(v + i) = \sigma(\mathfrak{a})(i), \qquad \mathfrak{c}(v + 4e + i) = \sigma(\mathfrak{b}_e)(i),$$

and $\mathfrak{a}$ and $\mathfrak{b}_e$ match in direction e.

Proof. For every k the pattern $\mathfrak{c}$ is legal on $w_k + Q + [-L_1, L_1]^2 = w_k + [-\varrho, \varrho]^2$, so by Lemma 7.15(a) its restriction to $w_k + Q$ is extendable. Let g_k be a tiling that agrees with $\mathfrak{c}$ on $w_k + Q$, and let π_k be its phase. By the remark after Lemma 7.15, π_k depends only on $\mathfrak{c}|_{w_k + \mathsf{D}_0}$. Since $w_{k+1} + \mathsf{D}_0 \subseteq w_k + Q$, the tiling g_k also agrees with $\mathfrak{c}$ on $w_{k+1} + \mathsf{D}_0$; hence $\pi_{k+1} = \pi_k$. So all π_k equal π_0.

The sites w_0, $w_0 + 4\mathbf{e}_1$ and $w_0 + 4\mathbf{e}_2$ lie in $w_0 + \mathsf{D}_0$ because $L_0 \geq 8$. Let p represent π_0. If $w_{0,1} \not\equiv p_1 \pmod 4$, then the columns $w_{0,1}$ and $w_{0,1} + 4$ have the same nonzero residue modulo 4 relative to p_1, so by Proposition 7.11(iii),(iv) they have the same column bit in g_0; this contradicts the hypothesis. So $w_{0,1} \equiv p_1$, and in the same way $w_{0,2} \equiv p_2$, using rows and row bits. Thus π_k is the class of w_0 for every k, and also the class of v.

Take k with $v + [0,8)^2 \subseteq w_k + \mathsf{D}_0$ and use the representative v: by Proposition 7.11(iii) there is a tiling y with $g_k(v + 4m + i) = \sigma(y(m))(i)$. Put $\mathfrak{a} = y(0)$ and $\mathfrak{b}_e = y(e)$. The cells $v + i$ and $v + 4e + i$ lie in $v + [0,8)^2$, where g_k agrees with $\mathfrak{c}$. Since y is a tiling, $\mathfrak{a}$ and $\mathfrak{b}_e$ match in direction e. □

The walk carries the phase step by step through overlapping windows. This is the only way the phase fixed at the anchor reaches a distant block, and every certificate below contains such a walk.

The parser. For $n \geq 0$ and $u, z \in \mathbb{Z}^2$, the *block* of z at depth n relative to u is $u + 4^n z + [0, 4^n)^2$, and its ports are $\mathrm{Port}_n(u + 4^n z)$.

Theorem 7.18 (Parser)**.** *Let B be as in Definition 7.16. For all $n \geq 0$, $u, z \in \mathbb{Z}^2$, $e \in \{\mathbf{e}_1, \mathbf{e}_2\}$ and $s \in u + 4^n z + [0, 4^n)^2$ there are three certificates:*

- *the* type certificate $\Gamma_n^{\mathrm{ty}}(u, z)$*, with boundary* $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u + 4^n z)$*;*
- *the* edge certificate $\Gamma_n^{\mathrm{ed}}(u, z, e)$*, with boundary* $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u + 4^n z) \cup \mathrm{Port}_n(u + 4^n(z + e))$*;*
- *the* point certificate $\Gamma_n^{\mathrm{pt}}(u, z, s)$*, with boundary* $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u + 4^n z) \cup \{s\}$.

Each has a rooted tree decomposition of width at most B. Let ψ be a homomorphism of one of them whose values on $\mathrm{Anc}_n(u)$ form a good n-anchor at u. Then:

(T) there is $\mathfrak{a} \in \mathbb{A}$ such that ψ reads $\mathfrak{a}$ on $\mathrm{Port}_n(u + 4^n z)$;

(E) for the edge certificate, moreover ψ reads some $\mathfrak{b}$ on $\mathrm{Port}_n(u + 4^n(z + e))$, and $\mathfrak{a}$ and $\mathfrak{b}$ match in direction e;

(P) for the point certificate, moreover $\psi(s) = \sigma^n(\mathfrak{a})(s - u - 4^n z)$.

The bound B does not depend on n, u, z, e, s, on the distance from the anchor to the block, or on any machine.

By Lemma 7.12(a) the letters $\mathfrak{a}$ and $\mathfrak{b}$ are unique. The query block may be arbitrarily far from the anchor.

Proof. We build the three certificates together, by induction on n, and check the decomposition and the conclusions at each step.

Depth 0. Here $P_0 = \{(0, 0)\}$, so the ports of z form the single site $u + z$, and $s = u + z$. The type and point certificates have one vertex for each boundary site and no edges. The edge certificate adds one edge from $u + z$ to $u + z + e$ with label e. One bag holds everything; it has at most 5 vertices. Conclusions (T) and (P) hold with $\mathfrak{a} = \psi(u + z)$, and (E) holds because of the edge. No anchor condition is needed at depth 0.

From depth n to $n + 1$: the skeleton. Fix u, z, and e or s as required. Coarse sites $x \in \mathbb{Z}^2$ stand for the depth-n blocks $u + 4^n x + [0, 4^n)^2$; these blocks tile the plane. The depth-$(n + 1)$ block of z is the union of the depth-n blocks of $4z + d$, $d \in [0, 4)^2$. Choose a corridor walk from $w_0 = (0, 0)$ to $w_m = 4z$, say first along $\mathbf{e}_1$ and then along $\mathbf{e}_2$, and let S be its corridor. The *skeleton* has one vertex for each site of

$$\bigcup_{x \in S} \mathrm{Port}_n(u + 4^n x),$$

and, for the point certificate, one more vertex for s if s is not already among these sites. Its grid map is the identity on sites. The skeleton contains the 25 *midpoint* sites $u + 4^n(\iota, \kappa_0)$, $\iota, \kappa_0 \in [0, 4]$, because each is the corner port of a coarse site of $[0, 4]^2 \subseteq S$. Write $\mathcal{M}_0$ for this set. The boundary lies in the skeleton: $\mathrm{Anc}_{n+1}(u) \subseteq \mathcal{M}_0$, and $\mathrm{Port}_{n+1}(u + 4^{n+1} z)$ is contained in the union of the ports of the coarse sites $4z + d$ by Lemma 7.12(b). The same holds for $z + e$ in the edge certificate, since $4z + 4e + [0, 4)^2 \subseteq 4z + \mathsf{D}_0 \subseteq S$.

Bit paths. For $\iota \in [0,4]$ and $\kappa_0 \in [0,3]$ attach a path of fresh vertices along the column of $u + 4^n(\iota, \kappa_0)$ from that site to $u + 4^n(\iota, \kappa_0 + 1)$, with edges labelled $\mathbf{e}_2$; its two end vertices are the skeleton vertices of these midpoints. For $\kappa_0 \in [0,4]$ and $\iota \in [0,3]$ attach in the same way a path along the row from $u + 4^n(\iota, \kappa_0)$ to $u + 4^n(\iota + 1, \kappa_0)$, with edges labelled $\mathbf{e}_1$. There are 40 paths.

The sixteen candidates. For each $c \in [0,4)^2$ put $v_c = u + 4^n c$, so that $\mathrm{Anc}_n(v_c) \subseteq \mathcal{M}_0$. For each c attach a fresh copy of each of the following depth-n certificates:

- $\Gamma_n^{\mathrm{ty}}(v_c, x - c)$ for every $x \in S$;
- $\Gamma_n^{\mathrm{ed}}(v_c, x - c, e')$ for every x and $e' \in \{\mathbf{e}_1, \mathbf{e}_2\}$ with $x, x + e' \in S$;
- for the point certificate only, $\Gamma_n^{\mathrm{pt}}(v_c, 4z + d_s - c, s)$, where $d_s \in [0,4)^2$ is the child whose block contains s.

The block of $x - c$ relative to v_c is the block of x relative to u. So the boundary of each copy consists of skeleton sites, and we identify each boundary vertex of the copy with the skeleton vertex of the same site. All other vertices of the copy are fresh. In particular, the sixteen candidates share the same port vertices: they read the same physical blocks.

The conclusions. Let ψ be a homomorphism of the depth-$(n+1)$ certificate whose values on $\mathrm{Anc}_{n+1}(u)$ form a good $(n+1)$-anchor at u. Write $m_{\iota\kappa_0} = \psi(u + 4^n(\iota, \kappa_0))$.

Step 1: one good candidate. Along a column path every edge joins letters that match in direction $\mathbf{e}_2$, so all letters on the path have the same column bit (Proposition 7.11(v)). Hence $\chi_1(m_{\iota\kappa_0}) = \chi_1(m_{\iota 0})$ for all $\iota, \kappa_0 \in [0,4]$, and in the same way $\chi_2(m_{\iota\kappa_0}) = \chi_2(m_{0\kappa_0})$. Since $\chi_1(m_{00}) \neq \chi_1(m_{40})$, some $\iota \in [0,3]$ has $\chi_1(m_{\iota 0}) \neq \chi_1(m_{\iota+1,0})$. Since $\chi_2(m_{00}) \neq \chi_2(m_{04})$, some $\kappa_0 \in [0,3]$ has $\chi_2(m_{0\kappa_0}) \neq \chi_2(m_{0,\kappa_0+1})$. Then $\chi_1(m_{\iota\kappa_0}) \neq \chi_1(m_{\iota+1,\kappa_0})$ and $\chi_2(m_{\iota\kappa_0}) \neq \chi_2(m_{\iota,\kappa_0+1})$. So ψ is a good n-anchor at v_c for $c = (\iota, \kappa_0)$. The construction attached all sixteen candidates in advance; which one is good depends on ψ, and the argument below uses only that one.

Step 2: a common coarse pattern. The restriction of ψ to a copy attached for c is a homomorphism of that copy, with the same values on its boundary. By the induction hypothesis (T), for every $x \in S$ the map ψ reads some letter on $\mathrm{Port}_n(u + 4^n x)$; by Lemma 7.12(a) it is unique, and we call it $\mathfrak{c}(x)$. It depends only on the values of ψ at the skeleton vertices of these ports, so every copy that meets these ports sees the same letter. By (E), $\mathfrak{c}(x)$ and $\mathfrak{c}(x + e')$ match in direction e' whenever $x, x + e' \in S$. So $\mathfrak{c}$ is a legal pattern on S.

Step 3: the phase. By Lemma 7.12(c), $m_{\iota\kappa_0} = \psi(u + 4^n(\iota, \kappa_0)) = \sigma^n(\mathfrak{c}(\iota, \kappa_0))(0,0)$ has the bits of $\mathfrak{c}(\iota, \kappa_0)$. The good $(n+1)$-anchor therefore gives $\chi_1(\mathfrak{c}(0,0)) \neq \chi_1(\mathfrak{c}(4,0))$ and $\chi_2(\mathfrak{c}(0,0)) \neq \chi_2(\mathfrak{c}(0,4))$. Apply Lemma 7.17 with $v = 4z$, which is congruent to $w_0 = (0,0)$ modulo 4 and satisfies $4z + [0,8)^2 \subseteq w_m + \mathsf{D}_0$ because $L_0 \geq 8$. It gives a letter $\mathfrak{a}$ with $\mathfrak{c}(4z + d) = \sigma(\mathfrak{a})(d)$ for $d \in [0,4)^2$, and, for the edge certificate, a letter $\mathfrak{b}$ with $\mathfrak{c}(4z + 4e + d) = \sigma(\mathfrak{b})(d)$, where $\mathfrak{a}$ and $\mathfrak{b}$ match in direction e.

Step 4: the readout. Let $y \in P_{n+1}$ and write $y = 4^n d + y'$ with $y' \in P_n$, by Lemma 7.12(b). Then

$$\begin{aligned} \psi(u + 4^{n+1}z + y) &= \psi\big(u + 4^n(4z + d) + y'\big) = \sigma^n\big(\mathfrak{c}(4z + d)\big)(y') \\ &= \sigma^n\big(\sigma(\mathfrak{a})(d)\big)(y') = \sigma^{n+1}(\mathfrak{a})(y). \end{aligned}$$

So ψ reads $\mathfrak{a}$ on $\mathrm{Port}_{n+1}(u + 4^{n+1}z)$, which is (T). The same computation for $z + e$ gives (E).

Step 5: the point. For the good candidate c, the induction hypothesis (P) for the attached copy of $\Gamma_n^{\mathrm{pt}}(v_c, 4z + d_s - c, s)$ gives a letter $\mathfrak{a}''$ read on the ports of $4z + d_s$, with $\psi(s) = \sigma^n(\mathfrak{a}'')(s - u - 4^n(4z + d_s))$. By uniqueness of the readout, $\mathfrak{a}'' = \mathfrak{c}(4z + d_s) = \sigma(\mathfrak{a})(d_s)$. Since

$s - u - 4^{n+1}z = 4^n d_s + \big(s - u - 4^n(4z + d_s)\big)$, Definition 7.8 gives $\psi(s) = \sigma^{n+1}(\mathfrak{a})(s - u - 4^{n+1}z)$, which is (P).

The decomposition. Let ∂ be the boundary of the depth-$(n+1)$ certificate. For $k \in [0, m]$ let β_k consist of the skeleton vertices at the sites of $\mathcal{M}_0$, of ∂, and of $\mathrm{Port}_n(u + 4^n x)$ for every $x \in S \cap \big(w_k + [-\varrho - 1, \varrho + 1]^2\big)$. Arrange $\beta_0, \dots, \beta_m$ in a path with root β_0. A skeleton vertex at a site of $\mathcal{M}_0 \cup \partial$ lies in every bag. Any other skeleton vertex lies in the ports of exactly one coarse site x, because the depth-n blocks are disjoint; the set of k with $x \in w_k + [-\varrho - 1, \varrho + 1]^2$ is an interval, because each coordinate of w_k is monotone in k. So the skeleton bags satisfy the connectivity condition. Every $x \in S$ lies in $w_k + [-\varrho, \varrho]^2$ for some k, and then its ports and those of every neighbour $x + e' \in S$ lie in β_k.

Attach the decomposition of each child copy by joining its root bag to a bag that contains the copy's boundary. For a type copy at x or an edge copy at $x, x + e'$, take β_k with $x \in w_k + [-\varrho, \varrho]^2$; the anchor sites of the copy lie in $\mathcal{M}_0$. For the point copy, take β_m: the coarse site $4z + d_s$ lies in $w_m + [-\varrho, \varrho]^2$, and $s \in \partial$. For a bit path with vertices $x_0, \dots, x_N$, use the bags $\{x_0, x_N, x_k, x_{k+1}\}$, $k < N$, in a path, and join its first bag to β_0, which contains the midpoints x_0 and x_N. The fresh vertices of a copy or a path lie only in its own bags, and each shared vertex lies in a connected set of skeleton bags together with the connected set of bags of each copy that contains it, joined at the attachment. Every edge lies inside a copy or a path, where a bag contains it. So this is a rooted tree decomposition, and $\beta_0 \supseteq \partial$.

Each β_k has at most $\ell(2\varrho + 3)^2$ port vertices, 25 midpoints, and at most $2\ell + 1$ further boundary vertices, so $|\beta_k| \le \ell(2\varrho + 3)^2 + 2\ell + 26 \le B$. The bags of the paths have 4 vertices, and the bags of the copies have at most B vertices by induction. So the width is at most B. The bound is a maximum over finitely many kinds of bags, not a sum over depths, and the length of the corridor changes only the number of bags. □

Two features of the construction carry the proof. First, all sixteen candidates are attached to the same port vertices, so the letter decoded at a coarse site is one letter, whatever certificate looks at it. Second, the point certificate contains the whole corridor from the anchor to the block that holds s. A window near the block alone would fix some phase of the decoded pattern, but nothing would tie that phase to the anchor's grid; inside the recursion the child block can lie at coarse distance $3, 15, 63, \dots$ from the candidate anchors.

The bit paths use the edgewise bit property of Proposition 7.11(v). If the bits were only known to be constant along the columns and rows of tilings, a bare path would force nothing, since a homomorphism of a certificate need not extend to a tiling. The literal bit colours of Definition 7.7 avoid this.

7.5 Transfer to pebble families, and anchors of every depth

We use E_K families for the Wang set $\mathbb{A}$ as in Definition 4.22: nonempty sets of legal patterns of $\mathbb{A}$ on finite subsets of $\mathbb{Z}^2$, each with at most K sites, closed under restriction and with the extension property for maps on fewer than K sites. By Remark 4.23, a member h extends to a member defined on $\mathrm{dom}\, h \cup U'$ whenever $|\,\mathrm{dom}\, h \cup U'| \le K$.

Lemma 7.19 (Transfer)**.** *Let H be an E_K family for $\mathbb{A}$. Let Γ be a certificate with a rooted tree decomposition of width at most B', and let $\Pi \subseteq \mathbb{Z}^2$ be a finite set of* pinned *sites with $|\Pi| \le K - B'$. Let $h \in H$ have domain $\partial\Gamma \cup \Pi$. Then there is a homomorphism ψ of Γ such that $\psi(x) = h(x)$ for every boundary site x, and $\psi(v) = h(\rho(v))$ for every vertex v with $\rho(v) \in \Pi$.*

So a use of a certificate of width at most B' may retain up to $K - B'$ pinned sites beyond its bags, and the certificate then respects the pins wherever it touches them. With $B' = B$ and

$K = j = 10(B + \ell + 100)$, as in Theorem 9.1, this allows $9B + 10\ell + 1000$ pinned sites. The applications in this paper need fewer: Proposition 7.21 and Corollary 7.20 below take $\Pi = \emptyset$, and Section 9 applies certificates only to restrictions of its maps to the certificate boundary, which is again the case $\Pi = \emptyset$. All maps there have at most $\ell + 6 \le B$ sites, so the single condition $K \ge B$ covers the fixed boxes of Proposition 7.21, the certificates, and the pins and sibling sites of Section 9.

Proof. Let (β_t) be the decomposition, with root t_0; it has width at most B', so $|\rho(\beta_t)| \le B'$. We choose maps $g_t \in H$ with domain $\rho(\beta_t) \cup \Pi$ and $g_t|_\Pi = h|_\Pi$, from the root down. For the root, $\partial\Gamma \subseteq \rho(\beta_{t_0})$, and h extends to g_{t_0} on $\rho(\beta_{t_0}) \cup \Pi$, a set of at most K sites. For a child t of a node t', restrict $g_{t'}$ to $\rho(\beta_t \cap \beta_{t'}) \cup \Pi$ and extend the restriction to $\rho(\beta_t) \cup \Pi$. Put $\psi(v) = g_t(\rho(v))$ for any t with $v \in \beta_t$. This does not depend on t: the nodes whose bags contain v form a subtree, and along each of its edges the two maps agree at $\rho(v)$ by construction. Every edge (v, v') of Γ with label e lies in a bag β_t; then $\psi(v) = g_t(\rho(v))$ and $\psi(v') = g_t(\rho(v) + e)$ are values of one legal pattern at sites $\rho(v)$ and $\rho(v) + e$, so they match in direction e. At a boundary site x, $\psi(x) = g_{t_0}(x) = h(x)$. If $\rho(v) \in \Pi$, then $\psi(v) = g_t(\rho(v)) = h(\rho(v))$. □

This is the standard argument showing that the existential K-pebble game decides homomorphism from structures of tree width less than K [28, 8]. Two vertices with one site in a common bag receive one letter. Vertices with one site in unrelated branches may receive different letters, and Theorem 7.18 never needs them to agree.

Corollary 7.20 (Readout in a family)**.** *Let H be an E_K family for $\mathbb{A}$ with $K \ge B$. Let $n \ge 0$, $u \in \mathbb{Z}^2$, and let $g \in H$ be defined on $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u)$ and be a good n-anchor at u. Then g reads a letter $\mathfrak{a}$ on $\mathrm{Port}_n(u)$, that is $g(y) = \sigma^n(\mathfrak{a})(y - u)$ for every $y \in \mathrm{Port}_n(u)$; the letter is unique. If g is also defined at a site $s \in u + [0, 4^n)^2$, then $g(s) = \sigma^n(\mathfrak{a})(s - u)$. More generally, for every $z \in \mathbb{Z}^2$: if g is defined on $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u + 4^n z)$ with a good anchor, it reads a letter on $\mathrm{Port}_n(u + 4^n z)$, and the value at any s in the block of z is the corresponding entry of σ^n of that letter.*

Proof. Restrict g to the boundary of the type certificate $\Gamma_n^{\mathrm{ty}}(u, z)$, or of the point certificate $\Gamma_n^{\mathrm{pt}}(u, z, s)$, and apply Lemma 7.19 with $\Pi = \emptyset$ and $B' = B$. The resulting homomorphism has a good anchor, and Theorem 7.18(T),(P) give the claims. The letter given by the point certificate is the one read on the ports, by Lemma 7.12(a). □

Proposition 7.21 (Anchors of every depth)**.** *Let H be an E_K family for $\mathbb{A}$ with $K \ge B$. For every $n \ge 0$ there are $u \in \mathbb{Z}^2$ and $h \in H$ with domain $\mathrm{Anc}_n(u)$ that is a good n-anchor at u.*

Proof. We induct on n. The site u may depend on H and on n.

Depth 0. The box $Y = [-L_2, 4 + L_2]^2$ has $(2L_2 + 5)^2 \le B \le K$ sites, so H has a member g with domain Y (Remark 4.23). It is a legal pattern, so by Lemma 7.15(a) its restriction to $[0, 5)^2$ is the restriction of a tiling f. Let p represent the phase of f, and let $x \in [0, 4)^2$ satisfy $x \equiv p + (1, 1) \pmod 4$. Then x, $x + \mathbf{e}_1$ and $x + \mathbf{e}_2$ lie in $[0, 5)^2$. By Proposition 7.11(iii),(iv), the column of x has residue 1 relative to p_1 and column bit 0, and the column of $x + \mathbf{e}_1$ has residue 2 and column bit 1. In the same way $\chi_2(f(x)) = 0$ and $\chi_2(f(x + \mathbf{e}_2)) = 1$. So $h = g|_{\mathrm{Anc}_0(x)}$ is a good 0-anchor at x, and $h \in H$.

From depth n to $n + 1$. Let $h_n \in H$ be a good n-anchor at u, with domain $\mathrm{Anc}_n(u)$. Let $U = [-L_3, 31 + L_3]^2$ and

$$\Sigma = \mathrm{Anc}_n(u) \cup \bigcup_{x \in U} \mathrm{Port}_n(u + 4^n x).$$

Then $|\Sigma| \le 3 + \ell(2L_3+32)^2 \le B \le K$, so h_n extends to a member $g \in H$ with domain Σ. By Corollary 7.20, for every $x \in U$ the map g reads a unique letter $\mathfrak{c}(x)$ on $\mathrm{Port}_n(u+4^n x)$. For $x, x+e \in U$, the edge certificate $\Gamma_n^{\mathrm{ed}}(u,x,e)$ with Lemma 7.19, applied to the restriction of g to its boundary, shows that $\mathfrak{c}(x)$ and $\mathfrak{c}(x+e)$ match in direction e. So $\mathfrak{c}$ is a legal pattern on U, and by Lemma 7.15(a) its restriction to $[0,32)^2$ is the restriction of a tiling f.

Apply Proposition 7.11(iii) to f and then to its parent: there are $o \in [0,16)^2$ and a tiling y' with $f(o+16m+i) = \sigma^2(y'(m))(i)$ for $i \in [0,16)^2$. By Proposition 7.11(iv) applied twice, every letter of $\sigma^2(\mathfrak{b})$ in column 4 has column bit $\chi_1(\sigma(\mathfrak{b})(1,\cdot)) = 0$, and every letter in column 8 has column bit $\chi_1(\sigma(\mathfrak{b})(2,\cdot)) = 1$; rows 4 and 8 have row bits 0 and 1. Put $w = o + (4,4)$. The sites w, $w + 4\mathbf{e}_1 = o + (8,4)$ and $w + 4\mathbf{e}_2 = o + (4,8)$ lie in $o + [0,16)^2 \subseteq [0,32)^2$. So

$$\chi_1(\mathfrak{c}(w)) = 0 \neq 1 = \chi_1(\mathfrak{c}(w+4\mathbf{e}_1)), \qquad \chi_2(\mathfrak{c}(w)) = 0 \neq 1 = \chi_2(\mathfrak{c}(w+4\mathbf{e}_2)).$$

By Lemma 7.12(c), $g(u+4^n x) = \sigma^n(\mathfrak{c}(x))(0,0)$ has the bits of $\mathfrak{c}(x)$. Put $u' = u + 4^n w$. Then $\mathrm{Anc}_{n+1}(u') = \{u+4^n w,\ u+4^n(w+4\mathbf{e}_1),\ u+4^n(w+4\mathbf{e}_2)\} \subseteq \Sigma$, and $h_{n+1} = g|_{\mathrm{Anc}_{n+1}(u')}$ is a good $(n+1)$-anchor at u'. It lies in H by restriction. □

The old anchor is dropped as soon as the new one is found, so no pins accumulate over the depths: each step uses one member of H on at most B sites. The anchor of depth n may sit anywhere, and different families may place it differently. Together, Proposition 7.21 and Corollary 7.20 give, in every E_K family with $K \ge B$ and at every depth n, a member on $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u)$, at most $\ell + 3$ sites, that reads one root letter $\mathfrak{a}$; and every member that agrees with it there and is defined at a site of the block $u + [0,4^n)^2$ takes the value prescribed by the word $\sigma^n(\mathfrak{a})$.

8 Routing a machine onto the scaffold

This section builds the family $M \mapsto R_M$. A tile of R_M sits on one quarter of a letter of the scaffold $\mathbb{A}$ of Section 7. It carries the scaffold letter and its quadrant, two flows that read the light red walls of the scaffold, and a payload. The flows mark the free rows and columns of nested boards, and the payload carries a computation of M along them. Routing a computation through the free rows and columns of nested boards is Robinson's method [43]; Berger embedded Turing machines in Wang sets to show that tiling is undecidable [3]. Our payload is a system of Horn clauses. Its local constraints are closed under intersection, the property that lets Section 9 rebuild a payload by local consistency [25, 17].

Section 8.1 classifies the red paths of the scaffold. Section 8.2 defines the quarters, the walls, the flows and the roles, and proves that walls force the carriers (Proposition 8.11). Section 8.3 studies the light frames. Section 8.4 builds a routing model on the whole plane, and Section 8.5 finds an initialized board inside every word $\sigma^n(\mathfrak{a})$ (Proposition 8.15). Section 8.6 builds the Horn payload (Proposition 8.16), and Section 8.7 defines R_M and shows that it tiles when M does not halt (Definition 8.17 and Proposition 8.18).

We keep the notation of Section 7: letters $\mathfrak{a} = (\theta, \kappa, \omega, b_1, b_2)$ of $\mathbb{A}$, column bit χ_1, row bit χ_2, the substitutions φ, φ' and $\sigma = (\varphi')^2$, and the endpoints $\partial_{\mathrm{N}}, \partial_{\mathrm{E}}, \partial_{\mathrm{S}}, \partial_{\mathrm{W}}$ of Table 4. Integer intervals are $[a,b] = \{a,\dots,b\}$, and $z \bmod m \in [0,m)$. An atom r_i is *red* and an atom g_i is *green*. The base tile $(\mathsf{a}, \angle\mathsf{b}, \mathsf{A})$ is the *index*-0 *tile*, and a letter with this base is an *index*-0 *letter*. Ruler bits $b_1 = b_2 = 0$ are *dark*.

8.1 Towers and red paths

Definition 8.1 (Towers)**.** Let $h \in \{1, 2, \dots\} \cup \{\infty\}$ and $v_m \in \{0,1\}^2$ for $0 \le m < h$. A *tower of height h with digits (v_m)* is a family of legal patterns f_m of $\mathbb{A}$ on sets $U_m \subseteq \mathbb{Z}^2$, for integers $0 \le m \le h$, such that for every $m < h$

$$U_m = v_m + 2U_{m+1} + [0,2)^2, \qquad f_m(v_m + 2v + \varsigma) = \varphi'\big(f_{m+1}(v)\big)(\varsigma) \quad (v \in U_{m+1},\ \varsigma \in [0,2)^2).$$

If $h = \infty$ we require $U_m = \mathbb{Z}^2$ for all m, so every f_m is a tiling. The *phase* of the tower is $\varpi = \sum_{m<h} 2^m v_m$, an element of $\mathbb{Z}^2$ if $h < \infty$ and of the 2-adic integers $\mathbb{Z}_2^2$ if $h = \infty$.

For a letter $\mathfrak{a}$ and $n \ge 1$, the patterns $f_m = (\varphi')^{2n-m}(\mathfrak{a})$ on $[0, 2^{2n-m})^2$, $0 \le m \le 2n$, form a tower of height $2n$ with all digits 0 and phase 0. Its bottom pattern is $f_0 = (\varphi')^{2n}(\mathfrak{a}) = \sigma^n(\mathfrak{a})$, because blockwise substitution of patterns composes. Its patterns are legal by Lemma 7.9.

Lemma 8.2 (Towers of every phase)**.** *For every $\varpi \in \mathbb{Z}_2^2$ there is a tower of height ∞ with phase ϖ.*

Proof. Write $\varpi = \sum_{m \ge 0} 2^m v_m$ with $v_m \in \{0,1\}^2$. For a tiling f of $\mathbb{A}$ and $v \in \{0,1\}^2$, the pattern $\mathrm{sub}_v(f)(v + 2v + \varsigma) = \varphi'(f(v))(\varsigma)$ is a tiling by Lemma 7.9(b). Let $\mathcal{T}_l$ be the set of sequences $(f_m)_{m \ge 0}$ of tilings with $f_m = \mathrm{sub}_{v_m}(f_{m+1})$ for $m < l$. It is nonempty: take any tiling f (Proposition 7.11), put $f_m = f$ for $m \ge l$ and $f_m = \mathrm{sub}_{v_m}(f_{m+1})$ for $m < l$. Each condition $f_m = \mathrm{sub}_{v_m}(f_{m+1})$ is a family of conditions on finitely many sites each, so $\mathcal{T}_l$ is closed in the product of the compact spaces $\mathbb{A}^{\mathbb{Z}^2}$. The sets $\mathcal{T}_l$ decrease, so their intersection is nonempty. An element of it is a tower with phase ϖ. □

Fix a tower. For $x \in U_0$ put $x^{(0)} = x$, and for $m < h$ let $\varsigma_m(x) \in \{0,1\}^2$ and $x^{(m+1)} \in U_{m+1}$ be the unique pair with

$$x^{(m)} = v_m + 2x^{(m+1)} + \varsigma_m(x).$$

Then $f_m(x^{(m)}) = \varphi'(f_{m+1}(x^{(m+1)}))(\varsigma_m(x))$, and by induction

$$x - \varpi \equiv \sum_{m<l} 2^m \varsigma_m(x) \pmod{2^l} \qquad (l \le h,\ l < \infty). \tag{39}$$

The first coordinate $\varsigma_{m,1}(x)$ depends only on x_1, and the second only on x_2. The *rank* of a column x_1 is the least $m < h$ with $\varsigma_{m,1}(x) = 1$, if there is one; the rank of a row x_2 is defined with $\varsigma_{m,2}$. By (39), a column of rank r satisfies $x_1 - \varpi_1 \equiv 2^r \pmod{2^{r+1}}$. If $h = \infty$ and $\varpi_1 \notin \mathbb{Z}$, every column has a finite rank, namely the 2-adic valuation of $x_1 - \varpi_1$.

Lemma 8.3 (Ruler bits)**.** *If the column x_1 has rank r, then $\chi_1(f_0(x)) \equiv r \pmod 2$ for every $x \in U_0$ in this column. If the row x_2 has rank r, then $\chi_2(f_0(x)) \equiv r \pmod 2$.*

Proof. By Table 6, a child ς of a letter has column bit 0 if $\varsigma_1 = 1$, and the complement of the column bit of the parent if $\varsigma_1 = 0$. The letter $f_r(x^{(r)})$ is a child with $\varsigma_{r,1}(x) = 1$, so its column bit is 0. For $m < r$ we have $\varsigma_{m,1}(x) = 0$, so each of the r steps down to $f_0(x)$ complements the bit. Rows are the same. □

A thick type has a *vertical red atom* if its north or its south endpoint is red, and a *horizontal red atom* if its east or its west endpoint is red. A line $\langle \alpha, \gamma \rangle$ has a vertical red atom exactly when α is red. Every corner has exactly one red endpoint in each direction (Table 4): $\angle$b has red north r_0 and red east r_1, $\angle$c red north r_2 and red west r_1, $\angle$a red south r_0 and red east r_3, and $\angle$d red south r_2 and red west r_3. So $\angle$b is the south-west corner of a red rectangle, $\angle$c the south-east, $\angle$a the north-west and $\angle$d the north-east corner.

Lemma 8.4 (Red paths). *Let (f_m) be a tower of height h, and let $r \geq 0$ with $r+2 \leq h$. For $i = 1, 2$ and $z \in \mathbb{Z}$ put $s_i(z) = (z - \varpi_i) \bmod 2^{r+2}$. Let x_1 be a column of rank r. Then $s_1(x_1) \in \{2^r, 3 \cdot 2^r\}$, and for every $x = (x_1, x_2) \in U_0$:*

(a) the thick type of $f_0(x)$ has a vertical red atom if and only if $s_2(x_2) \in [2^r, 3 \cdot 2^r]$;

(b) this atom is r_0 if $s_1(x_1) = 2^r$ and r_2 if $s_1(x_1) = 3 \cdot 2^r$;

(c) the thick type is the corner with red north endpoint ($\angle\mathsf{b}$ or $\angle\mathsf{c}$) if $s_2(x_2) = 2^r$, the corner with red south endpoint ($\angle\mathsf{a}$ or $\angle\mathsf{d}$) if $s_2(x_2) = 3 \cdot 2^r$, and a line if $2^r < s_2(x_2) < 3 \cdot 2^r$.

The same holds for a row x_2 of rank r with the coordinates exchanged: $s_2(x_2) \in \{2^r, 3 \cdot 2^r\}$; a horizontal red atom occurs if and only if $s_1(x_1) \in [2^r, 3 \cdot 2^r]$; it is r_1 if $s_2(x_2) = 2^r$ and r_3 if $s_2(x_2) = 3 \cdot 2^r$; and the thick type is the corner with red east endpoint ($\angle\mathsf{b}$ or $\angle\mathsf{a}$) if $s_1(x_1) = 2^r$, the corner with red west endpoint ($\angle\mathsf{c}$ or $\angle\mathsf{d}$) if $s_1(x_1) = 3 \cdot 2^r$, and a line in between.

Proof. By (39) with $l = r+2$, $s_1(x_1) = 2^r + 2^{r+1}\varsigma_{r+1,1}(x)$, which is 2^r or $3 \cdot 2^r$. We prove (a)–(c) by induction on r, for all towers of height at least $r+2$ at once.

Let $r = 0$, so $\varsigma_{0,1}(x) = 1$. The letter $f_0(x)$ is child $(1, \varsigma_{0,2})$ of $P = f_1(x^{(1)})$, and P is child $\varsigma_1(x)$ of $f_2(x^{(2)})$; this uses $h \geq 2$. By Table 6 the thin type θ of P is $\mathsf{a}, \mathsf{c}, \mathsf{d}, \mathsf{b}$ for $\varsigma_1(x) = (0,0), (1,0), (0,1), (1,1)$, and the thick type of $f_0(x)$ is the line $\langle \partial_{\mathrm{S}}\theta, \cdot \rangle$ if $\varsigma_{0,2} = 0$ and the corner $F(\theta)$ if $\varsigma_{0,2} = 1$. Here $s_2(x_2) = \varsigma_{0,2} + 2\varsigma_{1,2}$ and $s_1(x_1) = 1 + 2\varsigma_{1,1}$. Table 4 and $F(\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}) = (\angle\mathsf{b}, \angle\mathsf{d}, \angle\mathsf{c}, \angle\mathsf{a})$ give:

$s_2(x_2)$	0	1	2	3
$s_1(x_1) = 1$	$\partial_{\mathrm{S}}\mathsf{a} = \mathsf{g}_2$	$F(\mathsf{a}) = \angle\mathsf{b}$	$\partial_{\mathrm{S}}\mathsf{d} = \mathsf{r}_0$	$F(\mathsf{d}) = \angle\mathsf{a}$
$s_1(x_1) = 3$	$\partial_{\mathrm{S}}\mathsf{c} = \mathsf{g}_0$	$F(\mathsf{c}) = \angle\mathsf{c}$	$\partial_{\mathrm{S}}\mathsf{b} = \mathsf{r}_2$	$F(\mathsf{b}) = \angle\mathsf{d}$

This is (a)–(c) for $r = 0$.

Let $r \geq 1$, so $\varsigma_{0,1}(x) = 0$. The family $(f_{m+1})_m$ is a tower of height $h - 1 \geq r+1$ with digits (v_{m+1}) and phase ϖ', where $\varpi = v_0 + 2\varpi'$. In it, the column of $x^{(1)}$ has rank $r-1$, the digits of $x^{(1)}$ are $\varsigma_{m+1}(x)$, and $x - \varpi = 2(x^{(1)} - \varpi') + \varsigma_0(x)$. With $s'_i(z) = (z - \varpi'_i) \bmod 2^{r+1}$ this gives $s_i(x_i) = 2s'_i(x_i^{(1)}) + \varsigma_{0,i}(x)$ and $s_1(x_1) = 2s'_1(x_1^{(1)})$. The letter $f_0(x)$ is child $(0, \varsigma_{0,2})$ of $P = f_1(x^{(1)})$, with thick type κ_P. By Table 6, child $(0,0)$ has thick type κ_P, and child $(0,1)$ has the line $\langle \partial_{\mathrm{N}}\kappa_P, \cdot \rangle$. Apply the induction hypothesis to P. If $\varsigma_{0,2} = 0$, then $s_2 = 2s'_2$ and $f_0(x)$ has the thick type of P; (a)–(c) follow, since $s'_2 \in [2^{r-1}, 3 \cdot 2^{r-1}]$ if and only if $2s'_2 \in [2^r, 3 \cdot 2^r]$, and the end values correspond. If $\varsigma_{0,2} = 1$, then $s_2 = 2s'_2 + 1$ is odd, and $f_0(x)$ is a line whose vertical atom is $\partial_{\mathrm{N}}\kappa_P$. This endpoint is red exactly when κ_P is a line with a red vertical atom or the corner with red north endpoint, that is when $s'_2 \in [2^{r-1}, 3 \cdot 2^{r-1})$. Since $r \geq 1$, the odd numbers $2s'_2 + 1$ with this property are exactly the odd numbers in $[2^r, 3 \cdot 2^r]$, and they lie strictly between 2^r and $3 \cdot 2^r$. In both cases the atom is that of P, so (b) follows from $s_1(x_1) = 2s'_1(x_1^{(1)})$.

For rows the proof is the same with the coordinates exchanged. For $r = 0$ the letter is child $(\varsigma_{0,1}, 1)$ of P, that is the line $\langle \cdot, \partial_{\mathrm{W}}\theta \rangle$ or the corner $F(\theta)$, and the table reads, for $s_1(x_1) = 0, 1, 2, 3$: $\partial_{\mathrm{W}}\mathsf{a} = \mathsf{g}_3$, $F(\mathsf{a}) = \angle\mathsf{b}$, $\partial_{\mathrm{W}}\mathsf{c} = \mathsf{r}_1$, $F(\mathsf{c}) = \angle\mathsf{c}$ when $s_2(x_2) = 1$, and $\partial_{\mathrm{W}}\mathsf{d} = \mathsf{g}_1$, $F(\mathsf{d}) = \angle\mathsf{a}$, $\partial_{\mathrm{W}}\mathsf{b} = \mathsf{r}_3$, $F(\mathsf{b}) = \angle\mathsf{d}$ when $s_2(x_2) = 3$. In the induction step child $(1,0)$ is the line $\langle \cdot, \partial_{\mathrm{E}}\kappa_P \rangle$, and $\partial_{\mathrm{E}}\kappa_P$ is red exactly for a line with red horizontal atom or the corner with red east endpoint. □

So the red atoms in the columns and rows of rank r form the boundaries of squares: for a column x_1 with $s_1(x_1) = 2^r$, the column $x_1 + 2^{r+1}$ has $s_1 = 3 \cdot 2^r$ and the same rank, and the two sides run over the same rows. By Lemma 8.3 every letter on a vertical side of such a

boundary has column bit r mod 2, every letter on a horizontal side has row bit r mod 2, and the four corners have both bits equal to r mod 2.

8.2 Quarters, walls, flows and roles

Each letter of $\mathbb{A}$ is cut into four quarters. A *scaffold pair* is a pair $(\varsigma, \mathfrak{a}) \in \{0,1\}^2 \times \mathbb{A}$; the first entry is its *quadrant* and the second its *parent*. Let $\circ$ and $\bullet$ be two new symbols.

Definition 8.5 (Quarters)**.** The quarter set $\mathbb{A}^{\mathrm{q}}$ is the Wang set of the $4 \cdot 276$ scaffold pairs. The colours of $(\varsigma, \mathfrak{a})$, $\varsigma = (\varsigma_1, \varsigma_2)$, are

$$\begin{array}{lll} \text{east:} & (\circ, \mathfrak{a}, \varsigma_2) \text{ if } \varsigma_1 = 0, & (\bullet, \mathfrak{a}_{\mathrm{E}}, \varsigma_2) \text{ if } \varsigma_1 = 1; \\ \text{west:} & (\circ, \mathfrak{a}, \varsigma_2) \text{ if } \varsigma_1 = 1, & (\bullet, \mathfrak{a}_{\mathrm{W}}, \varsigma_2) \text{ if } \varsigma_1 = 0; \\ \text{north:} & (\circ, \mathfrak{a}, \varsigma_1) \text{ if } \varsigma_2 = 0, & (\bullet, \mathfrak{a}_{\mathrm{N}}, \varsigma_1) \text{ if } \varsigma_2 = 1; \\ \text{south:} & (\circ, \mathfrak{a}, \varsigma_1) \text{ if } \varsigma_2 = 1, & (\bullet, \mathfrak{a}_{\mathrm{S}}, \varsigma_1) \text{ if } \varsigma_2 = 0. \end{array}$$

Here $\mathfrak{a}_{\mathrm{E}}, \dots$ are the colours of $\mathfrak{a}$ in $\mathbb{A}$. The *quartering* of a pattern f of $\mathbb{A}$ on U is the pattern f^{q} on $2U + [0,2)^2$ with $f^{\mathrm{q}}(2x + \varsigma) = (\varsigma, f(x))$. Its value at ξ is the *quarter of f at ξ*.

Lemma 8.6 (Quarter encoding)**.** *(a) Scaffold pairs $(\varsigma, \mathfrak{a})$ and $(\varsigma', \mathfrak{a}')$ match horizontally if and only if $\varsigma' \equiv \varsigma + \mathbf{e}_1 \pmod 2$ and either $\varsigma_1 = 0$ and $\mathfrak{a}' = \mathfrak{a}$, or $\varsigma_1 = 1$ and $\mathfrak{a}, \mathfrak{a}'$ match horizontally in $\mathbb{A}$. They match vertically if and only if $\varsigma' \equiv \varsigma + \mathbf{e}_2 \pmod 2$ and either $\varsigma_2 = 0$ and $\mathfrak{a}' = \mathfrak{a}$, or $\varsigma_2 = 1$ and $\mathfrak{a}, \mathfrak{a}'$ match vertically in $\mathbb{A}$.*

(b) A pattern f of $\mathbb{A}$ is legal if and only if f^{q} is legal. A map $\mathcal{Y}$ on $\mathbb{Z}^2$ is a tiling of $\mathbb{A}^{\mathrm{q}}$ if and only if $\mathcal{Y}(\xi) = f^{\mathrm{q}}(\xi - c)$ for some $c \in \{0,1\}^2$ and some tiling f of $\mathbb{A}$.

Proof. (a) If $\varsigma_1 = 0$, the east colour $(\circ, \mathfrak{a}, \varsigma_2)$ equals a west colour only if that colour is internal; so $\varsigma'_1 = 1$, $\mathfrak{a}' = \mathfrak{a}$ and $\varsigma'_2 = \varsigma_2$. If $\varsigma_1 = 1$, the east colour is external; so $\varsigma'_1 = 0$, $\varsigma'_2 = \varsigma_2$ and $\mathfrak{a}_{\mathrm{E}} = \mathfrak{a}'_{\mathrm{W}}$. The converse is read off the same colours. The vertical case is the same.

(b) The first claim follows from (a): inside a block $2x + [0,2)^2$ the four quarters have one parent, and across blocks the external colours compare the parents. For the second, let $\mathcal{Y}$ be a tiling. By (a), the quadrant of $\mathcal{Y}(\xi)$ minus ξ is constant modulo 2 along rows and along columns, so it is a constant c modulo 2. By (a) again, the four quarters of each block $c + 2x + [0,2)^2$ have a common parent $f(x)$, and adjacent blocks have matching parents. The converse follows from the first claim. □

Definition 8.7 (Walls)**.** Let $(\varsigma, \mathfrak{a})$ be a scaffold pair whose parent has thick type κ.

- The pair lies on a *vertical red path* if κ has a vertical red atom, $\varsigma_1 = 1$ if this atom is r_0 and $\varsigma_1 = 0$ if it is r_2, and, when κ is a corner, $\varsigma_2 = 1$ if its north endpoint is red and $\varsigma_2 = 0$ if its south endpoint is red. It is then a *left side* if the atom is r_0 and a *right side* if it is r_2. The path is *light* at the pair if $\chi_1(\mathfrak{a}) = 1$ and *dark* if $\chi_1(\mathfrak{a}) = 0$. A pair on a light vertical red path lies on a *light vertical wall*.
- The pair lies on a *horizontal red path* if κ has a horizontal red atom, $\varsigma_2 = 1$ if this atom is r_1 and $\varsigma_2 = 0$ if it is r_3, and, when κ is a corner, $\varsigma_1 = 1$ if its east endpoint is red and $\varsigma_1 = 0$ if its west endpoint is red. It is then a *bottom side* if the atom is r_1 and a *top side* if it is r_3. The path is light if $\chi_2(\mathfrak{a}) = 1$, and the pair then lies on a *light horizontal wall*.

So the lane r_0 runs in the east half of a letter and r_2 in the west half; r_1 runs in the north half and r_3 in the south half. A corner turns in one quarter: $\angle\mathsf{b}$ in quadrant $(1,1)$, $\angle\mathsf{c}$ in $(0,1)$, $\angle\mathsf{a}$ in $(1,0)$ and $\angle\mathsf{d}$ in $(0,0)$. That quarter is on a vertical and on a horizontal path, and both

have the shade $b_1 = b_2$ by (R3). A pair lies on at most one vertical and at most one horizontal red path, so the side of a wall is well defined.

Definition 8.8 (Frames)**.** Let integers a_d, $d \geq 1$, satisfy $a_{d+1} \equiv a_d \pmod{4^{d+1}}$. The *frames of depth d* are the intervals

$$I_d(m) = \left[a_d + (4m+1)4^d + 1,\ a_d + (4m+3)4^d\right], \qquad m \in \mathbb{Z}.$$

The *left endpoint* of $I_d(m)$ is $a_d + (4m+1)4^d + 1$ and its *right endpoint* is $a_d + (4m+3)4^d$. We write $\mathring{I}$ for the interior of an interval I, that is I without its two endpoints.

For a tower whose phase (ϖ_0, ϖ_0) has equal coordinates we use the frames with $a_d \equiv 2\varpi_0 \pmod{4^{d+1}}$. They are the light red squares of the tower, seen in quarter coordinates.

Lemma 8.9 (Light walls)**.** *Let (f_m) be a tower of height h with phase (ϖ_0, ϖ_0), take the frames with $a_d \equiv 2\varpi_0 \pmod{4^{d+1}}$, and let $\xi \in 2U_0 + [0,2)^2$. Suppose the column $\lfloor \xi_1/2 \rfloor$ has rank at most $h-2$. Then the quarter of f_0 at ξ lies on a light vertical wall if and only if there are $d \geq 1$ and $m, m' \in \mathbb{Z}$ such that ξ_1 is an endpoint of $I_d(m)$ and $\xi_2 \in I_d(m')$. It is a left side when ξ_1 is the left endpoint and a right side when ξ_1 is the right endpoint, and the rank of the column is then $2d-1$. Dually, if the row $\lfloor \xi_2/2 \rfloor$ has rank at most $h-2$, the quarter lies on a light horizontal wall if and only if ξ_2 is an endpoint of some $I_d(m')$ and $\xi_1 \in I_d(m)$ for some m; it is a bottom side at the left endpoint and a top side at the right endpoint.*

Proof. Write $\xi = 2x + \varsigma$ and let $r \leq h-2$ be the rank of the column x_1. With s_i as in Lemma 8.4, put $\hat{s}_i = (\xi_i - 2\varpi_0) \bmod 2^{r+3}$, so that $\hat{s}_i = 2s_i(x_i) + \varsigma_i$. By Lemma 8.4(a)–(c) and Definition 8.7, the quarter lies on a vertical red path if and only if three conditions hold: $s_2(x_2) \in [2^r, 3\cdot 2^r]$; $\varsigma_1 = 1$ if $s_1(x_1) = 2^r$ and $\varsigma_1 = 0$ if $s_1(x_1) = 3\cdot 2^r$; and $\varsigma_2 = 1$ if $s_2(x_2) = 2^r$, $\varsigma_2 = 0$ if $s_2(x_2) = 3\cdot 2^r$. Since $s_1(x_1) \in \{2^r, 3\cdot 2^r\}$, this says

$$\hat{s}_1 \in \{2^{r+1}+1,\ 3\cdot 2^{r+1}\} \quad \text{and} \quad \hat{s}_2 \in [2^{r+1}+1,\ 3\cdot 2^{r+1}],$$

with a left side exactly when $\hat{s}_1 = 2^{r+1}+1$. By Lemma 8.3 the path is light if and only if r is odd. For $r = 2d-1$ we have $2^{r+1} = 4^d$, $2^{r+3} = 4^{d+1}$ and $\hat{s}_i \equiv \xi_i - a_d \pmod{4^{d+1}}$, and the two conditions say that ξ_1 is an endpoint of a frame of depth d and that ξ_2 lies in a frame of depth d. Conversely, if ξ_1 is an endpoint of a frame of depth d, then $\xi_1 - 2\varpi_0 \equiv 4^d + 1$ or $3\cdot 4^d \pmod{4^{d+1}}$. Hence $x_1 - \varpi_0 \equiv 2^{2d-1}$ or $3\cdot 2^{2d-1} \pmod{2^{2d+1}}$, the column has rank $2d-1$, and the first part applies. Rows are the same, with the second half of Lemma 8.4. □

Flows and roles. Each edge of a tile of R_M carries a *flow*: a subset of $\{\text{left}, \text{right}\}$ on the west and east edges, and a subset of $\{\text{down}, \text{up}\}$ on the south and north edges.

Definition 8.10 (Routed scaffold)**.** The *routed scaffold* $\mathbb{A}^{\mathrm{r}}$ is the Wang set of tuples

$$(\varsigma, \mathfrak{a}, \phi_{\mathrm{W}}, \phi_{\mathrm{E}}, \phi_{\mathrm{S}}, \phi_{\mathrm{N}})$$

with $(\varsigma, \mathfrak{a}) \in \mathbb{A}^{\mathrm{q}}$, $\phi_{\mathrm{W}}, \phi_{\mathrm{E}} \subseteq \{\text{left}, \text{right}\}$ and $\phi_{\mathrm{S}}, \phi_{\mathrm{N}} \subseteq \{\text{down}, \text{up}\}$ that satisfy the *flow rules*:

(F1) if $(\varsigma, \mathfrak{a})$ lies on no light vertical wall, then $\phi_{\mathrm{W}} = \phi_{\mathrm{E}}$;

(F2) on a left side of a light vertical wall, $\phi_{\mathrm{W}} \neq \emptyset$ and $\phi_{\mathrm{E}} \subseteq \{\text{left}\}$;

(F3) on a right side of a light vertical wall, $\phi_{\mathrm{E}} \neq \emptyset$ and $\phi_{\mathrm{W}} \subseteq \{\text{right}\}$;

(F4) if $(\varsigma, \mathfrak{a})$ lies on no light horizontal wall, then $\phi_{\mathrm{S}} = \phi_{\mathrm{N}}$;

(F5) on a bottom side of a light horizontal wall, $\phi_S \neq \emptyset$ and $\phi_N \subseteq \{\text{down}\}$;

(F6) on a top side of a light horizontal wall, $\phi_N \neq \emptyset$ and $\phi_S \subseteq \{\text{up}\}$.

The colour of a tile on a side $X \in \{N, E, S, W\}$ is the pair of the colour of $(\varsigma, \mathfrak{a})$ in $\mathbb{A}^q$ on side X and the flow ϕ_X.

A corner quarter on a light wall obeys one horizontal and one vertical rule, and the two concern different edges. A tile is a *horizontal carrier* if $\phi_W = \phi_E = \emptyset$, and a *vertical carrier* if $\phi_S = \phi_N = \emptyset$. Its *role* is *inactive* if it is neither, *horizontal* if it is only a horizontal carrier, *vertical* if it is only a vertical carrier, and *active* if it is both, except that an active tile whose scaffold pair is $((1,1), \mathfrak{a})$, with $\mathfrak{a}$ an index-0 letter with dark ruler bits, has the *seed* role. The role is a function of the tile.

Proposition 8.11 (Walls force carriers)**.** *Let P be a legal pattern of $\mathbb{A}^r$ whose domain contains the sites (x, z) with $w_1 \leq x \leq w_2$, where $w_1 < w_2$. Suppose that the scaffold pairs of $P(w_1, z)$ and $P(w_2, z)$ lie on light vertical walls and that those of $P(x, z)$, $w_1 < x < w_2$, do not. Then:*

(a) the flows $\phi_E(P(w_1, z))$, $\phi_W(P(x, z))$ and $\phi_E(P(x, z))$ for $w_1 < x < w_2$, and $\phi_W(P(w_2, z))$ are all equal to one set ψ;

(b) $\psi = \emptyset$ if and only if (w_1, z) carries a left side and (w_2, z) a right side;

(c) for $w_1 < x < w_2$, the tile $P(x, z)$ is a horizontal carrier if and only if (w_1, z) carries a left side and (w_2, z) a right side. No tile on a light vertical wall is a horizontal carrier.

The same holds for a vertical segment between two quarters on light horizontal walls, with the flows ϕ_S, ϕ_N, bottom and top sides, and vertical carriers. It also holds for legal patterns of R_M (Definition 8.17), whose tiles carry scaffold pairs and flows with the same rules.

Proof. (a) Matching edges carry equal flows, and (F1) makes the west and east flows of each tile strictly between the walls equal. (b) The set ψ is the east flow of the tile at w_1 and the west flow of the tile at w_2. If w_1 carries a left side then $\psi \subseteq \{\text{left}\}$ by (F2); if a right side, $\psi \neq \emptyset$ by (F3). If w_2 carries a right side then $\psi \subseteq \{\text{right}\}$ by (F3); if a left side, $\psi \neq \emptyset$ by (F2). So $\psi = \emptyset$ when the walls are a left side and a right side in this order, since $\{\text{left}\} \cap \{\text{right}\} = \emptyset$. In the three other cases one of the walls forces $\psi \neq \emptyset$. (c) The first claim is (a) and (b). A tile on a left side has $\phi_W \neq \emptyset$, and a tile on a right side has $\phi_E \neq \emptyset$. The vertical case is the same with (F4)–(F6). □

Only the emptiness of a flow is forced; a nonempty flow may take several values. We use an elementary fact about nested intervals.

Lemma 8.12 (Laminar walls)**.** *Let $\mathcal{J}$ be a set of integer intervals with at least two points each, such that any two distinct members are disjoint or one lies in the interior of the other. Let E be the set of their endpoints. Then every $e \in E$ is an endpoint of exactly one member, so it is a left or a right endpoint. If $w_1 < w_2$ are consecutive elements of E, w_1 is a left endpoint and w_2 is a right endpoint, then $[w_1, w_2] \in \mathcal{J}$. Conversely, if a member I contains no other member, its endpoints are consecutive in E.*

Proof. If e were an endpoint of two members, they would meet, so one would lie in the interior of the other, and that one would not have e as an endpoint. Now let $w_1 = \min I$ and $w_2 = \max I'$ be consecutive, with $I, I' \in \mathcal{J}$, and suppose $I \neq I'$. If $I \cap I' = \emptyset$, then $\max I < \min I'$ because $\min I < \max I'$, so $\max I$ lies strictly between w_1 and w_2. If I lies in the interior of I', then $\max I$ lies strictly between w_1 and w_2. If I' lies in the interior of I, then

$\min I'$ does. Each case contradicts consecutiveness, so $I = I' = [w_1, w_2]$. Conversely, let I contain no other member, and let $I' \neq I$. If $I' \cap I = \emptyset$, no endpoint of I' lies in I. Otherwise I lies in the interior of I', since the reverse is excluded, and then both endpoints of I' lie outside I. So no element of E lies strictly between the endpoints of I. □

8.3 Frames and free coordinates

Fix frames as in Definition 8.8. For $d \geq 1$ and $m \in \mathbb{Z}$, the *free coordinates* of $I_d(m)$ are the elements of

$$\mathcal{F}_d(m) = \big\{ z \in \mathring{I}_d(m) : \, z \notin I_e(m') \text{ for all } 1 \leq e < d \text{ and } m' \in \mathbb{Z} \big\},$$

and its *centre* is $\mathrm{cen}_d(m) = a_d + (4m+2)4^d + 1$. Lower frames are excluded together with their endpoints.

Lemma 8.13 (Frames). *(a) Frames of the same depth are disjoint. If $e < d$, no endpoint of a frame of depth d lies in a frame of depth e. Hence two frames of different depths are disjoint, or the one of smaller depth lies in the interior of the other.*

(b) If $e < d$, every frame of depth d contains a frame of depth e.

(c) $|\mathcal{F}_d(m)| = 2^{d+2} - 2$. The centre $\mathrm{cen}_d(m)$ lies in $\mathcal{F}_d(m)$, and $\mathcal{F}_d(m)$ has exactly $2^{d+1} - 1$ elements smaller than $\mathrm{cen}_d(m)$ and $2^{d+1} - 2$ elements larger.

(d) Suppose $a_d \equiv 2/3 \pmod{4^{d+1}}$ in $\mathbb{Z}_2$ for every d. Then every integer lies in the interior of some frame.

Proof. (a) The frame $I_d(m)$ lies in $[a_d + 4m \cdot 4^d,\ a_d + 4(m+1)4^d)$, and these windows are disjoint. Let $e < d$. Then $a_d \equiv a_e \pmod{4^{e+1}}$ and 4^{e+1} divides 4^d, so the endpoints of a frame of depth d are congruent to $a_e + 1$ and a_e modulo 4^{e+1}. An integer z lies in a frame of depth e exactly when $(z - a_e) \bmod 4^{e+1} \in [4^e + 1, 3 \cdot 4^e]$, which excludes 0 and 1. If a frame I of depth e meets a frame I' of depth d and contains neither endpoint of I', then I lies in the interior of I'.

(b) Let $I_d(m) = [L, L + 2 \cdot 4^d - 1]$. Left endpoints of frames of depth e form one residue class modulo 4^{e+1}, so one of them, p, lies in $[L, L + 4^{e+1} - 1]$. Its frame $[p, p + 2 \cdot 4^e - 1]$ ends before $L + 4^{e+1} + 2 \cdot 4^e - 1 \leq L + 2 \cdot 4^d - 1$, because $4^{e+1} + 2 \cdot 4^e \leq 2 \cdot 4^{e+1} \leq 2 \cdot 4^d$.

(c) Call a residue $s \in [0, 4^d)$ *allowed* if $s \bmod 4^{e+1} \notin [4^e + 1, 3 \cdot 4^e]$ for every $1 \leq e < d$, and let N_d be the number of allowed residues. Since $a_e \equiv a_d \pmod{4^{e+1}}$, an integer z lies in no frame of depth below d exactly when $(z - a_d) \bmod 4^d$ is allowed. Clearly $N_1 = 4$. Write $s \in [0, 4^{d+1})$ as $s = t \cdot 4^d + s'$ with $t \in [0, 4)$ and $s' \in [0, 4^d)$; the conditions for $e < d$ concern only s', and the condition for $e = d$ is $s \notin [4^d + 1, 3 \cdot 4^d]$. For $t = 0$ all N_d allowed s' remain; for $t = 1$ only $s' = 0$; for $t = 2$ none; for $t = 3$ all allowed s' except $s' = 0$. The residue 0 is allowed, so $N_{d+1} = N_d + 1 + 0 + (N_d - 1) = 2N_d$, and $N_d = 2^{d+1}$.

Now write $z = a_d + 4m \cdot 4^d + t \cdot 4^d + s'$. Then $z \in \mathring{I}_d(m)$ exactly when $t = 1$ and $s' \geq 2$, or $t = 2$. The residues $s' = 0$ and $s' = 1$ are both allowed, since $1 < 4^e + 1$. So $\mathcal{F}_d(m)$ has $N_d - 2$ elements with $t = 1$ and N_d elements with $t = 2$, in total $2^{d+2} - 2$. The centre has $t = 2$ and $s' = 1$, so it is free. The free coordinates below it are the $N_d - 2$ with $t = 1$ and the one with $t = 2$, $s' = 0$; those above it are the $N_d - 2$ with $t = 2$, $s' \geq 2$.

(d) In $\mathbb{Z}_2$ we have $-1/3 = \sum_{l \geq 0} 4^l$, so $-2/3 = \sum_{l \geq 0} 2 \cdot 4^l$ has every base-4 digit equal to 2. Adding a nonnegative integer z changes only finitely many digits, because a carry into a digit 2 stops there; subtracting a positive integer changes only finitely many digits, because a

borrow from a digit 2 stops there. So $z - 2/3$ has base-4 digit 2 in every position $d \geq d_0$. For such $d \geq 1$, $(z - a_d) \bmod 4^{d+1} \in [2 \cdot 4^d, 3 \cdot 4^d - 1]$, so z lies in the interior of a frame of depth d. □

When every integer lies in a frame, the *owner depth* $\mathrm{own}(z)$ of an integer z is the least $d \geq 1$ such that z lies in a frame of depth d. By definition, $z \in \mathcal{F}_d(m)$ if and only if $z \in \mathring{I}_d(m)$ and $\mathrm{own}(z) = d$. The centres are the only free coordinates with residue $a_d + 2 \cdot 4^d + 1$ modulo 4^{d+1}.

8.4 A routing model on the plane

We now take the phase $(1/3, 1/3)$. Both coordinates are 2-adic integers that are not integers, so every column and row has a finite rank, and $2\varpi_0 = 2/3$.

Lemma 8.14 (Routing on the plane)**.** *Let (f_m) be a tower of height ∞ with phase $(1/3, 1/3)$ (Lemma 8.2). Take the frames with $a_d \equiv 2/3 \pmod{4^{d+1}}$, and let $\mathcal{Y} = f_0^{\mathrm{q}}$, a tiling of $\mathbb{A}^{\mathrm{q}}$.*

(a) There are flows on the edges of $\mathbb{Z}^2$ such that $\mathcal{Y}$ together with these flows is a tiling ϑ of $\mathbb{A}^{\mathrm{r}}$.

(b) Let ϑ be any tiling of $\mathbb{A}^{\mathrm{r}}$ with scaffold pairs $\mathcal{Y}$. Then $\vartheta(\xi)$ is a horizontal carrier if and only if $\xi_1 \in \mathring{I}_d(m)$ for $d = \mathrm{own}(\xi_2)$ and some m, and it is a vertical carrier if and only if $\xi_2 \in \mathring{I}_{d'}(m')$ for $d' = \mathrm{own}(\xi_1)$ and some m'.

(c) In such a tiling, $\vartheta(\xi)$ is active or seed if and only if $\xi \in \mathcal{F}_d(m) \times \mathcal{F}_d(m')$ for some $d \geq 1$ and $m, m' \in \mathbb{Z}$. Every seed tile sits at a pair of centres $(\mathrm{cen}_d(m), \mathrm{cen}_d(m'))$.

Proof. By Lemma 8.13(d) the owner depth is defined everywhere. Fix a row ξ_2, let D be the set of depths d such that ξ_2 lies in a frame of depth d, and let $d_0 = \mathrm{own}(\xi_2) = \min D$. Let $\mathcal{J}$ be the set of all frames of the depths in D. By Lemma 8.9, applied with $h = \infty$, the quarters of the row ξ_2 that lie on light vertical walls are exactly the endpoints of the members of $\mathcal{J}$, left endpoints carrying left sides and right endpoints right sides. By Lemma 8.13(a), $\mathcal{J}$ satisfies the hypothesis of Lemma 8.12. Its members that contain no other member are exactly the frames of depth d_0: a frame of a larger depth in $\mathcal{J}$ contains a frame of depth d_0 by Lemma 8.13(b), and a member inside a frame of depth d_0 would have a smaller depth by Lemma 8.13(a), which D does not contain. The frames of depth d_0 are unbounded in both directions, so every horizontal edge of the row lies between two consecutive walls.

(a) Each horizontal edge of the row lies between a nearest wall to its west and a nearest wall to its east. Give it the flow $\emptyset$ if these are a left side and a right side, $\{\mathrm{left}\}$ if both are left sides, and $\{\mathrm{right}\}$ if the western one is a right side. A tile on no light vertical wall has both edges between the same two walls, so (F1) holds. At a left side, the edge to the west ends at a left side, so its flow is $\{\mathrm{left}\}$ or $\{\mathrm{right}\}$, and the edge to the east starts at a left side, so its flow is $\emptyset$ or $\{\mathrm{left}\}$; this is (F2). At a right side, the edge to the east starts at a right side and has flow $\{\mathrm{right}\}$, and the edge to the west has flow $\emptyset$ or $\{\mathrm{right}\}$; this is (F3). Columns are treated in the same way with (F4)–(F6). Matching edges carry one flow, and $\mathcal{Y}$ is legal by Lemma 8.6(b), so ϑ is a tiling.

(b) If ξ_1 is a wall of the row, $\vartheta(\xi)$ is not a horizontal carrier by Proposition 8.11(c), and ξ_1 lies in the interior of no frame of depth d_0, because the endpoints of such a frame are consecutive walls (Lemma 8.12). Otherwise let $w_1 < \xi_1 < w_2$ be the nearest walls. By Proposition 8.11(c), $\vartheta(\xi)$ is a horizontal carrier if and only if w_1 is a left side and w_2 a right side. By Lemma 8.12 this holds if and only if $[w_1, w_2]$ is a member of $\mathcal{J}$ whose endpoints are consecutive, that is a member containing no other member, that is a frame of depth d_0. The vertical claim is the same along the column ξ_1.

(c) If $\vartheta(\xi)$ is both a horizontal and a vertical carrier, then ξ_1 lies in a frame of depth $\mathrm{own}(\xi_2)$, so $\mathrm{own}(\xi_1) \le \mathrm{own}(\xi_2)$, and symmetrically $\mathrm{own}(\xi_2) \le \mathrm{own}(\xi_1)$. With d the common value, $\xi_1 \in \mathring{I}_d(m)$ and $\mathrm{own}(\xi_1) = d$, that is $\xi_1 \in \mathcal{F}_d(m)$; likewise $\xi_2 \in \mathcal{F}_d(m')$. The converse follows from (b).

Let $\vartheta(\xi)$ be a seed tile, with $\xi = 2x + (1,1)$ and $f_0(x)$ an index-0 letter with dark bits. Its thick type $\angle\mathfrak{b}$ has red north endpoint r_0. Let r be the rank of the column x_1. By Lemma 8.4(b),(c), $s_1(x_1) = 2^r$ and $s_2(x_2) = 2^r$. By Lemma 8.3, r is even. If $r = 0$, then $\varsigma_{0,1}(x) = 1$ and $f_0(x)$ is a child of thin type $\mathfrak{c}$ or $\mathfrak{b}$, not $\mathfrak{a}$. So $r = 2d$ with $d \ge 1$, and $x_i - 1/3 \equiv 4^d \pmod{2^{2d+2}}$ for $i = 1,2$. Hence $\xi_i - 2/3 \equiv 2 \cdot 4^d + 1 \pmod{4^{d+1}}$, so ξ_i is a centre of a frame of depth d. □

Distinct seed tiles therefore never share a product $\mathcal{F}_d(m) \times \mathcal{F}_d(m')$: each product contains at most one pair of centres.

8.5 The initialized board in every word

The refutation of Section 9 cannot use a plane tiling. It reads one finite word $\sigma^n(\mathfrak{a})$, with $\mathfrak{a}$ unknown, and needs a complete initialized board inside it. Every such word has one, at the same place.

Proposition 8.15 (The initialized board)**.** *Let $n \ge 2$, $\mathfrak{a} \in \mathbb{A}$ and $W = \sigma^n(\mathfrak{a})$, a pattern on $[0,4^n)^2$. Put $\Delta_n = 4^{n-1}$, $J = [\Delta_n + 1, 3\Delta_n]$ and $J^\circ = [\Delta_n + 2, 3\Delta_n - 1]$. Let $\mathcal{F}$ be the set of $z \in J^\circ$ that lie in no interval $[(4m+1)4^e + 1, (4m+3)4^e]$ with $1 \le e \le n-2$ and $m \in \mathbb{Z}$; these are the* free coordinates*. Then:*

(a) $2\Delta_n + 1 \in \mathcal{F}$ and $|\mathcal{F}| = 2^{n+1} - 2$; exactly $2^n - 1$ elements of $\mathcal{F}$ are smaller than $2\Delta_n + 1$ and $2^n - 2$ are larger. In particular at least $2^n - 1$ elements of $\mathcal{F}$ are $\ge 2\Delta_n + 1$.

(b) The quarter of W at $(2\Delta_n + 1, 2\Delta_n + 1)$ is $((1,1), \mathfrak{b})$, where $\mathfrak{b}$ is an index-0 letter with dark ruler bits.

(c) Let $Z_W = (\mathcal{F} \times J^\circ) \cup (J^\circ \times \mathcal{F})$ and $\xi \in Z_W$. The quarter of W at ξ lies on no light wall. In the row of ξ, the nearest quarters of W to the left and to the right of ξ that lie on a light vertical wall both lie in $J \times \{\xi_2\}$, and they are a left side and a right side, in this order, if and only if $\xi_2 \in \mathcal{F}$. In the column of ξ, the nearest quarters below and above ξ that lie on a light horizontal wall both lie in $\{\xi_1\} \times J$, and they are a bottom side and a top side, in this order, if and only if $\xi_1 \in \mathcal{F}$.

(d) Let P be a legal pattern of $\mathbb{A}^{\mathrm{r}}$ or of R_M on $J \times J$ whose scaffold pairs are the quarters of W. For $\xi \in Z_W$, $P(\xi)$ is a horizontal carrier if and only if $\xi_2 \in \mathcal{F}$, and a vertical carrier if and only if $\xi_1 \in \mathcal{F}$. So the role of $P(\xi)$ is seed at $(2\Delta_n + 1, 2\Delta_n + 1)$, active at the other sites of $\mathcal{F} \times \mathcal{F}$, horizontal if $\xi_2 \in \mathcal{F}$ and $\xi_1 \notin \mathcal{F}$, and vertical if $\xi_1 \in \mathcal{F}$ and $\xi_2 \notin \mathcal{F}$.

None of this depends on $\mathfrak{a}$.

Proof. Use the tower $f_m = (\varphi')^{2n-m}(\mathfrak{a})$ of height $h = 2n$, with all digits 0 and phase 0, so that $f_0 = W$ and $U_0 = [0,4^n)^2$. Take the frames with $a_d = 0$. Then $J = I_{n-1}(0)$, $\mathcal{F} = \mathcal{F}_{n-1}(0)$ and $2\Delta_n + 1 = \mathrm{cen}_{n-1}(0)$, so (a) is Lemma 8.13(c) with $d = n-1$.

Walls in the board. Let $\xi \in J \times J$. Its index coordinates $\lfloor \xi_i/2 \rfloor$ lie in $[2^{2n-3}, 3 \cdot 2^{2n-3}]$, so they are positive and below 2^{2n-1}. With phase 0, (39) shows that the rank of an index coordinate $x \in (0,4^n)$ is the 2-adic valuation of x; here it is at most $2n-2 = h-2$. So Lemma 8.9 applies to ξ in both directions. A light wall at ξ comes from a frame of a depth e with $2e - 1 \le 2n - 2$, that is $e \le n-1$; so every light wall in the board has rank at most $2n-3$, and it is determined by the word. A frame of depth $n-1$ that meets J is J itself, and a frame of smaller depth that

meets J lies in $\mathring{J}$, by Lemma 8.13(a). Hence, for $\xi_2 \in J$, the quarters (z, ξ_2) with $z \in J$ that lie on light vertical walls are exactly the endpoints of the members of

$$\mathcal{J}(\xi_2) = \{J\} \cup \{I_e(m) \subseteq J : \ e < n-1, \ \xi_2 \text{ lies in a frame of depth } e\},$$

with left endpoints carrying left sides and right endpoints right sides. The same holds for columns. This family satisfies the hypothesis of Lemma 8.12.

(b) The letter $W(x)$ at $x = (\Delta_n, \Delta_n)$ gives the quarter at $2x + (1,1)$. The column $x_1 = 2^{2n-2}$ has rank $r = 2n-2$, and $s_1(x_1) = s_2(x_2) = 2^r$ in Lemma 8.4. By Lemma 8.4(b),(c) the thick type of $W(x)$ is the corner with red north endpoint r_0, which is $\angle\mathsf{b}$. Since $\varsigma_0(x) = (0,0)$, the letter is a child $(0,0)$ and has thin type a. Every corner of $\mathbb{O}$ has black type A (Table 5). By Lemma 8.3 both bits equal $r \bmod 2 = 0$.

(c) Let $\xi \in Z_W$ and suppose first $\xi_1 \in \mathcal{F}$. A light vertical wall at ξ would make ξ_1 an endpoint of a frame of depth $e \le n-1$ meeting J. For $e < n-1$ this puts ξ_1 in a lower frame, which $\mathcal{F}$ excludes; for $e = n-1$ the frame is J, whose endpoints are not in J°. A light horizontal wall at ξ would need ξ_1 in a frame of depth e and ξ_2 an endpoint of a frame of the same depth. For $e < n-1$ the first is excluded; for $e = n-1$ the second frame is J and $\xi_2 \in J^\circ$ is not its endpoint. The case $\xi_2 \in \mathcal{F}$ is symmetric. So no light wall lies at ξ.

The endpoints $\Delta_n + 1$ and $3\Delta_n$ of J carry a left and a right side in every row $\xi_2 \in J$, and $\Delta_n + 1 < \xi_1 < 3\Delta_n$. So the nearest walls $w_1 < \xi_1 < w_2$ in the row exist and lie in J. If $\xi_2 \in \mathcal{F}$, no frame of depth below $n-1$ contains ξ_2, so $\mathcal{J}(\xi_2) = \{J\}$, and the nearest walls are $\Delta_n + 1$ and $3\Delta_n$: a left side and a right side. If $\xi_2 \notin \mathcal{F}$, then $\xi_1 \in \mathcal{F}$, and ξ_2 lies in a frame of some depth $e < n-1$. Suppose w_1 were a left side and w_2 a right side. By Lemma 8.12, $[w_1, w_2]$ is a member of $\mathcal{J}(\xi_2)$. It is not J, because J contains a frame of depth e (Lemma 8.13(b)), which is a member, and whose endpoints lie strictly between those of J. So $[w_1, w_2]$ is a frame of depth below $n-1$ containing ξ_1, which contradicts $\xi_1 \in \mathcal{F}$. The column statement is the same with rows and columns exchanged.

(d) By (c), the segment from ξ to its nearest walls in its row lies in $J \times \{\xi_2\}$, and Proposition 8.11(c) gives the horizontal claim; the vertical one is the same. Hence $P(\xi)$ is active or seed on $\mathcal{F} \times \mathcal{F}$, and its role is as stated off $\mathcal{F} \times \mathcal{F}$. At $(2\Delta_n + 1, 2\Delta_n + 1)$ it is seed by (b). Let $\xi \in \mathcal{F} \times \mathcal{F}$ be a quarter $((1,1), W(x))$ with $W(x)$ an index-0 letter with dark bits. As in the proof of Lemma 8.14(c), the column x_1 has an even rank $r = 2e \ge 2$ and $s_1(x_1) = s_2(x_2) = 2^r$; here $r \le 2n-2$, so $e \le n-1$. With phase 0 this gives $\xi_i \equiv 2 \cdot 4^e + 1 \pmod{4^{e+1}}$, so ξ_i is the centre of a frame of depth e. For $e < n-1$, ξ_i would lie in a lower frame; so $e = n-1$, and the only centre of depth $n-1$ in J is $2\Delta_n + 1$. Thus the seed role occurs on $\mathcal{F} \times \mathcal{F}$ only at $(2\Delta_n + 1, 2\Delta_n + 1)$. □

Every role in (d) is certified by two straight corridors from ξ to walls inside $J \times J$, together with the flow rules along them. The corridors may be long, but each step involves only two adjacent quarters. Section 9 walks along them with a bounded number of tokens.

8.6 The pointed Horn layer

Machines. A Turing machine M has finitely many states, among them an initial state and a set of halting states, a finite tape alphabet with a blank symbol $\sqcup$, and a transition function. It starts on a blank two-sided tape and *halts* if it enters a halting state. A *one-sided* machine has tape cells $1, 2, \ldots$, starts with its head on cell 1, moves its head one cell to the left or to the right at every step, and never moves left from cell 1. A standard simulation computes from M a one-sided machine M° with a nonhalting initial state that halts if and only if M halts: M°

keeps the two halves of the tape of M on two tracks, marks cell 1 at its first step, and performs a stay of M, or a change of track at cell 1, as a step to the right followed by a step to the left; Section B gives its transition table and proves these properties. Let $T(M) \in \{1, 2, \dots\} \cup \{\infty\}$ be the number of steps M° makes before it enters a halting state; we call $T(M)$ the *halting time* of M.

A cellular automaton. Let Γ, Q and tr be the tape alphabet, the state set and the transition function of M°, with initial state q_0. Let $\mathcal{S}_M = \Gamma \cup (Q \times \Gamma) \cup \{w\}$, where w marks the wall at cell 0. Define next: $\mathcal{S}_M^3 \to \mathcal{S}_M$ by

- $\mathrm{next}(a, w, c) = w$;
- $\mathrm{next}(a, (q, \gamma), c) = \gamma'$ if q is not halting and $\mathrm{tr}(q, \gamma) = (q', \gamma', \pm 1)$;
- for $\gamma \in \Gamma$: $\mathrm{next}((q, \gamma_a), \gamma, c) = (q', \gamma)$ if q is not halting and $\mathrm{tr}(q, \gamma_a) = (q', \cdot, +1)$; otherwise $\mathrm{next}(a, \gamma, (q, \gamma_c)) = (q', \gamma)$ if q is not halting and $\mathrm{tr}(q, \gamma_c) = (q', \cdot, -1)$;
- $\mathrm{next}(a, b, c) = b$ in all other cases.

Let $\mathcal{C}_t \colon \{0, 1, 2, \dots\} \to \mathcal{S}_M$ be given by $\mathcal{C}_t(0) = w$ for all t, $\mathcal{C}_0(1) = (q_0, \sqcup)$, $\mathcal{C}_0(x) = \sqcup$ for $x \geq 2$, and $\mathcal{C}_{t+1}(x) = \mathrm{next}(\mathcal{C}_t(x-1), \mathcal{C}_t(x), \mathcal{C}_t(x+1))$ for $x \geq 1$. For $t \leq T(M)$, $\mathcal{C}_t$ is the configuration of M° at time t: the head cell holds the pair of the state and the scanned symbol, and the other cells hold their symbols. Indeed, in a configuration with one head, the head cell writes, the neighbour in the direction of the move receives the head, and every other cell keeps its symbol; the head never reaches cell 0. If M halts, the head at time $T = T(M)$ is at a cell $x_T \leq T + 1$, and $\mathcal{C}_T(x_T) = (q, \gamma)$ with q halting.

Facts and clauses. Let $\mathcal{H}_M = \{\mathsf{Seed}, \mathsf{Start}, \mathsf{Ray}\} \cup \{\mathsf{P}_b : b \in \mathcal{S}_M\}$. A *fact field* on a set $V \subseteq \mathbb{Z}^2$ is a map $\mathsf{G} \colon V \to 2^{\mathcal{H}_M}$; we write $\mathsf{F}(\xi)$ for $\mathsf{F} \in \mathsf{G}(\xi)$. The clauses, for every $\xi \in \mathbb{Z}^2$, are:

$$
\begin{array}{ll}
\text{(H1) } \mathsf{Seed}(\xi) \to \mathsf{P}_w(\xi), & \text{(H2) } \mathsf{Seed}(\xi) \to \mathsf{Start}(\xi + \mathbf{e}_1), \\
\text{(H3) } \mathsf{Start}(\xi) \to \mathsf{P}_{(q_0, \sqcup)}(\xi), & \text{(H4) } \mathsf{Start}(\xi) \to \mathsf{Ray}(\xi + \mathbf{e}_1), \\
\text{(H5) } \mathsf{Ray}(\xi) \to \mathsf{P}_\sqcup(\xi), & \text{(H6) } \mathsf{Ray}(\xi) \to \mathsf{Ray}(\xi + \mathbf{e}_1), \\
\text{(H7) } \mathsf{P}_w(\xi) \to \mathsf{P}_w(\xi + \mathbf{e}_2), & \\
\text{(H8) } \mathsf{P}_a(\xi - \mathbf{e}_1) \wedge \mathsf{P}_b(\xi) \wedge \mathsf{P}_c(\xi + \mathbf{e}_1) \to \mathsf{P}_{\mathrm{next}(a,b,c)}(\xi + \mathbf{e}_2) \quad (a, b, c \in \mathcal{S}_M), & \\
\text{(H9) } \neg \mathsf{P}_b(\xi) \quad \text{for every } b = (q, \gamma) \text{ with } q \text{ halting}. &
\end{array}
$$

Each clause has at most one positive literal, so it is a Horn clause, and the sites it mentions lie in $\xi + \Lambda$, with Λ as in Section 4. A fact field on V *satisfies* a clause instance if the instance mentions only sites of V and holds there. A *model* is a fact field on $\mathbb{Z}^2$ that satisfies every instance.

The Wang presentation. A *legal fact block* is a fact field Θ on Λ that satisfies every clause instance mentioning only sites of Λ. Let $\mathcal{U}_M = \{h, v\} \times \Lambda \times \mathcal{H}_M$ and $C_M = 2^{\mathcal{U}_M}$, with $\wedge$ the intersection of sets. The tile of Θ is $\mathrm{tile}(\Theta) = (\Theta_W, \Theta_E, \Theta_S, \Theta_N) \in C_M^4$, where

$$
\begin{aligned}
\Theta_E &= \{(h, y, \mathsf{F}) : y \in \Lambda,\ y_1 \geq 0,\ \mathsf{F} \in \Theta(y)\}, \\
\Theta_W &= \{(h, y + \mathbf{e}_1, \mathsf{F}) : y \in \Lambda,\ y_1 \leq 0,\ \mathsf{F} \in \Theta(y)\}, \\
\Theta_N &= \{(v, y, \mathsf{F}) : y \in \Lambda,\ y_2 \geq 0,\ \mathsf{F} \in \Theta(y)\}, \\
\Theta_S &= \{(v, y + \mathbf{e}_2, \mathsf{F}) : y \in \Lambda,\ y_2 \leq 0,\ \mathsf{F} \in \Theta(y)\}.
\end{aligned}
$$

T_M is the Wang set of the tiles of all legal fact blocks, with these colours. Thus $\mathrm{tile}(\Theta)_{\mathrm{E}} = \mathrm{tile}(\Theta')_{\mathrm{W}}$ exactly when $\Theta(y) = \Theta'(y - \mathbf{e}_1)$ for every $y \in \Lambda$ with $y_1 \geq 0$: the two blocks agree on their overlap of two columns. The same holds vertically. Since the east and west colours together record all of Θ, the map $\Theta \mapsto \mathrm{tile}(\Theta)$ is injective.

The *least field* G_M on $\mathbb{Z}^2$ is $\mathsf{G}_M(0,0) = \{\mathsf{Seed}, \mathsf{P}_{\mathsf{w}}\}$, $\mathsf{G}_M(1,0) = \{\mathsf{Start}, \mathsf{P}_{(\mathsf{q}_0,\sqcup)}\}$, $\mathsf{G}_M(x,0) = \{\mathsf{Ray}, \mathsf{P}_{\sqcup}\}$ for $x \geq 2$, $\mathsf{G}_M(0,t) = \{\mathsf{P}_{\mathsf{w}}\}$ for $t \geq 1$, $\mathsf{G}_M(x,t) = \{\mathsf{P}_{\mathcal{C}_t(x)}\}$ for $x, t \geq 1$, and $\mathsf{G}_M(\xi) = \emptyset$ at all other sites. The *seed tile* is $s_M = \mathrm{tile}(\mathsf{G}_M|_\Lambda)$.

Proposition 8.16 (The pointed Horn layer)**.** *The map $M \mapsto (T_M, s_M)$ is total computable, C_M is a finite set with the associative, commutative and idempotent operation $\wedge$, and:*

(a) $T_M \subseteq C_M^4$ is closed under coordinatewise $\wedge$, and $s_M \in T_M$. The palettes C_M^4, $\{\mathbf{c} \in C_M^4 : \mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}\}$, $\{\mathbf{c} \in C_M^4 : \mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}}\}$, T_M and $\{s_M\}$ are all closed under $\wedge$.

(b) If M does not halt, T_M has a tiling Υ with $\Upsilon(0,0) = s_M$.

(c) If M halts, with halting time T, then for every $N \geq 2T + 2$ there is no legal pattern Ξ of T_M on $[0,N)^2$ with $\Xi(0,0) = s_M$. In particular there is none on $[0, 2T + 10)^2$.

Proof. Computability. M°, $\mathcal{S}_M$, next and $\mathcal{H}_M$ are computed from M. Legal fact blocks are found by checking the finitely many clause instances inside Λ for each of the finitely many fact fields on Λ, and s_M needs only $\mathcal{C}_0$ and $\mathcal{C}_1$ on cells $0, 1, 2$.

(a) Let Θ, Θ' be legal blocks and $\Theta \cap \Theta'$ their pointwise intersection. If the antecedents of an instance of (H1)–(H8) hold in $\Theta \cap \Theta'$, they hold in Θ and in Θ', so the consequent holds in both and in $\Theta \cap \Theta'$. A fact forbidden by (H9) is absent from Θ, hence from $\Theta \cap \Theta'$. So $\Theta \cap \Theta'$ is legal, and $\mathrm{tile}(\Theta \cap \Theta') = \mathrm{tile}(\Theta) \wedge \mathrm{tile}(\Theta')$ because every colour is a set of facts read off Θ. Thus T_M is closed under $\wedge$. For the palettes: C_M^4 is closed, $\{s_M\}$ is closed by idempotence, and if $\mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}$ and $\mathbf{c}'_{\mathrm{W}} = \mathbf{c}'_{\mathrm{E}}$ then $(\mathbf{c} \wedge \mathbf{c}')_{\mathrm{W}} = (\mathbf{c} \wedge \mathbf{c}')_{\mathrm{E}}$; the vertical palette is the same.

We show that G_M satisfies every instance of (H1)–(H8). The instances of (H1)–(H6) hold by the definition of row 0. (H7) holds because P_{w} occurs exactly at the sites $(0,t)$, $t \geq 0$. In (H8) the three antecedents hold only if $\xi = (x,t)$ with $x \geq 1$ and $t \geq 0$, since no site with negative coordinates carries a fact; then the antecedents are $\mathsf{P}_{\mathcal{C}_t(x-1)}$, $\mathsf{P}_{\mathcal{C}_t(x)}$, $\mathsf{P}_{\mathcal{C}_t(x+1)}$, and the consequent $\mathsf{P}_{\mathcal{C}_{t+1}(x)}$ holds at $(x, t+1)$. The block $\mathsf{G}_M|_\Lambda$ carries P-facts only for $\mathcal{C}_0(0) = \mathcal{C}_1(0) = \mathsf{w}$, $\mathcal{C}_0(1) = (\mathsf{q}_0, \sqcup)$ and $\mathcal{C}_1(1)$. The state q_0 is not halting, and $\mathcal{C}_1(1) \in \Gamma$ because the head leaves cell 1 at the first step. So (H9) holds on Λ, and $s_M \in T_M$.

(b) If M does not halt, no $\mathcal{C}_t(x)$ contains a halting state, so G_M also satisfies (H9) and is a model. Put $\Upsilon(u) = \mathrm{tile}(\Theta_u)$ with $\Theta_u(y) = \mathsf{G}_M(u + y)$ for $y \in \Lambda$. Each Θ_u is legal, and neighbouring blocks agree on their overlaps, so Υ is a tiling with $\Upsilon(0,0) = s_M$.

(c) Let $\Xi(u) = \mathrm{tile}(\Theta_u)$, $u \in [0,N)^2$, be a legal pattern with $\Xi(0,0) = s_M$. For a site ξ, the set of $u \in [0,N)^2$ with $\xi - u \in \Lambda$ is a product of two integer intervals; along each step inside it the two blocks agree at ξ, by the overlap property. So $\mathsf{G}(\xi) = \Theta_u(\xi - u)$ is a well defined fact field on $[-1,N]^2$, and G satisfies every instance that mentions only sites of $u + \Lambda$ for some $u \in [0,N)^2$. Every instance used below has this property with $u = \xi$ for its site $\xi \in [0, 2T+1] \times [0,T] \subseteq [0,N)^2$.

Since $s_M = \mathrm{tile}(\mathsf{G}_M|_\Lambda)$, $\mathsf{Seed}(0,0)$ holds in G. By (H1)–(H6) at the sites $(x,0)$ with $0 \leq x \leq 2T+1$, G contains $\mathsf{P}_{\mathsf{w}}(0,0)$ and $\mathsf{P}_{\mathcal{C}_0(x)}(x,0)$ for $1 \leq x \leq 2T+1$. By (H7), $\mathsf{P}_{\mathsf{w}}(0,t)$ holds for $0 \leq t \leq T$. By induction on $t \leq T$, $\mathsf{P}_{\mathcal{C}_t(x)}(x,t)$ holds for $1 \leq x \leq 2T+1-t$: the instance of (H8) at (x,t), with $x \leq 2T - t$, has its antecedents at $(x-1,t)$, (x,t) and $(x+1,t)$, which hold by the induction hypothesis and by $\mathsf{P}_{\mathsf{w}}(0,t)$. At $t = T$ this covers $1 \leq x \leq T+1$, and in particular x_T. So $\mathsf{P}_{\mathcal{C}_T(x_T)}(x_T, T)$ holds with a halting state, which violates (H9) at (x_T, T). Hence no such Ξ exists. Since $2T + 10 \geq 2T + 2$, the last claim follows. □

A whole-plane model always exists, the empty one; only the seed forces the computation. Relations closed under a semilattice operation such as $\wedge$ form a classical tractable class of constraints, decided by local consistency [25, 17, 2]. Section 9 uses this through the closure in (a).

8.7 The family

Definition 8.17 (The family R_M)**.** For a Turing machine M, R_M is the Wang set of tuples

$$\vartheta = \big(\mathrm{quad}(\vartheta), \mathrm{par}(\vartheta), \phi_{\mathrm{W}}(\vartheta), \phi_{\mathrm{E}}(\vartheta), \phi_{\mathrm{S}}(\vartheta), \phi_{\mathrm{N}}(\vartheta), \mathrm{pay}(\vartheta)\big)$$

such that the first six entries form a tile of the routed scaffold $\mathbb{A}^{\mathrm{r}}$ (Definition 8.10), and the payload $\mathrm{pay}(\vartheta) = (\mathbf{c}_{\mathrm{W}}, \mathbf{c}_{\mathrm{E}}, \mathbf{c}_{\mathrm{S}}, \mathbf{c}_{\mathrm{N}}) \in C_M^4$ lies in the palette of the role of ϑ:

$$\begin{array}{lll} \text{inactive: } C_M^4, & \text{horizontal: } \{\mathbf{c} : \mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}\}, & \text{vertical: } \{\mathbf{c} : \mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}}\}, \\ \text{active: } T_M, & \text{seed: } \{s_M\}. & \end{array}$$

The colour of ϑ on a side X is the pair of the colour of its $\mathbb{A}^{\mathrm{r}}$-tile on side X and $\mathbf{c}_X$.

So matching tiles of R_M have matching scaffold pairs (Lemma 8.6), equal flows on the shared edge, and equal payload colours on the shared edge. The scaffold $\mathbb{A}$, the quarters, the walls, the flow rules and the roles are fixed; only C_M, T_M and s_M depend on M. The map $M \mapsto R_M$ is total computable: $\mathbb{A}^{\mathrm{r}}$ is a fixed finite set, and R_M is obtained by listing, for each tile of $\mathbb{A}^{\mathrm{r}}$, the payloads in its palette, using Proposition 8.16.

Proposition 8.18 (Nonhalting members tile)**.** *If M does not halt, then R_M tiles.*

Proof. Let (f_m) be a tower with phase $(1/3, 1/3)$ (Lemma 8.2), take the frames with $a_d \equiv 2/3 \pmod{4^{d+1}}$, and let ϑ^{r} be a tiling of $\mathbb{A}^{\mathrm{r}}$ with scaffold pairs f_0^{q}, as in Lemma 8.14(a). By Proposition 8.16(b) there is a tiling Y of T_M with $\mathrm{Y}(0,0) = s_M$.

Addresses. For $d \geq 1$ and $m \in \mathbb{Z}$ let $\mathrm{ad}_{d,m} \colon \mathcal{F}_d(m) \to \mathbb{Z}$ be the increasing map onto an interval of integers with $\mathrm{ad}_{d,m}(\mathrm{cen}_d(m)) = 0$. For $z \in \mathring{I}_d(m)$ let $\mathrm{ad}^-_{d,m}(z) = \mathrm{ad}_{d,m}(z')$, where z' is the largest element of $\mathcal{F}_d(m)$ with $z' \leq z$, and $\mathrm{ad}^-_{d,m}(z) = \mathrm{ad}_{d,m}(\min \mathcal{F}_d(m)) - 1$ if there is none. Then $\mathrm{ad}^-_{d,m} = \mathrm{ad}_{d,m}$ on $\mathcal{F}_d(m)$; and for $z, z+1 \in \mathring{I}_d(m)$, $\mathrm{ad}^-_{d,m}(z+1) = \mathrm{ad}^-_{d,m}(z)$ if $z+1 \notin \mathcal{F}_d(m)$, while $\mathrm{ad}^-_{d,m}(z+1) = \mathrm{ad}^-_{d,m}(z) + 1$ if $z+1 \in \mathcal{F}_d(m)$.

Payloads. We give each site ξ a payload $\mathrm{pay}(\xi)$ according to the role of $\vartheta^{\mathrm{r}}(\xi)$, described by Lemma 8.14(b),(c). We drop the indices of ad and ad^-, which are determined by the frames containing the coordinates.

- Active or seed, $\xi \in \mathcal{F}_d(m) \times \mathcal{F}_d(m')$: $\mathrm{pay}(\xi) = \mathrm{Y}(\mathrm{ad}(\xi_1), \mathrm{ad}(\xi_2))$.
- Horizontal, with $d = \mathrm{own}(\xi_2)$ and $\xi_1 \in \mathring{I}_d(m)$: if $\xi_2 \in \mathcal{F}_d(m')$ for some m', put $\mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}} = \mathrm{Y}(\mathrm{ad}^-(\xi_1), \mathrm{ad}(\xi_2))_{\mathrm{E}}$; otherwise put $\mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}} = \emptyset$.
- Vertical, with $d = \mathrm{own}(\xi_1)$ and $\xi_2 \in \mathring{I}_d(m')$: if $\xi_1 \in \mathcal{F}_d(m)$ for some m, put $\mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}} = \mathrm{Y}(\mathrm{ad}(\xi_1), \mathrm{ad}^-(\xi_2))_{\mathrm{N}}$; otherwise $\mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}} = \emptyset$.

These sides are *fixed*: the west and east sides of horizontal carriers, and the south and north sides of vertical carriers. Every other side is *free*. Give a free side the colour of the facing side across its edge if that side is fixed, and $\emptyset$ otherwise. Each free side belongs to one tile and one edge, so this is well defined.

Palettes. Active payloads lie in T_M. A seed tile sits at a pair of centres by Lemma 8.14(c), where $\mathrm{ad} = 0$, so its payload is $\mathrm{Y}(0,0) = s_M$. Horizontal and vertical payloads satisfy $\mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}$,

respectively $\mathbf{c}_S = \mathbf{c}_N$, and inactive payloads are unconstrained. So $\vartheta(\xi) = (\vartheta^r(\xi), \mathrm{pay}(\xi))$ is a tile of R_M.

Matching. The $\mathbb{A}^r$-parts match because ϑ^r is a tiling. Consider a horizontal edge from ξ to $\xi + \mathbf{e}_1$. If at most one of its two sides is fixed, the free side copies the other or both are $\emptyset$. Otherwise both tiles are horizontal carriers. Let $d = \mathrm{own}(\xi_2)$. By Lemma 8.14(b), ξ_1 and $\xi_1 + 1$ lie in interiors of frames of depth d; these interiors are separated by endpoints, so both lie in one $\mathring{I}_d(m)$. If ξ_2 lies in no $\mathcal{F}_d(m')$, it is an endpoint of a frame of depth d, the row carries no active tile by Lemma 8.14(c), and both sides are $\emptyset$. If $\xi_2 \in \mathcal{F}_d(m')$, put $t = \mathrm{ad}(\xi_2)$. A tile on this row inside $\mathring{I}_d(m)$ is active or seed at the free coordinates and horizontal elsewhere. In both cases the east side of ξ is $\Upsilon(\mathrm{ad}^-(\xi_1), t)_E$. If $\xi_1 + 1$ is not free, the west side of $\xi + \mathbf{e}_1$ is $\Upsilon(\mathrm{ad}^-(\xi_1 + 1), t)_E = \Upsilon(\mathrm{ad}^-(\xi_1), t)_E$. If $\xi_1 + 1$ is free, it is $\Upsilon(\mathrm{ad}^-(\xi_1) + 1, t)_W$, which equals $\Upsilon(\mathrm{ad}^-(\xi_1), t)_E$ because Υ is a tiling. Vertical edges are the same with rows and columns exchanged. So ϑ is a tiling of R_M. □

Every product $\mathcal{F}_d(m) \times \mathcal{F}_d(m')$ carries its own copy of Υ, centred at its centre pair. Products without a seed tile receive a copy as well, and nothing links different products: the transverse sides of channels are free.

9 Every halting member has a bounded-width refutation

This section proves the second half of Theorem 9.8.

Theorem 9.1. *Let B be the constant of Theorem 7.18, let $\ell = \binom{276}{2} + 1 = 37951$, and put $j = 10(B + \ell + 100)$. If the machine M halts, then R_M is j-token refutable.*

The proof is by contradiction. Fix a machine M that halts, and let T be its halting time as in Proposition 8.16. Suppose that R_M is not j-token refutable. By Theorem 4.25(iii) there is an E_j family H for R_M. From H we build, in order: a family with one common quarter phase (Section 9.2); a family on the index grid with the same j (Section 9.3); a pin map p on at most $\ell + 3$ sites under which one whole substitution word is forced site by site (Section 9.4); the exact roles on one initialized carrier grid (Section 9.5); and, by a meet argument, a payload assignment on that grid that compresses to a legal T_M-pattern which Proposition 8.16 forbids (Sections 9.7 and 9.8). Section 9.9 checks that j sites suffice at every step. The constants B, ℓ and j do not depend on M, on the depth n chosen below, or on T.

9.1 Inputs from the setting, scaffold and family

Write $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ in $\mathbb{Z}^2$. Tiles ϑ, ϑ' of a Wang set *match horizontally* if $\vartheta_E = \vartheta'_W$, and *match vertically* if $\vartheta_N = \vartheta'_S$. Integer intervals are written $[a, b] = \{a, a + 1, \ldots, b\}$ and $[0, m) = \{0, \ldots, m - 1\}$.

Families. We use E_K families for a Wang set as in Definition 4.22: nonempty sets of legal patterns on finite subsets of $\mathbb{Z}^2$, with at most K sites each, closed under restriction and with the extension property for maps on fewer than K sites. If $h \in H$ and h' is supplied by (Extension) for a new site w, then $h'|_{\mathrm{dom}\, h \cup \{w\}}$ also extends h; so we may and do assume that an extension adds exactly the requested site. Repeating this, a map $h \in H$ extends to a map of H on $\mathrm{dom}\, h \cup U'$ whenever $|\mathrm{dom}\, h \cup U'| \le K$ (Remark 4.23). By Theorem 4.25(iii), a Wang set that is not j-token refutable has an E_j family.

The scaffold and the parser. Let $\mathbb{A}$ be the scaffold Wang set of 276 letters and $\sigma\colon \mathbb{A} \to \mathbb{A}^{[0,4)^2}$ its substitution, from Proposition 7.11. For $n \geq 0$ and $\mathfrak{a} \in \mathbb{A}$, $\sigma^n(\mathfrak{a})$ is a legal pattern on $[0,4^n)^2$. Let $\chi_1, \chi_2\colon \mathbb{A} \to \{0,1\}$ be the column bit and the row bit of the ruler. For $u \in \mathbb{Z}^2$ put $\mathrm{Anc}_n(u) = \{u,\, u + 4^n\mathbf{e}_1,\, u + 4^n\mathbf{e}_2\}$. A map g into $\mathbb{A}$ defined on $\mathrm{Anc}_n(u)$ is a *good n-anchor at u* if $\chi_1(g(u)) \neq \chi_1(g(u + 4^n\mathbf{e}_1))$ and $\chi_2(g(u)) \neq \chi_2(g(u + 4^n\mathbf{e}_2))$. The port set $\mathrm{Port}_n(u) \subseteq u + [0,4^n)^2$ has at most ℓ sites, and the readout $\mathfrak{a} \mapsto \big(\sigma^n(\mathfrak{a})(y-u)\big)_{y \in \mathrm{Port}_n(u)}$ is injective on $\mathbb{A}$. We use two consequences of Theorem 7.18, Lemma 7.19, and Proposition 7.21.

(P1) Anchors. For every E_K family G for $\mathbb{A}$ with $K \geq B$ and every $n \geq 0$, there are $u \in \mathbb{Z}^2$ and $g \in G$ with domain $\mathrm{Anc}_n(u)$ that is a good n-anchor at u (Proposition 7.21).

(P2) Type and point certificates. Let G be an E_K family for $\mathbb{A}$ with $K \geq B$, let $n \geq 0$ and $u \in \mathbb{Z}^2$, and let $g \in G$ be defined on $\mathrm{Anc}_n(u) \cup \mathrm{Port}_n(u)$, with $g|_{\mathrm{Anc}_n(u)}$ a good n-anchor at u. Then there is $\mathfrak{a} \in \mathbb{A}$ with $g(y) = \sigma^n(\mathfrak{a})(y-u)$ for every $y \in \mathrm{Port}_n(u)$. If g is also defined at a site $s \in u + [0,4^n)^2$, then $g(s) = \sigma^n(\mathfrak{a})(s-u)$. (The type and point certificates of Theorem 7.18 have rooted tree decompositions with bags of at most B sites and contain these sites in their boundary; Lemma 7.19 carries their conclusions to every E_K family with $K \geq B$.)

The family R_M. From Propositions 8.11, 8.15 and 8.16 and Definition 8.17 we use the following.

(Q1) Quarters. Every tile ϑ of R_M has a *quadrant* $\mathrm{quad}(\vartheta) \in \{0,1\}^2$ and a *parent* $\mathrm{par}(\vartheta) \in \mathbb{A}$; the pair $(\mathrm{quad}(\vartheta), \mathrm{par}(\vartheta))$ is its *scaffold pair*. Let ϑ, ϑ' match horizontally. Then $\mathrm{quad}(\vartheta') \equiv \mathrm{quad}(\vartheta) + \mathbf{e}_1 \pmod 2$. If the first coordinate of $\mathrm{quad}(\vartheta)$ is 0, the two quarters belong to one parent cell and $\mathrm{par}(\vartheta') = \mathrm{par}(\vartheta)$; if it is 1, then $\mathrm{par}(\vartheta)$ and $\mathrm{par}(\vartheta')$ match horizontally in $\mathbb{A}$. The same holds for vertical matching, with $\mathbf{e}_2$ and the second coordinate.

(Q2) Walls and flows. A scaffold pair may lie on a light vertical wall, as a *left side* or a *right side*, and it may lie on a light horizontal wall, as a *bottom side* or a *top side*; these attributes are functions of the scaffold pair. Each tile ϑ carries flows $\phi_{\mathrm{W}}(\vartheta), \phi_{\mathrm{E}}(\vartheta) \subseteq \{\mathrm{left}, \mathrm{right}\}$ and $\phi_{\mathrm{S}}(\vartheta), \phi_{\mathrm{N}}(\vartheta) \subseteq \{\mathrm{down}, \mathrm{up}\}$. If ϑ, ϑ' match horizontally then $\phi_{\mathrm{E}}(\vartheta) = \phi_{\mathrm{W}}(\vartheta')$, and if they match vertically then $\phi_{\mathrm{N}}(\vartheta) = \phi_{\mathrm{S}}(\vartheta')$. If the scaffold pair of ϑ lies on no light vertical wall, then $\phi_{\mathrm{W}}(\vartheta) = \phi_{\mathrm{E}}(\vartheta)$. On a left side, $\phi_{\mathrm{W}}(\vartheta) \neq \emptyset$ and $\phi_{\mathrm{E}}(\vartheta) \subseteq \{\mathrm{left}\}$. On a right side, $\phi_{\mathrm{E}}(\vartheta) \neq \emptyset$ and $\phi_{\mathrm{W}}(\vartheta) \subseteq \{\mathrm{right}\}$. The vertical rules are the same after rotation: with no light horizontal wall $\phi_{\mathrm{S}}(\vartheta) = \phi_{\mathrm{N}}(\vartheta)$; on a bottom side $\phi_{\mathrm{S}}(\vartheta) \neq \emptyset$ and $\phi_{\mathrm{N}}(\vartheta) \subseteq \{\mathrm{down}\}$; on a top side $\phi_{\mathrm{N}}(\vartheta) \neq \emptyset$ and $\phi_{\mathrm{S}}(\vartheta) \subseteq \{\mathrm{up}\}$.

(Q3) Roles and palettes. A tile ϑ is a *horizontal carrier* if $\phi_{\mathrm{W}}(\vartheta) = \phi_{\mathrm{E}}(\vartheta) = \emptyset$, and a *vertical carrier* if $\phi_{\mathrm{S}}(\vartheta) = \phi_{\mathrm{N}}(\vartheta) = \emptyset$. Its *role* is inactive, horizontal, vertical or active according as it is neither, only a horizontal carrier, only a vertical carrier, or both; an active tile whose scaffold pair is the north-east quarter of an index-0 letter with dark ruler bits has the *seed* role instead. Each tile carries a payload $\mathrm{pay}(\vartheta) \in C_M^4$, a tuple of colours indexed by $\mathrm{W}, \mathrm{E}, \mathrm{S}, \mathrm{N}$. Horizontally matching tiles have $\mathrm{pay}(\vartheta)_{\mathrm{E}} = \mathrm{pay}(\vartheta')_{\mathrm{W}}$, and vertically matching tiles have $\mathrm{pay}(\vartheta)_{\mathrm{N}} = \mathrm{pay}(\vartheta')_{\mathrm{S}}$. The payload lies in the palette of the role: all of C_M^4 for inactive, $\{\mathbf{c} : \mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}\}$ for horizontal, $\{\mathbf{c} : \mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}}\}$ for vertical, T_M for active, and $\{s_M\}$ for seed.

(Q4) The Horn layer. C_M is a finite set with a binary operation $\wedge$ (intersection of packets) that is associative, commutative and idempotent. Write $\wedge$ also coordinatewise on C_M^4.

The tile set $T_M \subseteq C_M^4$ is closed under $\wedge$, and $s_M \in T_M$. If M halts at time T, there is no legal T_M-pattern Ξ on $[0, 2T+10)^2$ with $\Xi(0,0) = s_M$ (Proposition 8.16).

(Q5) The initialized board. For a pattern W on $[0, 4^n)^2$ and $\xi \in [0, 2 \cdot 4^n)^2$, the *quarter of W at ξ* is the scaffold pair $(\xi \bmod 2, W(\lfloor \xi/2 \rfloor))$, coordinatewise. Let $n \geq 2$, $\mathfrak{a} \in \mathbb{A}$, $W = \sigma^n(\mathfrak{a})$ and $\Delta_n = 4^{n-1}$, and put $J = [\Delta_n + 1, 3\Delta_n]$ and $J^\circ = [\Delta_n + 2, 3\Delta_n - 1]$. By Proposition 8.15 there is a set $\mathcal{F} \subseteq J^\circ$, the *free coordinates*, such that:

(a) $2\Delta_n + 1 \in \mathcal{F}$, and at least $2^n - 1$ elements of $\mathcal{F}$ are $\geq 2\Delta_n + 1$;

(b) the quarter of W at $(2\Delta_n + 1, 2\Delta_n + 1)$ is the north-east quarter of an index-0 letter with dark ruler bits;

(c) for every $\xi = (\xi_1, \xi_2) \in Z_W = (\mathcal{F} \times J^\circ) \cup (J^\circ \times \mathcal{F})$, the quarter of W at ξ lies on no light wall. In the row of ξ, the nearest quarters of W to the left and to the right of ξ that lie on a light vertical wall both lie in $J \times \{\xi_2\}$, and they are a left side and a right side, in this order, if and only if $\xi_2 \in \mathcal{F}$. In the column of ξ, the nearest quarters below and above ξ that lie on a light horizontal wall both lie in $\{\xi_1\} \times J$, and they are a bottom side and a top side, in this order, if and only if $\xi_1 \in \mathcal{F}$.

Item (c) is the statement that every carrier role on the board is forced by a corridor between two walls inside the board. Since $3\Delta_n < 2 \cdot 4^n$, all these corridors lie in $[0, 2 \cdot 4^n)^2$.

9.2 One adjusted quarter phase

Lemma 9.2. *Let H be an E_K family for R_M with $K \geq 3$. For every $h \in H$ and all $v, w \in \operatorname{dom} h$,*

$$\operatorname{quad}(h(v)) - v \equiv \operatorname{quad}(h(w)) - w \pmod 2.$$

Proof. Choose a path $v = y_0, y_1, \ldots, y_m = w$ of distinct sites with $y_{l+1} - y_l \in \{\pm \mathbf{e}_1, \pm \mathbf{e}_2\}$. If $m \leq 1$, the claim is (Q1) applied to the legal pattern h. Let $m \geq 2$. Put $h_0 = h|_{\{y_0, w\}}$. For $0 \leq l \leq m-2$, extend h_l by the site y_{l+1} and restrict the result to $\{y_{l+1}, w\}$; call it h_{l+1}. Each step extends a map on two sites, which is possible since $K \geq 3$. Put $\vartheta_l = h_l(y_l)$, so that $\vartheta_0 = h(v)$, and note $h_l(w) = h(w)$ for every l. For $l \leq m-2$, the tiles ϑ_l and ϑ_{l+1} are values of one legal pattern at the adjacent sites y_l, y_{l+1}, so (Q1) gives $\operatorname{quad}(\vartheta_{l+1}) - \operatorname{quad}(\vartheta_l) \equiv y_{l+1} - y_l$. Finally ϑ_{m-1} and $h(w)$ are values of h_{m-1} at the adjacent sites y_{m-1}, w. Summing the increments gives the claim. □

For $\gamma \in (\mathbb{Z}/2\mathbb{Z})^2$ let H_γ be the set of $h \in H$ with $\operatorname{quad}(h(v)) - v \equiv \gamma$ for every $v \in \operatorname{dom} h$. We call it the *phase-$\gamma$ component* of H.

Lemma 9.3. *Let H be an E_K family for R_M with $K \geq 3$. Some γ has a map with nonempty domain in H_γ, and for each such γ the set H_γ is an E_K family for R_M.*

Proof. Extending the empty map gives $h \in H$ defined at some site v; with $\gamma = \operatorname{quad}(h(v)) - v$, Lemma 9.2 puts h in H_γ. Now fix such a γ. The set H_γ consists of legal patterns with at most K sites, and it is closed under restriction. Let $h \in H_\gamma$ with $|\operatorname{dom} h| < K$ and let $w \in \mathbb{Z}^2$. If $\operatorname{dom} h \neq \emptyset$, extend h in H by w; by Lemma 9.2 the new value has the phase of the old values, so the extension lies in H_γ. If h is empty, take a site v of some map of H_γ, extend its restriction to $\{v\}$ by w in H (possible since $K \geq 2$), and restrict to $\{w\}$; by Lemma 9.2 the result lies in H_γ. □

Apply this to the E_j family H of R_M; here $j \geq 3$. Fix γ as in Lemma 9.3, and let $\bar{\gamma} \in \{0,1\}^2$ be its representative. Then $\operatorname{quad}(h(v)) \equiv v - \bar{\gamma} \pmod 2$ for every $h \in H_\gamma$ and $v \in \operatorname{dom} h$. The phase is common to all maps of H_γ; no site is spent to fix it.

9.3 The index family

For $z \in \mathbb{Z}^2$ put $\iota(z) = \bar{\gamma} + 2z$. The four sites $\iota(z) + \{0,1\}^2$ form the *cell* of z, and in every map of H_γ the value at $\iota(z) + \varsigma$ has quadrant ς. For $h \in H_\gamma$ with $\operatorname{dom} h \subseteq \iota(\mathbb{Z}^2)$, define $\hat{h}$ on $\iota^{-1}(\operatorname{dom} h)$ by $\hat{h}(z) = \operatorname{par}(h(\iota(z)))$, and let $\widehat{H}$ be the set of these maps.

Lemma 9.4. *$\widehat{H}$ is an E_j family for $\mathbb{A}$.*

Proof. Restriction commutes with the hat, and extending h in H_γ by a site $\iota(w)$ extends $\hat{h}$ by w. It remains to show that each $\hat{h}$ is a legal pattern. Let $z, z + \mathbf{e}_1 \in \operatorname{dom} \hat{h}$ and $v = \iota(z)$, so that $\iota(z + \mathbf{e}_1) = v + 2\mathbf{e}_1$. Restrict h to the pair $\{v, v + 2\mathbf{e}_1\}$ and extend the restriction by the site $v + \mathbf{e}_1$; call the result h'. The values $h'(v)$ and $h'(v + \mathbf{e}_1)$ match horizontally and have quadrants $(0,0)$ and $(1,0)$, so (Q1) gives $\operatorname{par}(h'(v)) = \operatorname{par}(h'(v + \mathbf{e}_1))$. The values $h'(v + \mathbf{e}_1)$ and $h'(v + 2\mathbf{e}_1)$ match horizontally, and the first has quadrant $(1,0)$, so their parents match horizontally in $\mathbb{A}$. Since h' agrees with h at v and $v + 2\mathbf{e}_1$, the letters $\hat{h}(z)$ and $\hat{h}(z + \mathbf{e}_1)$ match horizontally. The vertical case is the same with $\mathbf{e}_2$. □

The check is binary: we restrict to the pair before adding the middle quarter, so we never extend a map that already has j sites.

9.4 Pins and the forced word

Fix $n \geq 2$ with $2^n - 2 \geq 2T + 10$. By (P1), applied to $\widehat{H}$ with $K = j \geq B$, there are $u \in \mathbb{Z}^2$ and $g_0 \in \widehat{H}$ with domain $\operatorname{Anc}_n(u)$ that is a good n-anchor at u. Extend g_0 in $\widehat{H}$ to a map g_1 on $\operatorname{Anc}_n(u) \cup \operatorname{Port}_n(u)$, a set of at most $\ell + 3 \leq j$ sites. By the first part of (P2) there is $\mathfrak{a} \in \mathbb{A}$ with $g_1(y) = \sigma^n(\mathfrak{a})(y - u)$ for $y \in \operatorname{Port}_n(u)$; it is unique because the readout is injective. Put

$$W = \sigma^n(\mathfrak{a}), \qquad o = \iota(u).$$

By the definition of $\widehat{H}$ there is $p \in H_\gamma$ with $\operatorname{dom} p = \iota(\operatorname{Anc}_n(u) \cup \operatorname{Port}_n(u))$ and $\hat{p} = g_1$. These are the *pins*. They occupy at most $\ell + 3$ sites, and every $h \in H_\gamma$ with $p \subseteq h$ projects to an index map that keeps the same good anchor and the same readout of W.

Lemma 9.5 (The forced word)**.** *Let $\xi \in [0, 2 \cdot 4^n)^2$, and let $h \in H_\gamma$ satisfy $p \subseteq h$ and $o + \xi \in \operatorname{dom} h$. Then the scaffold pair of $h(o + \xi)$ is the quarter of W at ξ.*

Proof. Write $x = o + \xi$, $\varsigma = \xi \bmod 2$ and $s = u + \lfloor \xi/2 \rfloor$, so that $x = \iota(s) + \varsigma$ lies in the cell of s. The quadrant of $h(x)$ is $x - \bar{\gamma} \equiv \xi \equiv \varsigma$. For the parent, let Y be the path from x to $\iota(s)$ inside the cell of s that has no step if $\varsigma = (0,0)$, one step if $\varsigma \in \{(1,0),(0,1)\}$, and the two steps $x \to \iota(s) + \mathbf{e}_1 \to \iota(s)$ if $\varsigma = (1,1)$. Restrict h to $\operatorname{dom} p \cup \{x\}$ and extend it by the at most two other sites of Y; call the result h'. It has at most $\ell + 6$ sites. In each step of Y, the western or southern of the two quarters has coordinate 0 in the direction of the step, so (Q1) gives equal parents along Y. Hence $\operatorname{par}(h(x)) = \operatorname{par}(h'(\iota(s)))$. Let $h'' = h'|_{\operatorname{dom} p \cup \{\iota(s)\}}$. Then $\hat{h}'' \in \widehat{H}$ is defined on $\operatorname{Anc}_n(u) \cup \operatorname{Port}_n(u) \cup \{s\}$ and agrees with g_1 on $\operatorname{Anc}_n(u) \cup \operatorname{Port}_n(u)$. So (P2) applies to $\hat{h}''$, and by injectivity of the readout the letter it provides is $\mathfrak{a}$. Since $s \in u + [0, 4^n)^2$, the second part of (P2) gives $\hat{h}''(s) = W(s - u) = W(\lfloor \xi/2 \rfloor)$, and $\hat{h}''(s) = \operatorname{par}(h'(\iota(s)))$. □

The lemma is pointwise. Each query uses the pins and at most three further sites, never all sites of W at once; and it holds for every map of H_γ that keeps the pins, not for one favourable map per query.

9.5 The forced role chart

Let Δ_n, J, J°, $\mathcal{F}$ and Z_W be as in (Q5) for the word W.

Lemma 9.6 (Corridors). *Let $\xi \in Z_W$, and let $h \in H_\gamma$ satisfy $p \subseteq h$ and $v = o + \xi \in \operatorname{dom} h$. Then $h(v)$ is a horizontal carrier if and only if $\xi_2 \in \mathcal{F}$, and a vertical carrier if and only if $\xi_1 \in \mathcal{F}$. Consequently the role of $h(v)$ is horizontal if $\xi_1 \notin \mathcal{F}$, vertical if $\xi_2 \notin \mathcal{F}$, active or seed if $\xi \in \mathcal{F}^2$, and seed if $\xi = (2\Delta_n + 1, 2\Delta_n + 1)$.*

Proof. By (Q5)(c), let $\xi + m\mathbf{e}_1$ and $\xi - m'\mathbf{e}_1$, with $m, m' \geq 1$, be the nearest quarters of W to the right and to the left of ξ in its row that lie on a light vertical wall. Both lie in $J \times \{\xi_2\}$.

We walk to the right. Put $h_0 = h$ and $v_l = v + l\mathbf{e}_1$. For $l = 1, \dots, m$, restrict h_{l-1} to $\operatorname{dom} p \cup \{v_{l-1}\}$, a set of at most $\ell + 4$ sites, and extend it by v_l; call the result h_l. Then $p \subseteq h_l$. Put $\vartheta_l = h_l(v_l)$, so that $\vartheta_0 = h(v)$. For $1 \leq l \leq m$, the tiles $\vartheta_{l-1} = h_l(v_{l-1})$ and ϑ_l are values of one legal pattern at horizontally adjacent sites, so $\phi_{\mathrm{E}}(\vartheta_{l-1}) = \phi_{\mathrm{W}}(\vartheta_l)$ by (Q2). By Lemma 9.5, the scaffold pair of ϑ_l is the quarter of W at $\xi + l\mathbf{e}_1$. For $0 \leq l < m$ this quarter lies on no light vertical wall: for $l = 0$ by (Q5)(c), and for $0 < l < m$ by the choice of m. So $\phi_{\mathrm{W}}(\vartheta_l) = \phi_{\mathrm{E}}(\vartheta_l)$ for $0 \leq l < m$. Writing $\psi = \phi_{\mathrm{W}}(\vartheta_0) = \phi_{\mathrm{E}}(\vartheta_0)$, we get $\psi = \phi_{\mathrm{W}}(\vartheta_m)$, where ϑ_m sits at the right wall. Walking to the left in the same way gives a tile ϑ' at the left wall, with the scaffold pair of the quarter of W at $\xi - m'\mathbf{e}_1$, and $\psi = \phi_{\mathrm{E}}(\vartheta')$.

If $\xi_2 \in \mathcal{F}$, the left wall is a left side and the right wall a right side. By (Q2), $\psi = \phi_{\mathrm{E}}(\vartheta') \subseteq \{\text{left}\}$ and $\psi = \phi_{\mathrm{W}}(\vartheta_m) \subseteq \{\text{right}\}$, so $\psi = \emptyset$ and $h(v)$ is a horizontal carrier. If $\xi_2 \notin \mathcal{F}$, then the left wall is not a left side or the right wall is not a right side. In the first case it is a right side and $\psi = \phi_{\mathrm{E}}(\vartheta') \neq \emptyset$; in the second it is a left side and $\psi = \phi_{\mathrm{W}}(\vartheta_m) \neq \emptyset$. Either way $h(v)$ is not a horizontal carrier.

The vertical statement is proved the same way along the column of ξ, with $\mathbf{e}_2$, the flows $\phi_{\mathrm{S}}, \phi_{\mathrm{N}}$, and bottom and top sides. The roles follow from (Q3). At $\xi = (2\Delta_n + 1, 2\Delta_n + 1)$ the tile is active, and by Lemma 9.5 and (Q5)(b) its scaffold pair is the north-east quarter of an index-0 letter with dark ruler bits, so its role is seed. □

The roles do not depend on the nonempty flow values that the family may choose: the walk only uses which flows are empty. An extra active crossing between two consecutive free coordinates would break a wire of the compression in Section 9.8; the lemma excludes it for every map that keeps the pins.

9.6 The selected component

Put $N = 2T + 10$. By (Q5)(a) and $2^n - 1 > N$, let $f_0 < f_1 < \cdots < f_{N-1}$ be the N smallest elements of $\mathcal{F}$ that are at least $2\Delta_n + 1$; thus $f_0 = 2\Delta_n + 1$. Let $F_N = \{f_0, \dots, f_{N-1}\}$ and

$$Z = \big\{\xi : \xi_2 \in F_N,\ f_0 \leq \xi_1 \leq f_{N-1}\big\} \cup \big\{\xi : \xi_1 \in F_N,\ f_0 \leq \xi_2 \leq f_{N-1}\big\}.$$

Since $f_0, f_{N-1} \in J^\circ$, we have $Z \subseteq Z_W$. An integer strictly between f_i and f_{i+1} is not in $\mathcal{F}$. The *channel edges* of Z are the pairs $\{\xi, \xi + \mathbf{e}_1\}$ with $\xi_2 \in F_N$ and $f_0 \leq \xi_1 < f_{N-1}$, and the pairs $\{\xi, \xi + \mathbf{e}_2\}$ with $\xi_1 \in F_N$ and $f_0 \leq \xi_2 < f_{N-1}$. Assign to each $\xi \in Z$ the palette

$$\operatorname{Pal}(\xi) = \begin{cases} \{s_M\} & \text{if } \xi = (f_0, f_0), \\ T_M & \text{if } \xi \in F_N^2 \setminus \{(f_0, f_0)\}, \\ \{\mathbf{c} \in C_M^4 : \mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}\} & \text{if } \xi_2 \in F_N,\ \xi_1 \notin F_N, \\ \{\mathbf{c} \in C_M^4 : \mathbf{c}_{\mathrm{S}} = \mathbf{c}_{\mathrm{N}}\} & \text{if } \xi_1 \in F_N,\ \xi_2 \notin F_N. \end{cases}$$

By Lemma 9.6 and (Q3), $\mathrm{pay}(h(o+\xi)) \in \mathrm{Pal}(\xi)$ for every $\xi \in Z$ and every $h \in H_\gamma$ with $p \subseteq h$ and $o+\xi \in \mathrm{dom}\, h$. On $F_N^2 \setminus \{(f_0, f_0)\}$ the role is active or seed, and both palettes lie in T_M because $s_M \in T_M$. Lower-depth initialized components may occur elsewhere in the board; they do not meet Z, and they are not used.

9.7 The projected meet

For $\xi \in Z$ and $v = o + \xi$ put

$$U_v = \big\{\mathrm{pay}(h(v)) : h \in H_\gamma,\ p \subseteq h,\ v \in \mathrm{dom}\, h\big\}, \qquad m_v = \bigwedge U_v.$$

For a channel edge $\{\xi, \xi'\}$ with $\xi' = \xi + \mathbf{e}_1$ or $\xi' = \xi + \mathbf{e}_2$, and $v = o + \xi$, $w = o + \xi'$, put

$$Q_{vw} = \big\{(\mathrm{pay}(h(v)), \mathrm{pay}(h(w))) : h \in H_\gamma,\ p \subseteq h,\ v, w \in \mathrm{dom}\, h\big\}.$$

Lemma 9.7 (Projected meet). *(a) Each U_v is a nonempty finite set, and $m_v \in \mathrm{Pal}(\xi)$.*

(b) The two coordinate projections of Q_{vw} are exactly U_v and U_w.

(c) If $\xi' = \xi + \mathbf{e}_1$ then $(m_v)_{\mathrm{E}} = (m_w)_{\mathrm{W}}$; if $\xi' = \xi + \mathbf{e}_2$ then $(m_v)_{\mathrm{N}} = (m_w)_{\mathrm{S}}$.

Proof. (a) Extending p by v gives an element of U_v, and $U_v \subseteq C_M^4$ is finite. Every palette is closed under $\wedge$: for T_M this is (Q4); for $\{s_M\}$ it is idempotence; and if $\mathbf{c}_{\mathrm{W}} = \mathbf{c}_{\mathrm{E}}$ and $\mathbf{c}'_{\mathrm{W}} = \mathbf{c}'_{\mathrm{E}}$, then $(\mathbf{c} \wedge \mathbf{c}')_{\mathrm{W}} = \mathbf{c}_{\mathrm{W}} \wedge \mathbf{c}'_{\mathrm{W}} = (\mathbf{c} \wedge \mathbf{c}')_{\mathrm{E}}$, and likewise for the vertical palette. Since $U_v \subseteq \mathrm{Pal}(\xi)$ by Section 9.6, the meet of its finitely many elements lies in $\mathrm{Pal}(\xi)$.

(b) A pair in Q_{vw} comes from a map that keeps p and is defined at v, so its first coordinate lies in U_v. Conversely, let $\mathbf{c} = \mathrm{pay}(h(v)) \in U_v$. Restrict h to $\mathrm{dom}\, p \cup \{v\}$ and extend by w; the result has at most $\ell + 5$ sites, keeps p, and has the value $h(v)$ at v, so it gives a pair of Q_{vw} with first coordinate $\mathbf{c}$. The second coordinate is symmetric.

(c) Let $\xi' = \xi + \mathbf{e}_1$ and list $Q_{vw} = \{(\mathbf{c}^1, \mathbf{d}^1), \dots, (\mathbf{c}^l, \mathbf{d}^l)\}$. Each pair consists of the payloads of one legal pattern at horizontally adjacent sites, so $\mathbf{c}^i_{\mathrm{E}} = \mathbf{d}^i_{\mathrm{W}}$ by (Q3). By (b), $\{\mathbf{c}^1, \dots, \mathbf{c}^l\} = U_v$ and $\{\mathbf{d}^1, \dots, \mathbf{d}^l\} = U_w$. As $\wedge$ acts coordinatewise,

$$(m_v)_{\mathrm{E}} = \textstyle\bigwedge_i \mathbf{c}^i_{\mathrm{E}} = \bigwedge_i \mathbf{d}^i_{\mathrm{W}} = (m_w)_{\mathrm{W}}.$$

The vertical case is the same. □

So the single assignment $v \mapsto m_v$ satisfies every palette and every channel equality at once. Neither H nor the sets U_v need to be closed under meets; only the palettes and the equality constraints are. This is the classical fact that constraints closed under a semilattice operation are solved by local consistency [25], applied to relations projected from H.

9.8 Compression and the contradiction

Proof of Theorem 9.1. With $N = 2T + 10$, define Ξ on $[0, N)^2$ by $\Xi(i, i') = m_{o+(f_i, f_{i'})}$. By Lemma 9.7(a), $\Xi(i, i') \in T_M$ for all i, i', and $\Xi(0, 0) = s_M$.

Fix i' and $i < N - 1$, and let $v_y = o + (y, f_{i'})$ for $f_i \le y \le f_{i+1}$. These sites lie in $o + Z$, and consecutive ones form channel edges. For $f_i < y < f_{i+1}$ the palette at $(y, f_{i'})$ is horizontal, so $(m_{v_y})_{\mathrm{W}} = (m_{v_y})_{\mathrm{E}}$. By Lemma 9.7(c), $(m_{v_y})_{\mathrm{E}} = (m_{v_{y+1}})_{\mathrm{W}}$ for $f_i \le y < f_{i+1}$. Chaining these equalities gives $\Xi(i, i')_{\mathrm{E}} = \Xi(i+1, i')_{\mathrm{W}}$. In the same way, along the column f_i through vertical palettes, $\Xi(i, i')_{\mathrm{N}} = \Xi(i, i'+1)_{\mathrm{S}}$.

So Ξ is a legal T_M-pattern on $[0, 2T + 10)^2$ with $\Xi(0, 0) = s_M$. Since M halts at time T, (Q4) says that no such pattern exists. This contradiction shows that R_M has no E_j family. By Theorem 4.25(iii), R_M is j-token refutable. □

The argument excludes every E_j family, so it excludes every possible surviving live symbol at the origin, not only a seeded one. The box radius and the refutation time of R_M depend on M; no bound on them is claimed.

9.9 The token budget

Every map used above has a bounded number of sites, independent of M, n and T:

- Lemmas 9.2 and 9.3: maps on at most 3 sites;
- Lemma 9.4: the binary sibling check uses 3 sites;
- (P1) and (P2) are applied to $\widehat{H}$, which needs $j \geq B$; the maps they are applied to have at most $\ell + 4$ sites, and their proofs use members of $\widehat{H}$ on at most B sites;
- the pins p: at most $\ell + 3$ sites;
- Lemma 9.5: at most $\ell + 6$ sites in H_γ, and at most $\ell + 4$ sites in $\widehat{H}$;
- Lemma 9.6: the walk uses maps on at most $\ell + 5$ sites, and each step calls Lemma 9.5;
- Lemma 9.7: at most $\ell + 5$ sites.

So $j \geq \max(B, \ell + 6)$ suffices, and $j = 10(B + \ell + 100)$ exceeds $B + \ell + 6$. The margin also covers a sibling check through two full parent cells, with 8 quarter sites, in place of the 3-site path of Lemma 9.4. The payload alphabet C_M grows with M; the number of sites does not.

Theorem 9.8 (Theorem C). *Let $j = 10(B + \ell + 100)$ be the integer of Theorem 9.1. The map $M \mapsto R_M$ of Definition 8.17 is total computable, R_M tiles the plane if and only if M does not halt, and R_M is j-token refutable whenever M halts.*

Proof of Theorem 9.8. The map $M \mapsto R_M$ is total computable (Definition 8.17). If M does not halt, R_M tiles (Proposition 8.18). If M halts, R_M is j-token refutable by Theorem 9.1, so it does not tile by Theorem 4.25(iv). Hence R_M tiles if and only if M does not halt, and every halting member is j-token refutable for the one integer $j = 10(B + \ell + 100)$. □

10 The machine-family theorem and its corollaries

Throughout this section $j = 10(B + \ell + 100)$ is the integer of Theorem 9.1, with $\ell = \binom{276}{2} + 1 = 37951$, and

$$k = \max(2, 2j) = 2j, \qquad q = k + 1, \qquad a = 2q^2 + 1, \qquad D_0 = 50q, \qquad \varepsilon_0 = \varepsilon_k = (4D_0^2)^{-a}.$$

The dimension is $d = 2 + k$. The family $M \mapsto R_M$ is the one of Theorem 9.8. We write $\delta_O(t)$ as in Definition 4.11, adding a superscript M when the Wang set is R_M.

Lemma 10.1 (One threshold). *Let $0 < \varepsilon \leq \varepsilon_0$. Then $\beta = 2D_0^2 \varepsilon^{1/a} \leq \frac{1}{2}$, the hypothesis of Theorem 6.2 holds, and $\varepsilon < 10^{-9}$. Both conditions on the rate have the form $\varepsilon \leq \varepsilon_0$, so they hold at every rate in $(0, \varepsilon_0]$.*

Proof. The map $\varepsilon \mapsto \beta$ is increasing, and at ε_0 it equals $2D_0^2(4D_0^2)^{-1} = \frac{1}{2}$. Theorem 6.2 asks for $0 < \varepsilon \leq (4D_0^2)^{-a} = \varepsilon_0$. Since $q \geq 3$, $D_0 \geq 150$ and $a \geq 19$, we get $\varepsilon_0 \leq (90000)^{-19} < 10^{-9}$. □

Theorem 10.2 (Machine-family theorem). *For every $\varepsilon \in (0, \varepsilon_0]$ and every Turing machine M, the noisy quorum rule $F_{R_M}^{[k]}$ on $\mathbb{Z}^{2+k}$ with uniform replacement at rate ε is a positive-rate probabilistic cellular automaton, and it is ergodic if and only if M halts. The map $M \mapsto F_{R_M}^{[k]}$ is total computable.*

Proof of Theorem 10.2. Fix $\varepsilon \in (0, \varepsilon_0]$ and use uniform replacement noise at rate ε. The map $M \mapsto R_M$ is total computable by Theorem 9.8, and T_R, $\mathcal{A}_R$ and the local rule of $F_R^{[k]}$ on $\mathcal{N} = \Lambda \times \{c_0, \dots, c_k\}$ are computed from R by finite enumeration. So $M \mapsto F_{R_M}^{[k]}$ is total computable.

Suppose M halts. By Theorem 9.8, R_M is j-token refutable; let $\tau = \tau(R_M)$ be its refutation time. The hypotheses of Theorem 5.1 hold: $j \geq 1$, $k \geq \max(2, 2j)$, $0 < \varepsilon < 1$, and $\beta \leq \frac{1}{2} < 1$ by Lemma 10.1. So for every window O and every $t \geq 0$,

$$\delta_O^M(t) \leq 2|O|\, \varepsilon^{-\tau/a}\, \frac{\beta^t}{1-\beta} \leq 4|O|\, \varepsilon^{-\tau/a}\, 2^{-t}, \tag{40}$$

which tends to 0. The process forgets, and it is ergodic by Lemma 4.12.

Suppose M does not halt. By Theorem 9.8, R_M tiles; let ξ be the cloned tiling configuration of a tiling. By Theorem 6.2 and Lemma 10.1, for every site v and every $t \geq 0$,

$$\mathbb{P}_\xi\big(X_t(v) \neq \xi(v)\big) < 1.01\,\varepsilon, \qquad \mathbb{P}_\perp\big(X_t(v) \text{ is live}\big) < 1.01\,\varepsilon. \tag{41}$$

Take the window $\{v\}$ and the event that v carries the live symbol $\xi(v)$. Then $\mathbb{P}_\xi(X_t(v) = \xi(v)) > 1 - 1.01\varepsilon$ and $\mathbb{P}_\perp(X_t(v) = \xi(v)) < 1.01\varepsilon$, so the process has memory with $\alpha = 1 - 2.02\varepsilon > 0$. By Lemma 4.13 it has two different invariant measures, and it is not ergodic.

Both cases hold for every $\varepsilon \in (0, \varepsilon_0]$, and j, k and ε_0 do not depend on M. □

Remark 10.3 (Scope of the constants). Since $B \geq 1$, we have $j \geq 10(1 + 37951 + 100) = 380520$, so the dimension $d = 2 + 2j$ given by Theorem 9.1 is a fixed integer, at least 761042. Its exact value depends on the constant B of Theorem 7.18, a fixed integer, shown to exist, that we do not evaluate. With the three-token refutation of Section 11, the same family gives $j = 3$ and $d = 8$ (Corollary 11.21). The rate ε_0 depends only on k, and β only on k and ε; the Wang set enters the forgetting bound only through $\tau(R_M)$. The alphabet $\mathcal{A}_{R_M}$ varies with M, so the least transition probability $\varepsilon/|\mathcal{A}_{R_M}|$ is not bounded below over the family. A fixed noise rate, as in the title, means a fixed replacement rate ε with uniform replacement on the alphabet of each member.

10.1 Corollaries

Corollary 10.4 (Every small rate)**.** *For every machine M the following are equivalent:*

(i) M halts;

(ii) $F_{R_M}^{[k]}$ with uniform replacement noise is ergodic at some rate $\varepsilon \in (0, \varepsilon_0]$;

(iii) it is ergodic at every rate $\varepsilon \in (0, \varepsilon_0]$.

Consequently, for each fixed $\varepsilon \in (0, \varepsilon_0]$, no algorithm decides, given a finite Wang set R, whether $F_R^{[k]}$ with uniform replacement noise at rate ε is ergodic.

Proof. By Lemma 10.1 the hypotheses of Theorems 5.1 and 6.2 used in the proof of Theorem 10.2 hold at every rate in $(0, \varepsilon_0]$. So (i) implies (iii), (iii) implies (ii), and if M does not halt the process is not ergodic at any rate in $(0, \varepsilon_0]$, which is (ii) implies (i). For the last claim, a decision procedure at rate ε, composed with the computable map $M \mapsto R_M$, would decide halting. □

The two cases are far apart, so the undecidability does not come from members near a boundary. At each $\varepsilon \in (0, \varepsilon_0]$, every halting member forgets exponentially fast, $\delta_O^M(t) \leq$

$4|O|\varepsilon^{-\tau/a}2^{-t}$ by (40), and every other member has memory with amplitude $1 - 2.02\varepsilon$ by (41); so no algorithm separates the members that forget exponentially fast from those that have memory with amplitude $1 - 2.02\varepsilon$, even when it is promised that the given member is of one of these two kinds.

Corollary 10.5 (Every Wang set at once)**.** *There is a total computable map $R \mapsto H_R$ from finite Wang sets to positive-rate probabilistic cellular automata on $\mathbb{Z}^{2+k}$, each with uniform replacement noise, such that for every $\varepsilon \in (0, \varepsilon_0]$ and every finite Wang set R, the automaton H_R at rate ε is ergodic if and only if R does not tile the plane.*

Proof. Let M_R be the machine that, for $n = 1, 2, \ldots$ in turn, checks whether R has a legal pattern on the $n \times n$ square, and halts at the first n for which it has none. The map $R \mapsto M_R$ is computable, and each check is finite. If R tiles, restrictions of a tiling give legal squares of every size, so M_R does not halt. If M_R does not halt, R has legal squares of every size; by compactness of $R^{\mathbb{Z}^2}$ (König's lemma on the finitely branching tree of legal centred squares) it tiles. So M_R halts if and only if R does not tile. Put $H_R = F^{[k]}_{R_{M_R}}$ with uniform replacement noise, and apply Theorem 10.2 to M_R. □

The automaton H_R is built from the recoded set R_{M_R}, not from R itself. So Corollary 10.5 gives one fixed dimension for every Wang set, whatever its token width, through the family of Theorem 9.8; this recoded construction by itself does not classify the thin-provider quorum rule of R. In contrast, Theorem 2.1 now classifies the raw clustered rule G_R directly, without this recoding.

The next corollary concerns other noise. For a probability law ν on $\mathcal{A}_R$, *ν-replacement noise* at rate ε is defined as in Definition 4.8, with the values $\zeta_t(v)$ independent with law ν in place of the uniform law. Uniform replacement is the case where ν is uniform. Lemmas 4.10, 4.12 and 4.13 and their proofs hold verbatim for ν-replacement noise: they use only the mixture formula and the fact that $X_1|_O$ depends on finitely many coordinates of X_0 and on independent noise.

Corollary 10.6 (Other replacement laws)**.** *For each machine M let ν_M be a probability law on $\mathcal{A}_{R_M}$, and let $\varepsilon \in (0, \varepsilon_0]$.*

(i) $F^{[k]}_{R_M}$ with ν_M-replacement noise at rate ε is ergodic if and only if M halts. Every transition probability of this process is at least $\varepsilon \min_s \nu_M(s)$, which is positive when ν_M has full support.

(ii) Suppose M does not halt. Let the replacement events be as in Definition 4.8, and let the values written at events be arbitrary random variables, which may depend on the event indicators, on the past of the run and on further randomness. Then the bounds (41) hold for the run from every cloned tiling configuration ξ and for the run from $\perp$, at every site and time.

Proof. (ii) The proof of Theorem 6.2 has a deterministic part and a counting part. The deterministic part (Lemma 6.5) holds for every realization of the events and of the values: a bad point that is not an event has two bad points one step earlier, in two different provider planes, and no point of time 0 is bad. The counting part uses the noise only through the bound $\mathbb{P}(\text{every point of } S \text{ is an event}) \le \varepsilon^{|S|}$ for finite sets S of points of positive time (Remark 6.9). Both parts hold under the stated hypotheses, so the bounds hold.

(i) Suppose M does not halt. Values drawn independently from ν_M are a special case of (ii), so (41) holds, the process has memory as in the proof of Theorem 10.2, and Lemma 4.13 shows that it is not ergodic.

Suppose M halts. In the proof of Theorem 5.1, the runs from x and from $\perp$ use the same indicators η and the same values ζ. The covering field E is a function of η alone (Definition 5.5). The transport and cover lemmas and Corollary 5.7 hold for every realization of η and ζ. The discrepancy path of Lemma 5.9 uses only that both runs take the common value at an event. The final count uses only the independence of the indicators and $\mathbb{P}(\eta_t(v) = 1) = \varepsilon$. None of these steps uses the law of ζ. So (29) and (30) hold for ν_M-replacement noise with the same constants, the process forgets, and Lemma 4.12 shows that it is ergodic.

Finally, given X_t, the value $X_{t+1}(v)$ equals s with probability $(1-\varepsilon)\mathbf{1}[F_{R_M}^{[k]}(X_t)(v) = s] + \varepsilon\,\nu_M(s) \geq \varepsilon\,\nu_M(s)$. □

The forgetting half needs the values to be drawn without looking at the configuration, because its coupling gives both runs the same value at each event. The memory half needs nothing about the values.

Corollary 10.7 (Mixing times are not computably bounded)**.** *Let $\varepsilon \in (0, \varepsilon_0]$ be rational, and use uniform replacement noise at rate ε.*

(i) For every machine M, window O and time t, the supremum in $\delta_O^M(t)$ is a maximum over finitely many restrictions of the start, and $\delta_O^M(t)$ is a rational number computable from M, O and t.

(ii) Let $t_{\mathrm{mix}}(M)$ be the least $t_0 \geq 0$ such that $\delta_{\{0\}}^M(t) < \frac{1}{2}$ for every $t \geq t_0$, and $t_{\mathrm{mix}}(M) = \infty$ if there is none. Then $t_{\mathrm{mix}}(M) < \infty$ if and only if M halts.

(iii) No computable function b satisfies $t_{\mathrm{mix}}(M) \leq b(M)$ for every halting M. No computable function b satisfies $\tau(R_M) \leq b(M)$ for every halting M.

Proof. (i) Let $\mathcal{N} = \Lambda \times \{c_0, \dots, c_k\}$ and $\mathcal{N}_t$ be as in Section 4, and write $\mathcal{A} = \mathcal{A}_{R_M}$ and $F = F_{R_M}^{[k]}$. Given X_s, the values $X_{s+1}(v)$ are independent, and $X_{s+1}(v) = s'$ with probability $(1-\varepsilon)\mathbf{1}[F(X_s)(v) = s'] + \varepsilon/|\mathcal{A}|$, where $F(X_s)(v)$ depends on X_s only on $v + \mathcal{N}$. Hence, for a window S and $w \in \mathcal{A}^S$,

$$\mathbb{P}_x\big(X_{s+1}|_S = w\big) = \sum_{y \in \mathcal{A}^{S+\mathcal{N}}} \mathbb{P}_x\big(X_s|_{S+\mathcal{N}} = y\big) \prod_{v \in S} \Big((1-\varepsilon)\mathbf{1}\big[F(y)(v) = w(v)\big] + \frac{\varepsilon}{|\mathcal{A}|} \Big).$$

where $F(y)(v)$ is the local rule applied to y on $v + \mathcal{N}$. Apply this t times, with $S = O + \mathcal{N}_{t-s-1}$ at step s, from $X_0|_{O+\mathcal{N}_t} = x|_{O+\mathcal{N}_t}$ to $X_t|_O$. Each step is a finite sum of products of rationals computable from M. So $\mathrm{Law}_x(X_t|_O)$ is a rational vector, computable from M, O, t and $x|_{O+\mathcal{N}_t}$, and it depends on x only through this restriction. The supremum over x is a maximum over the finite set $\mathcal{A}^{O+\mathcal{N}_t}$, and the total variation distance of two rational vectors is rational.

(ii) If M halts, (40) with $O = \{0\}$ gives $\delta_{\{0\}}^M(t) \to 0$, so $t_{\mathrm{mix}}(M) < \infty$. If M does not halt, let ξ be a cloned tiling configuration. By (41), for every t,

$$\delta_{\{0\}}^M(t) \geq \mathbb{P}_\xi\big(X_t(0) = \xi(0)\big) - \mathbb{P}_\perp\big(X_t(0) = \xi(0)\big) > 1 - 2.02\,\varepsilon > \tfrac{1}{2},$$

so $t_{\mathrm{mix}}(M) = \infty$.

(iii) Suppose b is computable and $t_{\mathrm{mix}}(M) \leq b(M)$ for every halting M. Given M, compute $\delta_{\{0\}}^M(b(M))$ by (i) and answer "halts" exactly when it is below $\frac{1}{2}$. If M halts, then $b(M) \geq t_{\mathrm{mix}}(M)$, so the value is below $\frac{1}{2}$. If M does not halt, the value exceeds $\frac{1}{2}$ by the proof of (ii). This decides the halting problem, which is impossible.

For the refutation time, let M halt and $\tau = \tau(R_M)$. By (40) with $|O| = 1$, $\delta_{\{0\}}^M(t) < \frac{1}{2}$ whenever $2^t > 8\varepsilon^{-\tau/a}$, that is, whenever $t > 3 + (\tau/a)\log_2(1/\varepsilon)$. So $t_{\mathrm{mix}}(M) \leq 4 + \tau L$ with $L = \lceil \log_2(1/\varepsilon) \rceil$, since $a \geq 1$. A computable b with $\tau(R_M) \leq b(M)$ for every halting M would give the computable bound $4 + b(M)L$ on $t_{\mathrm{mix}}(M)$, which we have excluded. □

The process and the all-bottom start are invariant under translations of $\mathbb{Z}^{2+k}$, so $\delta^M_{\{v\}} = \delta^M_{\{0\}}$ for every site v. Restricting to a smaller window does not increase total variation, so $\delta^M_O \geq \delta^M_{\{0\}}$ for every nonempty window O. The conclusions of Corollary 10.7 therefore hold for the mixing time of every nonempty window. Forgetting happens for every halting member, but its onset is not computably bounded. The bound (40) reads $\delta^M_O(t) \leq 2^{-(t-t_1)}$ with $t_1 = (\tau(R_M)/a)\log_2(1/\varepsilon) + \log_2(4|O|)$: a delay set by the refutation time, followed by decay at a rate that does not depend on M.

Ergodicity also has an equilibrium reading. A PCA on $\mathbb{Z}^d$ with a strictly positive transition table defines a Gibbs specification on $\mathbb{Z}^{d+1}$, its space-time interaction, in which time is one more lattice direction [21, 32, 9]. For $F^{[k]}_{R_M}$ with uniform replacement noise at rate ε, the table is

$$p_M(s \mid y) = (1-\varepsilon)\,\mathbf{1}\big[s = f_M(y)\big] + \frac{\varepsilon}{|\mathcal{A}_{R_M}|}, \qquad s \in \mathcal{A}_{R_M},\ y \in \mathcal{A}_{R_M}^{\mathcal{N}},$$

where f_M is the local rule of $F^{[k]}_{R_M}$ on $\mathcal{N}$. Each entry is at least $\varepsilon/|\mathcal{A}_{R_M}| > 0$. Write a space-time point as $(v,t) \in \mathbb{Z}^{2+k} \times \mathbb{Z}$; its parents are the points $(v+w, t-1)$, $w \in \mathcal{N}$. The *space-time interaction* Ψ_M assigns to each point (v,t) the energy $-\log p_M\big(\omega(v,t) \mid \omega \text{ at the parents of } (v,t)\big)$ of a configuration $\omega \in \mathcal{A}_{R_M}^{\mathbb{Z}^{3+k}}$. It has finite range and is invariant under all translations of $\mathbb{Z}^{3+k}$. Its specification $\gamma_M = (\gamma^\tau_{M,\Lambda})$ is the one of [9, eq. (6)]: for a finite set $\Lambda \subset \mathbb{Z}^{3+k}$ and a configuration τ outside Λ, the probability of a configuration ω on Λ is proportional to the product of the factors $p_M(\cdot \mid \cdot)$ of all points u such that u or a parent of u lies in Λ, read on ω inside Λ and on τ outside. A *Gibbs measure* for Ψ_M is a DLR state of this specification, as defined in [9, Section 2]: a probability measure on $\mathcal{A}_{R_M}^{\mathbb{Z}^{3+k}}$ whose conditional law on every finite Λ, given the configuration outside Λ, is $\gamma^\tau_{M,\Lambda}$ almost surely. No invariance is assumed. The specification has *exponential weak spatial mixing* [9, Definition 2.3] if there are $C, c > 0$ such that

$$d_{\mathrm{TV}}\big((\gamma^\tau_{M,\Lambda})_E, (\gamma^{\tau'}_{M,\Lambda})_E\big) \leq C \sum_{u \in E} e^{-c\,\mathrm{dist}(u,\Lambda^c)} \tag{42}$$

for every finite $\Lambda \subset \mathbb{Z}^{3+k}$, every $E \subseteq \Lambda$ and all configurations τ, τ' outside Λ. Here $(\cdot)_E$ is the marginal on E, distances are sup-norm distances, and Λ need not be connected.

Corollary 10.8 (Space-time Gibbs measures)**.** *Let $\varepsilon \in (0, \varepsilon_0]$.*

(i) If M halts, Ψ_M has exactly one Gibbs measure, among all probability measures and with no invariance assumed. It is the two-sided stationary trajectory law $\bar\pi_M$ of the invariant measure π_M of Theorem 10.2. Moreover there are $C_M, c_M > 0$, depending on M and ε, such that (42) holds with these constants, and $d_{\mathrm{TV}}\big((\gamma^\tau_{M,\Lambda})_E, (\bar\pi_M)_E\big) \leq C_M \sum_{u\in E} e^{-c_M \mathrm{dist}(u,\Lambda^c)}$, for every finite Λ, every $E \subseteq \Lambda$ and all τ, τ'.

(ii) If M does not halt, Ψ_M has at least two translation-invariant Gibbs measures.

(iii) Consequently each of the following properties holds for Ψ_M if and only if M halts: Ψ_M has exactly one Gibbs measure; Ψ_M has exactly one translation-invariant Gibbs measure; the specification of Ψ_M has exponential weak spatial mixing. For rational ε the table p_M is computable from M, so each of the three properties is undecidable on this computable family of strictly positive, finite-range, translation-invariant interactions on $\mathbb{Z}^{3+k}$.

Proof. We use five results of [9], stated there for every spatial dimension, every finite alphabet with at least two symbols, every finite neighbourhood containing the origin, and every strictly positive table. Our process fits: the alphabet is finite, $\mathcal{N}$ is finite and contains $0 = (0, c_0)$, and

every entry of p_M is positive. The five results are: the lemma that PCA histories satisfy the DLR equations [9, Lemma 2.4], so that every two-sided stationary trajectory law is a Gibbs measure; the corollary that periodic states are stationary trajectories [9, Corollary 10.4], which with that lemma identifies the translation-invariant Gibbs measures with the trajectory laws of translation-invariant stationary laws, the correspondence of Goldstein, Kuik, Lebowitz and Maes [21]; item (i) of the theorem on the Gibbs simplex of an ergodic PCA [9, Theorem 10.10]; the prefix-cap theorem [9, Theorem 13.3]; and its corollary for basic flows [9, Corollary 13.7], which we state in our notation. *Let a PCA with finite alphabet, finite neighbourhood containing the origin and strictly positive table have a local-map representation, a random map m from configurations on the neighbourhood to the alphabet with $\mathbb{P}(m(y) = s) = p(s \mid y)$ for all y and s, applied independently at every space-time point to every start at once, whose common-map coefficient q_t, the probability that the composite of these maps from time 0 to time t, read at the origin, is not constant on the configurations at time 0, satisfies $q_t \le Ce^{-\alpha t}$ for all $t \ge 0$ and some $C, \alpha > 0$. Then its space-time specification has exactly one DLR state, the two-sided stationary trajectory law of its unique stationary law, and it has exponential weak spatial mixing, together with the bound of Corollary 10.8(i) against that trajectory law.* The notion of ergodicity in [9], a stationary law that attracts every initial law, is the one of Definition 4.9.

If $\mathcal{A}_{R_M} = \{\perp\}$, then R_M has no legal 3×3 block, so it does not tile and M halts by Theorem 9.8. The configuration space is then one point, which carries exactly one measure, and (i) holds for every $C_M, c_M > 0$, since all the total-variation distances in (i) vanish. Otherwise $\mathcal{A}_{R_M}$ has at least three symbols.

(i). Suppose M halts, and let $\tau = \tau(R_M)$ be its refutation time (Theorem 9.8). Uniform replacement is a local-map representation of p_M: at the point (v, t) apply the local rule f_M if $\eta_t(v) = 0$, and the constant map with value $\zeta_t(v)$ if $\eta_t(v) = 1$. These maps are independent over space-time points, and $\mathbb{P}(\text{the map sends } y \text{ to } s) = (1-\varepsilon)\mathbf{1}[s = f_M(y)] + \varepsilon/|\mathcal{A}_{R_M}| = p_M(s \mid y)$. Applied to every start at once, they give the runs X^x of the common-noise coupling of all starts (Section 5.7). So the common-map coefficient is $q_t = \mathbb{P}(x \mapsto X_t^x(0) \text{ is not constant})$, and by Corollary 5.12 and Lemma 10.1,

$$q_t \le 2\varepsilon^{-\tau/a}\,\frac{\beta^t}{1-\beta} \le 4\varepsilon^{-\tau/a}\,2^{-t} \qquad (t \ge 0).$$

This is the hypothesis of [9, Corollary 13.7], with spatial dimension $2 + k$, $C = 4\varepsilon^{-\tau/a}$ and $\alpha = \log 2$. That corollary gives exactly one Gibbs measure, the two-sided stationary trajectory law of the unique stationary law, and the two bounds of (i). The unique stationary law is π_M, by Theorem 10.2.

(ii). Suppose M does not halt. By Theorem 9.8, R_M tiles, so the set $\mathcal{T}_M \subseteq R_M^{\mathbb{Z}^2}$ of tilings is nonempty. It is closed in the compact product space and invariant under the shifts of $\mathbb{Z}^2$. Fix $z_0 \in \mathcal{T}_M$ and let m_n be the average of the point masses at the translates $z_0(\cdot + w)$, $w \in [-n, n]^2$. A weak limit point m of (m_n) is a probability measure on $\mathcal{T}_M$, and it is shift invariant: for a unit vector e of $\mathbb{Z}^2$, the measures m_n and their translates by e differ by at most $2(2n+1)/(2n+1)^2$ in total variation. For a tiling z let ξ_z be its cloned tiling configuration. The map $z \mapsto \xi_z$ is continuous, since $\xi_z(u, P)$ depends on z only on $u + \Lambda$. It satisfies $\xi_{z(\cdot+w)} = \xi_z(\cdot + (w, 0))$, and ξ_z is invariant under translations of the transverse coordinate. So the image κ of m is a probability measure on Ω that is invariant under all translations of $\mathbb{Z}^{2+k}$.

Fix a site v. The function $x \mapsto \mathbb{P}_x(X_t(v) \text{ is live})$ depends on finitely many coordinates of x, and by (41) it exceeds $1 - 1.01\varepsilon$ at every ξ_z. By the mixture formula of Section 4, $\mathbb{P}_\kappa(X_t(v) \text{ is live}) \ge 1 - 1.01\varepsilon$ for every t. The laws $\kappa\Phi^t$ are translation invariant, because κ is and the process commutes with translations. By Lemma 4.10, the averages $\frac{1}{n}\sum_{t<n} \kappa\Phi^t$

have a weak limit point π^+, which is invariant. It is translation invariant, as a weak limit of translation-invariant measures, and $\pi^+(\omega(v) \text{ is live}) \geq 1 - 1.01\varepsilon$, because the cylinder is open and closed. In the same way the averages from the all-bottom configuration, which is translation invariant, have an invariant, translation-invariant limit point π^-, with $\pi^-(\omega(v) \text{ is live}) \leq 1.01\varepsilon$ by (41). As $1.01\varepsilon < 1 - 1.01\varepsilon$, the measures π^+ and π^- differ.

For an invariant π, its two-sided stationary trajectory law is the measure on $\mathcal{A}_{R_M}^{\mathbb{Z}^{3+k}}$ whose restriction to the times $t_0, \ldots, t_0 + n$ is the law of $(X_0, \ldots, X_n)$ under Law_π; these restrictions are consistent because π is invariant, and Kolmogorov's theorem gives the measure. For $\pi = \pi^\pm$ it is invariant under time translations, and under space translations because $\pi^\pm$ is and the process commutes with them. So both trajectory laws are translation-invariant Gibbs measures. Their marginals at time 0 are $\pi^+ \neq \pi^-$, so they differ.

(iii). Exponential weak spatial mixing allows at most one Gibbs measure. Let ν, ν' be Gibbs measures and E a finite set. For every finite $\Lambda \supseteq E$ the DLR equations give $\nu_E = \int (\gamma^\tau_{M,\Lambda})_E \, \nu(d\tau)$ and $\nu'_E = \int (\gamma^{\tau'}_{M,\Lambda})_E \, \nu'(d\tau')$, so by convexity of total variation $d_{\mathrm{TV}}(\nu_E, \nu'_E) \leq C \sum_{u \in E} e^{-c \operatorname{dist}(u, \Lambda^c)}$, which tends to 0 as Λ grows. So ν and ν' agree on every cylinder, and $\nu = \nu'$. If M halts, (i) gives the first and the third property; the unique Gibbs measure is translation invariant, since translates of Gibbs measures of the translation-invariant interaction Ψ_M are Gibbs measures, so the second property holds too. If M does not halt, (ii) gives two distinct translation-invariant Gibbs measures, so the first two properties fail, and the third fails by the argument just given. An algorithm deciding any of the three properties for the tables p_M would decide halting. □

Sandhu independently proves fixed-range, fixed-temperature Gibbs-uniqueness undecidability on $\mathbb{Z}^2$ by a finite-certificate construction [54, Theorem 1.1]. Its support shapes are fixed, the finite alphabet varies, the local energies lie in $\{0, 1, 20\}$ and $\beta = 1$: halting gives one Gibbs state, while nonhalting gives the full zero-field Ising simplex with two extremal states. Every state in that family is TI, so it also gives TI Gibbs-uniqueness undecidability.

The constants C_M and c_M depend on M, through $\tau(R_M)$ and $|\mathcal{A}_{R_M}|$, and on ε; no uniformity over the family is claimed. With $k = 6$ and $\varepsilon_0 = 490000^{-99}$, the space-time lattice is $\mathbb{Z}^9$ (Remark 11.22).

The translation-invariant part of (i) needs only ergodicity: by the correspondence of Goldstein, Kuik, Lebowitz and Maes [21], in the form of [9, Corollary 10.4 and Theorem 10.10(i)], an ergodic strictly positive automaton has exactly one translation-invariant Gibbs measure, and version 1.3.0 of this paper proved (i) in that form. As recorded in [9, Remark 13.9], it also follows from a theorem of van den Berg and Steif [49, Theorem 2.1] when the trajectory law is a finitary factor of an i.i.d. process, which coupling from the past provides for a coalescing representation. Uniqueness among all Gibbs measures, with weak mixing, is classical where the space-time specification satisfies the condition of Dobrushin [15] or the block condition of Dobrushin and Shlosman [16], in particular at high noise; we take this attribution from [9, Remark 13.9], and we have not checked the exact scope of the weak-mixing statements in the original papers. Here uniqueness and weak mixing come instead from the coalescence of Corollary 5.12, through the prefix-cap theorem [9, Theorem 13.3]. Undecidability of ergodicity for automata with some transition probabilities equal to zero is classical [29, 48, 5].

Remark 10.9 (Arithmetical placement). For a PCA with a strictly positive rational table, ergodicity is a Π^0_2 property of the table [9, Remark 10.9]. The reason is short. By Halawi's theorem [22], an ergodic positive PCA is uniformly ergodic. So ergodicity says: for every window O and every $m \geq 1$ there is t with $\Delta_O(t) < 1/m$, where $\Delta_O(t) = \sup_{x,y} d_{\mathrm{TV}}\big(\mathrm{Law}_x(X_t|_O), \mathrm{Law}_y(X_t|_O)\big)$. The argument of Corollary 10.7(i) computes $\Delta_O(t)$ exactly, and $\Delta_O(t)$ does not increase with

t [9, Remark 10.9]. Theorem 10.2 gives a lower bound at the first level. For each rational $\varepsilon \in (0, \varepsilon_0]$, the computable map $M \mapsto F^{[k]}_{R_M}$ reduces the halting problem to ergodicity at rate ε. So ergodicity of positive-rate PCA on $\mathbb{Z}^{2+k}$ at the fixed rate ε, with uniform replacement, is Σ^0_1-hard, and non-ergodicity is Π^0_1-hard; in particular the non-ergodic members do not form a computably enumerable set. Whether ergodicity at a fixed rate is Π^0_2-complete is not settled by Theorem 10.2.

10.2 Perturbed kernels

The machine-family theorem fixes one noise, uniform replacement at rate ε. The dichotomy does not depend on that choice. It holds for every probabilistic cellular automaton whose table lies within a fixed total-variation distance of the noiseless rule, with a radius that depends on k only. The proof reuses the proofs of Theorems 5.1 and 6.2. The memory half needs nothing beyond Remark 6.9. The forgetting half needs two changes: the discrepancy path of Lemma 5.9 may pass through faults, and a fault on that path receives a contour with no edges.

Throughout this subsection R is a Wang set, $k \geq 2$, $q = k + 1$, $a = 2q^2 + 1$, $D_0 = 50q$, $\mathcal{A} = \mathcal{A}_R$, $F = F^{[k]}_R$ and $\mathcal{N} = \Lambda \times \{c_0, \dots, c_k\}$. For $x \in \Omega$ and a site v let $x_v \in \mathcal{A}^{\mathcal{N}}$ be given by $x_v(w) = x(v + w)$, and let $f \colon \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$ be the local rule of F, so that $F(x)(v) = f(x_v)$.

Definition 10.10 (Local map fields)**.** A *local map field* consists of random variables $\chi_t(v) \in \{0, 1\}$ and random maps $f_{t,v} \colon \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$, for $t \geq 1$ and $v \in \mathbb{Z}^{2+k}$, all defined on one probability space, such that $f_{t,v} = f$ whenever $\chi_t(v) = 0$. A point (v, t) with $\chi_t(v) = 1$ is a *fault*. The *run* from $x \in \Omega$ is $X_0 = x$ and $X_t(v) = f_{t,v}\big((X_{t-1})_v\big)$ for $t \geq 1$. Runs from several starts driven by one field form the *common-map coupling*. For $0 < \gamma < 1$ the field is *γ-bounded* if

$$\mathbb{P}\big(\chi_t(v) = 1 \text{ for every } (v, t) \in S\big) \leq \gamma^{|S|}$$

for every finite set S of points of positive time.

Nothing else is assumed. Faults may be correlated in space and in time, and the map at a fault may be any function of its inputs, chosen at random jointly with all other maps. Bounds of this form are the noise condition of Toom's stability theorem [47]; see also [19]. Replacement noise at rate ε (Definition 4.8) is the ε-bounded field with $\chi_t(v) = \eta_t(v)$, in which $f_{t,v}$ is the constant map with value $\zeta_t(v)$ when $\eta_t(v) = 1$.

Lemma 10.11 (Common local maps)**.** *Let $0 < \gamma < 1$, let $(\chi, f_{t,v})$ be a γ-bounded local map field, and let E be the domination field of Definition 5.5 computed from χ in place of η.*

(i) *Let R tile, let ξ be a cloned tiling configuration, and suppose $\lambda = 2D_0^2\gamma^{1/q^2} < 1$. Let X be the run from ξ and Y the run from $\perp$. For every site v and every $t \geq 0$,*

$$\mathbb{P}\big(X_t(v) \neq \xi(v)\big) \leq \gamma\Big(1 + \frac{2\lambda}{1 - \lambda}\Big), \qquad \mathbb{P}\big(Y_t(v) \neq \perp\big) \leq \gamma\Big(1 + \frac{2\lambda}{1 - \lambda}\Big).$$

(ii) *Let $j \geq 1$, $k \geq \max(2, 2j)$, let R be j-token refutable with time τ, and suppose $\beta = 2D_0^2\gamma^{1/a} < 1$. Let X and Y be the runs from configurations x and y under the common-map coupling. For every site v and every $t \geq 0$,*

$$\mathbb{P}\big(X_t(v) \neq Y_t(v)\big) \leq 2\gamma^{-\tau/a}\,\frac{\beta^t}{1 - \beta}.$$

Proof. In both parts, read "event" as "fault" in Sections 5 and 6. At a point that is not a fault the run applies F: $X_t(v) = f\big((X_{t-1})_v\big) = F(X_{t-1})(v)$.

(i) The proof of Lemma 6.5 uses a point that is not an event only through $X_t(u,P) = F(X_{t-1})(u,P)$, and the start only at time 0. It never uses the value written at an event. So Lemma 6.5 and Lemma 6.6 hold for every realization of the field. The proof of Lemma 6.8 uses the indicators only through the recursive definition of E and through the bound that every finite set S of points of positive time consists of events with probability at most $\varepsilon^{|S|}$ (Remark 6.9). Here that bound holds with γ in place of ε. With $|S| = 1$ it also gives $\mathbb{P}(\chi_t(v) = 1) \le \gamma$. So the proof of Lemma 6.8 gives $\mathbb{P}(E_t(v) = 1) \le \gamma(1 + 2\lambda/(1-\lambda))$, and Lemma 6.6 gives (i).

(ii) If $L_R = \emptyset$, then $\mathcal{A} = \{\perp\}$ and there is nothing to prove; otherwise $\tau \ge 1$, as in the proof of Theorem 5.1. Fix r with $S_\tau(0,\dots,0) = \emptyset$ for the box $V = \{-r,\dots,r\}^2$.

Step 1: the cover. Lemma 5.4, Lemma 5.6 and Corollary 5.7 are statements about one run and one realization. Their proofs use a point that is not an event only through $X_t(v) = F(X_{t-1})(v)$ (Step 2 of the proof of Lemma 5.4), and an event only through $E = 1$ there. They never use the value written at an event. So they hold for every realization of the field and for the run from every start, with the one field E, which depends on χ alone. Hence, for $t \ge \tau$, every live value $X_t(v)$ or $Y_t(v)$ has $E_t(v) = 1$.

Step 2: the discrepancy path. Let $t \ge 1$ and $X_t(v) \ne Y_t(v)$. Since $X_t(v) = f_{t,v}\big((X_{t-1})_v\big)$ and $Y_t(v) = f_{t,v}\big((Y_{t-1})_v\big)$ with one and the same map $f_{t,v}$, the inputs differ: $X_{t-1}(v') \ne Y_{t-1}(v')$ for some $v' \in v + \mathcal{N} = \mathcal{N}(v)$. This holds whether or not (v,t) is a fault. Repeating down to time 0 gives sites $v_t = v, v_{t-1}, \dots, v_0$ with $v_{s-1} \in \mathcal{N}(v_s)$ and $X_s(v_s) \ne Y_s(v_s)$ for $0 \le s \le t$. So (v_s, s) and $(v_{s-1}, s-1)$ are joined in $\mathcal{G}$. This replaces Lemma 5.9. The points (v_s, s) may now be faults.

Step 3: marks on the path. Let $\tau \le s \le t$. If (v_s, s) is a fault, then $E_s(v_s) = 1$ by Definition 5.5. Otherwise, $X_s(v_s)$ and $Y_s(v_s)$ are different symbols of $L_R \cup \{\perp\}$, so one of them is live, and Step 1 gives $E_s(v_s) = 1$.

Step 4: packing. Run the proof of Lemma 5.10 with one change. Put $t_1 = t$. Given $t_l \ge \tau$ and $\rho_l = (v_{t_l}, t_l)$:

- If ρ_l is not a fault, apply Lemma 5.8, with $\mathcal{U}$ the marked points and $\mathcal{U}_\mathrm{f}$ the faults, at ρ_l. This gives m_l, Σ_l, K_l and s_l as there.
- If ρ_l is a fault, put $K_l = \emptyset$, $\Sigma_l = \{\rho_l\}$, $m_l = 1$ and $s_l = t_l$.

In both cases:

- Σ_l is a set of m_l faults with times in $[s_l, t_l]$, and every point of Σ_l is ρ_l or a vertex of K_l;
- every vertex of K_l has time in $[s_l, t_l]$;
- if $K_l \ne \emptyset$, its edges form a connected graph with ρ_l among its vertices (Lemma 5.8(i));
- and
$$|K_l| \le q^2(m_l - 1), \qquad t_l - s_l \le q^2(m_l - 1).$$

For the second case both sides of each inequality are 0. These are the only properties of the contours that the proof of Lemma 5.10 uses. Continue with $t_{l+1} = s_l - 1$ while $s_l - 1 \ge \tau$. The slabs $[s_l, t_l]$ are pairwise disjoint, so the sets Σ_l are pairwise disjoint. Moreover $t - \tau + 1 \le \sum_l (t_l - s_l + 1) \le q^2 m$ with $m = \sum_l m_l$. Let K consist of the t path edges together with all K_l. The path edges form a connected graph whose vertices include every ρ_l, since $t \ge \tau \ge 1$. Each nonempty K_l is connected and contains ρ_l. So K is connected, its vertices include (v,t) and $\Sigma = \bigcup_l \Sigma_l$, and $e = |K|$ satisfies $t \le e \le \tau - 1 + 2q^2 m$. Hence $|\Sigma| = m \ge (e - \tau)/a$. The

proof of Lemma 5.10 used the start $\perp$ only to know that one of two different values is live, and Step 3 supplies this for any two runs.

Step 5: counting. For $t < \tau$ the right side is at least 2, as in the proof of Theorem 5.1. For $t \geq \tau$, by Step 4 the event $X_t(v) \neq Y_t(v)$ is contained in the union, over $e \geq t$, over connected sets K of e edges of $\mathcal{G}$ with (v,t) among their vertices, and over sets Σ of at least $(e-\tau)/a$ of their vertices, of the events $\{\chi = 1 \text{ on } \Sigma\}$. These (K,Σ) form a countable family of fixed sets. Each event has probability at most $\gamma^{|\Sigma|} \leq \gamma^{(e-\tau)/a}$, by γ-boundedness; sets Σ that contain a point of time 0 have probability 0. By Lemma 5.11 there are at most $D_0^{2e}2^{e+1}$ pairs (K,Σ) for each e. Summing as in the proof of Theorem 5.1 gives the bound. No independence is used, only the union bound. □

A *local kernel* on $\mathcal{A}$ is a family $Q = (Q(\cdot \mid y))_{y \in \mathcal{A}^{\mathcal{N}}}$ of probability laws on $\mathcal{A}$. The *probabilistic cellular automaton with kernel* Q is the process with $X_0 = x$ in which, given $X_0, \dots, X_t$, the values $X_{t+1}(v)$ are independent and $X_{t+1}(v) = s$ with probability $Q\big(s \mid (X_t)_v\big)$. We write Law_x^Q, $\mathbb{P}_x^Q$ and $\delta_O^Q(t)$ as in Section 4. Lemmas 4.10, 4.12 and 4.13 and their proofs hold verbatim for it: they use only the mixture formula and the fact that $X_1|_O$ depends on X_0 only on $O + \mathcal{N}$ and on independent randomness. Uniform replacement at rate ε is the kernel $p_\varepsilon(s \mid y) = (1-\varepsilon)\mathbf{1}[s = f(y)] + \varepsilon/|\mathcal{A}|$. On local kernels over $\mathcal{A}$ put

$$d(Q,Q') = \max_{y \in \mathcal{A}^{\mathcal{N}}} d_{\mathrm{TV}}\big(Q(\cdot \mid y), Q'(\cdot \mid y)\big).$$

Then $d(Q, \delta_f) = \max_y \big(1 - Q(f(y) \mid y)\big)$, where δ_f is the noiseless rule $\delta_f(\cdot \mid y) = \delta_{f(y)}$.

Corollary 10.12 (Perturbed kernels)**.** *Let j, k, q, a, D_0 and ε_0 be as in this section, or as in Corollary 11.21. For a machine M write $\mathcal{A}_M = \mathcal{A}_{R_M}$, let f_M be the local rule of $F_{R_M}^{[k]}$, and let $p_{M,\varepsilon}$ be uniform replacement at rate ε on $\mathcal{A}_M$.*

(i) *Let $0 < \gamma \leq 2\varepsilon_0$ and let Q be a local kernel on $\mathcal{A}_M$ with*

$$Q\big(f_M(y) \mid y\big) \;\geq\; 1 - \gamma \qquad \text{for every } y \in \mathcal{A}_M^{\mathcal{N}}.$$

The probabilistic cellular automaton with kernel Q is ergodic if and only if M halts. More precisely, put $\beta = 2D_0^2\gamma^{1/a}$; then $\beta < 0.52$. If M halts and $\tau = \tau(R_M)$, then for every window O and every $t \geq 0$

$$\delta_O^Q(t) \leq 2|O|\,\gamma^{-\tau/a}\,\frac{\beta^t}{1-\beta}.$$

If M does not halt and ξ is a cloned tiling configuration, then for every site v and every $t \geq 0$

$$\mathbb{P}_\xi^Q\big(X_t(v) \neq \xi(v)\big) < 1.01\,\gamma, \qquad \mathbb{P}_\perp^Q\big(X_t(v) \text{ is live}\big) < 1.01\,\gamma,$$

and the process has memory with $\alpha = 1 - 2.02\gamma$.

(ii) *Let $\varepsilon \in (0, \varepsilon_0]$. Every local kernel Q on $\mathcal{A}_M$ with $d(Q, p_{M,\varepsilon}) \leq \varepsilon_0$ satisfies the hypothesis of (i) with $\gamma = 2\varepsilon_0$. So the probabilistic cellular automaton with kernel Q is ergodic if and only if M halts. The radius ε_0 depends on k only. It is the same for every machine M and every rate $\varepsilon \in (0, \varepsilon_0]$.*

(iii) *Let $M \mapsto Q_M$ be a total computable map to local kernels on $\mathcal{A}_M$ with rational entries, each satisfying the hypothesis of (i) with one fixed $\gamma \in (0, 2\varepsilon_0]$. Then Q_M is ergodic if and only if M halts. So no algorithm decides ergodicity for such a family.*

Proof. A map field for Q. Let Q and γ be as in (i), and put

$$Q'(s \mid y) = \frac{Q(s \mid y) - (1-\gamma)\mathbf{1}[s = f_M(y)]}{\gamma}.$$

By hypothesis $Q'(\cdot \mid y)$ is a probability law on $\mathcal{A}_M$ for every y. Enumerate $\mathcal{A}_M = \{s_1, \dots, s_n\}$. For $y \in \mathcal{A}_M^{\mathcal{N}}$ and $w \in (0,1]$ let $g(y,w) = s_r$ for the least r with $\sum_{r' \le r} Q'(s_{r'} \mid y) \ge w$. If W is uniform on $(0,1]$, then $g(y,W)$ has law $Q'(\cdot \mid y)$. Let $\chi_t(v)$ be Bernoulli with parameter γ and $W_t(v)$ uniform on $(0,1]$, all independent. Put $f_{t,v} = f_M$ if $\chi_t(v) = 0$ and $f_{t,v} = g(\cdot, W_t(v))$ if $\chi_t(v) = 1$. This is a local map field, and it is γ-bounded because $\mathbb{P}(\chi = 1 \text{ on } S) = \gamma^{|S|}$.

Let X be its run from x. Then $X_0, \dots, X_t$ are functions of x and of the variables of times at most t. The variables $(\chi_{t+1}(v), W_{t+1}(v))$, over all v, are independent of these and of each other. So, given $X_0, \dots, X_t$, the values $X_{t+1}(v)$ are independent, and

$$\mathbb{P}\big(X_{t+1}(v) = s \;\big|\; X_0, \dots, X_t\big) = (1-\gamma)\mathbf{1}\big[s = f_M((X_t)_v)\big] + \gamma\, Q'\big(s \mid (X_t)_v\big) = Q\big(s \mid (X_t)_v\big).$$

Hence the run from x has law Law_x^Q, and the common-map coupling of the runs from x and from $\perp$ is a coupling of Law_x^Q and $\mathrm{Law}_\perp^Q$.

Constants. Recall $\varepsilon_0 = (4D_0^2)^{-a}$, with $q \ge 3$, $D_0 \ge 150$ and $a \ge 19$ (Lemma 10.1). From $\gamma \le 2\varepsilon_0$ we get $\beta = 2D_0^2\gamma^{1/a} \le 2^{1/a}/2 \le 2^{-18/19} < 0.52$. Since $a/q^2 > 2$ and $4D_0^2 > 1$,

$$\gamma^{1/q^2} \le 2^{1/q^2}(4D_0^2)^{-2}, \qquad \text{so} \qquad \lambda = 2D_0^2\gamma^{1/q^2} \le \frac{2^{1/9}}{8D_0^2} < 10^{-5},$$

and $2\lambda/(1-\lambda) < 0.01$. Also $\gamma \le 2\varepsilon_0 < 10^{-8}$, so $1 - 2.02\gamma > 0$.

(i), M halts. By Theorem 9.8, or Theorem 11.1 with the constants of Corollary 11.21, R_M is j-token refutable with time $\tau = \tau(R_M)$, and $k \ge \max(2, 2j)$. For a window O and a start x, Lemma 10.11(ii) with $y = \perp$ gives

$$d_{\mathrm{TV}}\big(\mathrm{Law}_x^Q(X_t|_O), \mathrm{Law}_\perp^Q(X_t|_O)\big) \le \sum_{v \in O} \mathbb{P}\big(X_t(v) \ne Y_t(v)\big) \le 2|O|\,\gamma^{-\tau/a}\,\frac{\beta^t}{1-\beta},$$

which tends to 0. The process forgets, and it is ergodic by Lemma 4.12.

(i), M does not halt. R_M tiles (Theorem 9.8). By Lemma 10.11(i) and the constants, both probabilities are at most $\gamma(1 + 2\lambda/(1-\lambda)) < 1.01\gamma$. As in the proof of Theorem 10.2, $\mathbb{P}_\xi^Q(X_t(v) = \xi(v)) > 1 - 1.01\gamma$ and $\mathbb{P}_\perp^Q(X_t(v) = \xi(v)) < 1.01\gamma$ for every t. So the process has memory with $\alpha = 1 - 2.02\gamma > 0$, and Lemma 4.13 shows that it is not ergodic.

(ii) For every y,

$$d_{\mathrm{TV}}\big(Q(\cdot \mid y), p_{M,\varepsilon}(\cdot \mid y)\big) \ge p_{M,\varepsilon}(f_M(y) \mid y) - Q(f_M(y) \mid y) = 1 - \varepsilon + \frac{\varepsilon}{|\mathcal{A}_M|} - Q(f_M(y) \mid y).$$

So $Q(f_M(y) \mid y) \ge 1 - \varepsilon - d(Q, p_{M,\varepsilon}) \ge 1 - 2\varepsilon_0$. Apply (i) with $\gamma = 2\varepsilon_0$.

(iii) By (i), the set of M with Q_M ergodic is the set of halting machines, which is not decidable. □

Remark 10.13 (Scope of Corollary 10.12). (a) *Uniformity needs the growing alphabet.* The ball in (ii) lives in the space of local kernels on $\mathcal{A}_M$, whose dimension grows with M. This is unavoidable. For one fixed finite alphabet and neighbourhood, the local kernels form a compact metric space under d. Suppose the members of a computable family with rational entries kept their label on closed balls of one fixed rational radius r. Then members with different labels would lie at distance more than r. Take a finite net of

rational kernels with mesh less than $r/2$. Classify each member by a net point at distance less than $r/2$, which exists and is found by exact rational computation. Two members with the same net point lie within distance r of each other, so they share a label, and a finite table decides the labels.

(b) *Positive rates.* The ball in (ii) contains kernels with zero entries, among them the noiseless rule δ_{f_M}. Its positive-rate members, among them $p_{M,\varepsilon}$, are positive-rate probabilistic cellular automata in the sense of the machine-family theorem. As in Remark 10.3, their least transition probability is not bounded below over the family.

(c) *The joint bound is used.* What the proof of (i) uses is a γ-bounded common-map representation. For a kernel whose sites update independently given the past, the construction in the proof provides one. Closeness of the one-site conditional laws alone is not enough. Consider the transition that, with probability γ, sets every site to $\perp$ at once, and otherwise updates the sites independently by $p_{M,\varepsilon}$. Each one-site conditional law is within γ of $p_{M,\varepsilon}$. Couple the runs by sharing the reset coin and all replacement variables. Then the runs from any two starts agree after the first reset, so this process forgets and is ergodic for every M. When M does not halt, it has no γ'-bounded common-map representation for any $\gamma' < 1$: from a cloned tiling configuration ξ, a reset makes every site of any finite set wrong at once, with probability γ.

(d) *Site-dependent kernels.* The proof of (i) builds the map field site by site. So (i) holds verbatim when each site v has its own kernel Q_v satisfying the hypothesis with one γ.

(e) *Correlated faults.* Suppose the layers $(\chi_t(\cdot), f_{t,\cdot})$, $t = 1, 2, \ldots$, of a local map field are independent and identically distributed, and each layer is γ-bounded for sets of sites at one time, with arbitrary correlations inside a layer. Then:

- the field is γ-bounded, by independence of the layers;
- the runs form a time-homogeneous Markov process in which $X_1|_O$ depends on X_0 only on $O + \mathcal{N}$ and on randomness independent of X_0;
- so Lemmas 4.10, 4.12 and 4.13 apply verbatim;
- and with $\gamma \leq 2\varepsilon_0$, Lemma 10.11 and the constants above give the dichotomy of (i), with the same bounds, although the process is not a probabilistic cellular automaton.

For a γ-bounded field that is not time-homogeneous, Lemma 10.11 still gives the separation of (i) at every time and forgetting under the common-map coupling, but invariant measures are not defined. Forgetting needs the maps to read only $\mathcal{N}$ at the previous time and both runs to use the same map. Memory needs neither (Lemma 10.11(i), Remark 6.9).

Remark 10.14 (Gibbs measures for other noises). The proof of Corollary 10.8 uses the noise in two places: in (i), a coalescence bound for a local-map representation of the table; in (ii), the memory bounds (41) and translation invariance. Both are available in two further settings, so the classification of Corollary 10.8 holds there too, for the space-time interaction defined from the table as Ψ_M is defined from p_M.

(a) *Other replacement laws.* Let $\varepsilon \in (0, \varepsilon_0]$ and let ν_M have full support (Corollary 10.6). The table $(1-\varepsilon)\mathbf{1}[s = f_M(y)] + \varepsilon\,\nu_M(s)$ is strictly positive, and the maps of the proof of Corollary 10.8(i), with values $\zeta_t(v)$ of law ν_M, represent it. The proof of Corollary 5.12

never uses the law of ζ, so the bound on q_t is unchanged. The memory bounds hold by Corollary 10.6(ii), and the process commutes with translations.

(b) *Perturbed kernels.* Let Q be a strictly positive local kernel on $\mathcal{A}_{R_M}$ with $Q(f_M(y) \mid y) \geq 1-\gamma$ for every y, where $0 < \gamma \leq 2\varepsilon_0$ (Corollary 10.12(i)); by Corollary 10.12(ii) this includes every strictly positive kernel within distance ε_0 of a member $p_{M,\varepsilon}$. The map field built in the proof of Corollary 10.12 applies at each point the map f_M or the map $g(\cdot, W_t(v))$, independently over points, so it is a local-map representation of Q. In Steps 1–5 of the proof of Lemma 10.11(ii), the event $X_t(v) \neq Y_t(v)$ is contained in the union of the events $\{\chi = 1 \text{ on } \Sigma\}$ over a family of pairs (K, Σ) that does not depend on the two starts. The argument of Corollary 5.12 then gives $q_t \leq 2\gamma^{-\tau/a}\beta^t/(1-\beta)$, with $\beta = 2D_0^2\gamma^{1/a} < 0.52$. The memory bounds of Corollary 10.12(i), with $1.01\gamma < 1 - 1.01\gamma$, replace (41), and the process commutes with translations because one kernel acts at every site.

In both settings, if M halts, [9, Corollary 13.7] gives exactly one Gibbs measure and exponential weak spatial mixing; if M does not halt, the proof of Corollary 10.8(ii) gives two translation-invariant Gibbs measures. Kernels in the ball of Corollary 10.12(ii) that have zero entries do not define a space-time interaction of this kind. This remark has been checked less thoroughly than the corollaries of this section; see "Verification status."

10.3 Stacked automata: one phase or a continuum

The machine-family theorem separates automata with one invariant measure from automata with at least two. Running the members on parallel copies, each copy reading the next, widens the gap: the nonhalting members then have a continuum of extreme invariant measures. Throughout this subsection $k = 6$, $a = 99$ and $\varepsilon_0 = 490000^{-99}$, as in Corollary 11.21, and $\varepsilon \in (0, \varepsilon_0]$. For a machine M we write $\mathcal{A} = \mathcal{A}_{R_M}$, K_M for the automaton $F^{[6]}_{R_M}$ on $\mathbb{Z}^8$ with uniform replacement at rate ε, f_M for its local rule on $\mathcal{N} = \Lambda \times \{c_0, \dots, c_6\}$, and p_M for its table.

Definition 10.15 (The stack)**.** The *stack* G_M is the probabilistic cellular automaton on $\mathbb{Z}^8 \times \mathbb{Z} = \mathbb{Z}^9$, with alphabet $\mathcal{A}$, in which slice $h \in \mathbb{Z}$ at time $t+1$ is computed from slice $h+1$ at time t by the rule of K_M. For $x \in \mathcal{A}^{\mathbb{Z}^9}$ and $(v,h) \in \mathbb{Z}^9$ let $x_{(v,h)} \in \mathcal{A}^{\mathcal{N}}$ be the map $w \mapsto x(v+w, h+1)$. With independent variables $\eta_t(v,h)$ and $\zeta_t(v,h)$ as in Definition 4.8, the run from x is $X_0 = x$ and, for $t \geq 1$,

$$X_t(v,h) = \begin{cases} \zeta_t(v,h) & \text{if } \eta_t(v,h) = 1, \\ f_M\big((X_{t-1})_{(v,h)}\big) & \text{if } \eta_t(v,h) = 0. \end{cases}$$

Running all starts with the same variables is the common-noise coupling; X^x is the run from x.

Given X_t, the values $X_{t+1}(v,h)$ are independent with laws $p_M(\cdot \mid (X_t)_{(v,h)})$. So G_M is a positive-rate probabilistic cellular automaton whose neighbourhood consists of the 63 points $(d, c_i, 1)$, $d \in \{-1,0,1\}^2$, $0 \leq i \leq 6$, with the table p_M; every transition probability is at least $\varepsilon/|\mathcal{A}|$. For rational ε the map $M \mapsto G_M$ is total computable, with rational tables. Lemmas 4.10, 4.12 and 4.13 and their proofs hold verbatim for G_M, since $X_1|_O$ depends on X_0 only on a finite set and on independent noise. We write $\delta_O^{G_M}(t)$ as in Definition 4.11.

Lemma 10.16 (Characteristics)**.** *Fix a realisation of the noise and a start $x \in \mathcal{A}^{\mathbb{Z}^9}$. For $j \in \mathbb{Z}$, $s \geq 0$ and $v \in \mathbb{Z}^8$ put $Z^j_s(v) = X^x_s(v, j-s)$. Then Z^j is the run of K_M from the slice $x(\cdot, j)$, driven by*

the variables $\eta_s^j(v) = \eta_s(v, j-s)$ and $\zeta_s^j(v) = \zeta_s(v, j-s)$, $s \geq 1$. For distinct j these are disjoint subfamilies of the noise, and each has the law of the noise of K_M. In particular:

(a) $X_t^x(v,h)$ is the value at (v,t) of the run of K_M from $x(\cdot, h+t)$ driven by the noise of index $h+t$;

(b) if the start is random, independent of the noise, with independent slices $X_0(\cdot, j)$, $j \in \mathbb{Z}$, then the processes Z^j, $j \in \mathbb{Z}$, are independent, and Z^j is a run of K_M from a start with the law of $X_0(\cdot, j)$.

Proof. Let $s \geq 1$. If $\eta_s(v, j-s) = 1$, then $Z_s^j(v) = \zeta_s^j(v)$. Otherwise $Z_s^j(v) = f_M\big(w \mapsto X_{s-1}^x(v+w, j-s+1)\big) = f_M\big((Z_{s-1}^j)_v\big)$, in the notation of Section 10.2. This is the recursion of Definition 4.8 for K_M, and $Z_0^j = x(\cdot, j)$. The point (v,h,s) carries the noise of index $j = h+s$ only, so the subfamilies are disjoint, and each consists of independent variables with the laws of Definition 4.8. Part (a) is the case $j = h+t$. For (b), Z^j is a function of $X_0(\cdot, j)$ and of the noise of index j, and these pairs are independent over j. □

Corollary 10.17 (One phase or a continuum)**.** *Let $\varepsilon \in (0, \varepsilon_0]$ and let M be a machine.*

(i) *Suppose M halts, and let $\tau = \tau(R_M)$ be its refutation time for three tokens. For every site $w \in \mathbb{Z}^9$ and every $t \geq 0$,*

$$\mathbb{P}\big(x \mapsto X_t^x(w) \text{ is not constant}\big) \leq 4\varepsilon^{-\tau/99} 2^{-t}.$$

G_M has exactly one invariant probability measure π_M^G, and

$$d_{\mathrm{TV}}\big(\mu\Phi^t|_O, \pi_M^G|_O\big) \leq 8|O|\,\varepsilon^{-\tau/99} 2^{-t}$$

for every probability measure μ on $\mathcal{A}^{\mathbb{Z}^9}$, every window O and every $t \geq 0$, where Φ is the transition operator of G_M. Moreover $\delta_{\{(0,0)\}}^{G_M}(t) = \delta_{\{0\}}^{K_M}(t)$ for every t; so, for rational ε, no computable function of M bounds the mixing time of G_M, defined as in Corollary 10.7, over the halting machines.

(ii) *Suppose M does not halt. Let ν^+ and ν^- be the measures π^+ and π^- of the proof of Corollary 10.8(ii), for K_M: invariant and translation invariant, with $\nu^+(\omega(0) \text{ is live}) \geq 1 - 1.01\varepsilon$ and $\nu^-(\omega(0) \text{ is live}) \leq 1.01\varepsilon$. For $p \in [0,1]$ let $\rho_p = (1-p)\nu^- + p\,\nu^+$, and let μ_p be the probability measure on $\mathcal{A}^{\mathbb{Z}^9}$ under which the slices $\omega(\cdot, h)$, $h \in \mathbb{Z}$, are independent, each with law ρ_p. Then:*

(a) each μ_p is invariant for G_M and invariant under all translations of $\mathbb{Z}^9$;

(b) $p \mapsto \mu_p$ is a homeomorphism of $[0,1]$ onto its image, in the weak topology;

(c) each μ_p is an extreme point of the convex set of all invariant probability measures of G_M, translation invariant or not; in particular no μ_p is a proper convex combination of two others;

(d) the process with initial law μ_p is mixing under the time shift: if F and H are bounded functions of finitely many coordinates of $(X_0, \ldots, X_T)$, then for all large s

$$\mathbb{E}\big[F(X_0,\ldots,X_T)\,H(X_s,\ldots,X_{s+T})\big] = \mathbb{E}\big[F(X_0,\ldots,X_T)\big]\,\mathbb{E}\big[H(X_0,\ldots,X_T)\big];$$

(e) for $0 < p < 1$, μ_p is not ergodic under the translations of $\mathbb{Z}^8$ that act inside the slices.

(iii) *Let Ψ_M^G be the space-time interaction of G_M on $\mathbb{Z}^9 \times \mathbb{Z} = \mathbb{Z}^{10}$, defined from its table as Ψ_M is defined from p_M. If M halts, Ψ_M^G has exactly one Gibbs measure. If M does not halt, the*

two-sided stationary trajectory laws $\bar{\mu}_p$ of the measures μ_p are translation-invariant Gibbs measures, $p \mapsto \bar{\mu}_p$ is a homeomorphism of $[0,1]$ onto its image, each $\bar{\mu}_p$ is an extreme point of the convex set of translation-invariant Gibbs measures, and for $0 < p < 1$ no $\bar{\mu}_p$ is an extreme point of the convex set of all Gibbs measures.

(iv) *For rational ε, no algorithm decides, given M, whether G_M has exactly one invariant measure or a continuum of extreme invariant measures as in (ii); the same holds for the translation-invariant Gibbs measures of Ψ^G_M. With the machine M_R of Corollary 10.5 in place of M, the same alternative holds for every finite Wang set R, with one invariant measure exactly when R does not tile the plane.*

Proof. If $\mathcal{A} = \{\perp\}$, then M halts, as in the proof of Corollary 10.8, and every claim about the halting case is trivial on the one-point space. So let $\mathcal{A}$ have at least three symbols.

(i). Let $w = (v,h)$. By Lemma 10.16(a), $X^x_t(w)$ is the value at (v,t) of the run of K_M from $x(\cdot, h+t)$ driven by the noise of index $h+t$, which has the law of the noise of K_M. As x ranges over $\mathcal{A}^{\mathbb{Z}^9}$, the slice $x(\cdot, h+t)$ ranges over $\mathcal{A}^{\mathbb{Z}^8}$. So the event that $x \mapsto X^x_t(w)$ is not constant is the event of Corollary 5.12 for K_M at the site v. By Theorem 11.1, R_M is 3-token refutable with time τ, and with the constants of Corollary 11.21, $\beta \le \frac{1}{2}$. Corollary 5.12 gives the first bound. For a start x and a window O it gives $d_{\mathrm{TV}}(\mathrm{Law}_x(X_t|_O), \mathrm{Law}_\perp(X_t|_O)) \le \mathbb{P}(X^x_t|_O \neq X^\perp_t|_O) \le 4|O|\varepsilon^{-\tau/99}2^{-t}$, so G_M forgets, and Lemma 4.12 gives the invariant measure and the bound with the factor 8. By Lemma 10.16(a) with $w = (0,0)$, the law of $X^x_t(0,0)$ under G_M is the law of $X_t(0)$ under K_M from the start $x(\cdot, t)$; for the all-bottom start both are the law from the all-bottom start. Taking the supremum over x gives $\delta^{G_M}_{\{(0,0)\}}(t) = \delta^{K_M}_{\{0\}}(t)$. So the mixing times of G_M and K_M coincide, and Corollary 10.7(iii), which applies with $k = 6$ by Remark 11.22, gives the last claim.

(ii)(a). Let X_0 have law μ_p, independent of the noise. By Lemma 10.16, the slice $X_1(\cdot, h) = Z_1^{h+1}$ is one step of K_M from $X_0(\cdot, h+1)$, and these slices are independent over h. Each has law $\rho_p \Phi_{K_M} = \rho_p$, since ν^+ and ν^- are invariant and one step is linear in the initial law; here Φ_{K_M} is the transition operator of K_M. So X_1 has law μ_p. The measure ρ_p is invariant under the translations of $\mathbb{Z}^8$, so μ_p is invariant under them, and μ_p is invariant under the shift of the slice index because its factors are equal.

(ii)(b). An event that depends on a finite window $O \subset \mathbb{Z}^9$ is a finite union of points of $\mathcal{A}^O$, and the probability of each point under μ_p is a product, over the slices that meet O, of ρ_p-probabilities, each affine in p. So the probability of every such event is a polynomial in p, and $p \mapsto \mu_p$ is continuous, since weak convergence is convergence on these events (Section 4). The probability that the site $(0,0)$ is live is $(1-p)a_- + p\,a_+$, with $a_- = \nu^-(\omega(0) \text{ is live}) \le 1.01\varepsilon$ and $a_+ = \nu^+(\omega(0) \text{ is live}) \ge 1 - 1.01\varepsilon$; it is strictly increasing in p, so the map is injective. A continuous injective map from the compact space $[0,1]$ to the Hausdorff space of probability measures is a homeomorphism onto its image.

(ii)(c). Let g depend on finitely many coordinates, lying in the slices of a finite set $J \subset \mathbb{Z}$, and let $t \ge 0$. By Lemma 10.16(a), $x \mapsto \Phi^t g(x) = \mathbb{E}_x\, g(X_t)$ depends only on the slices $J + t$ of x. Let η' be an invariant probability measure and $0 < \lambda \le 1$ with $\lambda\eta' \le \mu_p$. We show $\eta' = \mu_p$; this gives (c). The measure η' has a density φ with respect to μ_p, and $0 \le \varphi \le 1/\lambda$. Let $\delta > 0$. Functions that depend on finitely many coordinates are dense in $L^1(\mu_p)$, so there is such a function f with $\int |\varphi - f|\, d\mu_p < \delta$; let I be the finite set of slices that it reads. For t so large that I and $J + t$ are disjoint, the slices being independent under μ_p,

$$\eta'(g) = \eta'(\Phi^t g) = \int \varphi\, \Phi^t g\, d\mu_p, \qquad \left| \int \varphi\, \Phi^t g\, d\mu_p - \mu_p(f)\, \mu_p(\Phi^t g) \right| \le \delta\, \|g\|_\infty,$$

and $\mu_p(\Phi^t g) = \mu_p(g)$ by (a), while $|\mu_p(f) - 1| = |\mu_p(f) - \mu_p(\varphi)| \le \delta$. Hence $|\eta'(g) - \mu_p(g)| \le 2\delta\|g\|_\infty$ for every $\delta > 0$, so $\eta'(g) = \mu_p(g)$ for every such g, and $\eta' = \mu_p$. If $\mu_p = \lambda\eta_1 + (1-\lambda)\eta_2$ with invariant η_1, η_2 and $0 < \lambda < 1$, both are equal to μ_p, so μ_p is extreme. By (b), distinct parameters give distinct measures.

(ii)(d). By Lemma 10.16(b), under the process with initial law μ_p the processes Z^j, $j \in \mathbb{Z}$, are independent. Every coordinate $X_t(v,h)$ is the coordinate $Z_t^{h+t}(v)$, so F is a function of the processes Z^j with j in a finite set J_F, and $H(X_s, \dots, X_{s+T})$ is a function of those with j in $J_H + s$, for a finite set J_H that depends only on H. For s so large that J_F and $J_H + s$ are disjoint the two functions are independent, and stationarity gives the claim.

(ii)(e). The law of the slice $\omega(\cdot, 0)$ under μ_p is ρ_p, a proper convex combination of the measures ν^- and ν^+, which are invariant under the translations of $\mathbb{Z}^8$ and distinct. A measure that is ergodic under a group action is an extreme point of the invariant measures, so ρ_p is not ergodic under $\mathbb{Z}^8$: some event $\mathcal{E}$ of the slice, invariant under $\mathbb{Z}^8$, has $0 < \rho_p(\mathcal{E}) < 1$. The event that $\omega(\cdot, 0) \in \mathcal{E}$ is invariant under the translations of $\mathbb{Z}^8$ inside the slices and has μ_p-probability strictly between 0 and 1.

(iii). The neighbourhood of G_M does not contain the origin, and [9] asks for one that does. Add the origin as a parent that the table ignores: the table $p'(s \mid y) = p_M\big(s \mid y \text{ restricted to } \mathcal{N} \times \{1\}\big)$ on the neighbourhood $\{0\} \cup (\mathcal{N} \times \{1\})$. This changes neither the process nor the specification, since in [9, eq. (6)] a factor that meets Λ only through the added parent does not depend on the configuration in Λ and cancels with the normalisation. So the Gibbs measures of Ψ_M^G are the DLR states of [9] for p'.

Suppose M halts. Uniform replacement is a local-map representation of p', as in the proof of Corollary 10.8(i), and its common-map coefficient is the probability in (i) at the site $(0,0)$, at most $4\varepsilon^{-\tau/99}2^{-t}$. [9, Corollary 13.7], in spatial dimension 9, gives exactly one Gibbs measure.

Suppose M does not halt. By [9, Lemma 2.4], $\bar\mu_p$ is a Gibbs measure, and it is translation invariant, because μ_p is invariant and translation invariant and G_M commutes with translations. By [9, Corollary 10.4], the translation-invariant Gibbs measures are exactly the trajectory laws $\bar\alpha$ of the invariant, translation-invariant measures α of G_M; the map $\alpha \mapsto \bar\alpha$ is affine, and the marginal at time 0 inverts it. If $\bar\mu_p = \lambda\mathsf{A} + (1-\lambda)\mathsf{B}$ with translation-invariant Gibbs measures $\mathsf{A} = \bar\alpha$, $\mathsf{B} = \bar\beta$ and $0 < \lambda < 1$, then the marginals at time 0 give $\mu_p = \lambda\alpha + (1-\lambda)\beta$, so $\alpha = \beta = \mu_p$ by (ii)(c), and $\mathsf{A} = \mathsf{B} = \bar\mu_p$. The probability of an event on finitely many space-time points under $\bar\mu_p$ is the μ_p-integral of a function of finitely many coordinates of the start, hence a polynomial in p; so $p \mapsto \bar\mu_p$ is continuous, injective because $p \mapsto \mu_p$ is, and a homeomorphism onto its image.

It remains to show that $\bar\mu_p$ is not extreme among all Gibbs measures when $0 < p < 1$. Every factor of Ψ_M^G involves a point (v,h,t) and its parents $(v+w, h+1, t-1)$, so it lies in one hyperplane $H_j = \{(v,h,t) : h + t = j\}$. The map $\iota_j(v,t) = (v, j-t, t)$ is a bijection from $\mathbb{Z}^8 \times \mathbb{Z}$ onto H_j that carries the parents of (v,t) for Ψ_M, the space-time interaction of K_M, to the parents of $\iota_j(v,t)$ for Ψ_M^G, with the same table. Hence, for a finite set Λ, the specification kernel of Ψ_M^G on Λ is the product over j of the kernels of Ψ_M on $\iota_j^{-1}(\Lambda \cap H_j)$, each with the exterior configuration read on H_j: the product of factors in [9, eq. (6)] splits over the hyperplanes, and so does its normalisation. Consequently, if ν_j, $j \in \mathbb{Z}$, are Gibbs measures of Ψ_M, the measure under which the restrictions to the hyperplanes H_j are independent, with the images of the measures ν_j under the maps ι_j, is a Gibbs measure of Ψ_M^G: given the configuration outside Λ, independence makes the conditional law on Λ the product over j of the conditional laws on $\Lambda \cap H_j$ given the configuration on $H_j \setminus \Lambda$, which are the kernels of Ψ_M.

Now take independent two-sided stationary runs Z^j, $j \in \mathbb{Z}$, of K_M with marginal ρ_p, and put $X_t(v,h) = Z_t^{h+t}(v)$. For each t, the variables X_s, $s \le t$, determine and are determined by

the variables Z_s^j, $s \le t$, $j \in \mathbb{Z}$. Given them, $X_{t+1}(\cdot, h) = Z_{t+1}^{h+t+1}$ has the conditional law of one step of K_M from $Z_t^{h+t+1} = X_t(\cdot, h+1)$, independently over h and over sites. So X is a stationary run of G_M, its marginal is μ_p, and its law is $\bar{\mu}_p$. In other words, $\bar{\mu}_p$ is the measure of the previous paragraph with every ν_j equal to the trajectory law $\bar{\rho}_p$ of ρ_p under K_M. The trajectory law is affine in the invariant initial law, so $\bar{\rho}_p = (1-p)\bar{\nu}^- + p\,\bar{\nu}^+$, and

$$\bar{\mu}_p = (1-p)\,\mathsf{A} + p\,\mathsf{B},$$

where A and B are the measures of the previous paragraph with $\nu_0 = \bar{\nu}^-$, respectively $\nu_0 = \bar{\nu}^+$, and $\nu_j = \bar{\rho}_p$ for $j \neq 0$. By [9, Lemma 2.4], $\bar{\nu}^\pm$ and $\bar{\rho}_p$ are Gibbs measures of Ψ_M, so A and B are Gibbs measures of Ψ_M^G. They differ: the origin of $\mathbb{Z}^{10}$ is $\iota_0(0,0)$, and it is live with probability $a_- \le 1.01\varepsilon$ under A and $a_+ \ge 1 - 1.01\varepsilon$ under B. For $0 < p < 1$, $\bar{\mu}_p$ is therefore a proper convex combination of two distinct Gibbs measures.

(iv). By (i)–(iii) the alternatives are exclusive, and the first holds exactly when M halts. A decision procedure would decide halting. For a Wang set R, the machine M_R halts exactly when R does not tile (Corollary 10.5). □

So a nonhalting member, stacked, has an arc of invariant measures, each extreme among all invariant measures and mixing in time, while a halting member, stacked, has one invariant measure, reached exponentially fast. By (ii)(e), the mixing in (d) does not extend to the translations inside the slices. The construction is a standard device: along each set $h + t = j$ the stack runs an independent copy of K_M, so mixing in time comes from independence, and products of a two-point mixture give the arc. The Gibbs measures $\bar{\mu}_p$, $0 < p < 1$, are extreme among the translation-invariant Gibbs measures but not among all Gibbs measures, because the interaction is a sum of decoupled layers. That an equilibrium state can be ergodic, and even mixing, without being extreme among all Gibbs measures is classical; Slawny [44] constructs Gibbs states that are mixing but not extreme (we have read only the abstract of [44]).

11 Three tokens and dimension eight

This section proves that three tokens refute every halting member of the family of Definition 8.17. The machine-family theorem then holds on $\mathbb{Z}^8$, with an explicit rate (Corollary 11.21).

Theorem 11.1 (Three tokens)**.** *Let $M \mapsto R_M$ be the family of Definition 8.17. If the machine M halts, then R_M is 3-token refutable.*

The family is the one of Definition 8.17, without change. With Proposition 8.18 and Theorem 4.25(iv), Theorem 11.1 gives Theorem C with $j = 3$: R_M tiles if and only if M does not halt, and every halting member is 3-token refutable. By Corollary 4.27, the integer 3 cannot be replaced by 1. Whether it can be replaced by 2 is open.

The proof keeps the plan of Section 9 and changes how the scaffold is read. Two tokens stay fixed as *pins*, and the third token walks. The letters of $\mathbb{A}$ that the walker can carry form arc-consistent lists (Section 11.1). Once a corner letter with ruler bits 0 fixes the phase, arc consistency alone coarsens such lists by one level of the substitution (Section 11.2). Two corner letters with opposite bits, at a displacement divisible by 4^n, then force a whole word $\sigma^n(\mathfrak{a})$. Such pairs exist at every scale, by repeated coarsening and one finite lemma, Lemma 11.12, whose proof is a computation (Section 11.3). The roles, the projected meet and the compression then run as in Section 9 (Section 11.4).

From Sections 7 and 8 the proof uses the tables of Section A, Proposition 7.11, and the items (Q1)–(Q5) of Section 9.1, which come from Lemma 8.6, Propositions 8.11, 8.15 and 8.16, and Definition 8.17. The Horn layer enters only through Proposition 8.16 and its proof in Section 8. In particular, the halting time T below is the halting time $T(M)$ of the one-sided machine M° built from M in Section 8.6, the machine that the Horn clauses simulate.

Fix a machine M. Until the proof of Theorem 11.1, $\mathcal{W}$ is a nonempty plane-safe set of 3-token states for R_M (Definition 4.17). By Theorem 4.25(i), such a set exists exactly when R_M is not 3-token refutable. Live symbols, compatibility and the centre ctr are those of Definition 4.3, for the Wang set R_M.

11.1 Lists read by one moving token

Lemma 11.2 (Moves). *Let s, s' be live symbols of R_M, let $d \in D$, and let s' be compatible with s at displacement d. Let λ be the block of s. Then* $\mathrm{ctr}(s') = \lambda(d)$. *So* $\mathrm{ctr}(s)$ *and* $\mathrm{ctr}(s')$ *are the values at* 0 *and at d of the legal pattern λ, and* $\mathrm{quad}(\mathrm{ctr}(s')) \equiv \mathrm{quad}(\mathrm{ctr}(s)) + d \pmod 2$.

Proof. Let λ' be the block of s'. Compatibility gives $\lambda(w) = \lambda'(w - d)$ for every $w \in \Lambda \cap (d + \Lambda)$. The cell $w = d$ lies in this set, so $\lambda(d) = \lambda'(0,0) = \mathrm{ctr}(s')$. The cells 0 and d of Λ are joined inside Λ by a path of at most two steps in $\{\pm\mathbf{e}_1, \pm\mathbf{e}_2\}$. The pattern λ is legal, so (Q1) of Section 9.1 adds each step to the quadrant modulo 2. □

Lemma 11.3 (One quarter phase). *For every state $(\mathbf{u}, \mathbf{s}) \in \mathcal{W}$ and all indices i, i',*

$$\mathrm{quad}(\mathrm{ctr}(s_i)) - u_i \equiv \mathrm{quad}(\mathrm{ctr}(s_{i'})) - u_{i'} \pmod 2.$$

For $\gamma \in (\mathbb{Z}/2\mathbb{Z})^2$, let $\mathcal{W}_\gamma$ be the set of states of $\mathcal{W}$ in which this common class is γ. Each $\mathcal{W}_\gamma$ is plane-safe, and some $\mathcal{W}_\gamma$ is nonempty.

Proof. Let $i \neq i'$. Choose Moore steps leading from u_i to $u_{i'}$, and apply (W3) to token i once for each step. This gives states of $\mathcal{W}$ in which only token i has moved. By Lemma 11.2, the class of $\mathrm{quad}(\mathrm{ctr}(s_i)) - u_i$ does not change along the way. At the end tokens i and i' share a site, so by (W1) they carry one symbol, and the claim follows. A clone copies the position and the symbol of a token, and a move keeps the class of the moved token, by Lemma 11.2. So the clones and moves that (W2) and (W3) supply in $\mathcal{W}$ for a state of $\mathcal{W}_\gamma$ lie in $\mathcal{W}_\gamma$, and (W1) is inherited. Every state of $\mathcal{W}$ lies in one set $\mathcal{W}_\gamma$. □

We fix γ with $\mathcal{W}_\gamma \neq \emptyset$ and replace $\mathcal{W}$ by $\mathcal{W}_\gamma$. Let $\bar\gamma \in \{0,1\}^2$ represent γ, and put $\iota(z) = \bar\gamma + 2z$ for $z \in \mathbb{Z}^2$, as in Section 9.3. In every state of $\mathcal{W}$, a token at a site x carries a symbol s with $\mathrm{quad}(\mathrm{ctr}(s)) \equiv x - \bar\gamma \pmod 2$. The *cell* of z is $\iota(z) + \{0,1\}^2$; the quadrant at $\iota(z)$ is $(0,0)$.

Definition 11.4 (Pins and walker). A *pin pair* is a tuple $\mathbf{p} = (p_1, s_1; p_2, s_2)$ of sites $p_1, p_2 \in \mathbb{Z}^2$ and live symbols s_1, s_2 such that $\mathcal{W}$ contains a state $((p_1, p_2, x), (s_1, s_2, s))$ for some x and s. Its *walker set* is

$$\mathcal{X}_\mathbf{p} = \{(x, s) \in \mathbb{Z}^2 \times L_{R_M} : ((p_1, p_2, x), (s_1, s_2, s)) \in \mathcal{W}\},$$

and its *index lists* are $\mathcal{L}_\mathbf{p}(z) = \{\mathrm{par}(\mathrm{ctr}(s)) : (\iota(z), s) \in \mathcal{X}_\mathbf{p}\}$ for $z \in \mathbb{Z}^2$.

Tokens 1 and 2 hold the pins, and token 3, the *walker*, moves. A pin pair may have $p_1 = p_2$ and $s_1 = s_2$; it is then a single pin held by two tokens.

Definition 11.5 (List families)**.** Letters $\mathfrak{a}, \mathfrak{b} \in \mathbb{A}$ *fit in direction* $e \in \{\pm\mathbf{e}_1, \pm\mathbf{e}_2\}$ if the pattern with $\mathfrak{a}$ at 0 and $\mathfrak{b}$ at e is legal. A *list family* $\mathcal{L}$ assigns to every $z \in \mathbb{Z}^2$ a nonempty set $\mathcal{L}(z) \subseteq \mathbb{A}$. It is *arc consistent* if, for every $z \in \mathbb{Z}^2$, every $\mathfrak{a} \in \mathcal{L}(z)$ and every $e \in \{\pm\mathbf{e}_1, \pm\mathbf{e}_2\}$, some $\mathfrak{b} \in \mathcal{L}(z+e)$ is such that $\mathfrak{a}, \mathfrak{b}$ fit in direction e.

So $\mathfrak{a}, \mathfrak{b}$ fit in direction $\mathbf{e}_1$ exactly when they match in direction $\mathbf{e}_1$, and in direction $-\mathbf{e}_1$ exactly when $\mathfrak{b}, \mathfrak{a}$ match in direction $\mathbf{e}_1$; the same holds for $\mathbf{e}_2$. This is arc consistency as in the paragraph after Proposition 4.26, for the grid $\mathbb{Z}^2$ with its four cardinal relations and the target $\mathbb{A}$.

Lemma 11.6 (Walker lists)**.** *Let* $\mathbf{p} = (p_1, s_1; p_2, s_2)$ *be a pin pair.*

(a) For every $x \in \mathbb{Z}^2$ *some* s *has* $(x, s) \in \mathcal{X}_\mathbf{p}$*. If* $(x, s) \in \mathcal{X}_\mathbf{p}$ *and* $d \in D$*, then* $(x + d, s') \in \mathcal{X}_\mathbf{p}$ *for some* s' *compatible with* s *at displacement* d*.*

(b) If $i \in \{1, 2\}$ *and* $(p_i, s) \in \mathcal{X}_\mathbf{p}$*, then* $s = s_i$*.*

(c) $\mathcal{L}_\mathbf{p}$ *is an arc-consistent list family.*

(d) If $i \in \{1, 2\}$ *and* $p_i = \iota(y_i)$ *with* $y_i \in \mathbb{Z}^2$*, then* $\mathcal{L}_\mathbf{p}(y_i) = \{\mathrm{par}(\mathrm{ctr}(s_i))\}$*.*

Proof. (a) The second claim is (W3) for token 3; the move leaves tokens 1 and 2 in place. Start from a state of Definition 11.4. Finitely many such moves bring token 3 to any site x.

(b) In the state $((p_1, p_2, p_i), (s_1, s_2, s))$, tokens i and 3 share a site, so $s = s_i$ by (W1).

(c) The lists are nonempty by (a). Let $\mathfrak{a} \in \mathcal{L}_\mathbf{p}(z)$, given by $(\iota(z), s) \in \mathcal{X}_\mathbf{p}$. Apply (a) twice with $d = \mathbf{e}_1$. This gives $(\iota(z) + \mathbf{e}_1, s')$ and $(\iota(z) + 2\mathbf{e}_1, s'')$ in $\mathcal{X}_\mathbf{p}$, each symbol compatible with the previous one at displacement $\mathbf{e}_1$. By Lemma 11.2, $\mathrm{ctr}(s)$ and $\mathrm{ctr}(s')$ match horizontally, and so do $\mathrm{ctr}(s')$ and $\mathrm{ctr}(s'')$. The quadrants of $\mathrm{ctr}(s)$ and $\mathrm{ctr}(s')$ are $(0, 0)$ and $(1, 0)$. By (Q1), $\mathrm{par}(\mathrm{ctr}(s')) = \mathfrak{a}$, and $\mathrm{par}(\mathrm{ctr}(s'))$ and $\mathrm{par}(\mathrm{ctr}(s''))$ match horizontally in $\mathbb{A}$. Since $\iota(z) + 2\mathbf{e}_1 = \iota(z + \mathbf{e}_1)$, the letter $\mathrm{par}(\mathrm{ctr}(s''))$ lies in $\mathcal{L}_\mathbf{p}(z + \mathbf{e}_1)$, and $\mathfrak{a}$ and this letter fit in direction $\mathbf{e}_1$. For $-\mathbf{e}_1$, apply (a) twice with $d = -\mathbf{e}_1$. Now $\mathrm{ctr}(s')$ lies west of $\mathrm{ctr}(s)$ and has quadrant $(1, 0)$, so by (Q1) the letters $\mathrm{par}(\mathrm{ctr}(s'))$ and $\mathfrak{a}$ match horizontally. Next, $\mathrm{ctr}(s'')$ lies west of $\mathrm{ctr}(s')$ and has quadrant $(0, 0)$, so $\mathrm{par}(\mathrm{ctr}(s'')) = \mathrm{par}(\mathrm{ctr}(s'))$. Hence $\mathfrak{a}$ and $\mathrm{par}(\mathrm{ctr}(s'')) \in \mathcal{L}_\mathbf{p}(z - \mathbf{e}_1)$ fit in direction $-\mathbf{e}_1$. The directions $\pm\mathbf{e}_2$ are the same, with vertical matching and the second coordinate of the quadrant.

(d) By (a) some s has $(p_i, s) \in \mathcal{X}_\mathbf{p}$, and by (b) every such s equals s_i. □

11.2 Arc consistency on the scaffold

This subsection is about list families on $\mathbb{A}$ alone. We write a letter as $\mathfrak{a} = (\theta, \kappa, \omega, b_1, b_2)$, as in Definition 7.7, and use the tables of Section A. For a thin type θ let $\theta^{\leftrightarrow}$ and $\theta^{\updownarrow}$ be given by

$$\mathsf{a}^{\leftrightarrow} = \mathsf{c},\ \mathsf{c}^{\leftrightarrow} = \mathsf{a},\ \mathsf{b}^{\leftrightarrow} = \mathsf{d},\ \mathsf{d}^{\leftrightarrow} = \mathsf{b}, \qquad \mathsf{a}^{\updownarrow} = \mathsf{d},\ \mathsf{d}^{\updownarrow} = \mathsf{a},\ \mathsf{b}^{\updownarrow} = \mathsf{c},\ \mathsf{c}^{\updownarrow} = \mathsf{b}.$$

Both maps are involutions, and they commute. By Section A, thin types θ, θ' match horizontally, θ to the west, exactly when $\theta' = \theta^{\leftrightarrow}$, and vertically, θ to the south, exactly when $\theta' = \theta^{\updownarrow}$. The corner $F(\theta)$ has the four endpoints of θ, so $\partial_\mathrm{E} F(\theta) = \partial_\mathrm{W} F(\theta')$ exactly when $\theta' = \theta^{\leftrightarrow}$, and $\partial_\mathrm{N} F(\theta) = \partial_\mathrm{S} F(\theta')$ exactly when $\theta' = \theta^{\updownarrow}$. For every θ, the four types θ, $\theta^{\leftrightarrow}$, $\theta^{\updownarrow}$ and $(\theta^{\leftrightarrow})^{\updownarrow}$ are distinct.

A *corner letter* is a letter whose thin type is a and whose thick type is a corner. By Table 5 its black type is A, and by (R3) its two bits are equal; so it is $(\mathsf{a}, \kappa, \mathsf{A}, b, b)$, and we call b its *bit*. Two corner letters are *opposite* if their bits differ. A letter of thin type b is $(\mathsf{b}, F(\theta), \mathsf{A}, 0, 0)$ for exactly one thin type θ (Tables 5 and 8). We use five facts, read off Tables 4 and 5 and Definition 7.7.

(T1) Every corner has exactly one red endpoint among its north and south endpoints, and exactly one among its east and west endpoints.

(T2) The endpoint $\partial_{\mathrm{W}}\theta$ is green exactly for $\theta \in \{\mathsf{a}, \mathsf{d}\}$, and $\partial_{\mathrm{S}}\theta$ is green exactly for $\theta \in \{\mathsf{a}, \mathsf{c}\}$.

(T3) A letter of thin type c is a line with black B or C and column bit 0. Black B excludes the vertical atom r_0, black C excludes r_2, and B and C have equal east and west bits, 0 and 1 respectively. A letter of thin type d is a line with black D or E and row bit 0. Black D excludes the horizontal atom r_3, black E excludes r_1, and D and E have equal north and south bits, 1 and 0 respectively.

(T4) By (R4), a line with equal bits is not a red crossing: if one of its atoms is red, the other is green.

(T5) The only black type with north and east bits 0 is A. The black types with north bit 1 and south bit 0 are B and C. The black types with east bit 1 and west bit 0 are D and E. A letter of thin type a with black A is a corner; with any other black type it is a line.

A list family $\mathcal{L}$ has *thin phase* $c \in \mathbb{Z}^2$ if, for every $z \in \mathbb{Z}^2$, every letter of $\mathcal{L}(z)$ has thin type a, c, d or b according as $z - c$ is congruent to $(0,0)$, $(1,0)$, $(0,1)$ or $(1,1)$ modulo 2. Only the class of c modulo 2 matters. This rule is the pattern of Lemma 7.2: the type it assigns to $z + \mathbf{e}_1$ is the image under $\leftrightarrow$ of the type at z, and the type at $z + \mathbf{e}_2$ is the image under $\updownarrow$. We call it the *thin type of $\mathcal{L}$ at z*.

Lemma 11.7 (Thin phase)**.** *Let $\mathcal{L}$ be arc consistent, and suppose that for some z_0 all letters of $\mathcal{L}(z_0)$ have one thin type. Then $\mathcal{L}$ has a thin phase.*

Proof. Let θ_0 be the thin type at z_0, and choose c so that the rule assigns θ_0 to z_0. Let $\mathfrak{a} \in \mathcal{L}(z)$. Choose a path $z = x_0, x_1, \ldots, x_r = z_0$ with steps in $\{\pm\mathbf{e}_1, \pm\mathbf{e}_2\}$. By arc consistency there are letters $\mathfrak{a}_0 = \mathfrak{a}$ and $\mathfrak{a}_l \in \mathcal{L}(x_l)$ such that $\mathfrak{a}_{l-1}, \mathfrak{a}_l$ fit in direction $x_l - x_{l-1}$, for $1 \le l \le r$. Along a horizontal step the thin type changes by $\leftrightarrow$, and along a vertical step by $\updownarrow$. The thin type of $\mathfrak{a}_r$ is θ_0. The two maps are commuting involutions, so the thin type of $\mathfrak{a}$ is the image of θ_0 under $\leftrightarrow$ applied a number of times congruent to $(z - z_0)_1$ modulo 2, and $\updownarrow$ applied a number of times congruent to $(z - z_0)_2$. The rule assigns the same type to z. □

Let $\mathcal{L}$ have thin phase c. For $z \in \mathbb{Z}^2$ we call $c + 2z$, $c + 2z + \mathbf{e}_1$, $c + 2z + \mathbf{e}_2$ and $c + 2z + (1,1)$ the *south-west, south-east, north-west and north-east sites of z*; their thin types are a, c, d and b. Each letter at the north-east site of z is $(\mathsf{b}, F(\theta), \mathsf{A}, 0, 0)$ for one thin type θ. Let $\mathcal{L}^{\angle}(z)$ be the set of these θ; it is nonempty.

Lemma 11.8 (The corner grid)**.** *Let $\mathcal{L}$ be arc consistent with thin phase c. For every $z \in \mathbb{Z}^2$,*

$$\mathcal{L}^{\angle}(z + \mathbf{e}_1) = \{\theta^{\leftrightarrow} : \theta \in \mathcal{L}^{\angle}(z)\}, \qquad \mathcal{L}^{\angle}(z + \mathbf{e}_2) = \{\theta^{\updownarrow} : \theta \in \mathcal{L}^{\angle}(z)\}.$$

If $\mathcal{L}^{\angle}(z_0)$ is a singleton for some z_0, then every $\mathcal{L}^{\angle}(z)$ is a singleton $\{\theta_z\}$, and $\theta_{z+\mathbf{e}_1} = \theta_z^{\leftrightarrow}$ and $\theta_{z+\mathbf{e}_2} = \theta_z^{\updownarrow}$ for every z.

Proof. Let $\theta \in \mathcal{L}^{\angle}(z)$, given by a letter $\mathfrak{n}$ at the north-east site x of z. By arc consistency there are $\mathfrak{l} \in \mathcal{L}(x + \mathbf{e}_1)$ and $\mathfrak{n}' \in \mathcal{L}(x + 2\mathbf{e}_1)$ such that $\mathfrak{n}, \mathfrak{l}$ and $\mathfrak{l}, \mathfrak{n}'$ fit in direction $\mathbf{e}_1$. The site $x + \mathbf{e}_1$ is the north-west site of $z + \mathbf{e}_1$, so $\mathfrak{l}$ is a line (T3), and its east and west endpoints are both its horizontal atom. The site $x + 2\mathbf{e}_1$ is the north-east site of $z + \mathbf{e}_1$, so $\mathfrak{n}' = (\mathsf{b}, F(\theta'), \mathsf{A}, 0, 0)$ with $\theta' \in \mathcal{L}^{\angle}(z + \mathbf{e}_1)$. Matching gives $\partial_{\mathrm{E}}F(\theta) = \partial_{\mathrm{W}}F(\theta')$, so $\theta' = \theta^{\leftrightarrow}$. Hence the image of $\mathcal{L}^{\angle}(z)$ under $\leftrightarrow$ lies in $\mathcal{L}^{\angle}(z + \mathbf{e}_1)$. The same argument in direction $-\mathbf{e}_1$ shows that the image of

$\mathcal{L}^{\angle}(z+\mathbf{e}_1)$ under $\leftrightarrow$ lies in $\mathcal{L}^{\angle}(z)$; applying the involution gives equality. The vertical case is the same: the south-east site of $z+\mathbf{e}_2$ carries lines, whose north and south endpoints are both the vertical atom, and $\partial_{\mathrm{N}}F(\theta)=\partial_{\mathrm{S}}F(\theta')$ gives $\theta'=\theta^{\updownarrow}$. The last claim follows because both maps are bijections. □

Lemma 11.9 (A corner letter with bit 0 fixes the coarse phase)**.** *Let $\mathcal{L}$ be arc consistent with thin phase c, let $z_0\in\mathbb{Z}^2$, and let $\mathcal{L}(c+2z_0)=\{\mathfrak{p}\}$ for a corner letter $\mathfrak{p}$ with bit 0. Then every $\mathcal{L}^{\angle}(z)$ is a singleton $\{\theta_z\}$, and $\theta_z=\mathsf{a}$ exactly when $z\in z_0+2\mathbb{Z}^2$.*

Proof. Write $x=c+2z_0$ and $\mathfrak{p}=(\mathsf{a},\kappa,\mathsf{A},0,0)$. We show $\mathcal{L}^{\angle}(z_0)\subseteq\{\mathsf{a},\mathsf{d}\}$ and $\mathcal{L}^{\angle}(z_0)\subseteq\{\mathsf{a},\mathsf{c}\}$.

By (T1), $\partial_{\mathrm{N}}\kappa$ or $\partial_{\mathrm{S}}\kappa$ is red. Suppose that $\partial_{\mathrm{N}}\kappa$ is red, and let $\theta\in\mathcal{L}^{\angle}(z_0)$. By arc consistency, the letter $(\mathsf{b},F(\theta),\mathsf{A},0,0)\in\mathcal{L}(x+(1,1))$ and some $\mathfrak{l}\in\mathcal{L}(x+\mathbf{e}_2)$ fit in direction $-\mathbf{e}_1$, and $\mathfrak{l}$ and the only letter $\mathfrak{p}$ of $\mathcal{L}(x)$ fit in direction $-\mathbf{e}_2$. The site $x+\mathbf{e}_2$ has thin type d, so $\mathfrak{l}$ is a line with row bit 0 (T3). The first fit makes its horizontal atom $\partial_{\mathrm{W}}F(\theta)=\partial_{\mathrm{W}}\theta$. The second makes its vertical atom $\partial_{\mathrm{N}}\kappa$, which is red, and its column bit $\chi_1(\mathfrak{p})=0$. By (T4) the horizontal atom is green, and by (T2) $\theta\in\{\mathsf{a},\mathsf{d}\}$. Suppose instead that $\partial_{\mathrm{S}}\kappa$ is red, and let $\theta\in\mathcal{L}^{\angle}(z_0-\mathbf{e}_2)$. The same argument with the sites $x+(1,-1)$, $x-\mathbf{e}_2$ and x, where now $\mathfrak{p}$ lies above the line, gives $\theta\in\{\mathsf{a},\mathsf{d}\}$. By Lemma 11.8, $\mathcal{L}^{\angle}(z_0)$ is the image of $\mathcal{L}^{\angle}(z_0-\mathbf{e}_2)$ under $\updownarrow$, which maps $\{\mathsf{a},\mathsf{d}\}$ onto itself.

By (T1), $\partial_{\mathrm{E}}\kappa$ or $\partial_{\mathrm{W}}\kappa$ is red. Suppose that $\partial_{\mathrm{E}}\kappa$ is red, and let $\theta\in\mathcal{L}^{\angle}(z_0)$. The letter $(\mathsf{b},F(\theta),\mathsf{A},0,0)$ and some $\mathfrak{l}\in\mathcal{L}(x+\mathbf{e}_1)$ fit in direction $-\mathbf{e}_2$, and $\mathfrak{l}$ and $\mathfrak{p}$ fit in direction $-\mathbf{e}_1$. The site $x+\mathbf{e}_1$ has thin type c, so $\mathfrak{l}$ is a line with column bit 0. Its vertical atom is $\partial_{\mathrm{S}}F(\theta)=\partial_{\mathrm{S}}\theta$, its horizontal atom is $\partial_{\mathrm{E}}\kappa$, which is red, and its row bit is $\chi_2(\mathfrak{p})=0$. By (T4) and (T2), $\theta\in\{\mathsf{a},\mathsf{c}\}$. If instead $\partial_{\mathrm{W}}\kappa$ is red, the sites $x+(-1,1)$, $x-\mathbf{e}_1$ and x give $\mathcal{L}^{\angle}(z_0-\mathbf{e}_1)\subseteq\{\mathsf{a},\mathsf{c}\}$, and $\leftrightarrow$ maps $\{\mathsf{a},\mathsf{c}\}$ onto itself.

Hence $\mathcal{L}^{\angle}(z_0)\subseteq\{\mathsf{a}\}$, and this set is nonempty. Lemma 11.8 gives the singletons, and θ_{z_0+w} is a, c, d or b according as w is congruent to $(0,0)$, $(1,0)$, $(0,1)$ or $(1,1)$ modulo 2. □

Lemma 11.10 (Coarsening)**.** *Let $\mathcal{L}$ be arc consistent with thin phase c, and suppose that every $\mathcal{L}^{\angle}(z)$ is a singleton $\{\theta_z\}$. For $z\in\mathbb{Z}^2$ and $\mathfrak{a}=(\mathsf{a},\kappa,\omega,b_1,b_2)\in\mathcal{L}(c+2z)$ put*

$$\mathfrak{a}^{\uparrow}=(\theta_z,\kappa,\omega,1-b_1,1-b_2),\qquad \mathcal{L}^{\uparrow}(z)=\{\mathfrak{a}^{\uparrow}:\mathfrak{a}\in\mathcal{L}(c+2z)\}.$$

(a) Every $\mathfrak{a}^{\uparrow}$ is a letter of $\mathbb{A}$, and the map $\mathfrak{a}\mapsto\mathfrak{a}^{\uparrow}$ is injective on $\mathcal{L}(c+2z)$.

(b) $\mathcal{L}^{\uparrow}$ is an arc-consistent list family, and every letter of $\mathcal{L}^{\uparrow}(z)$ has thin type θ_z. In particular $\mathcal{L}^{\uparrow}$ has a thin phase.

(c) If $\mathcal{L}^{\uparrow}(z)=\{\mathfrak{c}\}$, then $\mathcal{L}(c+2z+\varsigma)=\{\varphi'(\mathfrak{c})(\varsigma)\}$ for every $\varsigma\in\{0,1\}^2$.

The letters of $\mathcal{L}(c+2z)$ have thin type a, so $\mathfrak{a}^{\uparrow}$ is defined. By Table 6, $\mathfrak{a}^{\uparrow}$ is the letter of thin type θ_z whose child $(0,0)$ under φ' is $\mathfrak{a}$. We call $\mathcal{L}^{\uparrow}$ the *coarsening* of $\mathcal{L}$.

Proof. For every z, the list at the north-east site of z is $\{(\mathsf{b},F(\theta_z),\mathsf{A},0,0)\}$. Each letter at the south-east site of z fits this letter in direction $\mathbf{e}_2$, and each letter at the north-west site of z fits it in direction $\mathbf{e}_1$. These letters are lines (T3), so:

(∗) every letter at the south-east site of z has vertical atom $\partial_{\mathrm{S}}\theta_z$, and every letter at the north-west site of z has horizontal atom $\partial_{\mathrm{W}}\theta_z$.

By Lemma 11.8, $\theta_{z-\mathbf{e}_1}=\theta_z^{\leftrightarrow}$ and $\theta_{z-\mathbf{e}_2}=\theta_z^{\updownarrow}$.

(a) Fix z, write $x=c+2z$, and let $\mathfrak{a}=(\mathsf{a},\kappa,\omega,b_1,b_2)\in\mathcal{L}(x)$. We first show $(\theta_z,\kappa,\omega)\in\mathbb{O}$. If $\theta_z=\mathsf{a}$, this holds because $\mathfrak{a}\in\mathbb{A}$.

Let $\theta_z = \mathsf{b}$. Then $\partial_S\theta_z = \mathsf{r}_2$ and $\partial_W\theta_z = \mathsf{r}_3$. By $(*)$ and (T3), every letter of $\mathcal{L}(x+\mathbf{e}_1)$ has black B, and every letter of $\mathcal{L}(x+\mathbf{e}_2)$ has black E. The letter $\mathfrak{a}$ fits a letter of $\mathcal{L}(x+\mathbf{e}_1)$ in direction $\mathbf{e}_1$ and a letter of $\mathcal{L}(x+\mathbf{e}_2)$ in direction $\mathbf{e}_2$. So $\partial_E\omega = \partial_W\mathsf{B} = 0$ and $\partial_N\omega = \partial_S\mathsf{E} = 0$. By (T5), $\omega = \mathsf{A}$ and κ is a corner, so $(\mathsf{b},\kappa,\mathsf{A}) \in \mathbb{O}$.

Let $\theta_z = \mathsf{c}$. Then $\partial_W\theta_z = \mathsf{r}_1$, so every letter of $\mathcal{L}(x+\mathbf{e}_2)$ has black D, by $(*)$ and (T3), and $\partial_N\omega = \partial_S\mathsf{D} = 1$. The letter $\mathfrak{a}$ and some $\mathfrak{n} \in \mathcal{L}(x-\mathbf{e}_2)$ fit in direction $-\mathbf{e}_2$. The site $x-\mathbf{e}_2$ is the north-west site of $z-\mathbf{e}_2$, and $\theta_{z-\mathbf{e}_2} = \mathsf{c}^{\updownarrow} = \mathsf{b}$. By $(*)$ the horizontal atom of $\mathfrak{n}$ is $\partial_W\mathsf{b} = \mathsf{r}_3$, so its black is E, and $\partial_S\omega = \partial_N\mathsf{E} = 0$. By (T5), $\omega \in \{\mathsf{B},\mathsf{C}\}$ and κ is a line. For $\omega \in \{\mathsf{B},\mathsf{C}\}$ the conditions of Table 5 on $(\mathsf{a},\kappa,\omega)$ and on $(\mathsf{c},\kappa,\omega)$ are the same, so $(\mathsf{c},\kappa,\omega) \in \mathbb{O}$.

Let $\theta_z = \mathsf{d}$. Then $\partial_S\theta_z = \mathsf{r}_0$, so every letter of $\mathcal{L}(x+\mathbf{e}_1)$ has black C, and $\partial_E\omega = \partial_W\mathsf{C} = 1$. The letter $\mathfrak{a}$ and some $\mathfrak{l} \in \mathcal{L}(x-\mathbf{e}_1)$ fit in direction $-\mathbf{e}_1$. The site $x-\mathbf{e}_1$ is the south-east site of $z-\mathbf{e}_1$, and $\theta_{z-\mathbf{e}_1} = \mathsf{d}^{\leftrightarrow} = \mathsf{b}$. So the vertical atom of $\mathfrak{l}$ is $\partial_S\mathsf{b} = \mathsf{r}_2$, its black is B, and $\partial_W\omega = \partial_E\mathsf{B} = 0$. By (T5), $\omega \in \{\mathsf{D},\mathsf{E}\}$ and κ is a line, and as before $(\mathsf{d},\kappa,\omega) \in \mathbb{O}$.

Next the bits. Conditions (R3) and (R4) for $\mathfrak{a}^{\uparrow}$ follow from the same conditions for $\mathfrak{a}$, since they depend only on κ and on whether the two bits are equal. Condition (R1) asks $b_1 = 1$ when $\theta_z \in \{\mathsf{b},\mathsf{c}\}$, and (R2) asks $b_2 = 1$ when $\theta_z \in \{\mathsf{b},\mathsf{d}\}$.

Let $\theta_z = \mathsf{b}$; then κ is a corner. The letter $\mathfrak{a}$ and some $\mathfrak{n} \in \mathcal{L}(x+\mathbf{e}_2)$ fit in direction $\mathbf{e}_2$, and $\mathfrak{a}$ and some $\mathfrak{n}' \in \mathcal{L}(x-\mathbf{e}_2)$ fit in direction $-\mathbf{e}_2$. Both are lines of thin type d, with row bit 0 and column bit b_1, and their vertical atoms are $\partial_N\kappa$ and $\partial_S\kappa$. By $(*)$, the horizontal atom of $\mathfrak{n}$ is $\partial_W\mathsf{b} = \mathsf{r}_3$, and that of $\mathfrak{n}'$ is $\partial_W\theta_{z-\mathbf{e}_2} = \partial_W\mathsf{c} = \mathsf{r}_1$; both are red. If $b_1 = 0$, then (T4) makes $\partial_N\kappa$ and $\partial_S\kappa$ green, against (T1). So $b_1 = 1$. In the same way, the letters of $\mathcal{L}(x+\mathbf{e}_1)$ and $\mathcal{L}(x-\mathbf{e}_1)$ that fit $\mathfrak{a}$ are lines of thin type c with column bit 0 and row bit b_2. Their vertical atoms are $\partial_S\mathsf{b} = \mathsf{r}_2$ and $\partial_S\theta_{z-\mathbf{e}_1} = \partial_S\mathsf{d} = \mathsf{r}_0$, and their horizontal atoms are $\partial_E\kappa$ and $\partial_W\kappa$. So $b_2 = 1$.

Let $\theta_z = \mathsf{c}$. Then $\theta_{z+\mathbf{e}_2} = \mathsf{c}^{\updownarrow} = \mathsf{b}$. The letter $\mathfrak{a}$ fits a letter of $\mathcal{L}(x+\mathbf{e}_2)$ in direction $\mathbf{e}_2$, which fits a letter $\mathfrak{a}' \in \mathcal{L}(x+2\mathbf{e}_2)$ in direction $\mathbf{e}_2$. Vertical matching keeps the column bit, so $\chi_1(\mathfrak{a}') = b_1$. The site $x+2\mathbf{e}_2$ is the south-west site of $z+\mathbf{e}_2$, so the case $\theta = \mathsf{b}$ gives $\chi_1(\mathfrak{a}') = 1$. Hence $b_1 = 1$. Let $\theta_z = \mathsf{d}$. Then $\theta_{z+\mathbf{e}_1} = \mathsf{b}$, and the same argument along the row gives $b_2 = 1$.

So $\mathfrak{a}^{\uparrow} \in \mathbb{A}$. The letter $\mathfrak{a}^{\uparrow}$ determines κ, ω and the bits of $\mathfrak{a}$, whose thin type is a; so the map is injective.

(b) The lists are nonempty. Let $\mathfrak{a} = (\mathsf{a},\kappa,\omega,b_1,b_2) \in \mathcal{L}(x)$. It fits some $\mathfrak{l} \in \mathcal{L}(x+\mathbf{e}_1)$ in direction $\mathbf{e}_1$, and $\mathfrak{l}$ fits some $\mathfrak{a}' = (\mathsf{a},\kappa',\omega',b_1',b_2') \in \mathcal{L}(x+2\mathbf{e}_1)$ in direction $\mathbf{e}_1$; here $x+2\mathbf{e}_1$ is the south-west site of $z+\mathbf{e}_1$. We check that $\mathfrak{a}^{\uparrow}$ and $\mathfrak{a}'^{\uparrow}$ match in direction $\mathbf{e}_1$. Their thin types θ_z and $\theta_{z+\mathbf{e}_1} = \theta_z^{\leftrightarrow}$ match. The line $\mathfrak{l}$ has equal east and west endpoints, so $\partial_E\kappa = \partial_W\kappa'$. Its black has equal east and west bits (T3), so $\partial_E\omega = \partial_W\omega'$. Its row bit equals b_2 and b_2', so $1-b_2 = 1-b_2'$. Direction $-\mathbf{e}_1$ is the same with the sites $x-\mathbf{e}_1$ and $x-2\mathbf{e}_1$. The directions $\pm\mathbf{e}_2$ use the lines at $x \pm \mathbf{e}_2$, whose blacks D, E have equal north and south bits, and the column bits. So $\mathcal{L}^{\uparrow}$ is arc consistent. Its letters at z have thin type θ_z, and by Lemma 11.8 the map $z \mapsto \theta_z$ follows the rule of a thin phase.

(c) Let $\mathcal{L}^{\uparrow}(z) = \{\mathfrak{c}\}$ with $\mathfrak{c} = (\theta_z,\kappa,\omega,c_1,c_2)$. By Table 6, the four children of $\mathfrak{c}$ under φ' are

$$\varphi'(\mathfrak{c})(0,0) = (\mathsf{a},\kappa,\omega,1-c_1,1-c_2), \qquad \varphi'(\mathfrak{c})(1,0) = (\mathsf{c},\langle\partial_S\theta_z,\partial_E\kappa\rangle,\mathrm{se}(\omega),0,1-c_2),$$
$$\varphi'(\mathfrak{c})(1,1) = (\mathsf{b},F(\theta_z),\mathsf{A},0,0), \qquad \varphi'(\mathfrak{c})(0,1) = (\mathsf{d},\langle\partial_N\kappa,\partial_W\theta_z\rangle,\mathrm{nw}(\omega),1-c_1,0).$$

By (a), $\mathcal{L}(x)$ is the singleton of $\mathfrak{a}_0 = \varphi'(\mathfrak{c})(0,0)$. The list at $x+(1,1)$ is the singleton of $(\mathsf{b},F(\theta_z),\mathsf{A},0,0) = \varphi'(\mathfrak{c})(1,1)$. Let $\mathfrak{l} \in \mathcal{L}(x+\mathbf{e}_1)$. By (T3) and $(*)$ it is a line of thin type c, with column bit 0 and vertical atom $\partial_S\theta_z$. It fits $\mathfrak{a}_0$ in direction $-\mathbf{e}_1$, so its horizontal atom is $\partial_E\kappa$, its row bit is $1-c_2$, and its black is the one of B, C whose west bit is $\partial_E\omega$, namely $\mathrm{se}(\omega)$ (Section A). So $\mathfrak{l} = \varphi'(\mathfrak{c})(1,0)$. In the same way, every letter of $\mathcal{L}(x+\mathbf{e}_2)$ is a line of thin type

d with row bit 0 and horizontal atom $\partial_W\theta_z$; it fits $\mathfrak{a}_0$ in direction $-\mathbf{e}_2$, so its vertical atom is $\partial_N\kappa$, its column bit is $1-c_1$, and its black is the one of D, E whose south bit is $\partial_N\omega$, namely $\mathrm{nw}(\omega)$. So it equals $\varphi'(\mathfrak{c})(0,1)$. The lists are nonempty, so all four are singletons. □

Lemma 11.11 (Repeated coarsening)**.** *Let $m \geq 1$, and let $\mathcal{L}$ be an arc-consistent list family with $\mathcal{L}(y_1) = \{\mathfrak{p}_1\}$ and $\mathcal{L}(y_2) = \{\mathfrak{p}_2\}$, where $\mathfrak{p}_1, \mathfrak{p}_2$ are opposite corner letters and $y_2 - y_1 \in 2^m\mathbb{Z}^2$. Put $w = (y_2 - y_1)/2^m$, $z_1 = 0$ and $z_2 = w$. There are arc-consistent list families $\mathcal{L}_0, \dots, \mathcal{L}_m$, with $\mathcal{L}_0(z) = \mathcal{L}(y_1 + z)$ for every z, such that:*

(a) for $0 \leq k < m$, $\mathcal{L}_k$ has thin phase 0, every $\mathcal{L}_k^{\swarrow}(z)$ is a singleton, and $\mathcal{L}_{k+1} = \mathcal{L}_k^{\uparrow}$;

(b) for every z, the letters of $\mathcal{L}(y_1 + 2^m z)$ have thin type a, and the letters of $\mathcal{L}_m(z)$ are obtained from them by complementing both bits m times and replacing the thin type by the thin type of $\mathcal{L}_m$ at z;

(c) for some $r \in \{1,2\}$, $\mathcal{L}_m$ has thin phase z_r and $\mathcal{L}_m(z_r) = \{(\mathsf{a}, \kappa_r, \mathsf{A}, 1, 1)\}$, where κ_r is the thick type of $\mathfrak{p}_r$;

(d) if $\mathcal{L}_m(z) = \{\mathfrak{c}\}$ for some z and $\mathfrak{c}$, then $\mathcal{L}(y_1 + 2^m z + y) = \{(\varphi')^m(\mathfrak{c})(y)\}$ for every $y \in [0, 2^m)^2$.

Proof. For $0 \leq k \leq m$ and $i \in \{1,2\}$, let $\mathfrak{p}_i^{(k)}$ be the corner letter with the thick type of $\mathfrak{p}_i$ whose bit is the bit of $\mathfrak{p}_i$ complemented k times. We show by induction on $k \in [0, m-1]$ that $\mathcal{L}_k$ can be defined as in (a), with thin phase 0, $\mathcal{L}_k(0) = \{\mathfrak{p}_1^{(k)}\}$ and $\mathcal{L}_k(2^{m-k}w) = \{\mathfrak{p}_2^{(k)}\}$. For $k = 0$, $\mathcal{L}_0$ is arc consistent, and its lists at 0 and at $2^m w = y_2 - y_1$ are $\{\mathfrak{p}_1\}$ and $\{\mathfrak{p}_2\}$. Since $\mathfrak{p}_1$ has thin type a, Lemma 11.7 gives thin phase 0.

Let $0 \leq k \leq m-1$ and assume the claim for k. The letters $\mathfrak{p}_1^{(k)}$ and $\mathfrak{p}_2^{(k)}$ are opposite; let $x \in \{0, 2^{m-k}w\}$ be the site of the one with bit 0. Since $m - k \geq 1$, $x \in 2\mathbb{Z}^2$. By Lemma 11.9, with $c = 0$ and $z_0 = x/2$, every $\mathcal{L}_k^{\swarrow}(z)$ is a singleton $\{\theta_z\}$, and $\theta_z = \mathsf{a}$ exactly for $z \in x/2 + 2\mathbb{Z}^2$. Put $\mathcal{L}_{k+1} = \mathcal{L}_k^{\uparrow}$. By Lemma 11.10(b) it is arc consistent and has thin phase $x/2$. By Lemma 11.10(a), its lists at 0 and at $2^{m-k-1}w$ are the singletons of $(\mathfrak{p}_1^{(k)})^{\uparrow}$ and $(\mathfrak{p}_2^{(k)})^{\uparrow}$. These letters have the thick types of $\mathfrak{p}_1$ and $\mathfrak{p}_2$, black A, complemented bits, and the thin types θ_0 and $\theta_{2^{m-k-1}w}$. If $k + 1 \leq m - 1$, then 0, $2^{m-k-1}w$ and $x/2$ lie in $2\mathbb{Z}^2$, so the thin phase is 0, both thin types are a, and the claim holds for $k+1$. If $k + 1 = m$, then $x/2 \in \{0, w\}$; let r be the index with $x/2 = z_r$. The letter $(\mathfrak{p}_r^{(m-1)})^{\uparrow}$ has thin type $\theta_{z_r} = \mathsf{a}$, thick type κ_r, black A and both bits 1. This proves (a) and (c).

For (b) we show by induction on $k \in [1, m]$ that the letters of $\mathcal{L}(y_1 + 2^k z)$ have thin type a and that the letters of $\mathcal{L}_k(z)$ are obtained from them by complementing both bits k times and putting the thin type of $\mathcal{L}_k$ at z. The site $y_1 + 2^k z$ lies in $y_1 + 2\mathbb{Z}^2$, where $\mathcal{L}$ has thin type a. For $k = 1$ the claim is the definition of $\mathcal{L}_0^{\uparrow}$. Suppose it holds for some $k \leq m - 1$. The letters of $\mathcal{L}_k(2z)$ have thin type a, because $\mathcal{L}_k$ has thin phase 0, and by the claim for k they come from the letters of $\mathcal{L}(y_1 + 2^{k+1}z)$. The map $\mathfrak{a} \mapsto \mathfrak{a}^{\uparrow}$ complements both bits once more and puts the thin type of $\mathcal{L}_{k+1}$ at z.

For (d) we show by induction on $l \in [0, m]$ that $\mathcal{L}_{m-l}(2^l z + y) = \{(\varphi')^l(\mathfrak{c})(y)\}$ for every $y \in [0, 2^l)^2$. For $l = 0$ this is the hypothesis. Let $l < m$, and let $y \in [0, 2^{l+1})^2$, written $y = 2y' + \varsigma$ with $y' \in [0, 2^l)^2$ and $\varsigma \in \{0,1\}^2$. By the induction hypothesis, $\mathcal{L}_{m-l}(2^l z + y')$ is the singleton of $(\varphi')^l(\mathfrak{c})(y')$. Since $\mathcal{L}_{m-l}$ is the coarsening of $\mathcal{L}_{m-l-1}$ with $c = 0$, Lemma 11.10(c) gives

$$\mathcal{L}_{m-l-1}(2^{l+1}z + y) = \{\varphi'((\varphi')^l(\mathfrak{c})(y'))(\varsigma)\} = \{(\varphi')^{l+1}(\mathfrak{c})(y)\},$$

the last step by the blockwise definition of φ' (Definition 7.8). For $l = m$, $\mathcal{L}_0(2^m z + y) = \mathcal{L}(y_1 + 2^m z + y)$. □

11.3 Corner pairs at every scale

The next lemma is a finite statement. Let $V_{16} = \{-16, \dots, 16\}^2$, as in Section 4. For a corner κ, define lists $\mathcal{D}_\kappa(v) \subseteq \mathbb{A}$ for $v \in V_{16}$ as follows. The list $\mathcal{D}_\kappa(0)$ is $\{(\mathsf{a}, \kappa, \mathsf{A}, 1, 1)\}$. For $v \neq 0$, the list $\mathcal{D}_\kappa(v)$ consists of the letters whose thin type is a, c, d or b according as v is congruent to $(0,0)$, $(1,0)$, $(0,1)$ or $(1,1)$ modulo 2, except the four corner letters with bit 0, namely $(\mathsf{a}, \kappa', \mathsf{A}, 0, 0)$ for the four corners κ'. Every other letter of $\mathbb{A}$ with the right thin type is kept, lines and letters with unequal bits included. A family $\mathcal{K}$ of sets $\mathcal{K}(v) \subseteq \mathcal{D}_\kappa(v)$, $v \in V_{16}$, is *consistent on the box* if, for every $v \in V_{16}$, every $\mathfrak{a} \in \mathcal{K}(v)$ and every $e \in \{\pm\mathbf{e}_1, \pm\mathbf{e}_2\}$ with $v + e \in V_{16}$, some $\mathfrak{b} \in \mathcal{K}(v + e)$ is such that $\mathfrak{a}, \mathfrak{b}$ fit in direction e. Neighbours outside V_{16} impose nothing: the box does not wrap around.

Lemma 11.12 (Corner lemma)**.** *Let κ be one of the four corners. Every family $\mathcal{K}$ that is consistent on the box and contained in $\mathcal{D}_\kappa$ has $\mathcal{K}(v) = \emptyset$ for every $v \in V_{16}$.*

Proof. This is a finite arc-consistency computation, carried out by two independent programs with positive controls. The programs are the files

- `verification/corner-lemma/ENGINE-PALETTE-A.py`,
- `verification/corner-lemma/CORNERS-B.py`

in the repository of this paper. Each program builds the 276 letters of $\mathbb{A}$ and their edge colours from the rules of Section A and Definition 7.7. It sets up the lists $\mathcal{D}_\kappa$ on the box $[0, 32]^2$ with centre $(16, 16)$, the translate of V_{16} by the even vector $(16, 16)$. It then deletes from a list each letter that has no fitting letter in the current list at some neighbour inside the box, until no deletion applies. A deleted letter lies in no family that is consistent on the box and contained in the current lists. So, by induction on the number of deletions, every family that is consistent on the box and contained in $\mathcal{D}_\kappa$ stays inside the current lists. The final lists form a family that is consistent on the box. Hence both programs compute the largest such family, whatever the order of the deletions. For each of the four corners κ, both programs find that every list of this largest family is empty.

The two programs encode the colours differently and delete in different orders: the first revisits one site at a time from a queue, the second sweeps the whole box synchronously. As positive controls, both run the same computation on the lists $\mathcal{D}_\kappa^+$ in which the four corner letters with bit 0 are kept. For each κ the final lists then have no empty site and 57290 letters in total, and they contain all 1089 values of a legal pattern of $\mathbb{A}$ on the box whose value at the centre is $(\mathsf{a}, \kappa, \mathsf{A}, 1, 1)$; the pattern is a restriction of a word $\sigma^3(\mathfrak{o})$. The program `verification/corner-lemma/CORNERS-B-CONTROLS.py` is the second program with this last check added. A third implementation, written separately, is included in the same folder and gives the same results. The file `verification/corner-lemma/VERIFY.md` says how to run the programs and what they print. □

Corollary 11.13 (A corner letter with bit 0 nearby)**.** *Let $\mathcal{L}$ be an arc-consistent list family, and let $\mathcal{L}(z_0) = \{(\mathsf{a}, \kappa, \mathsf{A}, 1, 1)\}$ for some $z_0 \in \mathbb{Z}^2$ and some corner κ. Then some $z \in z_0 + V_{16}$ has a corner letter with bit 0 in $\mathcal{L}(z)$.*

Proof. Suppose not. By Lemma 11.7, $\mathcal{L}$ has thin phase z_0. Put $\mathcal{K}(v) = \mathcal{L}(z_0 + v)$ for $v \in V_{16}$. Then $\mathcal{K}(v) \subseteq \mathcal{D}_\kappa(v)$ for every v: the thin types follow the rule of $\mathcal{D}_\kappa$, the list $\mathcal{K}(0)$ is the given

singleton, and no corner letter with bit 0 occurs. The family $\mathcal{K}$ is consistent on the box, because $\mathcal{L}$ is arc consistent. By Lemma 11.12, $\mathcal{K}(0) = \emptyset$, which is false. □

Remark (A smaller box). The proof of Corollary 11.13 uses Lemma 11.12 only at the centre: it needs $\mathcal{K}(0) = \emptyset$, not that every list is empty. That weaker statement already holds on the box V_6, of side 13. With V_6 in place of V_{16} in the definitions of $\mathcal{D}_\kappa$ and of consistency on the box, every family that is consistent on V_6 and contained in $\mathcal{D}_\kappa$ has $\mathcal{K}(0) = \emptyset$, for each of the four corners κ. The third program in `verification/corner-lemma/`, `ENGINE-C.py`, computes the largest such family for the centred boxes of odd side $3, 5, \ldots$ and finds that 13 is the least side for which the list at the centre empties, for each corner (`minimal_emptying_side` in its output). That no smaller side suffices is reported by this program alone and is used nowhere. The Lean formalization of Section C proves the same statement on V_6 with a certificate checked by the Lean kernel. A family that is consistent on a box stays consistent when it is restricted to a smaller box with the same centre, so the statement on V_6 gives the same conclusion on every box V_r with $r \geq 6$. Hence Corollary 11.13 holds with $z_0 + V_6$ in place of $z_0 + V_{16}$. Nothing else in the paper changes; we keep V_{16}, which is the box on which the two programs cited in the proof of Lemma 11.12 show that every list is empty.

Definition 11.14 (Corner pairs)**.** Let $m \geq 1$. A pin pair $(\iota(y_1), s_1; \iota(y_2), s_2)$ is a *corner pair of order* m if $\mathrm{par}(\mathrm{ctr}(s_1))$ and $\mathrm{par}(\mathrm{ctr}(s_2))$ are opposite corner letters and $y_2 - y_1 \in 2^m\mathbb{Z}^2$.

Lemma 11.15 (A first corner pair)**.** *$\mathcal{W}$ has a corner pair of order* 1.

Proof. A pin on a cell. Take a state of $\mathcal{W}$. Move token 1 by Moore steps to $\iota(0)$, using (W3), and then clone it onto tokens 2 and 3, using (W2). The result is a state $((\iota(0), \iota(0), \iota(0)), (s, s, s)) \in \mathcal{W}$, so $\mathbf{p}_0 = (\iota(0), s; \iota(0), s)$ is a pin pair. By Lemma 11.6, $\mathcal{L}_{\mathbf{p}_0}$ is arc consistent and $\mathcal{L}_{\mathbf{p}_0}(0)$ is a singleton. By Lemma 11.7 it has a thin phase c.

A pin of thin type b. Let $y = c + (1, 1)$. By Lemma 11.6(a) there is $(\iota(y), s') \in \mathcal{X}_{\mathbf{p}_0}$, and $\mathrm{par}(\mathrm{ctr}(s')) \in \mathcal{L}_{\mathbf{p}_0}(y)$ has thin type b. In the state $((\iota(0), \iota(0), \iota(y)), (s, s, s')) \in \mathcal{W}$, clone token 3 onto tokens 1 and 2. So $\mathbf{p}_1 = (\iota(y), s'; \iota(y), s')$ is a pin pair, and $\mathcal{L}_{\mathbf{p}_1}(y) = \{(\mathsf{b}, F(\theta^*), \mathsf{A}, 0, 0)\}$ for one thin type θ^*.

A corner pin with bit 1. The family $\mathcal{L} = \mathcal{L}_{\mathbf{p}_1}$ has thin phase $c' = y - (1, 1)$ by Lemma 11.7, and y is the north-east site of 0. So $\mathcal{L}^{\angle}(0) = \{\theta^*\}$, and by Lemma 11.8 every $\mathcal{L}^{\angle}(z)$ is a singleton $\{\theta_z\}$. The four values θ_0, $\theta_{\mathbf{e}_1}$, $\theta_{\mathbf{e}_2}$, $\theta_{(1,1)}$ are distinct, so $\theta_{z_\mathsf{b}} = \mathsf{b}$ for some z_b. By Lemma 11.10(a),(b), every letter of $\mathcal{L}^{\uparrow}(z_\mathsf{b})$ is a letter of $\mathbb{A}$ with thin type b, so it is $(\mathsf{b}, F(\theta), \mathsf{A}, 0, 0)$ for some thin type θ. These letters are the letters $\mathfrak{a}^{\uparrow}$ for $\mathfrak{a} \in \mathcal{L}(y_\mathsf{b})$, where $y_\mathsf{b} = c' + 2z_\mathsf{b}$, and $\mathfrak{a}^{\uparrow} = (\mathsf{b}, F(\theta), \mathsf{A}, 0, 0)$ forces $\mathfrak{a} = (\mathsf{a}, F(\theta), \mathsf{A}, 1, 1)$. So every letter of $\mathcal{L}(y_\mathsf{b})$ is a corner letter with bit 1. By Lemma 11.6(a) there is $(\iota(y_\mathsf{b}), s'') \in \mathcal{X}_{\mathbf{p}_1}$. Clone token 3 onto tokens 1 and 2. Then $\mathbf{p}_2 = (\iota(y_\mathsf{b}), s''; \iota(y_\mathsf{b}), s'')$ is a pin pair, and $\mathcal{L}_{\mathbf{p}_2}(y_\mathsf{b}) = \{(\mathsf{a}, \kappa, \mathsf{A}, 1, 1)\}$ for a corner κ.

An opposite corner letter. By Corollary 11.13, applied to $\mathcal{L}_{\mathbf{p}_2}$, some $z' \in y_\mathsf{b} + V_{16}$ has a corner letter $\mathfrak{b}$ with bit 0 in $\mathcal{L}_{\mathbf{p}_2}(z')$. So there is $(\iota(z'), s''') \in \mathcal{X}_{\mathbf{p}_2}$ with $\mathrm{par}(\mathrm{ctr}(s''')) = \mathfrak{b}$. In the state $((\iota(y_\mathsf{b}), \iota(y_\mathsf{b}), \iota(z')), (s'', s'', s''')) \in \mathcal{W}$, clone token 3 onto token 2. The result shows that $(\iota(y_\mathsf{b}), s''; \iota(z'), s''')$ is a pin pair. Its two letters are opposite corner letters. Both have thin type a, and $\mathcal{L}_{\mathbf{p}_2}$ has thin phase y_b by Lemma 11.7, so $z' - y_\mathsf{b} \in 2\mathbb{Z}^2$. This is a corner pair of order 1. □

Lemma 11.16 (Promotion)**.** *If $\mathcal{W}$ has a corner pair of order $m \geq 1$, then it has a corner pair of order $m + 1$.*

Proof. Let $\mathbf{p} = (\iota(y_1), s_1; \iota(y_2), s_2)$ be a corner pair of order m, with letters $\mathfrak{p}_i = \mathrm{par}(\mathrm{ctr}(s_i))$. By Lemma 11.6, $\mathcal{L} = \mathcal{L}_\mathbf{p}$ is arc consistent and $\mathcal{L}(y_i) = \{\mathfrak{p}_i\}$ for $i = 1, 2$. Let $\mathcal{L}_0, \dots, \mathcal{L}_m$, r and z_r be as in Lemma 11.11, and note that $y_1 + 2^m z_r = y_r$. By Lemma 11.11(c) and Corollary 11.13, applied to $\mathcal{L}_m$ at z_r, some $z \in z_r + V_{16}$ has a corner letter $\mathfrak{b}$ with bit 0 in $\mathcal{L}_m(z)$. Since $\mathcal{L}_m$ has thin phase z_r and $\mathfrak{b}$ has thin type a, $z - z_r \in 2\mathbb{Z}^2$.

By Lemma 11.11(b), $\mathfrak{b}$ is obtained from a letter $\mathfrak{b}_0 \in \mathcal{L}(y')$, $y' = y_1 + 2^m z$, by complementing both bits m times and changing the thin type. So $\mathfrak{b}_0$ is a corner letter with the thick type of $\mathfrak{b}$, and its bit is 0 complemented m times. In the same way, $\mathcal{L}_m(z_r) = \{(\mathsf{a}, \kappa_r, \mathsf{A}, 1, 1)\}$ comes from $\mathcal{L}(y_r) = \{\mathfrak{p}_r\}$, so the bit of $\mathfrak{p}_r$ is 1 complemented m times. Hence $\mathfrak{b}_0$ and $\mathfrak{p}_r$ are opposite corner letters, and $y' - y_r = 2^m(z - z_r) \in 2^{m+1}\mathbb{Z}^2$.

By the definition of $\mathcal{L}_\mathbf{p}$ there is $(\iota(y'), s) \in \mathcal{X}_\mathbf{p}$ with $\mathrm{par}(\mathrm{ctr}(s)) = \mathfrak{b}_0$, so that

$$((\iota(y_1), \iota(y_2), \iota(y')), (s_1, s_2, s)) \in \mathcal{W}.$$

If $r = 1$, clone token 3 onto token 2; if $r = 2$, clone token 3 onto token 1. The resulting state shows that $(\iota(y_1), s_1; \iota(y'), s)$, respectively $(\iota(y'), s; \iota(y_2), s_2)$, is a pin pair. Its letters are $\mathfrak{p}_r$ and $\mathfrak{b}_0$, and its displacement is $\pm(y' - y_r)$. It is a corner pair of order $m + 1$. □

The witness s is a symbol that occurs in $\mathcal{X}_\mathbf{p}$; the argument chooses a state of $\mathcal{W}$, not a reply to a move.

Lemma 11.17 (The forced word)**.** *Let $n \geq 1$, and let $\mathbf{p} = (\iota(y_1), s_1; \iota(y_2), s_2)$ be a corner pair of order $2n$. There are $r \in \{1, 2\}$ and a corner letter $\mathfrak{a}$ with bit 1 such that*

$$\mathcal{L}_\mathbf{p}(y_r + y) = \{\sigma^n(\mathfrak{a})(y)\} \qquad \textit{for every } y \in [0, 4^n)^2.$$

Proof. Apply Lemma 11.11 to $\mathcal{L}_\mathbf{p}$, with $m = 2n$; this is possible by Lemma 11.6. Let r be as there and $\mathfrak{a} = (\mathsf{a}, \kappa_r, \mathsf{A}, 1, 1)$, so that $\mathcal{L}_{2n}(z_r) = \{\mathfrak{a}\}$. By Lemma 11.11(d), $\mathcal{L}_\mathbf{p}(y_1 + 4^n z_r + y) = \{(\varphi')^{2n}(\mathfrak{a})(y)\}$ for $y \in [0, 4^n)^2$. Here $y_1 + 4^n z_r = y_r$, and $(\varphi')^{2n}(\mathfrak{a}) = \sigma^n(\mathfrak{a})$ because blockwise substitution of patterns composes, as noted after Definition 8.1. □

11.4 The forced word and the contradiction

In this subsection M halts, and $T = T(M)$ is its halting time, the number of steps of M° before it halts (Section 8.6). Fix $n \geq 2$ with $2^n - 2 \geq 2T + 10$. By Lemma 11.15, and by Lemma 11.16 applied $2n - 1$ times, $\mathcal{W}$ has a corner pair $\mathbf{p} = (\iota(y_1), s_1; \iota(y_2), s_2)$ of order $2n$. Let r and $\mathfrak{a}$ be given by Lemma 11.17, and put

$$W = \sigma^n(\mathfrak{a}), \qquad o = \iota(y_r).$$

Let Δ_n, J, J°, $\mathcal{F}$ and Z_W be as in (Q5) of Section 9.1 for this word, and let $N = 2T + 10$, $f_0 < \cdots < f_{N-1}$, F_N, Z, the channel edges and the palettes $\mathrm{Pal}(\xi)$ be as in Section 9.6. They depend only on W and T. Tokens 1 and 2 hold $\mathbf{p}$ throughout, and all statements below concern the walker set $\mathcal{X}_\mathbf{p}$.

Lemma 11.18 (Forced quarters)**.** *Let $\xi \in [0, 2 \cdot 4^n)^2$ and $(o + \xi, s) \in \mathcal{X}_\mathbf{p}$. Then the scaffold pair of $\mathrm{ctr}(s)$ is the quarter of W at ξ.*

Proof. Let $\varsigma = \xi \bmod 2$ and $z = \lfloor \xi/2 \rfloor$, coordinatewise, so that $o + \xi = \iota(y_r + z) + \varsigma$ and $z \in [0, 4^n)^2$. The quadrant of $\mathrm{ctr}(s)$ is $o + \xi - \bar{\gamma} \equiv \xi \equiv \varsigma \pmod 2$. For the parent, if $\varsigma_1 = 1$, move the walker by $-\mathbf{e}_1$ (Lemma 11.6(a)). By Lemma 11.2, the new centre lies west of the old one in

a legal pattern and has first quadrant coordinate 0, so by (Q1) the two centres have the same parent. If $\varsigma_2 = 1$, move the walker by $-\mathbf{e}_2$ in the same way. This reaches $(\iota(y_r + z), s'') \in \mathcal{X}_{\mathbf{p}}$ with $\mathrm{par}(\mathrm{ctr}(s'')) = \mathrm{par}(\mathrm{ctr}(s))$. By Lemma 11.17, $\mathrm{par}(\mathrm{ctr}(s'')) \in \mathcal{L}_{\mathbf{p}}(y_r + z) = \{W(z)\}$. So the scaffold pair of $\mathrm{ctr}(s)$ is $(\varsigma, W(z))$, the quarter of W at ξ. □

Lemma 11.19 (Corridors). *Let $\xi \in Z_W$ and $(o + \xi, s) \in \mathcal{X}_{\mathbf{p}}$. Then $\mathrm{ctr}(s)$ is a horizontal carrier if and only if $\xi_2 \in \mathcal{F}$, and a vertical carrier if and only if $\xi_1 \in \mathcal{F}$. Consequently its role is horizontal if $\xi_1 \notin \mathcal{F}$, vertical if $\xi_2 \notin \mathcal{F}$, active or seed if $\xi \in \mathcal{F}^2$, and seed if $\xi = (2\Delta_n + 1, 2\Delta_n + 1)$.*

Proof. By (Q5)(c), let $\xi + m\mathbf{e}_1$ and $\xi - m'\mathbf{e}_1$, with $m, m' \geq 1$, be the nearest quarters of W to the right and to the left of ξ in its row that lie on a light vertical wall. Both lie in $J \times \{\xi_2\} \subseteq [0, 2 \cdot 4^n)^2$.

Walk to the right: put $s_0 = s$, and for $l = 1, \ldots, m$ use Lemma 11.6(a) to get $(o + \xi + l\mathbf{e}_1, s_l) \in \mathcal{X}_{\mathbf{p}}$ with s_l compatible with s_{l-1} at displacement $\mathbf{e}_1$. By Lemma 11.2, $\mathrm{ctr}(s_{l-1})$ and $\mathrm{ctr}(s_l)$ match horizontally, so $\phi_{\mathrm{E}}(\mathrm{ctr}(s_{l-1})) = \phi_{\mathrm{W}}(\mathrm{ctr}(s_l))$ by (Q2). By Lemma 11.18, the scaffold pair of $\mathrm{ctr}(s_l)$ is the quarter of W at $\xi + l\mathbf{e}_1$. For $0 \leq l < m$ this quarter lies on no light vertical wall: for $l = 0$ by (Q5)(c), and for $0 < l < m$ by the choice of m. So $\phi_{\mathrm{W}}(\mathrm{ctr}(s_l)) = \phi_{\mathrm{E}}(\mathrm{ctr}(s_l))$ for $0 \leq l < m$, and $\psi = \phi_{\mathrm{W}}(\mathrm{ctr}(s))$ satisfies $\psi = \phi_{\mathrm{E}}(\mathrm{ctr}(s)) = \phi_{\mathrm{W}}(\mathrm{ctr}(s_m))$, where s_m sits at the right wall. Walking to the left in the same way gives a symbol s' at the left wall with $\psi = \phi_{\mathrm{E}}(\mathrm{ctr}(s'))$.

If $\xi_2 \in \mathcal{F}$, the left wall is a left side and the right wall a right side, so by (Q2) $\psi \subseteq \{\mathrm{left}\} \cap \{\mathrm{right}\} = \emptyset$, and $\mathrm{ctr}(s)$ is a horizontal carrier. If $\xi_2 \notin \mathcal{F}$, then the left wall is a right side, which gives $\psi = \phi_{\mathrm{E}}(\mathrm{ctr}(s')) \neq \emptyset$, or the right wall is a left side, which gives $\psi = \phi_{\mathrm{W}}(\mathrm{ctr}(s_m)) \neq \emptyset$; so $\mathrm{ctr}(s)$ is not a horizontal carrier. The vertical statement is proved in the same way along the column of ξ, with $\mathbf{e}_2$, the flows $\phi_{\mathrm{S}}, \phi_{\mathrm{N}}$, and bottom and top sides. The roles follow from (Q3). At $\xi = (2\Delta_n + 1, 2\Delta_n + 1)$ the tile is active, and by Lemma 11.18 and (Q5)(b) its scaffold pair is the north-east quarter of an index-0 letter with dark ruler bits, so its role is seed. □

The lemma holds for every symbol of the walker set at $o + \xi$, not only for one choice of flows. In this sense the corridors of (Q5)(c) force the roles on the lists themselves: each corridor is a chain of equalities between neighbouring flows, closed by the two wall conditions of (Q2), and the walk checks it for every value that the walker can carry.

By Lemma 11.19 and (Q3), $\mathrm{pay}(\mathrm{ctr}(s)) \in \mathrm{Pal}(\xi)$ for every $\xi \in Z$ and every $(o + \xi, s) \in \mathcal{X}_{\mathbf{p}}$. On $F_N^2 \setminus \{(f_0, f_0)\}$ the role is active or seed, and both palettes lie in T_M, because $s_M \in T_M$.

Lemma 11.20 (Projected meet). *For $\xi \in Z$ and $v = o + \xi$ put*

$$U_v = \{\mathrm{pay}(\mathrm{ctr}(s)) : (v, s) \in \mathcal{X}_{\mathbf{p}}\}, \qquad m_v = \bigwedge U_v.$$

(a) Each U_v is a nonempty finite set, and $m_v \in \mathrm{Pal}(\xi)$.

(b) Let $\{\xi, \xi + e\}$ be a channel edge with $e \in \{\mathbf{e}_1, \mathbf{e}_2\}$, and let $w = v + e$. If $e = \mathbf{e}_1$ then $(m_v)_{\mathrm{E}} = (m_w)_{\mathrm{W}}$, and if $e = \mathbf{e}_2$ then $(m_v)_{\mathrm{N}} = (m_w)_{\mathrm{S}}$.

Proof. (a) U_v is nonempty by Lemma 11.6(a), and $U_v \subseteq C_M^4$ is finite. Every palette is closed under $\wedge$, as in the proof of Lemma 9.7(a), and $U_v \subseteq \mathrm{Pal}(\xi)$. So $m_v \in \mathrm{Pal}(\xi)$.

(b) Let $e = \mathbf{e}_1$, and let Q_{vw} be the set of pairs $(\mathrm{pay}(\mathrm{ctr}(s)), \mathrm{pay}(\mathrm{ctr}(s')))$ with $(v, s), (w, s') \in \mathcal{X}_{\mathbf{p}}$ and s' compatible with s at displacement $\mathbf{e}_1$. Its first projection is U_v: for $(v, s) \in \mathcal{X}_{\mathbf{p}}$, Lemma 11.6(a) gives $(w, s') \in \mathcal{X}_{\mathbf{p}}$ with s' compatible with s at displacement $\mathbf{e}_1$. Its second projection is U_w: for $(w, s') \in \mathcal{X}_{\mathbf{p}}$, Lemma 11.6(a) gives $(v, s) \in \mathcal{X}_{\mathbf{p}}$ with s compatible with

s' at displacement $-\mathbf{e}_1$, and compatibility is symmetric (Definition 4.3). For each pair, $\mathrm{ctr}(s)$ and $\mathrm{ctr}(s')$ match horizontally by Lemma 11.2, so their payloads satisfy $\mathrm{pay}(\mathrm{ctr}(s))_{\mathrm{E}} = \mathrm{pay}(\mathrm{ctr}(s'))_{\mathrm{W}}$ by (Q3). List $Q_{vw} = \{(\mathbf{c}^1, \mathbf{d}^1), \dots, (\mathbf{c}^l, \mathbf{d}^l)\}$. Since $\wedge$ acts coordinatewise,

$$(m_v)_{\mathrm{E}} = \bigwedge_i \mathbf{c}^i_{\mathrm{E}} = \bigwedge_i \mathbf{d}^i_{\mathrm{W}} = (m_w)_{\mathrm{W}}.$$

The case $e = \mathbf{e}_2$ is the same. □

Proof of Theorem 11.1. Let M halt, with halting time $T = T(M)$, and suppose that R_M is not 3-token refutable. By Theorem 4.25(i) there is a nonempty plane-safe set of 3-token states for R_M. Let $\mathcal{W}$ be a nonempty set $\mathcal{W}_\gamma$ obtained from it by Lemma 11.3, and let n, $\mathbf{p}$, W, o and the meets m_v be as above. Define Ξ on $[0, N)^2$ by $\Xi(i, i') = m_{o+(f_i, f_{i'})}$. By Lemma 11.20(a), $\Xi(i, i') \in T_M$ for all i, i', and $\Xi(0, 0) = s_M$.

Fix i' and $i < N - 1$, and let $v_y = o + (y, f_{i'})$ for $f_i \le y \le f_{i+1}$. These sites lie in $o + Z$, and consecutive ones form channel edges. For $f_i < y < f_{i+1}$ the palette at $(y, f_{i'})$ is horizontal, so $(m_{v_y})_{\mathrm{W}} = (m_{v_y})_{\mathrm{E}}$. By Lemma 11.20(b), $(m_{v_y})_{\mathrm{E}} = (m_{v_{y+1}})_{\mathrm{W}}$ for $f_i \le y < f_{i+1}$. Chaining these equalities gives $\Xi(i, i')_{\mathrm{E}} = \Xi(i + 1, i')_{\mathrm{W}}$. In the same way, along the column f_i through vertical palettes, $\Xi(i, i')_{\mathrm{N}} = \Xi(i, i' + 1)_{\mathrm{S}}$.

So Ξ is a legal T_M-pattern on $[0, 2T + 10)^2$ with the seed tile s_M at its south-west corner $(0, 0)$. By Proposition 8.16(c), applied with the halting time $T = T(M)$, no such pattern exists. This contradiction shows that no nonempty plane-safe set of 3-token states exists, and by Theorem 4.25(i) R_M is 3-token refutable. □

11.5 The three tokens

Every state used above has three tokens, at the following sites.

- Lemma 11.3: the three tokens of a given state; one of them walks onto another.
- Lemma 11.6: tokens 1 and 2 at the pins p_1, p_2, and token 3 at the site where a list is read. A step of the walker is one move of token 3.
- Lemma 11.15: first all three tokens at $\iota(0)$. Then tokens $1, 2$ at $\iota(0)$ and token 3 at $\iota(y)$, after which token 3 is cloned onto tokens 1 and 2. The same happens with $\iota(y_{\mathrm{b}})$. Finally tokens $1, 2$ at $\iota(y_{\mathrm{b}})$ and token 3 at $\iota(z')$, after which token 3 is cloned onto token 2.
- Lemma 11.16: tokens $1, 2$ at the old pins and token 3 at $\iota(y')$; then token 3 is cloned onto the token of the pin that is dropped.
- Lemmas 11.18 to 11.20: tokens $1, 2$ at the corner pair $\mathbf{p}$ of order $2n$, and token 3 walking inside the board $o + [0, 2 \cdot 4^n)^2$.

Section 11.2, Lemma 11.12 and Corollary 11.13 concern list families and finite lists, and use no tokens.

A fourth token is not needed, for two reasons. First, a step between neighbouring sites is a single move of the walker. The move rule compares the old and the new symbol of that one token, through the overlap of their blocks (Lemma 11.2). This makes the lists arc consistent (Lemma 11.6(c)), and it supplies the pairs of Q_{vw} in Lemma 11.20. In the pebble families of Section 9, the same step keeps both endpoints together with the pins. Second, the scale is carried by the two pins alone. A corner pair of order m is a property of two letters and of their

displacement, so when Lemma 11.16 drops a pin, nothing that the next step uses is lost.

11.6 Dimension eight

Corollary 11.21 (The machine-family theorem on $\mathbb{Z}^8$)**.** *Put $j = 3$, $k = \max(2, 2j) = 6$, $q = k + 1 = 7$, $a = 2q^2 + 1 = 99$, $D_0 = 50q = 350$ and*

$$\varepsilon_0 = (4D_0^2)^{-a} = (490000)^{-99}.$$

For every $\varepsilon \in (0, \varepsilon_0]$ and every Turing machine M, the noisy quorum rule $F_{R_M}^{[6]}$ on $\mathbb{Z}^8$, with uniform replacement at rate ε, is a positive-rate probabilistic cellular automaton, and it is ergodic if and only if M halts. The map $M \mapsto F_{R_M}^{[6]}$ is total computable.

Proof. The arithmetic is $2 \cdot 7^2 + 1 = 99$, $50 \cdot 7 = 350$, $350^2 = 122500$ and $4 \cdot 122500 = 490000$; the dimension is $2 + k = 8$. The process is a positive-rate probabilistic cellular automaton by Definition 4.8, and $M \mapsto F_{R_M}^{[6]}$ is total computable as in the proof of Theorem 10.2. Let $\varepsilon \in (0, \varepsilon_0]$. Then $0 < \varepsilon < 1$ and, as in Remark 5.2, $\beta = 2D_0^2 \varepsilon^{1/a} \le 2D_0^2 (4D_0^2)^{-1} = \frac{1}{2}$.

Suppose that M halts. By Theorem 11.1, R_M is 3-token refutable with some time τ. The hypotheses of Theorem 5.1 hold with $j = 3$: $k = 6 \ge \max(2, 2j)$, the constants q, a and D_0 of that theorem are the ones above, and $\beta < 1$. So for every window O and every $t \ge 0$,

$$\delta_O^M(t) \le 2|O|\, \varepsilon^{-\tau/99} \frac{\beta^t}{1 - \beta} \le 4|O|\, \varepsilon^{-\tau/99}\, 2^{-t},$$

the process forgets, and it is ergodic by Lemma 4.12.

Suppose that M does not halt. By Proposition 8.18, R_M tiles. The hypotheses of Theorem 6.2 hold: $k = 6 \ge 2$, and $0 < \varepsilon \le (4D_0^2)^{-a}$ with the constants q, a and D_0 of that theorem. As in the proof of Theorem 10.2, the process has memory, with a cloned tiling configuration and the all-bottom configuration as the two starts, and it is not ergodic by Lemma 4.13. □

Remark 11.22. The proof of Lemma 10.1 uses only $q \ge 3$, $D_0 \ge 150$ and $a \ge 19$, which hold for $q = 7$, $D_0 = 350$ and $a = 99$. So the proofs of Corollaries 10.4 to 10.8 apply with $k = 6$ and $\varepsilon_0 = (490000)^{-99}$, with Theorem 11.1 in place of Theorem 9.1 and Corollary 11.21 in place of Theorem 10.2. In Corollary 10.7, $\tau(R_M)$ is then the refutation time for 3 tokens, and in Corollary 10.8 the space–time lattice is $\mathbb{Z}^9$.

12 Lower dimensions

The natural raw-Wang rule can be realized on $\mathbb{Z}^4$, and a transverse-quorum lift of Kari's nilpotency reduction gives fixed-rate undecidability on $\mathbb{Z}^3$. The original classification statements below are supplemented by the exponential all-input estimate of Theorem 12.12 at the end of this section. Corollary 12.14 then gives all-Gibbs uniqueness and exponential weak spatial mixing for the normalized space–time interactions on $\mathbb{Z}^5$ and $\mathbb{Z}^4$. The qualitative $\mathbb{Z}^4$ classification has its own Lean core (Section C) on the version 2.2.0 interval; the enlarged interval is proved on paper with exact arithmetic; the $\mathbb{Z}^3$ classification, the rates and the Gibbs corollary are not formalized. The robustness theorem and the scope of the $\mathbb{Z}^6$ Lean core remain separate. The change is to pair provider offsets only when neither of the two vertices has a marked parent, and to control all backward paths at once at the other steps.

12.1 Why the budget-only estimate stops at six

The obstruction concerns a specified auxiliary cover and exclusion certificate, not every cellular automaton in a given dimension. The convex geometry underlying it is the classical erosion geometry of Toom's method; compare [47, 46]. We prove the needed form here.

Theorem 12.1 (Saturation of weighted covers)**.** *Let $k \geq 0$, let $a_1, \dots, a_c$ be distinct points of $\mathbb{Z}^k$, where $c \geq 1$, and give them real weights $w_i > 0$. Put $q = \sum_i w_i$. Let $W \geq 0$ and let $F : 2^{\{1,\dots,c\}} \to \{0,1\}$ be monotone, with*

$$F(A) = 0 \implies \sum_{i \in A} w_i \leq W.$$

If $q > (k+1)W$ and $F(\varnothing) = 0$, there are nonempty finite sets $E_t \subset \mathbb{Z}^k$ for all $t \in \mathbb{Z}$, an integer $L \geq 1$ and $d \in \mathbb{Z}^k$ such that

$$E_{t+L} = E_t - d, \qquad F(\{i : x + a_i \in E_{t-1}\}) = 1 \quad (x \in E_t).$$

There is thus a finite periodically translating immortal nucleus. Set $J = |E_0| \geq 1$. For every $0 < e \leq 1$, start $M_0 = \varnothing$ and, for integers $t \geq 1$, set

$$x \in M_t \iff \eta(x,t) = 1 \text{ or } F(\{i : x + a_i \in M_{t-1}\}) = 1,$$

where the η are independent Bernoulli(e) variables. For every x and $T \geq 1$,

$$\mathbb{P}(x \notin M_T) \leq (1 - e^J)^T. \tag{43}$$

The least two-sided closure is full almost surely. If $F(\varnothing) = 1$, fullness is immediate. The strict inequality on q is sharp for the universal finite-nucleus conclusion.

Proof. Call $Z \subset \{1, \dots, c\}$ zero-forcing if $F(Z^c) = 0$. There is at least one such set, the full set, and none is empty because $q > W$. For any $r \leq k+1$ zero-forcing sets, the union of their complements has weight at most $rW < q$. Hence those r sets share an index. The convex hulls

$$C_Z = \operatorname{conv}\{a_i : i \in Z\}$$

therefore intersect in every subfamily of at most $k+1$ members.

Here is the finite Helly argument needed to conclude that all C_Z intersect. If a finite convex family has empty intersection, choose a minimal empty subfamily $C_1, \dots, C_n$. If every at

most $k+1$ members intersect, then $n \geq k+2$. Choose $x_i \in \bigcap_{j \neq i} C_j$. Among $k+2$ of the vectors $(x_i, 1)$ in $\mathbb{R}^{k+1}$ there is a nonzero linear dependence. Its coefficients sum to zero, so its positive and negative parts, normalized, give a point in the convex hulls of two disjoint nonempty groups of the x_i. Each C_j contains at least one whole group, and hence that point, a contradiction. In dimension zero the assertion is immediate.

Choose a common point z of the C_Z with rational barycentric coefficients. Such a choice exists: the equations specifying a common point and the nonnegative barycentric coefficients form a feasible rational linear system. At a real feasible point retain the inequalities that are equalities there. Gaussian elimination parametrizes their rational affine solution space over the rationals; rational parameters sufficiently close to the given ones preserve all remaining strict inequalities. This supplies a rational feasible point, even though the weights themselves need not be rational.

List the zero-forcing sets as $Z_1, \ldots, Z_m$. Choose a common denominator $L \geq 1$ for coefficients $\lambda_{j,i} \geq 0$, supported in Z_j, with

$$\sum_i \lambda_{j,i} = 1, \qquad \sum_i \lambda_{j,i} a_i = z.$$

For each j, form a periodic list $(b_r^j)_{r \in \mathbb{Z}}$ of period L using a_i exactly $L\lambda_{j,i}$ times per period. All period sums equal $d = Lz \in \mathbb{Z}^k$. Define $A_0^j = 0$ and $A_{r+1}^j - A_r^j = b_r^j$ for every integer r. Thus $A_{r+L}^j = A_r^j + d$. Set

$$E_t = \left\{ \sum_{j=1}^m A_{r_j}^j : \sum_j r_j = -t, \quad 0 \leq r_j < L \ (j < m) \right\}.$$

The last index is determined by the others, so $1 \leq |E_t| \leq L^{m-1}$, including $m = 1$. Also $E_{t+L} = E_t - d$.

Fix a representation of $x \in E_t$. Incrementing any r_j supplies the parent $x + b_{r_j}^j \in E_{t-1}$. If $j < m$ wraps from $L-1$ to L, subtract L from that index and add L to the last one; the spatial sum is unchanged. The active provider indices at x therefore meet every Z_j. If their set A had $F(A) = 0$, then A^c would be a zero-forcing set disjoint from A, a contradiction. This proves the nucleus property. Colliding representations only reduce the physical set size.

Fix x, T. Choose $u_n \in E_n$ for $n \geq 0$. At each source time $s = 1, \ldots, T$, require flags on every site of $x - u_{T-s} + E_0$. This event has probability e^J and, by the nucleus property and monotonicity, forces $x \in M_T$. These events use disjoint time slices and are independent. Their simultaneous failure bounds the omission probability by (43). Sending the starting time to $-\infty$ gives zero omission probability for any fixed space–time vertex; countability gives simultaneous fullness. If $F(\varnothing) = 1$, monotonicity makes F identically one.

For sharpness, for every $k \geq 1$ and $W > 0$ use providers $0, e_1, \ldots, e_k$, each of weight W, and activate on two or more providers. Then $q = (k+1)W$ and every finite deterministic island erodes. To see this, put $\ell_0(x) = \sum_i x_i$, $\ell_i(x) = -x_i$, and

$$\phi_0(x,t) = \ell_0(x) + \frac{k}{k+1}t, \qquad \phi_i(x,t) = \ell_i(x) - \frac{1}{k+1}t \quad (1 \leq i \leq k).$$

They sum to zero. From any two supporting providers, charge zero can choose a nonzero one and charge i can choose one other than e_i, gaining at least $1/(k+1)$ per backward step. If a marked site survives to time t from a finite nonempty initial set A, follow such a path for each charge to time zero and sum the gains. Then

$$t \leq \sum_{i=0}^k \max_{x \in A} \ell_i(x) < \infty.$$

No marked site survives beyond this bound. Empty A is trivial. For $k = 0$, one provider with weight $W > 0$ and F identically zero gives an equality example that erodes immediately. Thus equality cannot be substituted in the universal assertion. □

Corollary 12.2 (Scope of the budget obstruction)**.** *Consider a planar checker and pairwise disjoint transverse provider clusters in $\mathbb{Z}^k$, with positive cluster masses summing to q. Let its unflagged cluster rule be a monotone Boolean F satisfying $F(A) = 0 \Rightarrow \sum_{i\in A} w_i \le W$ for every cluster subset A. Suppose its quorum-coherence certificate requires*

$$q - h \ge W, \qquad q > 2[(q-h) + W]. \tag{44}$$

If $k \le 3$, that marked field saturates at every positive independent flag rate. Thus this certificate cannot supply a sparse cover in total spatial dimension $2 + k \le 5$.

Proof. The two inequalities imply $q > 4W \ge (k+1)W$. Choose an integer representative from each cluster and the zero planar offset of the checker. The induced monotone rule on these parents has $F(A) = 0$ only when the contaminated mass is at most W; otherwise the asserted unmarked guarantee fails. Apply Theorem 12.1. Its finite nucleus is also supported by the full marked rule, so fullness transfers to that cover. □

This says nothing about the actual noisy automaton's ergodicity. Correlations, other pairings, realizable-error restrictions or another cover can evade (44). The following construction does so at the equality $q = 3W$ in transverse dimension two.

12.2 Three boxes and an all-path density lemma

All constants in the rest of this section are local to the lower-dimensional constructions; they do not replace the constants of Section 2. Put

$$\begin{gathered} K = 16^2 = 256, \qquad q = 3K = 768, \qquad h = 2K = 512, \\ S_0 = \{0,\dots,15\}^2, \quad S_1 = 160e_1 + S_0, \quad S_2 = 160e_2 + S_0, \quad S = S_0 \cup S_1 \cup S_2, \end{gathered} \tag{45}$$

$$D = 350, \qquad \delta = \frac{1}{4(D+1)} = \frac{1}{1404}, \qquad \alpha_4 = \frac{96}{13}, \qquad e_0 = 10^{-30}. \tag{46}$$

Here $D = \operatorname{diam}_{\ell^1}(S)$: the coordinate bounds give at most 350, attained between $(175, 0)$ and $(0, 175)$. (Version 2.2.0 used $e_0 = 382205952^{-33720}$, obtained from $B = 2(18q)^2$; the Lean statement of Section C keeps that value.)

For the common noise argument take a logical coordinate $u \in \mathbb{Z}^r$ and a finite nonempty logical parent set $J \subseteq \{-1,0,1\}^r$. An ordinary parent of (u, P, t) is

$$(u + j, P + a, t - 1), \qquad j \in J,\ a \in S.$$

We will use $r = 2$, $J = \{-1,0,1\}^2$ for the planar checker and $r = 1$, $J = \{0,1\}$ for the generator; for every $J \subseteq \{-1,0,1\}^r$, $r \in \{1,2\}$, Lemma 2.3A(b) bounds the spheres of the undirected ordinary graph, and we write $Z(y) \le \widehat{Z}_4(y)$ for its generating function. (Version 2.2.0 allowed any J with $|J| \le 9$; the constructions here, and the reflected generator $J = \{-1,0\}$, satisfy $J \subseteq \{-1,0,1\}^r$.) Fix any $0 < e \le e_0$ and independent Bernoulli(e) flags η at every integer time. Let M be the least set containing the flags and closed whenever a vertex has two marked ordinary parents in distinct clusters S_i, S_j. Let

$$H(v) \iff v \text{ has some ordinary parent in } M.$$

The recurrence for M still requires two distinct clusters; H is only an auxiliary event.

Lemma 12.3 (Finite contours and sparse roots)**.** *Every mark has a finite derivation DAG ending at flags. An unmarked vertex is unflagged and has marked parents in at most one cluster, so at most K provider offsets have a marked input. Every marked root has a connected physical witness with $f \geq 1$ distinct flagged terminals and an abstract arm realization with at most $\alpha_4(f-1)$ edge occurrences, $\alpha_4 = 96/13 < 12$. All its vertices lie between its earliest terminal time and its root time, and its physical edge count and time span are at most $\alpha_4(f-1)$. A flagged root can use $f = 1$ and no edges. Moreover, for every vertex v and every $0 < e \leq e_0$,*

$$\mathbb{P}(v \in M) \leq e\Big(1 + \widehat{Z}_4(\tfrac{1}{2})\frac{\lambda}{1-\lambda}\Big) < 2e, \qquad \lambda = 2^{15} e\, \widehat{Z}_4(\tfrac{1}{2}) < 10^{-17}. \tag{47}$$

Proof. Iterate the finitary two-parent closure starting from all flags. A supporting pair in the union belongs to some common finite rank, so the union is closed and least. Each finite-rank mark has a finite derivation; choosing smaller-rank supports gives a finite acyclic graph. This is also the union of closures started empty at arbitrarily early times, by the finite earliest time of each derivation. The exclusion budget is exactly the argument of Lemma 2.2, with one-cluster size K.

The charge-arm construction of Lemma 2.3 transfers with three balanced potentials

$$\psi_0(P,t) = 3(P_1 + P_2) + 350t, \qquad \psi_i(P,t) = -3P_i - 175t \quad (i = 1,2).$$

For charge zero a backward step in S_1 or S_2 gains between 130 and 220; for charge i a step outside S_i gains between 130 and 175. Every possible gain is at least -350. Each distinct-cluster support pair offers a gain of at least 130 to each charge.

For clarity, the merged-support accounting is as follows. Give frontier vertices their descendant physical terminal-name sets and join them when those sets overlap. Exposing a vertex replaces its set by the union of its two support sets; it cannot join an outside component and can split its component into at most two, since each remaining overlap branch attaches to one support. Begin one good arm of each charge at the root, track its eventual terminal name, and maintain one endpoint name of each charge per overlap component. At a split add a source copy at the exposed vertex; for each charge send an arm to the daughter missing its old endpoint, then follow good supports to a terminal. Only the initial edge of such a new arm may be bad. At the terminal frontier each physical terminal name has one endpoint of each charge. Copy arm interiors, identify each source's three starts and each terminal name's endpoints, and retain the component of the original root source. Each charge gives a bijection of sources and terminals there. If there are f terminals, there are f sources, of which $f-1$ are new. The facts (A1)–(A9) listed in the proof of Lemma 2.3 hold here with three charges and three clusters (the logical offset $j \in J$ plays the role of $d \in \Lambda$ in (A6)). Only the initial edge of a new arm can be bad. At a split with selected supports in clusters $S_i \neq S_j$, the initial edge of each charge goes to one of the two supports; for the one charge outside $\{i,j\}$ both are good and gain at least 130, and each of the two other initial edges gains at least -350. So the three initial edges at a new source gain at least $130 - 700 = -570$, and every other edge occurrence gains at least 130. Balance at every source and terminal yields, for the edge count E with arm multiplicity,

$$0 \;\geq\; 130\big(E - 3(f-1)\big) - 570(f-1), \qquad E \leq \frac{96}{13}(f-1) = \alpha_4(f-1).$$

Projection only reduces the edge count. As in the span paragraph of Lemma 2.3, all sources and arms are at or below the root time, and a minimum-time vertex of the abstract arm graph must be a terminal. Thus every contour vertex lies in the closed interval from the earliest

terminal time to the root time; connectedness bounds the span by E. This time bracket is needed below.

For (47), a flagged root costs e. An unflagged marked root has a witness with $f \geq 2$ flags and at most $\alpha_4(f-1)$ edges; by Lemma 2.3C its flags lie on a closed walk from v of length at most $2\alpha_4(f-1) \leq 15(f-1)$. Lemma 2.3B with $\Phi = \{2,3,\dots\}$, $L(f) = 15(f-1)$, $y = 1/2$ and Lemma 2.3A(b) give

$$\mathbb{P}(v \in M) \leq e + \sum_{f\geq 2} e^f 2^{15(f-1)} \widehat{Z}_4(\tfrac{1}{2})^f = e\Big(1 + \widehat{Z}_4(\tfrac{1}{2})\frac{\lambda}{1-\lambda}\Big).$$

Here $\widehat{Z}_4(\frac{1}{2}) = 265\,090\,000$ exactly, so for $e \leq 10^{-30}$ we have $\lambda \leq 2^{15} \cdot 265\,090\,000 \cdot 10^{-30} < 10^{-17}$ and $\widehat{Z}_4(\frac{1}{2})\lambda/(1-\lambda) < 1$, which is (47). This counts distinct physical flags even when arms or support branches merge; no conditional independence of selected paths is used.

□

Lemma 12.4 (All-path marked-parent density)**.** *For each root v and integer $N \geq 1$, let $A_N(v)$ be the event that some ordinary backward path of N edges from v has H at at least δN of its initial N vertices. For every $0 < y < 1$ with $\mu := e\,\widehat{Z}_4(y)\,y^{-15} < 1$,*

$$\mathbb{P}(A_N(v)) \;\leq\; y^{-N} \sum_{f \geq \lceil N/10368 \rceil} \mu^f \;\leq\; \frac{\rho^N}{1-\mu}, \qquad \rho := \big(e\,\widehat{Z}_4(y)\,y^{-10383}\big)^{1/10368}. \tag{48}$$

In particular, $\mathbb{P}(A_N(v)) \leq 2\exp(-N/6000)$ for every $0 < e \leq e_0 = 10^{-30}$, and $\mathbb{P}(A_N(v)) \leq \frac{16}{15}\exp(-N) \leq \frac{16}{3}\exp(-N)$ for every $0 < e \leq 10^{-4535}$. Almost surely, simultaneously at every root v, there is a finite $N_0(v)$ such that for every $N \geq N_0(v)$ and every ordinary backward N-edge path from v, fewer than δN of those vertices have H. In particular there is no infinite all-M ordinary backward path.

Proof. Given a path in $A_N(v)$, select its latest remaining H occurrence w_j. Attach a spur to one marked parent p_j and a full contour W_j of Lemma 12.3 at that parent, with f_j flags. If the contour's earliest terminal time is τ, skip all remaining occurrences whose parent time is at least τ, then repeat. Retain the entire final contour even if it extends below the path endpoint. By the contour time bracket, successive closed terminal slabs are disjoint. Their selected flags are therefore distinct.

A selected contour with f_j terminals has span at most $\alpha_4(f_j - 1)$ and accounts for at most $\alpha_4(f_j-1)+1 \leq \alpha_4 f_j$ path time positions. Thus $\delta N \leq \alpha_4 f$ for $f = \sum_j f_j$, that is, $f \geq \lceil \delta N/\alpha_4 \rceil = \lceil N/10368 \rceil$, since $\alpha_4/\delta = \frac{96}{13} \cdot 1404 = 10368$.

Form a walk from v: follow the path; at each selected w_j step to p_j, perform the closed tour of a spanning tree of W_j from p_j (length at most $2\alpha_4(f_j-1)$ by Lemma 2.3C), step back to w_j and continue. Its length is at most $N + \sum_j\big(2 + 2\alpha_4(f_j-1)\big) \leq N + 15f$, because $2\alpha_4 = 192/13 < 15$, and it visits all f distinct flags. Lemma 2.3C, Lemma 2.3B with $\Phi = \{f : f \geq \lceil N/10368 \rceil\}$, $L(f) = N + 15f$ and Lemma 2.3A(b) give

$$\mathbb{P}(A_N(v)) \leq \sum_{f \geq \lceil N/10368 \rceil} e^f y^{-N-15f} \widehat{Z}_4(y)^f = y^{-N} \sum_{f \geq \lceil N/10368 \rceil} \mu^f.$$

Since $\mu < 1$, the last sum is at most $\mu^{N/10368}/(1-\mu)$, and $y^{-N}\mu^{N/10368} = \big(\mu\, y^{-10368}\big)^{N/10368} = \rho^N$. This proves (48). The path itself is never encoded; δ enters only through $\alpha_4/\delta = 10368$.

For $e \leq 10^{-30}$ take $y = 99942/100000$. Then $\widehat{Z}_4(y) \leq 3.857 \cdot 10^{26}$, $y^{-10383} \leq 413.19$ and $y^{-15} \leq 1.01$, so $e\,\widehat{Z}_4(y) y^{-10383} \leq 0.1594$ and $\mu \leq 1/2$. To bound $\exp(1)$, put $S =$

$\sum_{j=0}^{16} 1/j!$. Its factorial-series tail is at most $(1/17!)/(1-1/18)$, so $\exp(1) < S + (1/17!)/(1-1/18) < 27183/10000$. Since $0.1594^{6000} \cdot 2.7183^{10368} < 1$ and $2.7183 > \exp(1)$, we get $\rho < \exp(-1/6000)$, hence $\mathbb{P}(A_N(v)) \le 2\exp(-N/6000)$. For $e \le 10^{-4535}$ take $y = 999422/10^6$. The exact-arithmetic certificate in `verification-thresholds/` gives

$$e\,\widehat{Z}_4(y)y^{-10383} \le (10000/27183)^{10368},$$

so $\rho \le \exp(-1)$, and $\mu \le 1/16$; hence $\mathbb{P}(A_N(v)) \le \frac{16}{15}\exp(-N)$. All these bounds increase with e. The bound is summable in N because $\rho < 1$. Borel–Cantelli and countably many roots give the asserted event. It quantifies over all paths, even paths chosen using all innovations and noncausally selected symbol histories. On an infinite all-M path each vertex has its next marked vertex as a parent and thus has H, contradicting that event. □

The coefficient $\alpha_4 = 96/13$ replaces the rounded 12 of version 2.2.0 (any bound below 12 suffices for every qualitative use, and the arm count "at most $12(f-1)$ edges" remains true). $\delta = 1/[4(D+1)]$ leaves a factor-two margin for two paths. In particular, with $B_0 = 382205952$, the old condition $0 < e \le B_0^{-33720}$ implies $e \le 10^{-4535}$ and hence the retained bound $\mathbb{P}(A_N(v)) \le (16/3)\exp(-N)$. This holds for both the checker and the reflected generator $J = \{-1, 0\}$.

12.3 The natural rule on four spatial dimensions

For every finite raw Wang set R, let $A_R = R \sqcup \{\perp\}$, let $u \in \mathbb{Z}^2$, $P \in \mathbb{Z}^2$, and put $\Gamma = \{0, \pm e_1, \pm e_2\}$ and $\Lambda = \{-1, 0, 1\}^2$. The cardinal checker C_R keeps its central tile precisely when all five tiles of its cardinal star are live and their four shared edges match; otherwise it returns $\perp$. For each $a \in S$, the provider at (u, P) is $C_R(x(\cdot, P + a))(u)$. Define $G_R^{(4)}$ to output the live tile occurring at least h times among these q providers, or $\perp$ if no tile qualifies. Since $2h > q$, this is unambiguous. Its actual neighbourhood is $\Gamma \times S$, independent of R.

Theorem 12.5 (Natural raw-Wang classification on $\mathbb{Z}^4$)**.** *For every finite Wang set R and every $0 < e \le e_0$, independent uniform whole-symbol replacement with probability e after $G_R^{(4)}$ is ergodic if and only if R does not tile $\mathbb{Z}^2$. If R tiles, there are two distinct invariant probability laws. If R does not tile, almost surely every finite observation window couples from the past in finite time, simultaneously over all initial configurations; every initial law converges weakly to the unique invariant law. All constants and the neighbourhood are common to every R.*

Use the common marked field with logical parents $J = \Lambda$. At every flagged vertex use the same replacement value in A_R for all starts. A complete history means a configuration at every integer time satisfying every prescribed noisy update. Fix the probability-one flag event of Lemma 12.4; all following deterministic assertions hold for every choice of replacement values and every complete history on it. Call a provider clean if all its Λ inputs are unmarked. An unmarked live output is unflagged, has at least h successful providers, and loses at most K to contamination, leaving at least $h - K = K$ clean successful providers.

Lemma 12.6 (Amortized same-site coherence)**.** *In every complete history, at every u, t, all unmarked live readings over all $P \in \mathbb{Z}^2$ carry the same tile.*

Proof. Suppose two such readings at P, Q carry different tiles s, s'. Their distance $d = \|Q - P\|_1$ is positive. If neither current vertex has H, all provider inputs are unmarked. Choose a coordinate and sign σ with which to reduce $Q - P$. Pair offsets $a, a' = a - \sigma e_j$ inside each box.

Both projections are injective, and there are $3 \cdot 15 \cdot 16 = 720$ pairs. Each root excludes at most $q - h = K$ failed providers. Since $720 > 2K = 512$, a pair of successful providers remains. Their clean central parents carry s, s' at the same u, one time earlier, with distance exactly $d - 1$.

If at least one current vertex has H, choose any clean successful provider at each root. Their central parents again carry the respective different tiles, and the distance grows by at most D. These two paths are ordinary and remain live and unmarked. If a chosen parent distance were zero, the different readings would already contradict equality at one vertex. Otherwise continue indefinitely.

After N steps let b_N count steps with at least one H root. The positive distances satisfy

$$1 \le d_N \le d_0 - N + (D+1)b_N. \tag{49}$$

For N at least both roots' N_0, Lemma 12.4 bounds each path's H count by δN, so $b_N < 2\delta N$. Hence $d_N < d_0 - N/2$, impossible for $N \ge 2d_0$. This proves the assertion simultaneously for every complete history on the same flag event. □

Lemma 12.7 (Planar edge matching)**.** *In any complete history, two unmarked live readings at cardinal-neighbour planar sites u, v at the same time, with arbitrary transverse positions, match on their shared edge.*

Proof. A clean successful provider for the root at u supplies one time earlier an unmarked live checker star, whose central tile is the root's tile and whose neighbour at v matches it. A clean successful provider for the root at v supplies an unmarked live central tile at v one time earlier equal to that root's tile. Apply Lemma 12.6 at v and that earlier time to identify the two v readings. This is a separate edge assertion; same-site coherence alone does not assert it. □

Lemma 12.8 (Planar extraction)**.** *If R does not tile, every live symbol of every complete history belongs to M.*

Proof. Let V_t be the planar sites admitting an unmarked live reading at time t. Give each its common tile from Lemma 12.6. These labels are locally compatible by Lemma 12.7. A clean successful checker supplies the whole cardinal star one time earlier, so $V_t + \Gamma \subset V_{t-1}$. If $u \in V_t$, iteration gives a legal patch on $u + \{z : \|z\|_1 \le n\}$ at time $t - n$ for every $n \ge 0$. Such patches contain squares of unbounded side. Recenter them and use compactness of the product of the finite tile alphabet to obtain a legal full-plane tiling. Temporal consistency of the patches is unnecessary. This contradicts non-tiling, so every V_t is empty. □

We give the full compactness passage once, to reuse it for the three-dimensional rule.

Lemma 12.9 (Complete histories and all-start forgetting)**.** *For either finite-alphabet rule in this section, with its ordinary graph containing every actual dependency, complete histories exist for every fixed two-sided innovation field. If, on the event of Lemma 12.4, every live reading of every complete history lies in M, the noisy rule is ergodic and has almost-sure finite-window coupling from the past uniformly over all starts.*

Proof. The space of all symbol histories is a product of finite discrete spaces and is compact. Each local update constraint is closed. Any finite set of constraints can be satisfied by starting before its earliest time and evolving the global rule. The finite intersection property gives a complete history.

If two complete histories for the same innovations disagree at a vertex, at least one reading is live, so that vertex lies in M. A flagged vertex receives the same replacement value and

cannot be discrepant. At an unflagged discrepancy at least one actual parent is discrepant, because the deterministic rule is a function. Repeating gives an infinite ordinary discrepancy path, all in M, prohibited by Lemma 12.4. Thus the complete history is unique.

Fix a time t. The image E_n at time t of all row configurations at time $t-n$ is nonempty and compact, since each fixed finite-range noisy map is continuous. These images are nested. Every row in $\bigcap_n E_n$ extends to a complete past by compactness of compatible finite histories, and its future is determined. Their intersection is therefore the singleton row of the unique complete history. For any finite window F, projection commutes with this nested compact intersection: if a projected value occurs in every E_n, intersect each E_n with its closed cylinder and apply compactness. The finite nested sets $\pi_F(E_n)$ thus have singleton intersection and eventually are singletons. This is the asserted uniform all-start coupling from the past.

The ordered length-n past maps have the same law as the ordered length-n future maps, by the iid innovations in time. The probability $p_n(F)$ of a nonconstant future image on F consequently tends to zero. The finite-range noisy kernel is Feller on compact configuration space. Cesàro averages of any initial law have an invariant subsequential limit: on each continuous test the difference between its average after one transition and before it telescopes and tends to zero. Couple any initial law with one such invariant law using the same future innovations. Their F-marginal total variation distance is at most $p_n(F)$. It tends to zero for every finite F, proving weak convergence from every initial law and uniqueness of the invariant law. □

Proof of Theorem 12.5. Non-tiling gives the hypothesis of Lemma 12.9 by Lemma 12.8. For the converse, let θ be a tiling and clone it transversely. It and the all-$\perp$ configuration are fixed by the noiseless rule. Start the noisy processes X from the clone and Y from all $\perp$. Let $M^{[0]}$ be the finite-past marked closure started empty at time zero; it is contained in M. Induction covers all errors of X relative to θ and all live sites of Y by $M^{[0]}$. Indeed, flags are covered. At an unflagged wrong X output there must be at least $q-h+1=K+1$ providers with erroneous checker inputs, occupying two distinct clusters and supplying two marked ordinary parents. At an unflagged live Y output at least $h=2K>K$ providers are live, again in two clusters, each supplying a marked live central parent. The closure therefore covers both kinds of error.

At a fixed spatial site with planar coordinate u, Equation (47) gives, for every time $n \geq 0$,

$$\mathbb{P}(X_n = \theta(u)) \geq 1-2e, \qquad \mathbb{P}(Y_n = \theta(u)) \leq 2e.$$

Temporal Cesàro subsequential limits of the two sequences of laws are invariant by the Feller telescoping argument in Lemma 12.9. The fixed cylinder event $\{x(u,P)=\theta(u)\}$ is clopen, so the gap at least $1-4e>0$ persists in those limits. They are distinct invariant laws. If $R=\varnothing$, the alphabet is a singleton and ergodicity is immediate. Finally $R \mapsto G_R^{(4)}$ is a finite algorithm with one common neighbourhood, so the raw domino undecidability theorem [3, 43] implies fixed-rate undecidability at $e=e_0$ on this family. □

12.4 Three dimensions via Kari's two-input generator

Call a finite Wang set T north/east deterministic if, for all $s,t \in T$,

$$s_N = t_N \text{ and } s_E = t_E \quad \Longrightarrow \quad s=t. \tag{50}$$

The imported classical theorem is undecidability of unseeded full-plane tiling for finite north/west deterministic sets, with no rotations of individual copies: [26], as explicitly stated

in [51, Section 2.8, Theorem 5]. The global design reflection

$$\rho(N,E,S,W) = (N,W,S,E), \qquad \tau'(x,y) = \rho(\tau(-x,y))$$

turns north/west determinism into (50) and bijects the legal tilings in both directions. Thus the same decision problem is undecidable on north/east deterministic sets. This is an imported classical undecidability result, not a new tiling reduction.

For $A_T = T \sqcup \{\perp\}$ define $f_T(a,b) = c$ if $a, b \in T$ and there is a tile $c \in T$ with

$$c_N = a_S, \qquad c_E = b_W;$$

otherwise define $f_T(a,b) = \perp$. Condition (50) makes c unique. In particular either $\perp$ input forces $\perp$, and missing colour pairs are allowed. Put

$$D_T(x)(i) = f_T(x(i), x(i+1)). \tag{51}$$

Up to reflection and the corresponding orientation of the diagonal, this is exactly Kari's 1992 nilpotency cellular automaton; see [51, discussion after Theorem 15, p. 24]. The rule below is a noise-robust transverse-quorum lift of Kari's nilpotency reduction. The dynamical argument added here is amortized coherence and all-start forgetting under uniform noise.

Lemma 12.10 (Space–time geometry)**.** *The map $z(i,t) = (i, -t-i)$ is a bijection from logical space–time to the Wang plane, with*

$$z(i,t-1) = z(i,t) + (0,1), \qquad z(i+1,t-1) = z(i,t) + (1,0).$$

A live depth-n backward cone of (51) *gives a legal triangle $\{(x,y) \in \mathbb{Z}^2 : x,y \geq 0,\ x+y \leq n\}$ after translation, containing a square of side $\lfloor n/2 \rfloor + 1$. Conversely, every legal tiling τ gives a complete all-live orbit*

$$\theta_t(i) = \tau(i, -t-i), \qquad t \in \mathbb{Z}. \tag{52}$$

Proof. The displayed identities follow by substitution; the inverse is $i = x$, $t = -x-y$. A live output forces both inputs live. Iterating this fact gives every vertex of the cone. The two matching equations for f_T check every internal north/east adjacency, whose lower-left endpoint is at a level strictly before the cone boundary. At depth r the logical sites are $i, \dots, i+r$, which map to the triangle's diagonal $x+y = r$. The indicated square is contained because $2\lfloor n/2 \rfloor \leq n$. Substitution of a legal tiling in the matching equations proves (52); the orbit need not be fixed in time. □

On spatial coordinates $(i,P) \in \mathbb{Z} \times \mathbb{Z}^2$, compute at each provider $a \in S$

$$y_a = f_T(x(i, P+a), x(i+1, P+a)).$$

Define $G_T^{(3)}$ to output a live tile occurring at least h times among the y_a, and $\perp$ otherwise. This is a total computable rule because $2h > q$. Its common neighbourhood is $\{0,1\} \times S$, contained in the radius-175 cube. After each update replace the whole A_T symbol uniformly with probability e, independently in space and time. Every output letter has probability at least $e/(|T|+1) > 0$.

Theorem 12.11 (Fixed-rate undecidability on $\mathbb{Z}^3$)**.** *For every finite north/east deterministic Wang set T and every $0 < e \leq e_0$, this perturbation of $G_T^{(3)}$ is ergodic if and only if T does not tile the plane. Non-tiling gives finite-window all-start coupling from the past and attraction from every initial law. Tiling gives two distinct invariant laws. The family is total computable with one common finite neighbourhood. In particular ergodicity is undecidable across it at the single rational rate*

$$e_0 = 10^{-30}.$$

Proof. Use the common marked construction with logical parents $J = \{0,1\}$. Its logical parent set $J = \{0,1\}$ is covered by Lemma 2.3A(b) with $m = 1$. The potentials ignore the logical coordinate; hence Lemmas 12.3 and 12.4 apply with precisely the constants (46). At an unmarked live output at least K successful providers are clean, meaning both inputs are unmarked.

We prove same-site coherence in every complete history on the probability-one density event. Suppose two unmarked live readings at one logical i, t, at transverse P, Q, have different tiles. At a step where neither vertex has H, use the same 720 injective pairs as in Lemma 12.6; at most $2K = 512$ are killed by unsuccessful providers. The surviving providers have different live outputs under the *same function* f_T. Their input pairs are live, since a $\perp$ input would force $\perp$ output. Different outputs force a common index $j \in \{0,1\}$ at which the two input tiles differ. Follow that same j on both sides. These are unmarked live readings at the same logical site $i + j$ and time $t - 1$, at transverse distance $d - 1$. At a step with at least one H vertex, take arbitrary clean successful providers on both sides and again follow a common differing input index. The distance is at most $d + D$.

Difference is preserved although the tile identities and logical coordinate may change. Both ancestries are ordinary paths and are aligned at each logical site and time. A zero-distance pair is impossible. If they continue forever, (49) holds and the two all-path density bounds contradict it for N at least both N_0 values and $2d_0$. Thus all unmarked live readings at a logical site and time agree, for every complete history and every replacement-value field.

Let V_t be the logical sites with an unmarked live reading, labelled by this common tile $a_t(i)$. A clean successful provider at $i \in V_t$ supplies both unmarked live inputs, so

$$i, i+1 \in V_{t-1}, \qquad a_t(i) = f_T(a_{t-1}(i), a_{t-1}(i+1)).$$

If any V_t is nonempty, iteration gives live legal cones of every finite depth. By Lemma 12.10 these contain legal squares of unbounded side. Recenter and use finite-alphabet compactness to get a full-plane tiling. Therefore for non-tiling T every live reading of every complete history lies in M, and Lemma 12.9 proves all asserted forgetting and ergodicity conclusions.

For tiling T, choose τ and clone the time-dependent orbit (52) transversely. Every provider computes $\theta_t(i)$ from time $t - 1$, so this is an exact orbit of the noiseless quorum rule. Start X at its time-zero row and Y at all $\perp$. The finite-past error-cover induction in the proof of Theorem 12.5 applies with two-input providers: a wrong unflagged X output needs at least $K + 1$ erroneous providers in two clusters, each with an erroneous input; a live unflagged Y output needs at least $h > K$ live providers in two clusters, each with live inputs. Thus both error sets are covered by $M^{[0]} \subset M$.

At any fixed spatial root and every $n \geq 0$,

$$\mathbb{P}(X_n \text{ is live}) \geq 1 - 2e, \qquad \mathbb{P}(Y_n \text{ is live}) \leq 2e.$$

Use the fixed clopen event LIVE, not agreement with the moving tile $\theta_n(i)$. The Feller Cesàro argument of Lemma 12.9 gives invariant subsequential limits of both sequences, and their LIVE probabilities retain a gap at least $1 - 4e > 0$. They are distinct. For empty T the singleton alphabet is trivially ergodic.

Finally enumerate finite sets of colour quadruples and retain those passing the decidable pairwise test (50). This gives an effective enumeration; canonical finite colour renamings are computable, so every input has a searchable finite code. The map from such a code to the alphabet, f_T table and quorum rule is finite and total. A decision algorithm for ergodicity of these indexed rules at e_0 would decide their restricted tiling problem, contradicting Kari's

theorem and the reflection above. The rate is independent of the instance and its alphabet size. □

The two classifications above assert ergodicity and two invariant laws on the respective sides. The following quantitative supplement gives their own exponential all-input bound on the non-tiling side. Corollary 12.14 transfers all-DLR uniqueness and exponential weak spatial mixing to the associated space–time interactions. The robustness theorem and the scope of the $\mathbb{Z}^6$ Lean proofs are not transferred; Section C describes the separate qualitative $\mathbb{Z}^4$ core. No rule on every raw Wang set on $\mathbb{Z}^3$, no result on $\mathbb{Z}^2$ or $\mathbb{Z}$, and no token-width-two theorem is proved here.

12.5 Exponential all-input coalescence for non-tiling inputs

This supplement uses the geometry of (45)–(46) and the density bounds of Lemma 12.4. The noise interval, provider geometry and both local rules are those of Theorems 12.5 and 12.11.

Theorem 12.12 (Uniform exponential rate on $\mathbb{Z}^4$ and $\mathbb{Z}^3$)**.** *Let $\mathcal{R}$ be either a finite raw Wang set for $G^{(4)}_{\mathcal{R}}$ of Theorem 12.5, or a finite north/east deterministic Wang set for $G^{(3)}_{\mathcal{R}}$ of Theorem 12.11. Suppose $\mathcal{R}$ does not tile the plane. There is an effectively findable integer $C_{\mathcal{R}} \geq 1$, depending only on this input, such that for every $0 < \varepsilon \leq e_0 = 10^{-30}$, every integer $t \geq 0$ and every finite spatial observation set F,*

$$\mathbb{P}\{\textit{the } t\textit{-step shared random map from all starts is nonconstant on } F\} \leq |F|\, C_{\mathcal{R}} \exp(-t/24000),$$

and, if moreover $0 < \varepsilon \leq 10^{-4535}$, the same probability is at most $|F|\, C_{\mathcal{R}} \exp(-t/4)$. The exponents and noise intervals are common to both families; $C_{\mathcal{R}}$ is independent of ε. The prefactor search is required to terminate only on non-tiling inputs.

Proof. We give finite-history versions of the coherence and extraction arguments, then control full discrepancies. All initial rows are arbitrary. Use one common flag and replacement letter at every update, and extend the flags independently to all integer times to define the noise-only M, H, A_N of Lemmas 12.3 and 12.4. This auxiliary extension does not change any finite-time random map. No time-zero symbol is marked for free.

The quantitative density bound. Fix one of the two cases: (I) $0 < \varepsilon \leq 10^{-30}$, $\rho_* = \exp(-1/6000)$, $C_A = 2$; or (II) $0 < \varepsilon \leq 10^{-4535}$, $\rho_* = e^{-1}$, $C_A = 16/3$. Put $\gamma = \log(1/\rho_*)$, so $\gamma = 1/6000$ in case (I) and $\gamma = 1$ in case (II). By Lemma 12.4,

$$\mathbb{P}(A_N(v)) \leq C_A\, \rho_*^N, \qquad N \geq 1.$$

The distinct-terminal contour and flag-position counting have already been paid for in Lemmas 12.3–12.4; the bound includes every ordinary path, even one selected using all innovations and initial symbols.

Finite-time coherence. Consider one history begun at time zero. Suppose two unmarked live readings at the same logical site and time s differ, at transverse distance $d_0 > 0$. Follow their paired ancestry for any $N \leq s$ steps. When neither root has H, use the 720 injectively paired offsets of Lemma 12.6; at most $2(q-h) = 512$ pairs are lost, so successful providers reduce the distance by one. For the checker follow their central inputs, preserving their different tile labels. For the generator, distinct successful outputs of the same function $f_{\mathcal{R}}$ force a common index $j \in \{0,1\}$ at which the inputs differ. Follow that same index on both sides. These inputs are live and unmarked. Their labels may change, but inequality and alignment of the logical sites persist, exactly as in the proof of Theorem 12.11.

When either root has H, at least $h - K = K$ clean successful providers remain at each root by Lemma 12.3. Choose any such providers and again follow central checker inputs or a common differing generator input index. Distance grows by at most $D = 350$. All selected vertices lie in the *same* symbol history. Their distance cannot become zero, including at time zero, because one physical vertex has one symbol. If b_N counts paired steps with at least one H root, the finite version of (49) is

$$1 \le d_N \le d_0 - N + 351 b_N.$$

For $N \ge 2d_0$, this forces $b_N \ge N/702 = 2\delta N$. At least one of the two ordinary paths then has δN occurrences of H, so A_N occurs at one of its starting roots. Consequently two such readings agree whenever their distance is at most L, their common time satisfies $s \ge N \ge 2L$, and neither starting root belongs to A_N. This deterministic implication holds simultaneously over every initial row; it stops at time zero and uses no initial-row assumption.

An effective finite obstruction. If the input is empty, its alphabet is the singleton $\{\perp\}$ and the theorem is immediate, with $C_{\mathcal{R}} = 1$. Otherwise find $r \ge 1$ as follows. For the checker, search for a planar diamond

$$D_r = \{u \in \mathbb{Z}^2 : \|u\|_1 \le r\}$$

with no legal $\mathcal{R}$-labeling. For the generator, search for a triangle

$$T_r = \{(x, y) \in \mathbb{Z}^2 : x, y \ge 0,\ x + y \le r\}$$

with no legal labeling. Each search tests finitely many assignments at each stage. It terminates on a non-tiling input: if every centered diamond were legal, the finitely branching tree of legal diamond assignments, with restriction as parent, would have an infinite branch and hence a full-plane tiling. To construct the branch, repeatedly choose a child having descendants at arbitrarily large levels; one exists because there are finitely many children. For triangles, depth $2m$ contains a square of side $m + 1$, by Lemma 12.10. If all triangles were legal, recentering and restricting these squares would give legal centered diamonds of every radius, and the same tree argument would give a tiling. No bound on the running time of this search is needed.

Put $b = r + 1$ and define the effective integers

$$J_b = (b+1)(2b+1)^2(175b+1)^2, \qquad U_{\mathcal{R}} = \begin{cases} 2J_b\, 2^b & \text{in case (I),} \\ 6J_b\, 3^{701b} & \text{in case (II).} \end{cases}$$

In both cases $U_{\mathcal{R}} \ge C_A J_b \rho_*^{-b}$ and $U_{\mathcal{R}} \ge \rho_*^{-701b}$. In case (I) this holds because $\rho_*^{-701b} = \exp(701b/6000) < 2^b$, as $\exp(701/6000) \le 1/(1 - 701/6000) = 6000/5299 < 2$; in case (II) because $3 > \exp(1)$ and $6 > 16/3$.

Let $L_t(v)$ be the event that v at time t is unmarked and live in at least one history begun at time zero. The existential quantifier is over *all* initial rows, including a row chosen after seeing the noise.

Uniform unmarked-liveness tail. For $t \ge 701b$ set $N = t - b \ge 700b$. From an unmarked live root recursively choose clean successful providers through b levels. The candidate points at depth $j \le b$ have transverse positions in $P + [0, 175j]^2$ and logical positions in a coordinate box of radius j. For both families the deterministic union of these possible points contains at most J_b physical space–time vertices. At a common depth its transverse ℓ^1 span is at most $350b$; every candidate time is at least $t - b = N$. If A_N is absent at all candidates, finite-time coherence identifies every pair of unmarked live readings needed at the same logical site and time, since $N \ge 2 \cdot 350b$.

For the checker, expansion to depth r gives a live reading at each site of D_r at time $t-r$. Give it its coherent label. For adjacent sites, a clean successful provider of the first reading supplies a matching neighbor one time earlier; a clean successful provider of the second reading supplies its own central tile one time earlier. As in Lemma 12.7, coherence at that earlier level identifies the two readings at the second site. Their central labels are preserved by the checker. Thus every internal edge of D_r matches. The extra level $b = r+1$ includes every reading used for this edge check, contradicting the choice of r.

For the generator, at depth $k \le r$ identify the labels at logical positions $i, \dots, i+k$. Each selected clean successful provider gives

$$a_k(j) = f_{\mathcal{R}}(a_{k+1}(j), a_{k+1}(j+1)) \qquad (0 \le j \le k < r),$$

because coherence identifies its inputs with the chosen next-level labels. Lemma 12.10 maps $(i+j, t-k)$ to the translated triangle position $(j, k-j)$; these relations check all internal north/east edges of T_r. This also contradicts the chosen obstruction. The argument uses inequality, rather than preservation of tile identities, in generator ancestry.

Therefore $L_t(v)$ implies A_N at one of at most J_b fixed candidate points. A finite union bound gives

$$\mathbb{P}(L_t(v)) \le C_A\, J_b\, \rho_*^{t-b} \le U_{\mathcal{R}}\, \rho_*^t \qquad (t \ge 701b).$$

For smaller t the same bound follows from probability at most one and $U_{\mathcal{R}}\rho_*^t \ge U_{\mathcal{R}}\rho_*^{701b} \ge 1$. This is a uniform existential event bound; there is no union over initial configurations.

Full all-input coupling. Let $D_t(v)$ mean that the shared t-step map is nonconstant at v. Fix any two witnessing starts. At each discrepancy the common replacement flag is absent, since a replacement writes the same letter. A differing deterministic output has a differing actual parent. Trace such parents for $k = \lfloor t/2 \rfloor$ edges, where $t \ge 2$. Each of the first k vertices is discrepant and has a live reading in at least one of the two histories, since $\perp$ is the only non-live symbol.

If all first k vertices are marked, they are *nonflagged* marks and hence each has marked parents by the closure defining M. Thus each has H, and this path witnesses $A_k(v)$. This does not require its selected discrepant parent itself to be marked. Otherwise a first-k vertex w at depth $j < k$ is unmarked and live in one witnessing history, so $L_{t-j}(w)$ occurs. All candidates lie in the ordinary backward cone, whose vertices through depth k number at most

$$Q_k = (k+1)(2k+1)^2(175k+1)^2 \le 30625(t+1)^5.$$

Indeed $k+1 \le t+1$, $2k+1 \le t+1$ and $175k+1 \le 175(t+1)$; the generator cone is smaller. Since $t-j \ge t-k \ge t/2$, the preceding bounds imply

$$\mathbb{P}(D_t(v)) \le C_A\, \rho_*^k + 30625\, U_{\mathcal{R}}\, (t+1)^5 \rho_*^{t/2}.$$

Use $k = \lfloor t/2 \rfloor \ge (t-1)/2$, so $C_A\rho_*^k \le C_A\rho_*^{-1/2}\rho_*^{t/2}$, where $C_A\rho_*^{-1/2} = 2\exp(1/12000) < 4$ in case (I) and $\frac{16}{3}e^{1/2} < 16$ in case (II). Also $(t+1)^5\rho_*^{t/4} = \exp\bigl(5\log(t+1) - \gamma t/4\bigr) \le (20/\gamma)^5$ for $t \ge 0$: the exponent has derivative $5/(t+1) - \gamma/4$, so its maximum is at $t+1 = 20/\gamma$ and equals $5\log(20/\gamma) - 5 + \gamma/4 \le 5\log(20/\gamma)$. Thus $(t+1)^5\rho_*^{t/4} \le 120000^5$ in case (I) and $\le 20^5$ in case (II). With $\rho_*^{t/2} \le \rho_*^{t/4}$, the explicitly effective integers

$$\begin{aligned} C_{\mathcal{R}}^{(\mathrm{I})} &= 4 + 30625 \cdot 120000^5 \cdot 2J_b 2^b, \\ C_{\mathcal{R}}^{(\mathrm{II})} &= 16 + 30625 \cdot 20^5 \cdot 6J_b 3^{701b}, \\ C_{\mathcal{R}} &= \max\bigl(C_{\mathcal{R}}^{(\mathrm{I})}, C_{\mathcal{R}}^{(\mathrm{II})}\bigr). \end{aligned}$$

give $\mathbb{P}(D_t(v)) \le C_{\mathcal{R}}\rho_*^{t/4} = C_{\mathcal{R}}\exp(-\gamma t/4)$ for $t \ge 2$, that is $C_{\mathcal{R}}\exp(-t/24000)$ in case (I) and $C_{\mathcal{R}}\exp(-t/4)$ in case (II); for $t = 0, 1$ the bound is trivial since $C_{\mathcal{R}}\rho_*^{1/4} \ge 4e^{-1/4} > 1$.

All witness implications hold simultaneously for every pair of starts. Integrating out the auxiliary nonpositive-time flags leaves the same bound for the actual finite-time maps. Finally, nonconstancy on F entails nonconstancy at one of its coordinates, so a union bound supplies $|F|$. Past and future blocks have the same law under iid innovations. This proves the claimed all-input rate throughout the common noise interval. □

Example 12.13 (Why non-tiling is necessary)**.** Let $T = \{a\}$, with all four edge colors of a equal to zero. This is north/east deterministic and tiles constantly. Its generator has $f_T(a, a) = a$ and returns $\perp$ on every other input pair. For every $0 < \varepsilon \le e_0$, the all-input coefficient at a single site stays at least $1 - 4\varepsilon > 0$ for all time, so no finite prefactor can give an exponentially decaying all-input bound on this tileable input.

Proof. Use the same flags and replacement letters to start X from all a and Y from all $\perp$. Let $M^{[0]}$ be the marked closure begun empty at time zero. The error-cover induction from Theorem 12.11 gives

$$\{X_t \ne a\} \subset M^{[0]}, \qquad \{Y_t = a\} \subset M^{[0]}.$$

For completeness, outside $M^{[0]}$ the output is unflagged and at most $K = 256$ providers have marked inputs. The other 512 providers in X have both inputs a and reach the quorum; in Y they have $\perp$ inputs, leaving at most 256 possible live providers, below quorum. The initial error sets are empty, proving both inclusions by induction. Since $M^{[0]} \subset M$, Lemma 12.3 and (47) give at any fixed site v

$$\mathbb{P}(X_t(v) = a) \ge 1 - 2\varepsilon, \qquad \mathbb{P}(Y_t(v) = a) \le 2\varepsilon.$$

Thus $\mathbb{P}(X_t(v) \ne Y_t(v)) \ge 1 - 4\varepsilon$, and this discrepancy is contained in all-input nonconstancy. Any finite multiple of $\exp(-t/24000)$, indeed any exponentially decaying bound, tends to zero, contradicting this uniform positive lower bound. The non-tiling hypothesis in Theorem 12.12 is therefore essential even within the one-tile NE family. □

12.6 All space–time Gibbs measures in lower dimensions

Corollary 12.14 (All space–time Gibbs measures on $\mathbb{Z}^5$ and $\mathbb{Z}^4$)**.** *Let $\mathcal{R}$ and $G_{\mathcal{R}}^{(d)}$ be as in Theorem 12.12, with $d = 4$ for the checker family of Theorem 12.5 and $d = 3$ for the generator family of Theorem 12.11. Let N_d be their common neighbourhood, respectively $\Gamma \times S$ and $\{0, 1\} \times S$. Fix $0 < \varepsilon \le e_0$ and put*

$$p_{\mathcal{R},\varepsilon}(c \mid y) = (1 - \varepsilon)\mathbf{1}_{\{c = G_{\mathcal{R}}^{(d)}(y)\}} + \frac{\varepsilon}{|A_{\mathcal{R}}|}, \qquad y \in A_{\mathcal{R}}^{N_d}.$$

On $\mathbb{Z}^d \times \mathbb{Z} = \mathbb{Z}^{d+1}$ define the normalized finite-range interaction

$$\Psi_{\mathcal{R},\varepsilon,(v,t)}(\sigma) = -\log p_{\mathcal{R},\varepsilon}\big(\sigma(v, t) \mid (\sigma(v + a, t - 1))_{a \in N_d}\big).$$

Its finite-volume specification γ_V^τ is that of [9, equation (6)]: multiply all factors whose supports $\{(v, t)\} \cup \{(v + a, t - 1) : a \in N_d\}$ meet V, including factors of exterior children, and normalize with exterior τ. A Gibbs measure means a DLR state, with no invariance assumed.

(i) *If $\mathcal{R}$ does not tile the plane, this interaction has exactly one Gibbs measure $\bar{\pi}_{\mathcal{R},\varepsilon}$, the two-sided stationary trajectory law of the unique invariant law. There are $C_*, c_* > 0$, depending on $\mathcal{R}$ and ε, such that for every finite $V \subset \mathbb{Z}^{d+1}$, every $E \subset V$ and all exteriors τ, τ',*

$$d_{\mathrm{TV}}((\gamma_V^{\tau})_E, (\bar{\pi}_{\mathcal{R},\varepsilon})_E) \le C_* \sum_{z \in E} e^{-c_* \operatorname{dist}_\infty(z, V^c)},$$
$$d_{\mathrm{TV}}((\gamma_V^{\tau})_E, (\gamma_V^{\tau'})_E) \le 2C_* \sum_{z \in E} e^{-c_* \operatorname{dist}_\infty(z, V^c)}.$$

(ii) *If $\mathcal{R}$ tiles the plane, the interaction has at least two Gibbs measures invariant under every translation of $\mathbb{Z}^{d+1}$.*

Thus at every fixed $0 < \varepsilon \le e_0$, uniqueness, uniqueness among fully translation-invariant Gibbs measures, and exponential weak spatial mixing each hold exactly on non-tiling inputs. At $\varepsilon = e_0$ each is undecidable across either family. For the $\mathbb{Z}^4$ checker family these decision properties are Σ_1^0-complete, with Π_1^0-complete complements; for the $\mathbb{Z}^3$ generator family only undecidability is asserted.

Proof. Suppose $\mathcal{R} \neq \varnothing$. Its alphabet is finite with at least two symbols. The finite neighbourhood N_d contains zero: $0 \in \Gamma$ (respectively $0 \in \{0,1\}$), and $0 \in S_0$. Both neighbourhoods have sup-radius 175. Every table entry is at least $\varepsilon/|A_{\mathcal{R}}| > 0$, and rows update independently given the preceding row.

At each space–time update take an independent pair consisting of a Bernoulli(ε) flag and a uniform letter in $A_{\mathcal{R}}$, with the two entries independent. It defines the local map $m(y)$ equal to that letter when the flag is set and to $G_{\mathcal{R}}^{(d)}(y)$ otherwise. For each c, y, $\mathbb{P}(m(y) = c) = p_{\mathcal{R},\varepsilon}(c \mid y)$, so these maps are an iid local-map representation in the sense of [9, Section 2.2]. Their chronological composites on $\mathbb{Z}^d$ are exactly the shared random maps of Theorem 12.12, using one common flag and replacement letter for every input row. Hence the common-map coefficient of [9, equation (4)] is

$$q_t = \mathbb{P}\{\eta \mapsto F_t(\eta)(0) \text{ is not constant}\} \le C_{\mathcal{R}}\, e^{-t/24000} \qquad \text{for every integer } t \ge 0.$$

The map law is the same at every site, so this bound is uniform in the root. These check every hypothesis of [9, Corollary 13.7], in spatial dimension d. That corollary, through its prefix-cap Theorem 13.3, gives (i), with its equations (8) and (7), respectively. The general transfer is the result of [9]; the assertion here is its application to these two families. If $\mathcal{R} = \varnothing$, there is one space–time configuration, one DLR state, and both total-variation distances are zero; this handles the singleton alphabet separately from the cited theorem.

For (ii), take the invariant laws μ_X, μ_Y obtained as temporal Cesàro limits in the proofs of Theorems 12.5 and 12.11, using the Feller argument of Lemma 12.9. At every spatial site v, write $\mathrm{LIVE}_v = \{x : x(v) \neq \perp\}$. The error covers and the sparse-root bound of Lemma 12.3 give

$$\mu_X(\mathrm{LIVE}_v) \ge 1 - 2\varepsilon, \qquad \mu_Y(\mathrm{LIVE}_v) \le 2\varepsilon.$$

For the checker, agreement with the cloned tile implies LIVE, and every live site of the blank-started process is covered; for the generator these are the LIVE bounds already used in its proof. The blank-started law μ_Y is spatially invariant, because its start is and the dynamics commute with translations.

The clone-started law need not be spatially invariant. Every translate of μ_X is stationary, and stationary laws form a weakly closed convex set. Average its translates over boxes $[-n, n]^d \cap \mathbb{Z}^d$ and take a weak limit point $\widehat{\mu}_X$. The boundary-to-volume ratio tends to zero, so this limit is invariant under every spatial translation, and it is stationary. At every site

the LIVE lower bound survives the averages and their weak limit because LIVE is clopen. The two-sided stationary trajectories of $\widehat{\mu}_X$ and μ_Y are DLR states by [9, Lemma 2.4], which assumes no spatial invariance. They are invariant under all space–time translations. Their time-zero LIVE probabilities differ by at least $1 - 4\varepsilon > 0$, so they are distinct. This is consistent with the classical translation-invariant correspondence of [21].

Weak spatial mixing forces uniqueness: integrate its bound over the exteriors of two DLR states and let the distance of a fixed finite E to V^c tend to infinity. It therefore fails on tiling inputs. The map from a tile input to its table at e_0 is total computable. The raw domino problem [3, 43] and Kari's restricted tiling theorem with the reflection used in Theorem 12.11 [26, 51] give the two undecidability assertions. For the checker family, non-tiling is computably enumerable by finite-patch search and Σ_1^0-complete by the recalled Berger–Robinson reduction, as in Corollary 3.7; the identity classification gives the asserted degrees. No many-one completeness claim is made for the generator family. □

Compare Sandhu's independent fixed-temperature Gibbs-uniqueness result on $\mathbb{Z}^2$ [54, Theorem 1.1], described after Corollary 10.8.

Distance is the sup norm, as in [9, Definition 2.3]. Volumes may have holes and need not be connected. Distance is to the entire complement, so strong spatial mixing is not claimed. No uniformity of the Gibbs constants as $\varepsilon \downarrow 0$ is claimed. The robustness theorem and the scope of the $\mathbb{Z}^6$ Lean core are not transferred. The separate $\mathbb{Z}^4$ Lean core in Section C does not formalize this corollary.

13 Open problems

Dimension and the old token question. Fixed-rate positive-noise undecidability now holds on $\mathbb{Z}^3$ by Theorem 12.11, and the natural rule on every raw Wang set is classified on $\mathbb{Z}^4$ by Theorem 12.5. The old three-token family and its $\mathbb{Z}^8$ realization remain formally checked. One token never suffices for an undecidable tiling family (Corollary 4.27); the existence of a two-token family is still open. Such a family would give another $\mathbb{Z}^6$ route at the thin-provider rate. A forgetting mechanism using fewer than $2j$ transverse coordinates would improve the inspector route. The $\mathbb{Z}^6$ natural rule's equilibrium statements and its stack live on $\mathbb{Z}^7$; the retained machine route's corresponding dimensions are nine and nine. For the lower-dimensional rules, Corollary 12.14 gives the corresponding Gibbs dichotomies on $\mathbb{Z}^5$ and $\mathbb{Z}^4$; no lower-dimensional stacking or robustness theorem is asserted.

Open problem 13.1. What is the least dimension $d \leq 3$ for which ergodicity of positive-rate cellular automata on $\mathbb{Z}^d$ at one fixed replacement rate is undecidable? In particular, can this be done on $\mathbb{Z}^2$ or $\mathbb{Z}$? What is the least token width of a computable family with undecidable tiling problem: two or three?

The exact degree. On the $\mathbb{Z}^6$ natural family ergodicity, Gibbs uniqueness and exponential weak spatial mixing have exact degree Σ_1^0, with Π_1^0 complements (Corollary 3.7). This settles the raw-family question. For the translation-invariant variants the degree is now known under the source condition of Lemma 14.2: by Theorem 14.1, uniqueness among translation-invariant stationary laws on $\mathbb{Z}^6$, and among fully translation-invariant Gibbs measures on $\mathbb{Z}^7$, is Π_2^0-complete at one fixed rational rate. For rational positive tables, ergodicity has the Π_2^0 upper bound of [9, Remark 10.9]; the old hardness conclusions of Remark 10.9 and Corollary 10.8 remain valid.

Open problem 13.2. Is ergodicity of positive-rate cellular automata at one fixed noise rate Π_2^0-complete on the full class of tables? Its translation-invariant variant is answered by Theorem 14.1 under that theorem's source condition; the problem as stated remains open. What are the exact degrees of uniqueness among all stationary laws, all-DLR uniqueness and exponential weak spatial mixing for the normalized space–time interactions?

Natural rules below four dimensions. The raw-Wang classification on $\mathbb{Z}^4$ answers the dimension-four existence instance of the earlier natural-rule question. It uses thick boxes and amortized coherence, rather than the thin-provider $F_R^{[2]}$. The saturation theorem Theorem 12.1 and Corollary 12.2 explains why the specified budget-only certificate stops at six; it does not exclude a different mechanism. The $\mathbb{Z}^3$ construction here assumes north/east determinism.

Open problem 13.3. Is there a natural fixed rule shape on $\mathbb{Z}^3$, applied directly to every finite raw Wang set, whose fixed-rate ergodicity is equivalent to non-tileability? What can be done on $\mathbb{Z}^2$ or $\mathbb{Z}$? Separately, is the original $k = 2$ thin-provider quorum $F_R^{[2]}$ on $\mathbb{Z}^4$ ergodic exactly for non-tiling R throughout one common positive noise interval?

The unstacked continuum gap. Theorem 3.10 needs uncountably many tilings. Corollary 3.11 gives an unconditional arc after one extra coordinate, but does not answer the question for the base rule.

Open problem 13.4. If R tiles but has only countably many tilings, does the noisy G_R itself have continuum many extreme stationary laws or continuum many extremal Gibbs measures? This is open here even for one tile whose opposite edges match.

Quantitative lower-dimensional results. Theorem 12.12 proves exponential all-input forgetting on the non-tiling side of both lower-dimensional families, with a universal exponent and an effectively findable input-dependent prefactor. Corollary 12.14 gives all-Gibbs uniqueness and arbitrary-volume exponential weak spatial mixing. Stronger forms of mixing and robustness remain questions. The common interval is now $0 < \varepsilon \leq 10^{-30}$, with exponent $1/24000$ and exponent $1/4$ retained for $\varepsilon \leq 10^{-4535}$. The all-Gibbs clause of the following version 2.1.0 problem is answered by Corollary 12.14; its other clauses remain open.

Open problem 13.5. For the rules of Theorems 12.5 and 12.11, can one prove stronger forms of mixing, transfer uniqueness to all space–time Gibbs measures, or establish robustness under perturbation of the local table? Can their common sufficient noise interval be enlarged substantially beyond 10^{-30}, or the exponent $1/4$ retained on the whole interval?

14 The degree of translation-invariant uniqueness

Even at one fixed positive replacement rate, a uniqueness question can be Π^0_2-complete. Here the property is uniqueness among spatially translation-invariant stationary laws, abbreviated TI stationary laws. The distinction from ergodicity matters: uniqueness in this class does not assert attraction from arbitrary starts. We use a different rule from the raw-Wang families of the preceding sections.

Theorem 14.1 (The TI uniqueness degree)**.** *Assume the source condition C_T stated in Lemma 14.2. There is a total computable family $e \mapsto G_e$ of deterministic finite-alphabet rules on $\mathbb{Z}^6$, of maximum-norm range at most 11, such that at each fixed rational rate $0 < \varepsilon \leq 10^{-22}$ uniform whole-symbol replacement gives exactly one TI stationary law if φ_e is total, and at least two otherwise. At each such rate TI stationary uniqueness is Π^0_2-complete, and TI coexistence is Σ^0_2-complete, under computable many-one reductions. The normalized space–time interactions have the same degrees for fully TI Gibbs uniqueness and coexistence on $\mathbb{Z}^7$. Their local Boltzmann factors are strictly positive rationals.*

Thus $\varepsilon = 10^{-22}$ may be fixed once for the whole decision problem. The alphabets vary effectively; the dimension, range and rate do not. Our Rice paper already contains this construction at the rate 160000^{-101}: its Lemma 6.1 gives the same six-dimensional family and the same totality dichotomy under the same source condition, and its Theorem C with $r = 1$ makes TI stationary uniqueness Π^0_2-complete on $\mathbb{Z}^7$ [52, Lemma 6.1 and Theorem C]. This section extends the dichotomy to every fixed rational rate in $(0, 10^{-22}]$, states the six-dimensional degree, adds the translation-invariant Gibbs degree on $\mathbb{Z}^7$, and writes the proof out in full. Törmä established the deterministic unique-ergodicity degree [55, Proposition 48]. Sandhu's independent fixed-temperature family [54, Theorem 1.1] has a halting-versus-uniqueness dichotomy on $\mathbb{Z}^2$. Its promised-family uniqueness degree is Σ^0_1, as is that of our natural raw-Wang family on $\mathbb{Z}^6$ (Corollary 3.7). These results concern different families and predicates.

Family	Decision property	Degree
Törmä [55]	deterministic unique ergodicity	Π^0_2-complete
Sandhu [54]	fixed-temperature Gibbs uniqueness on its family, $\mathbb{Z}^2$	Σ^0_1-complete
Natural rule on $\mathbb{Z}^6$, Corollary 3.7	fixed-noise ergodicity and Gibbs uniqueness	Σ^0_1-complete
Rice, Lemma 6.1, Theorem C [52]	TI stationary uniqueness at 160000^{-101}, $\mathbb{Z}^7$ ($\mathbb{Z}^6$ dichotomy), under C_T	Π^0_2-complete
Theorem 14.1	fixed-noise TI uniqueness, $\mathbb{Z}^6$, under C_T	Π^0_2-complete

14.1 The source and its precise dependency

Condition C_T is that the statements of [55, Lemma 25, Corollary 26 and Lemmas 27–29], on macrocells, absence of overlap, strong rigidity, strong connection and live coverage of the universal simulator, hold as published for the universal simulators used below, and that their blank state is quiescent. Their printed proofs are used without re-derivation. This is the published-source dependency of the condition of [52, Section 6.1].

Lemma 14.2 (The totality source under C_T)**.** *Assume C_T. There is a total computable family of one-dimensional radius-one CA U_e, with finite alphabet D_e, blank 0 and local map $u_e : D_e^3 \to D_e$, satisfying:*

(i) $u_e(0,0,0) = 0$. *If φ_e is total, every fully $\mathbb{Z}^2$-TI law on the two-sided input-erasure trajectory SFT of U_e is all blank.*

(ii) *If φ_e is nontotal, the ordinary two-sided trajectory SFT of U_e has a fully $\mathbb{Z}^2$-TI law of positive nonblank activity.*

Input erasure replaces any selected subset of the three inputs by 0, from one common preceding row. The repaired source reserves zero for every bookkeeping field's blank or no-kind code, normalizes invalid own bookkeeping to blank, leaves simulated data unrestricted, and increments age only while live. Field widths are fixed before self-reference.

The proof is in Section D.

14.2 A finite checker and five providers

Set

$$L_e = \{0\} \sqcup \{(a,w) : a \neq 0,\ w \in D_e^3,\ u_e(w) = a\}, \qquad \pi(0) = 0, \quad \pi(a,w) = a.$$

On $u = (i,s) \in \mathbb{Z}^2$, let $H_{0,e}$ retain a nonblank full center $b = (a,w)$ precisely when

$$\pi(x(i+r, s-1)) = w_r \quad \text{for each } r \in \{-1,0,1\} \text{ with } w_r \neq 0.$$

Otherwise it outputs 0; a blank center stays blank. This rule only deletes labels. Its stencil is $N_0 = \{0,(-1,-1),(0,-1),(1,-1)\}$, where s is a spatial source coordinate, not stochastic evolution time. Every fixed field projects to an input-erasure trajectory: use its stored witness at a nonblank output and erase every input at a blank one. Conversely an ordinary trajectory lifts by storing its actual predecessor triple at each nonblank output. The lift is local and translation covariant. Since no nonblank full label projects to zero, Lemma 14.2 gives only the blank TI fixed-field law on totality and a positive-activity TI fixed-field law on nontotality.

The physical sites are $(u,P) \in \mathbb{Z}^2 \times \mathbb{Z}^4$. In addition to the $H_{0,e}$ tests at the same P, require all eight transverse unit neighbours to equal the same full center label. Call this deletion rule H_e. Its twelve read offsets are

$$N_* = (N_0 \times \{0\}) \cup \{(0, \pm e_j) : 1 \le j \le 4\}.$$

Put $c_0 = 0$, $c_j = 10e_j$ for $1 \le j \le 4$. The rule G_e returns a nonblank label if at least four of the five values $H_e(x)(u, P + c_j)$ equal it, and returns 0 otherwise. The read offsets are fixed and have range at most 11. The finite alphabet and table are total computable from e. The noisy table is

$$p_e(a \mid x) = (1-\varepsilon)\mathbf{1}_{\{a = G_e(x)\}} + \varepsilon/|L_e|.$$

For the fixed rational ε this table is positive and rational.

14.3 The two noise estimates

Use physical evolution time τ. Mark a space–time vertex if it is a replacement flag, or if two distinct providers each have some marked N_* parent one time earlier. Empty starts receding into the past give an increasing minimal mark field E^*. It is a stationary, translation-covariant, causal factor of the iid flags. Finitary closure implies that every mark already occurs in a finite-start field and has a finite explanation ending at physical flags.

The sixty parent steps have the form $(d, c_j + n, -1)$ with $(d, n) \in N_*$. Providers have disjoint stencils: their centers have transverse ℓ^1 distance at least 10 and the checker offsets differ by at most 2. The undirected degree is at most 120. For the contour use

$$\psi_j(P, \tau) = -P_j - 2\tau \ (1 \le j \le 4), \qquad \psi_0(P, \tau) = \sum_j P_j + 8\tau.$$

Their sum is zero. Through provider k, the gains are $2 - 10\mathbf{1}_{\{k=j\}} - n_j$ and $10\mathbf{1}_{\{k\neq 0\}} + \sum_j n_j - 8$. Outside its excluded provider every charge gains at least 1; all gains are at least -9.

For completeness, the copied-arm construction applies with these gains as follows. Choose lower-rank supports at every nonflag mark in a finite explanation and stop at flags. Give each vertex its descendant physical terminal set. Expose nonterminals in root-to-parent order, retaining terminals on the frontier. Replacing one set by the union of its two support sets cannot merge overlap components: an intersection with a support already intersected its parent. Each remaining vertex of the old component has an overlap path to one of the two supports, so a component can split into at most two daughters with disjoint terminal unions. Maintain one eventual arm endpoint of each charge per component. At a split put a new source at the exposed vertex and send each charge toward its missing daughter; after this possibly bad initial edge, follow good supports. The original source starts with good edges. At the terminal frontier each physical terminal has one endpoint of each charge.

Copy arm interiors and identify starts belonging to a source and ends with the same terminal name. A new source need not lie on an old arm. Retain only the abstract component of the original root. Each charge's source-to-terminal bijection restricts to that component, giving m sources and m distinct flags. Only the five initial edges at each of its $m - 1$ nonroot sources can be bad. Thus, with multiplicities,

$$0 \ge e_{\mathrm{arm}} - 10 \cdot 5(m-1), \qquad e_{\mathrm{arm}} \le 50(m-1).$$

Projection gives a connected physical contour with at most this many edges. A minimum-time vertex of the abstract component is a terminal, because every other vertex has an outgoing arm edge to an earlier time. All vertices lie between the earliest selected flag and the root, with time span at most $50(m-1)$. An unflagged root has $m \ge 2$; a flagged root may use its singleton contour.

We now count flag positions with Lemmas 2.3B and 2.3C. For one graph edge each logical coordinate changes by at most one, physical time by one and transverse ℓ^1 distance by at most 11. The radius-r ball is therefore contained in

$$[-r, r]^3 \times \{P \in \mathbb{Z}^4 : |P|_1 \le 11r\}.$$

The four-dimensional integer ℓ^1 ball has size

$$L_4(R) = \sum_{j=0}^{4} 2^j \binom{4}{j} \binom{R}{j}.$$

Choose its j nonzero coordinates, their signs and positive magnitudes of total at most R; the latter choices number $\binom{R}{j}$. Consequently the graph ball is bounded by the degree-seven polynomial $b(r) = (2r+1)^3 L_4(11r)$. Its Newton coefficients are

$$(D_0, \dots, D_7) = (1, 323162, 20858800, 225345520, \\ 881336720, 1558189600, 1275204480, 393550080).$$

Taking successive differences proves $b(r) = \sum_{j=0}^{7} D_j \binom{r}{j}$ for every integer $r \geq 0$. The ball-to-sphere identity and the binomial generating series give the explicit bound

$$Z_G(y) = (1-y) \sum_{r \geq 0} |\operatorname{Ball}_G(r)| y^r \leq Z_b(y) := \sum_{j=0}^{7} D_j \left(\frac{y}{1-y} \right)^j, \qquad 0 < y < 1.$$

A contour tour has length at most $100(m-1)$. If $\lambda = \varepsilon Z_b(y) y^{-100} < 1$, the flag-position bound gives

$$p := \mathbb{P}(E^*(0) = 1) \leq \varepsilon + \sum_{m \geq 2} \varepsilon^m Z_b(y)^m y^{-100(m-1)} = \varepsilon + \varepsilon Z_b(y) \frac{\lambda}{1-\lambda}.$$

For a backward N-edge path whose first N vertices are marked and nonflagged, select whole contours in successive disjoint terminal slabs, restarting one time below the preceding earliest terminal. Keep the final contour whole and truncate only its connector. A contour with m_i terminals covers at most $50(m_i - 1) + 1 \leq 50 m_i$ of the prescribed vertices. All $m = \sum_i m_i$ selected flags are distinct and $m \geq \lceil N/50 \rceil$. Tour the contours and follow the path connectors. The walk length is at most

$$\sum_i [100(m_i - 1) + 50(m_i - 1) + 1] \leq 150 m.$$

The last connector stops at the prescribed endpoint; a possible final step is already included in the span-plus-one budget. With $\kappa_y = \varepsilon Z_b(y) y^{-150} < 1$ we obtain

$$\mathbb{P}\{\text{such a path from } v\} \leq \sum_{m \geq \lceil N/50 \rceil} \kappa_y^m \leq \frac{\kappa_y^{N/50}}{1 - \kappa_y}.$$

For any finite path every selected mark has a finite-start explanation; take a common sufficiently early start. Thus the same proof applies to the minimal field. The tail excludes infinite marked nonflagged paths, simultaneously at all roots.

Here are the endpoint inequalities, stated as rational comparisons. At $\varepsilon = 10^{-22}$,

$$y = \frac{1193}{1250}: \qquad 10^{-22} Z_b(y) y^{-150} \leq \frac{885}{10000} < \frac{1}{2},$$
$$y = \frac{9322}{10000}: \qquad \lambda = 10^{-22} Z_b(y) y^{-100} < 1, \qquad 10^{-22} + 10^{-22} Z_b(y) \frac{\lambda}{1-\lambda} \leq \frac{24}{10^9} < \frac{1}{2}.$$

The independent exact-arithmetic certificate in `verification-thresholds/` derives the Newton coefficients and checks these inequalities. All expressions increase with ε on their domains, proving both noise estimates throughout the stated interval. In particular the density bound used below is $p \leq 2.4 \cdot 10^{-8}$.

14.4 Clean coherence and exhaustion on totality

Call a vertex clean when $E^* = 0$. In every two-sided noisy trajectory, two clean nonblank full labels at the same source coordinate u and time τ, at transverse positions P, Q, agree. If they differ, choose a unit direction taking P toward Q. Each root has at least four supporting providers; their support sets meet in at least three indices. At each clean root at most one provider has any marked parent. Excluding the pair's at most two dirty indices leaves a common supporter with wholly clean neighbourhoods at both roots. The transverse equality

test supplies the first label at $P + c_k + a$ one time earlier and the second at $Q + c_k$, reducing their distance by one. Repetition would put two different labels at one physical vertex. This finite descent proves coherence without a bound on how far back the trajectory extends.

Let ρ be any TI stationary law. Form its two-sided stationary Markov trajectory with fresh independent innovations: the consistent finite-interval laws start from ρ at their left endpoint. Append the causal E^*. For every u, τ define $y_\tau(u)$ to be $b \neq 0$ if some clean transverse site carries b, and 0 otherwise. Coherence makes the nonblank choice unique. The countable existential definition is measurable; its law is source-TI and stationary in physical time.

A nonblank $y_\tau(u) = b$ has a clean witness. A supporting provider with wholly clean neighbourhood gives the same old center b and realizes every nonzero stored predecessor in $y_{\tau-1}$. In particular $y_\tau \leq y_{\tau-1}$ in the flat order, in which 0 is below each nonblank label. Stationarity makes the probabilities of nonblank activity equal at successive times; deletion therefore has probability zero. Countably many sites and times give $y_\tau = y_{\tau-1}$ everywhere almost surely. All old parent assertions now hold in this same field, which is $H_{0,e}$-fixed. On totality Lemma 14.2 makes it blank. Thus every nonblank symbol of every TI stationary law is marked by E^*.

For any two TI stationary margins, evolve their product with common innovations and take a Cesàro limit. The pair kernel is Feller; its margins remain the prescribed laws and its limit is TI and stationary. In a two-sided realization the extraction above applies to both margins with the same E^*. A discrepancy cannot occur at a replacement, since the written symbol is common. A deterministic discrepancy has a parent discrepancy in the finite read stencil. Choosing such parents gives an infinite nonflagged discrepancy path. Every discrepancy has a nonblank margin and hence is marked. The path estimate excludes it. The joining is diagonal, so the margins agree. Existence follows by Feller compactness and Cesàro averaging from the blank TI start. This proves exactly one TI stationary law on totality.

14.5 Two stationary laws on nontotality

Choose a TI $H_{0,e}$-fixed source law m, either blank or the positive-activity law. Clone its sample as $\mathrm{Ref}(u, P) = Z(u)$; this is fixed by G_e. Consider the finite-alphabet joint process (Ref, X, E), with the reference fixed, X updated by the noisy rule and E by the mark recurrence with the same flags. Start $X = \mathrm{Ref}$ and give E its stationary minimal marginal, independently of the reference. This extended process is Feller and translation covariant.

The relation $X_v \neq \mathrm{Ref}_v \Rightarrow E_v = 1$ is preserved. At a nonflagged unmarked output, at most one provider has a marked parent, so four fully correct providers pass the reference checks and force its label. This holds also for a blank reference. All joint iterates retain the reference marginal, the stationary E^* marginal and their independence. Cesàro compact limits are stationary and retain the closed coordinatewise error relation. Their X marginals are TI stationary laws of the target rule.

For fixed u, the transverse averages of $E^*(u, P)$ concentrate below $1/2$. Indeed approximate its indicator in L^1 by a finite-start mark, which reads finitely many iid flags. Sufficiently separated transverse translates of that approximant are independent. Hence the covariance of the limiting indicators tends to zero at large separation. Averaging covariances over growing transverse cubes $Q_L = [-L, L]^4 \cap \mathbb{Z}^4$ gives

$$\mathrm{Var}\left(\frac{1}{|Q_L|} \sum_{P \in Q_L} E^*(u, P) \right) \longrightarrow 0.$$

The mean $p < 1/2$, so Chebyshev gives $\delta_L := \mathbb{P}\{|Q_L|^{-1} \sum_P E^*(u, P) \geq 1/2\} \to 0$. Let

$D_L(X)(u)$ be the strict majority full label in that cube, choosing 0 when there is none. Failure to recover the constant reference label requires at least half the sites to be errors. Thus its error probability is at most δ_L, uniformly in the reference law.

Let $a > 0$ be the positive source activity. For the positive-reference and blank-reference stationary lifts,

$$\mathbb{P}_1\{D_L(X)(u) \neq 0\} \geq a - \delta_L, \qquad \mathbb{P}_0\{D_L(X)(u) \neq 0\} \leq \delta_L.$$

Some finite L has $\delta_L < a/3$, and the resulting cylinder observable separates the two laws. The reduction stays effective although a and L need not be computable.

14.6 Arithmetic upper bounds and the Gibbs consequence

We finish the proof of Theorem 14.1. Totality, the set

$$\mathrm{TOT} = \{e : \forall n\ \exists s\ [\varphi_e(n) \text{ halts by } s]\},$$

is Π^0_2-complete: from a decidable relation $S(a, n, s)$ construct the program on input n searching for an s with $S(a, n, s)$. It is total exactly when $\forall n \exists s\, S(a, n, s)$. The preceding construction reduces TOT to TI stationary uniqueness and its complement to coexistence.

For membership, let a rational finite-range PCA table be given. Marginal probability tables on increasing boxes, satisfying all TI and stationarity cylinder equations whose supports fit, form nonempty rational outer polytopes. Their restrictions preserve earlier constraints. Stationarity equations are linear because the expected updated cylinder indicator is a computable rational cylinder function of the preceding row. Exact rational linear optimization, for example by finite vertex enumeration, computes the two extrema of every fixed cylinder probability on each polytope. These nested extrema converge to the genuine extrema over TI stationary laws: extend optimizing finite tables arbitrarily to full configurations, take a weakly convergent subsequence, and pass each fixed equation once its finite support fits. The converse follows by restricting any genuine law.

The nonempty stationary TI set is a singleton precisely when, for every cylinder C and integer m, some finite outer polytope has cylinder width below 2^{-m}. The width is a rational computable number; the formula is $\forall(C, m)\exists n$ with a decidable matrix. This proves Π^0_2 membership and Σ^0_2 membership of coexistence.

Finally put the normalized factors $p_e(\text{child} \mid \text{parents})$ on $\mathbb{Z}^6 \times \mathbb{Z}$. Finite-volume specifications include every touching factor, including factors of exterior children. Positivity, finite alphabet, finite range and independent-site updates check the hypotheses of the classical TI space–time correspondence [21], in the form [9, Lemma 2.4, Theorem 10.3 and Corollary 10.4]. It is an affine bijection between TI stationary row laws and fully TI Gibbs laws. The stationary trajectory satisfies the DLR equations by local cancellation. Conversely the TI Gibbs law has the PCA's product conditional transitions: the normalized causal law and the full touching-factor law differ by a bounded log density per boundary site; the chronological relative-entropy chain rule bounds the sum of nonnegative causal conditional defects by a boundary term. Translation invariance and a vanishing boundary-to-volume ratio force each defect to zero. This is exactly the TI converse of the cited theorem. The new degree conclusion uses its existing statement.

The Gibbs upper bound can also be read directly from rational constraints. Test every finite-volume DLR identity against every finite outside cylinder, together with TI and marginal equations. With positive rational factors each conditional kernel is a rational cylinder function.

Finite-volume limits followed by spatial averaging give nonempty TI Gibbs sets. The same nested-polytopes argument gives a Π^0_2 singleton test. For translation-invariant Gibbs sets this upper bound has an effective antecedent in Gayral, Sablik and Taati [53, Propositions 14–15]: they compute pressure and semidecide avoidance of a closed rational ball by the union of translation-invariant equilibrium sets over a rational inverse-temperature interval, so at one rational inverse temperature this compact set is uniformly Π_1-computable and singleton testing gives a Π^0_2 description. Their realization theorem concerns zero-temperature accumulation sets. The affine correspondence supplies the lower bound and proves the stated Gibbs degrees. The rational objects here are the Boltzmann factors; their potentials $-\log p_e$ are real. The theorem concerns TI uniqueness. All-law ergodicity, uniqueness among all Gibbs laws, exact nontotal phase counts and the smallest possible dimension remain separate questions.

How this work was done

AI note. The authors chose the question, directed the work and judged every result. AI models from two developers explored the problem, derived and drafted the arguments, and challenged each other's results: OpenAI's GPT-6.1 Sol and Anthropic's Claude Fable 5.2 and Claude Opus 5.5. The forgetting theorem and the construction of Theorem C were each derived by one model and then re-derived refutation-first, step by step, by a model of the other developer, which did not write them. The memory theorem rests on the contour lemma of the forgetting proof, which was cross-checked in this way; its remaining steps were written and checked by models of one developer, the check in a separate run that did not write them, and were checked again by a model of the other developer in the review of version 1.3.0 described below. These reviews found one gap in the parser of Section 7 (a missing transport of the phase from the anchor to a queried block in the point certificate) and one constant in the memory bound that rested on an unproved property; both were repaired and re-checked. The three-token theorem of Section 11 was derived by GPT-6.1 Sol and re-derived refutation-first by Claude Opus 5.5. That review found that the source of the argument used the halting time of M where the halting time of the simulating machine M° is needed; Section 11 uses the latter. Its one computed lemma was checked by two independent programs with positive controls, and the reviewer re-ran both and wrote a third. The perturbed-kernel corollary of Section 10.2, added in version 1.2.0, was derived by GPT-6.1 Sol and re-derived refutation-first, in two rounds, by Claude Opus 5.5. The reconstruction of the manuscript from the proofs was itself read refutation-first by a reviewer of the other developer; for Section 11, added in version 1.1.0, and for Section 10.2, added in version 1.2.0, it was read refutation-first by a model of one developer, in a separate run that did not write it. For version 1.3.0, a separate run of GPT-6.1 Sol, which had not seen the work, reviewed the whole manuscript refutation-first, including the memory theorem, compared it with the source derivations, and re-ran the corner-lemma programs. It found no false step in the main results, one omitted case in item (c) of Remark 10.13 and one credit sentence, both corrected in version 1.3.0. It also tested the forced-word lemma (Lemma 11.17) for $n = 1, 2$ with two independently written programs.

Version 1.4.0 added the following. The observation that the forgetting bound holds for all starts at once (Corollary 5.12) and its application to all Gibbs measures of the space–time interaction (Corollary 10.8), through the prefix-cap theorem of [9], were derived by GPT-6.1 Sol and checked cross-family, refutation-first, by Claude Opus 5.5; a separate run of the same model attacked the load-bearing steps without seeing that check. Neither found a false step. The extension in Remark 10.14 was derived during that check by Claude Opus 5.5 and derived

again, in writing version 1.4.0, by the same model. The stacked automata of Section 10.3 were derived by GPT-6.1 Sol and checked refutation-first by Claude Opus 5.5. That check found one clause stated without proof, the bound on the mixing time in Corollary 10.17(i), and supplied the proof used here; it also proposed the last claim of Corollary 10.17(ii)(c), item (ii)(e) and the last claim of (iii), which were derived again, in writing version 1.4.0. The halting case of Corollary 10.17(iii) was derived in writing version 1.4.0. The Lean proofs of Section C were written by agents running Claude Opus 5.5, directed by a lead agent running the same model, which also read the challenge statement of the comparator check against Section 4 and Corollary 11.21; the formalization led to Section B and to the remark after Corollary 11.13. A run of GPT-6.1 Sol ported the proofs to the toolchain pinned for release, built the comparator challenge, and ran the comparator, the two kernel checkers and the axiom audit. A separate run of GPT-6.1 Sol, which had not seen the work, then checked the new material of the written version 1.4.0 refutation-first, including Remark 10.14, the halting case of Corollary 10.17(iii), Section B and the fidelity of the formal statement described in Section C, which it read in the Lean sources without building them. It found no failure in the new arguments. It found two statements to correct: Section C claimed the classification for the family of Definition 8.17 itself, which the development proves for an encoded realization once the convention for encoded machines stated there is added; and one sentence of Section B misstated which moves of M take a single step of M°. It also asked for corrections of the wording of the one-symbol case in the proof of Corollary 10.8 and of the last clause of the abstract, for the removal of a sentence about earlier work, and for explicit notes where it could not confirm the evidence. All of this was done in version 1.4.0. Claude Opus 5.5 wrote version 1.4.0. Models are not listed as authors; the authors take responsibility for every claim.

Version 2.0.0 (AI note continued). The natural-Wang classification and its transfer results were derived by GPT-6.1 Sol in research runs and checked refutation-first by Claude Opus 5.5. An independent blind GPT-6.1 Sol replication, which had not seen the derivation, rederived the central natural-rule classification. The transfer verification by Claude Opus 5.5 rederived the load-bearing all-input, Gibbs, robustness and complexity arguments and fixed their scopes. The cross-family checks also contributed: the five-cluster count of Remark 3.12 was first recorded in the check of the classification, and the transfer verification proposed the port of the stacking corollary to the natural rule (Corollary 3.11) and the identification of the dimension count with Toom's eroder criterion. The version 1.1.0 noise-diagram package, whose Theorem 9.5 and Appendix A provide the finite-history antecedents, was directed by Nidhal Mghirbi and developed with GPT-6 Astra and GPT-6.1 Sol, with the verification recorded in that paper's AI note. GPT-6.1 Sol wrote version 2.0.0, and a separate run of Claude Opus 5.5 read the written version refutation-first. This writing revision integrates the resulting proofs and checks; it is not a further independent research replication. The Lean development of the natural rule was written by Claude Opus 5.5 agents, and its statements were read blind by a further Claude Opus 5.5 agent before the statement file was frozen. On the interval $0 < \varepsilon \leq (2(18q)^2)^{-200}$ it verifies both sentences of Theorem D, Theorem A on $\mathbb{Z}^6$, Corollary 3.1, the domino reduction and the undecidability of the domino problem; the existing $\mathbb{Z}^8$ core is retained (Section C). The research and manuscript reviews described above are internal model-based reviews, not human peer review. The Lean proofs and their checks are described separately in Section C. The authors remain responsible for the claims.

Lower dimensions (AI note continued). The lower-dimensional theorems were found in a GPT-6.1 Sol research run and checked refutation-first by Claude Opus 5.5, including an

independent blind reconstruction of the key lemmas by a further Claude Opus 5.5 agent. The in-run audits were from one model family; the shared amortized-coherence argument's originator also audited its use in the three-dimensional construction. The cross-family review and its blind reconstruction supply separate checks of that shared argument. This writing revision integrates the checked proofs and wording repairs; it is not a further independent research replication. An independent blind GPT-6.1 Sol replication, given the statements but not the proofs, proved both lower-dimensional theorems again. Models are not authors; the authors remain responsible for the claims.

Version 2.1.0 (AI note continued). The lower-dimensional rate was found in a GPT-6.1 Sol research run and proved with the sharper exponent $1/4$ in an independent blind GPT-6.1 Sol replication. The replication's separate in-run audit rederived the finite-history, distinct-terminal, changing-label and full-discrepancy arguments. It also exhibited the tileable one-tile NE example that requires the non-tiling restriction in both families. This writing revision integrates those proofs into Theorem 12.12 and Example 12.13; it is not a further independent research replication. The new rate is not formalized. Claude Opus 5.5 then checked the rate refutation-first, re-deriving it from the source papers; it confirmed the exponent $1/4$, the input-dependent prefactor and the necessity of the non-tiling restriction. Models are not authors; the authors remain responsible for the claims.

Version 2.2.0 (AI note continued). The lower-dimensional all-Gibbs result was found in a GPT-6.1 Sol research run with two in-run audits. Claude Opus 5.5 then checked it refutation-first and returned the verdict *PROVED*. The new Corollary 12.14 is a hypothesis check of the shared dependency [9, Corollary 13.7], through its Theorem 13.3, rather than an independent transfer proof. The check rederived the finite-history argument of Theorem 12.12 modulo its contour and density inputs, and the Corollary 13.7 chain, but did not fully rederive the arm construction of Lemma 12.3. The run's two audits were non-blind on the key transfer step, and the checks share the published proof path. The $\mathbb{Z}^4$ Lean core was written by GPT-6.1 Sol agents in Codesk sessions. Its frozen statement passed a blind fidelity read by Claude Opus 5.5 with verdict *FAITHFUL*; the recorded package passed the 133-module replay, sandboxed comparator, `leanchecker`, `nanoda` with a negative control, and the axiom audit described in Section C. This revision integrates those records and adds no Lean replay. Models are not authors; the authors remain responsible for the claims.

Version 3.0.0 (AI note continued). The enlarged noise intervals were found in a campaign of AI agents directed by Nidhal Mghirbi, with in-campaign audits. GPT-6.1 Sol agents in a Codesk session checked the threshold arguments as a different model family, re-derived both charge-arm constructions and the flag-position and rate bounds, and wrote independent exact arithmetic. Their review required the distinction between these arithmetic certificates and the older corner-configuration computation, which is made here. The threshold campaign's own audits preceded that cross-family check. Claude Opus 5.5 re-derived the threshold arguments in its refutation-first review of this written version. The degree construction was found by GPT-6.1 Sol agents in a Codesk session at the rate 160000^{-101} and checked there, refutation-first, by Claude Opus 5.5 as a different model family; the same construction appears in [52, Lemma 6.1]. Its extension to every rational rate in $(0, 10^{-22}]$ uses noise estimates found in the threshold campaign and re-derived in the GPT-6.1 Sol threshold check, which also confirmed that the rest of the proof uses the rate only through these two estimates and

positivity; that check is by the model family that found the construction. Claude Opus 5.5 also re-derived the enlarged-rate estimates and checked their use in the complete proof in its review of this written version. The source conventions and the review's proof expansions are included in Lemma 14.2 and Section D. The authors have not independently verified the new mathematical arguments; verification consists of these model checks and the exact-arithmetic certificates. This writing revision integrates those sources and repairs, runs the independent arithmetic again and checks the build; it is not a further independent mathematical audit. GPT-6.1 Sol agents in a Codesk session wrote version 3.0.0; Claude Opus 5.5 read the written version refutation-first. Models are not authors; the authors remain responsible for the claims.

Verification status

- **Natural rule.** Theorem 2.1 and its consequences are proved analytically in Sections 2 and 3, using the explicitly cited classical and space–time inputs. The enlarged natural-rule interval uses finitely many explicit rational inequalities certified in `verification-thresholds/`; its raw-tile reduction is the classical domino reduction. The corner-configuration computation belongs to the retained machine route. Both sentences of Theorem D (Theorem 2.1 without its coalescence bound), Theorem A on $\mathbb{Z}^6$, all of Corollary 3.1, the domino reduction and the undecidability of the domino problem are formalized without named hypotheses (Section C) on the version 2.2.0 interval $0 < \varepsilon \leq (2(18q)^2)^{-200}$. The enlarged interval is proved on paper with exact arithmetic. The coalescence bound, nilpotency and the other consequences of Section 3 are not formalized.
- **Lower dimensions.** Theorems E and F (Theorems 12.5 and 12.11) and the saturation, contour, all-path density, coherence, edge-matching and compactness arguments in Section 12 are proved analytically, using the classical domino imports with their stated scopes. The exponential all-input estimate of Theorem 12.12, including its effective prefactor and essential non-tiling restriction, is proved analytically. Corollary 12.14 is a hypothesis check of the shared dependency [9, Corollary 13.7], through Theorem 13.3, with equations (7)–(8) and Lemma 2.4. The refutation-first check rederived the argument of Theorem 12.12 modulo its contour and density inputs and the published transfer chain, but not the full arm construction of Lemma 12.3. Its two preceding in-run audits were non-blind on the key step; these checks follow the same transfer proof path. For version 3.0.0 both arm constructions and their improved accounting were independently re-derived in the cross-family threshold review. Robustness has not been transferred. The qualitative $\mathbb{Z}^4$ classification and its direct undecidability corollaries are formalized (Section C) on the version 2.2.0 interval $0 < \varepsilon \leq 382205952^{-33720}$; the enlarged interval is proved on paper with exact arithmetic; the $\mathbb{Z}^3$ classification, the lower-dimensional rates and the new Gibbs corollary are not.
- **Proved in the text.** Theorems A, B, C and D and every lemma and corollary of Sections 4 to 11 and of Section B, given the cited inputs below; in Section 11 the one exception is the corner lemma Lemma 11.12, which is checked by machine.
- **Imported with stated scope.** From [9], where they are proved: Lemma 2.4, Corollary 10.4 and Theorem 10.10(i), that is, the correspondence of [21] between translation-invariant Gibbs measures of the space–time interaction and stationary trajectories, and the uniqueness of the periodic Gibbs measure of an ergodic positive-rate automaton,

used in Corollaries 10.8 and 10.17; the prefix-cap theorem and its corollary for basic flows, Theorem 13.3 and Corollary 13.7, used in Corollary 10.8(i), Remark 10.14 and Corollary 10.17(iii); and, for the remark on the arithmetical degree, Remark 10.9 with Halawi's theorem [22]. The natural-rule results import, from [9], Lemma 2.4 (Corollary 3.3 and Theorem 3.10), Corollary 10.4 (Corollary 3.11) and Theorem 13.3 with Corollary 13.7 (Corollaries 3.3, 3.8 and 3.11 and the Gibbs remark after Corollary 3.5); from [12], the conversion of Corollary B (Remark 3.4) and the cardinality argument of Theorem 9.6 (Theorem 3.10); and the classical domino reduction [3, 43] (Corollaries 3.1 and 3.7). The lower-dimensional Gibbs result Corollary 12.14 imports [9, Corollary 13.7, equations (7)–(8), Theorem 13.3 and Lemma 2.4]; its decision clauses use the raw domino problem and the restricted tiling import of Theorem 12.11, with their stated scopes. Section 14 imports [9, Lemma 2.4, Theorem 10.3 and Corollary 10.4] and, under C_T, [55, Lemma 25, Corollary 26 and Lemmas 27–29].

- **Checked less thoroughly.** Remark 10.14 and the halting case of Corollary 10.17(iii) were derived by models of one developer and checked by a model of the other developer only in the review of the written version 1.4.0 (see "How this work was done").
- **Existence only.** The parser constant B of Theorem 7.18 is shown to exist and was not computed. The machine route's explicit constants come from Corollary 11.21, which does not use it.
- **Checked by machine.** One finite statement, the corner lemma Lemma 11.12, is a computation: an arc-consistency deletion on a box of side 33 for four roots, with four positive controls. Two independently written programs carry it out and agree, and a third, written separately during review, agrees with them. The programs, their outputs and run receipts are in `verification/corner-lemma/` of the repository. The review of version 1.3.0 re-ran the programs; the review of version 1.4.0 did not. The third program also finds that the list at the centre empties already on the box of side 13, the statement used in the remark after Corollary 11.13. The new threshold and degree-rate bounds also use explicit rational inequalities, checked by an independently written exact-arithmetic program (75 checks, integers and fractions only), released with its output and run instructions in `verification-thresholds/`. It checks finite arithmetic only; the combinatorial and probabilistic proofs are in the text. The forced-word lemma (Lemma 11.17), which is proved in the text, was also tested for $n = 1$ and $n = 2$ by two further independently written programs, for all sixteen pairs of corner letters, with legal-pattern and ablation controls (`verification/forced-word/`). These tests support the proof and are not part of it.
- **Formalized, natural-rule core.** These targets retain only the version 2.2.0 interval $0 < \varepsilon \leq (2(18q)^2)^{-200}$. The five targets are `FNR.natural_classification` (Theorem D, both sentences), `FNR.theorem_A` (Theorem A on $\mathbb{Z}^6$), `FNR.machine_corollary` (Corollary 3.1), `FNR.domino_reduction` and `FNR.domino_undecidable`. Their only axioms are `propext`, `Classical.choice` and `Quot.sound`. The statement file was corrected after a blind fidelity read found one mismatch and was then frozen; the package passed a sandboxed comparator, `leanchecker`, `nanoda`, the negative control and the axiom audit. Sources and reproduction are in `lean-natural/` of the release repository (Section C).
- **Formalized, lower-dimensional core.** These targets retain only the version 2.2.0 interval $0 < \varepsilon \leq 382205952^{-33720}$. The three targets `FLD.z4_classification`, `FLD.z4_machine_corollary` and `FLD.z4_raw_undecidable` prove the qualitative $\mathbb{Z}^4$ classification of The-

orem 12.5 (Theorem E) and its direct machine and raw-family undecidability corollaries, with no named hypotheses. The frozen statement passed a blind fidelity read with verdict *FAITHFUL*. The recorded package passed a fresh 133-module replay, a sandboxed comparator, `leanchecker`, `nanoda` with a negative control, and an axiom audit admitting only `propext`, `Classical.choice` and `Quot.sound`. It will ship as `lean-lower/`; the $\mathbb{Z}^3$ family, the lower-dimensional rates and Corollary 12.14 are excluded (Section C).

- **Formalized, unchanged machine-family core.** Corollary 11.21, the machine-family theorem on $\mathbb{Z}^8$, and the undecidability of ergodicity in the family are proved in Lean 4 with Mathlib, with no unproved step and only the three standard axioms (Section C). The corner lemma enters there in the form of the remark after Corollary 11.13, checked by the Lean kernel. Theorem 9.8, Theorem 10.2, the other corollaries of Section 10 and Corollary 5.12 are not formalized. The Lean sources are in `lean/` of the repository. A comparator check confirms that the proved statements are those of a challenge file that imports only Mathlib, and two kernel checkers re-check the proofs (Section C).
- **Citations.** Of the references added in version 1.4.0, [44] was read only in its abstract, and [15, 16, 32, 49] are cited for statements recorded in [9]; we have not checked them against their full texts. For version 2.0.0, the primary texts of [19, 5, 34, 43, 46] were consulted for the cited arguments; these were targeted checks, not full-text reviews. No full-text check is recorded for [6, 27, 42, 23, 24, 41, 4].
- **Lower-dimensional citation check.** The SIAM publisher record for [26] and the primary survey [51, Section 2.8, Theorem 5; discussion after Theorem 15, p. 24] were checked for the restricted unseeded tiling import and the two-input nilpotency construction. This was a targeted reading, not a full reconstruction of the 1992 undecidability proof.
- **Degree.** The proof of Theorem 14.1, including the enlarged fixed-rate estimates, is a paper proof with exact arithmetic under the source condition in Lemma 14.2; it has no Lean formalization. The condition retains the published simulator structure lemmas, whose internal proofs were used as published.
- **New citation checks.** The public Zenodo record and TeX of [54] were checked for Theorem 1.1, its fixed support shapes, integer energies and statement that every Gibbs state is TI. [53, Propositions 14–15] and [55, Proposition 48] were checked in the primary author manuscripts. These were targeted statement checks. Lemma 6.1, Theorem C with $r = 1$, and the source condition of [52] were checked in the supplied version 1.0.0 PDF.
- **Internal only.** Every review is internal; none of it is journal peer review.

A Tables for the scaffold

This appendix lists the finite data of Section 7: Ollinger's substitution [40], in the transcription of [31], and its ruler decoration. Every entry below was derived again from the definitions in that section; the counts are obtained by listing, not by computer search. Vertical lane atoms are $\mathsf{r}_0, \mathsf{r}_2, \mathsf{g}_0, \mathsf{g}_2$ and horizontal lane atoms are $\mathsf{r}_1, \mathsf{r}_3, \mathsf{g}_1, \mathsf{g}_3$; the letter r means red and g means green.

The left table gives the following consequences, used in Lemma 7.4.

- Each corner has exactly the four endpoints of one thin type: $F(\mathsf{a}) = \angle\mathsf{b}$, $F(\mathsf{b}) = \angle\mathsf{d}$, $F(\mathsf{c}) = \angle\mathsf{c}$, $F(\mathsf{d}) = \angle\mathsf{a}$.

	∂_N	∂_S	∂_E	∂_W
thin a	r_0	g_2	r_1	g_3
thin b	g_0	r_2	g_1	r_3
thin c	r_2	g_0	g_3	r_1
thin d	g_2	r_0	r_3	g_1
corner $\angle$a	g_2	r_0	r_3	g_1
corner $\angle$b	r_0	g_2	r_1	g_3
corner $\angle$c	r_2	g_0	g_3	r_1
corner $\angle$d	g_0	r_2	g_1	r_3
line $\langle\alpha,\gamma\rangle$	α	α	γ	γ

	∂_N	∂_S	∂_E	∂_W
black A	0	1	0	1
black B	1	0	0	0
black C	1	0	1	1
black D	1	1	1	0
black E	0	0	1	0

Table 4: Endpoints of the thin and thick components, and bits of the black component. A line $\langle\alpha,\gamma\rangle$ has a vertical atom α and a horizontal atom γ; all sixteen lines occur. The edge colour of a tile (θ,κ,ω) on a side is the triple of the three entries on that side.

- Thin types match horizontally exactly in the pairs (a,c), (c,a), (b,d), (d,b) (west tile first), and vertically exactly in the pairs (a,d), (d,a), (b,c), (c,b) (south tile first).
- ∂_S of the thin types a,b,c,d is g_2, r_2, g_0, r_0; ∂_N is r_0, g_0, r_2, g_2; ∂_W is g_3, r_3, r_1, g_1; ∂_E is r_1, g_1, g_3, r_3.

The black maps of the substitution are

ω	A	B	C	D	E
$\mathrm{se}(\omega)$	B	B	C	C	C
$\mathrm{nw}(\omega)$	E	D	D	D	E

They are characterized by $\partial_W \mathrm{se}(\omega) = \partial_E\omega$ with $\mathrm{se}(\omega) \in \{B,C\}$, and $\partial_S \mathrm{nw}(\omega) = \partial_N\omega$ with $\mathrm{nw}(\omega) \in \{D,E\}$. They also satisfy $\partial_E \mathrm{se}(\omega) = \partial_E\omega$, $\partial_N \mathrm{nw}(\omega) = \partial_N\omega$, $\partial_S \mathrm{se}(\omega) = 0$, $\partial_W \mathrm{nw}(\omega) = 0$, $\partial_N \mathrm{se}(\omega) = 1$ and $\partial_E \mathrm{nw}(\omega) = 1$ for every ω.

thin	thick	black	condition	tiles
a	corner	A	none	4
a	line $\langle\alpha,\gamma\rangle$	B	$\alpha \neq r_0$	12
a	line $\langle\alpha,\gamma\rangle$	C	$\alpha \neq r_2$	12
a	line $\langle\alpha,\gamma\rangle$	D	$\gamma \neq r_3$	12
a	line $\langle\alpha,\gamma\rangle$	E	$\gamma \neq r_1$	12
b	corner	A	none	4
c	line $\langle\alpha,\gamma\rangle$	B	$\alpha \neq r_0$	12
c	line $\langle\alpha,\gamma\rangle$	C	$\alpha \neq r_2$	12
d	line $\langle\alpha,\gamma\rangle$	D	$\gamma \neq r_3$	12
d	line $\langle\alpha,\gamma\rangle$	E	$\gamma \neq r_1$	12
			total	104

Table 5: The 104 tiles of $\mathbb{O}$. A line is excluded in a black class by one of four red atoms.

child	thin	thick	black
$(0,0)$	a	κ	ω
$(1,0)$	c	$\langle \partial_{\mathrm{S}}\theta,\ \partial_{\mathrm{E}}\kappa \rangle$	$\mathrm{se}(\omega)$
$(0,1)$	d	$\langle \partial_{\mathrm{N}}\kappa,\ \partial_{\mathrm{W}}\theta \rangle$	$\mathrm{nw}(\omega)$
$(1,1)$	b	$F(\theta)$	A

child	χ_1	χ_2
$(0,0)$	$1-b_1$	$1-b_2$
$(1,0)$	0	$1-b_2$
$(0,1)$	$1-b_1$	0
$(1,1)$	0	0

Table 6: The substitution φ of a tile (θ,κ,ω) (left), and the bits of the decorated substitution φ' of a letter with bits (b_1,b_2) (right). Child $(1,0)$ is south-east and child $(0,1)$ is north-west.

θ	blacks allowed by children $(1,0)$, $(0,1)$	formal parents	in $\mathbb{O}$	excluded
a	all	52	52	0
b	A	4	4	0
c	B, C, D	36	24	12: black D, line with $\gamma \neq \mathsf{r}_3$
d	C, D, E	36	24	12: black C, line with $\alpha \neq \mathsf{r}_2$
total		128	104	24

Table 7: Formal parents of an aligned legal 2×2 block (Lemma 7.5). The count of formal parents uses the 52 pairs (κ,ω) allowed at child $(0,0)$.

thin	base tiles	bit rule	variants each	letters
a	4 corners	$b_1 = b_2$	2	8
a	8 red crossings	$b_1 \neq b_2$	2	16
a	40 other lines	none	4	160
b	4 corners	$b_1 = b_2 = 0$	1	4
c	4 red crossings	$b_1 = 0, b_2 = 1$	1	4
c	20 other lines	$b_1 = 0$	2	40
d	4 red crossings	$b_1 = 1, b_2 = 0$	1	4
d	20 other lines	$b_2 = 0$	2	40
			total	276

Table 8: The 276 letters of $\mathbb{A}$. A red crossing is a line whose two atoms are red. Each black class B, C, D, E admits exactly two of the four red crossings (Table 5), which gives 8 red crossings with thin type a and 4 each with thin types c and d.

position	thin	thick	red crossing exactly when
$(2,1)$	d	$\langle \partial_{\mathrm{S}}\theta,\ \mathsf{r}_1 \rangle$	$\theta \in \{\mathsf{b},\mathsf{d}\}$
$(2,3)$	d	$\langle \partial_{\mathrm{N}}\theta,\ \mathsf{r}_3 \rangle$	$\theta \in \{\mathsf{a},\mathsf{c}\}$
$(1,2)$	c	$\langle \mathsf{r}_0,\ \partial_{\mathrm{W}}\theta \rangle$	$\theta \in \{\mathsf{b},\mathsf{c}\}$
$(3,2)$	c	$\langle \mathsf{r}_2,\ \partial_{\mathrm{E}}\theta \rangle$	$\theta \in \{\mathsf{a},\mathsf{d}\}$

Table 9: Four cells of the depth-two block $\varphi^2(\theta,\kappa,\omega)$, for every base tile. Position $(2,1)$ is child $(0,1)$ of child $(1,0)$; $(2,3)$ is child $(0,1)$ of child $(1,1)$; $(1,2)$ is child $(1,0)$ of child $(0,1)$; $(3,2)$ is child $(1,0)$ of child $(1,1)$.

B The one-sided machine

This appendix constructs the one-sided machine M° of Section 8.6 and proves the properties used there and in Proposition 8.16: M° moves its head at every step, never moves left from cell 1, starts in a state that is not halting, leaves cell 1 at its first step, and halts if and only if M halts. The construction and these properties are part of the Lean formalization of Section C.

The machine M. Let M have the finite state set Q, the initial state q_{in}, the set $Q_{\mathrm{h}} \subseteq Q$ of halting states, the finite tape alphabet Σ with blank $\sqcup$, and the transition function $\delta\colon Q \times \Sigma \to Q \times \Sigma \times \{-1, 0, 1\}$. A configuration is a triple (q, p, y) of a state q, a head position $p \in \mathbb{Z}$ and a tape $y\colon \mathbb{Z} \to \Sigma$. If $\delta(q, y(p)) = (q', a', d)$, the next configuration is $(q', p + d, y')$, where y' equals y except that $y'(p) = a'$. The run starts from $C_0 = (q_{\mathrm{in}}, 0, \sqcup)$, the blank tape, and $C_n = (q_n, p_n, y_n)$ is its configuration after n steps. M halts if $q_n \in Q_{\mathrm{h}}$ for some n. If δ is only partially defined, a pair (q, a) without a value is treated as $\delta(q, a) = (q, a, 0)$, a step that changes nothing; the construction below does not change.

Tracks. Write U and L for the upper and the lower track. For a cell $x \geq 1$ put $\mathrm{pos}(\mathrm{U}, x) = x - 1$ and $\mathrm{pos}(\mathrm{L}, x) = -x$. The map $(\ell, x) \mapsto \mathrm{pos}(\ell, x)$ is a bijection from $\{\mathrm{U}, \mathrm{L}\} \times \{1, 2, \dots\}$ onto $\mathbb{Z}$. Its inverse sends a position p to $(\ell(p), c(p))$, with $\ell(p) = \mathrm{U}$ and $c(p) = p + 1$ if $p \geq 0$, and $\ell(p) = \mathrm{L}$ and $c(p) = -p$ if $p < 0$. Thus the upper track of cell x holds the position $x - 1$ of M, and the lower track holds the position $-x$.

The machine M°. Its tape alphabet is $\Gamma = \{0, 1\} \times \Sigma \times \Sigma$. A symbol $\gamma = (b, a_{\mathrm{U}}, a_{\mathrm{L}})$ has a *marker* b and the entries a_{U} and a_{L} on the two tracks; the blank of M° is $(0, \sqcup, \sqcup)$, which Section 8.6 also writes $\sqcup$. For $\ell \in \{\mathrm{U}, \mathrm{L}\}$ and $a \in \Sigma$ let $\gamma[\ell \leftarrow a]$ be γ with the entry on track ℓ replaced by a. The states are the initial state q_0 and the states $\mathsf{sim}(q, \ell)$ and $\mathsf{back}(q, \ell)$ for $q \in Q$ and $\ell \in \{\mathrm{U}, \mathrm{L}\}$. The halting states are the states $\mathsf{sim}(q, \ell)$ with $q \in Q_{\mathrm{h}}$; in particular q_0 is not halting. The transition function tr gives, for a state and a scanned symbol, the new state, the symbol written and the move, $+1$ to the right and -1 to the left:

(T1) $\mathrm{tr}(\mathsf{q}_0, (b, a_{\mathrm{U}}, a_{\mathrm{L}})) = (\mathsf{back}(q_{\mathrm{in}}, \mathrm{U}), (1, a_{\mathrm{U}}, a_{\mathrm{L}}), +1)$: mark the cell, and step right.

(T2) $\mathrm{tr}(\mathsf{back}(q, \ell), \gamma) = (\mathsf{sim}(q, \ell), \gamma, -1)$: step left, writing nothing new.

(T3) $\mathrm{tr}(\mathsf{sim}(q, \ell), \gamma)$, where γ has marker b, a is its entry on track ℓ, $\delta(q, a) = (q', a', d)$ and $\gamma' = \gamma[\ell \leftarrow a']$, is

$$
\begin{array}{ll}
(\mathsf{back}(q', \ell), \gamma', +1) & \text{if } d = 0, \\
(\mathsf{sim}(q', \ell), \gamma', +1) & \text{if } (\ell, d) = (\mathrm{U}, +1) \text{ or } (\ell, d) = (\mathrm{L}, -1), \\
(\mathsf{back}(q', \mathrm{L}), \gamma', +1) & \text{if } (\ell, d) = (\mathrm{U}, -1) \text{ and } b = 1, \\
(\mathsf{sim}(q', \mathrm{U}), \gamma', -1) & \text{if } (\ell, d) = (\mathrm{U}, -1) \text{ and } b = 0, \\
(\mathsf{back}(q', \mathrm{U}), \gamma', +1) & \text{if } (\ell, d) = (\mathrm{L}, +1) \text{ and } b = 1, \\
(\mathsf{sim}(q', \mathrm{L}), \gamma', -1) & \text{if } (\ell, d) = (\mathrm{L}, +1) \text{ and } b = 0.
\end{array}
$$

M° has $1 + 4|Q|$ states and $2|\Sigma|^2$ symbols, and it is computed from M. It starts in q_0 with its head on cell 1 of a blank tape on the cells $1, 2, \dots$. Every transition moves the head. A nonzero move of M that does not change tracks is simulated by one step of M°. A stay, and a move between the positions 0 and -1, which changes the track at cell 1, are *bounces*: a step right into a state $\mathsf{back}(q', \cdot)$ and a step back.

Simulation. The machine M° *simulates* the configuration (q, p, y) of M if it is in the state $\mathsf{sim}(q, \ell(p))$ with its head on cell $c(p)$, cell 1 carries the marker 1 and every other cell the marker 0, and for every cell $x \geq 1$ the entries on the tracks U and L are $y(x-1)$ and $y(-x)$.

Lemma B.1 (The one-sided machine). *(a) At time* 2, M° *simulates* C_0. *At time* 1 *it is in the state* $\mathsf{back}(q_{\mathrm{in}}, \mathrm{U})$ *on cell* 2.

(b) If M° *simulates* C_n *at time* t, *then it simulates* C_{n+1} *at time* $t+1$, *or it simulates* C_{n+1} *at time* $t+2$ *and is in a state* $\mathsf{back}(\cdot,\cdot)$ *on a cell* $x \geq 2$ *at time* $t+1$.

(c) There are times $2 = t_0 < t_1 < \dots$ *with* $t_{n+1} - t_n \in \{1, 2\}$ *such that* M° *simulates* C_n *at time* t_n. *At every time* $t \geq 1$ *that is not one of the* t_n, M° *is in a state* $\mathsf{back}(\cdot,\cdot)$ *on a cell* $x \geq 2$.

(d) M° *enters a halting state if and only if* M *halts. If* $q_n \in Q_{\mathrm{h}}$, *the halting time satisfies* $T(M) \leq t_n \leq 2n + 2$.

(e) At every time the head of M° *is on a cell* $x \geq 1$, *and whenever it is on cell* 1 *it moves right. Its initial state is not halting, and its head is on cell* 2 *after the first step.*

Proof. (a) By (T1), at time 1 the cell 1 is marked and the head is on cell 2 in the state $\mathsf{back}(q_{\mathrm{in}}, \mathrm{U})$. By (T2), at time 2 the head is on cell 1 in the state $\mathsf{sim}(q_{\mathrm{in}}, \mathrm{U})$. All entries are blank, only cell 1 is marked, and $\ell(0) = \mathrm{U}$, $c(0) = 1$. So M° simulates C_0.

(b) Let $C_n = (q, p, y)$, $\ell = \ell(p)$ and $x = c(p)$. The scanned symbol γ has on track ℓ the entry $y(\mathrm{pos}(\ell, x)) = y(p)$, and its marker b is 1 exactly when $x = 1$. Let $\delta(q, y(p)) = (q', a', d)$. In every case of (T3), M° writes a' on track ℓ of cell x. Since $(\ell, x) \mapsto \mathrm{pos}(\ell, x)$ is injective, this changes the entry for the position p and no other, as the step of M does; the markers do not change. It remains to check the state and the head. If $d = 0$, the head moves to cell $x + 1 \geq 2$ in the state $\mathsf{back}(q', \ell)$, and by (T2) returns to cell x in the state $\mathsf{sim}(q', \ell)$; the position p is unchanged. If $(\ell, d) = (\mathrm{U}, +1)$, then $p \geq 0$, the new position $p + 1$ is on track U at cell $x + 1$, and the head moves there in $\mathsf{sim}(q', \mathrm{U})$. If $(\ell, d) = (\mathrm{L}, -1)$, then $p < 0$, the new position $p - 1$ is on track L at cell $-(p-1) = x + 1$, and the head moves there. If $(\ell, d) = (\mathrm{U}, -1)$ and $b = 0$, then $x \geq 2$, the new position $p - 1 \geq 0$ is on track U at cell $x - 1$, and the head moves there. If $b = 1$, then $p = 0$ and the new position -1 is on track L at cell 1: the head moves to cell 2 in $\mathsf{back}(q', \mathrm{L})$ and returns to cell 1 in $\mathsf{sim}(q', \mathrm{L})$. The case $(\ell, d) = (\mathrm{L}, +1)$ is symmetric: if $b = 0$, then $p \leq -2$, the new position $p + 1 < 0$ is at cell $x - 1$ on track L; if $b = 1$, then $p = -1$ and the new position 0 is on track U at cell 1, reached by a bounce through $\mathsf{back}(q', \mathrm{U})$ on cell 2. A bounce passes through a state $\mathsf{back}(\cdot,\cdot)$ on the cell $x + 1 \geq 2$, and (T2) writes nothing new.

(c) This follows from (a) and (b) by induction on n.

(d) If $q_n \in Q_{\mathrm{h}}$, then at time $t_n \leq 2 + 2n$ the machine M° is in the state $\mathsf{sim}(q_n, \cdot)$, which is halting; so $T(M) \leq t_n$. Conversely, suppose M° is in a halting state $\mathsf{sim}(q, \ell)$ at time t. Then $t \geq 1$, because q_0 is not halting, and t is one of the t_n by (c), because the state is not a state $\mathsf{back}(\cdot,\cdot)$. So $q = q_n \in Q_{\mathrm{h}}$, and M halts.

(e) At time 0 the head is on cell 1 in q_0, and (T1) moves it right, to cell 2. At a time t_n the head is on the cell $c(p_n) \geq 1$ in a state $\mathsf{sim}(\cdot,\cdot)$; if that cell is cell 1, its marker is 1, and every case of (T3) with $b = 1$ moves right. At any other time $t \geq 1$ the head is on a cell $x \geq 2$ in a state $\mathsf{back}(\cdot,\cdot)$, by (c), and (T2) moves it to the cell $x - 1 \geq 1$. The initial state q_0 is not halting by definition. □

By (e), the head of M° never reaches the wall at cell 0, so the cellular automaton $\mathcal{C}_t$ of Section 8.6 describes M° up to time $T(M)$, and $\mathcal{C}_1(1) \in \Gamma$, as used in the proof of Proposition 8.16(a). By (d), $T(M) < \infty$ exactly when M halts, as used in Proposition 8.16(b),(c) and Proposition 8.18.

C Formal proofs in Lean 4

The machine-family theorem on $\mathbb{Z}^8$

Corollary 11.21, the machine-family theorem on $\mathbb{Z}^8$, is proved in the proof assistant Lean 4 [30] with its mathematical library Mathlib [33], together with the statement that ergodicity in the family is undecidable. This appendix says what is stated, which choices a reader must know to compare the formal statement with the printed one, and what is proved.
This machine-family core is unchanged from version 1.4.0. The natural-rule cores on $\mathbb{Z}^6$ and $\mathbb{Z}^4$ described below retain only the intervals of version 2.2.0. The enlarged version 3.0.0 intervals are paper proofs with exact arithmetic. No Lean build was performed for this text pass; the checks reported below are those of the verification packages.

The statement. A statement file defines, with no proofs of substance, every object that occurs in Corollary 11.21: Wang tiles and Wang sets, legal 3×3 blocks, live symbols and compatibility (Definition 4.3), the checker (Definition 4.5), the quorum rule (Definition 4.7), the replacement noise and the process (Definition 4.8), invariant measures, weak convergence and ergodicity (Definition 4.9), Turing machines and halting (Section 8.6), and the rate $\varepsilon_0 = 490000^{-99}$. Each definition carries a comment that names the definition of this paper it mirrors. The file then states Corollary 11.21 as one proposition: there is a computable map $M \mapsto R_M$ from Turing machines to finite Wang sets such that, for every real ε with $0 < \varepsilon \leq \varepsilon_0$ and every machine M, the noisy quorum rule $F^{[6]}_{R_M}$ on $\mathbb{Z}^2 \times \mathbb{Z}^6$ with uniform replacement at rate ε is a positive-rate probabilistic cellular automaton, and it is ergodic if and only if M halts. The statement file was written before the proof. Its only later change that is not a comment turned the stated theorem into the definition of this proposition, with the text of the statement unchanged; the proof is a theorem about the proposition, in a file that imports the whole development.

Choices in the statement. The formal statement renders the printed one with the following choices.

- The family is quantified existentially, as in the existential machine-family statement of version 1.4.0 and the realization in Corollary 11.21. The development also proves the classification, and the undecidability below, for an encoded realization of the construction of Definition 8.17, for machines encoded as in the third item.
- A Wang set is computed as a finite list of tiles whose colours are natural numbers. Computability of $M \mapsto F^{[6]}_{R_M}$ is rendered as computability of $M \mapsto R_M$; the local rule is then an explicit finite construction from R_M.
- A Turing machine is encoded finitely: its states and tape symbols are natural numbers, the initial state and the blank are 0, the transition function is a finite list, and a pair without an entry is a step that changes nothing. A machine halts if, started on a blank two-sided tape, it enters a halting state. For a machine encoded in this way, the finite state set and the finite tape alphabet from which Section 8.6 and Section B build M°, and with it the payload of Definition 8.17, are both the set of labels below one plus the largest target-state or written-symbol label in the transition table, the maximum being 0 for an empty table. Labels outside this bound, which the run from the blank tape never reaches, are omitted, and unused labels below it are included. This convention preserves the run from the blank tape and its halting.

- "Positive-rate probabilistic cellular automaton" is rendered by four clauses: the rule commutes with the translations of $\mathbb{Z}^8$; it is local, with neighbourhood $\Lambda \times \{c_0, \dots, c_6\}$; given $X_0 = x$, the values $X_1(v)$ are independent with the laws $(1-\varepsilon)\delta_{F(x)(v)} + \varepsilon \cdot$ uniform; and every symbol has positive probability under each of these laws.
- The process is defined through a realisation of the noise, as in Definition 4.8: the law of X_t is the image of the product of the initial law and the law of the noise under the map that runs the process for t steps. Ergodicity is: there is exactly one invariant probability measure π, and for every initial probability measure μ the law of X_t converges to π against every bounded continuous function.
- "Undecidable" is Mathlib's notion: for each fixed $\varepsilon \in (0, \varepsilon_0]$, the property "$F_{R_M}^{[6]}$ is ergodic" is not a computable predicate of M, with machines coded as above. The development proves, in addition, that the machines for which the rule is not ergodic do not form a recursively enumerable set.
- Smaller encodings are recorded in the statement file: $\mathbb{Z}^8$ is written $\mathbb{Z}^2 \times \mathbb{Z}^6$, the provider indices are $\{0, \dots, 6\}$ written as an option type, the checker bit is a Boolean, and the noise of time $t+1$ has index t.

What is proved. The proof follows the paper. It formalizes every step that Corollary 11.21 uses: Section 4, the forgetting and memory theorems of Sections 5 and 6 (Theorem B in the generality of the introduction), the scaffold of Section 7 with the tables of Section A, the family of Section 8, for machines encoded as above, with the machine M° of Section B, the items of Section 9 used in Section 11, Section 11, and the computability of $M \mapsto R_M$. The corner lemma enters in the form of the remark after Corollary 11.13, $\mathcal{K}(0) = \varnothing$ on V_6, proved by a certificate that the Lean kernel checks by evaluation, without compiled code; the form on V_{16} follows. The undecidability is derived from Mathlib's theorem that the halting problem for its partial recursive codes is undecidable, through a compiler from those codes to the machines above. The compiler is built from Mathlib's simulation of partial recursive functions by Turing machines, and its computability is proved. The proof contains no unproved step (`sorry`). An audit of the axioms of the final theorems shows that they use only the three standard axioms of Lean and Mathlib, `propext`, `Classical.choice` and `Quot.sound`, and no external theorem enters as a hypothesis. The 104 base tiles of Table 5 were also compared once with the transcription [31], and they agree.

The formalization found one place where the paper was less detailed than a formal proof needs: the "standard simulation" M° of Section 8.6, now given in Section B. It found no false step. Theorem C in the general form of Theorem 9.8, Theorem 10.2, the other corollaries of Section 10 and Corollary 5.12 are proved in the text only.

Toolchain and release. The proof was developed with Lean v4.30.0 and a Mathlib snapshot of May 2026. For release it was ported to Lean 4.35.0 and Mathlib, with the commits

Lean: `eb6cafb7290e8bd5f2937522d8d804f58f02bafc`,
Mathlib: `56a005855ea17e0b0958f73ec801ae140b0a1679`.

The port changed proof text in 28 of the 1084 theorem and lemma blocks of the source, spread over 14 files, and one import, as the newer Mathlib required. It changed no statement and no definition, and the statement file is byte for byte that of the development. The sources, the pins of every dependency and the records of the checks below are in the directory `lean/` of release v1.4.0 of the repository,

https://github.com/Apsiape/fixed-noise-ergodicity/tree/v1.4.0/lean,

whose `README.md` says how to reproduce them.

- *The comparator.* A challenge file imports only Mathlib. It contains the statement file verbatim, with one more Mathlib import and two commands that set the module name under which its definitions are recorded, so that their internal names agree with those of the development, which it does not import. It then states two theorems and leaves their proofs open: Corollary 11.21, as the proposition of the statement file, and Corollary 11.21 together with its undecidability, in the vocabulary of the statement file and of Mathlib: for every $\varepsilon \in (0, \varepsilon_0]$, ergodicity of the automata of the family is not a computable predicate of M, and non-ergodicity is not recursively enumerable. The comparator was run on Linux inside a bubblewrap sandbox, with the command

  ```
  lake comparator --config comparator.json.
  ```

  It made its own exports of the challenge and of a solution file that imports the proofs, checked that the two proved theorems have exactly the statements of the challenge, with the same definitions, checked that their proofs use no axiom other than `propext`, `Classical.choice` and `Quot.sound`, and replayed the proofs in Lean's kernel. It accepted both theorems.

- *Two kernel checkers.* An export of the solution contains the two theorems of the challenge, the theorems that for the encoded family R_M of the development ergodicity is not a computable predicate and non-ergodicity is not recursively enumerable, the theorem that halting of encoded machines is not a computable predicate, and every declaration on which they depend, those of Mathlib included. Two kernel checkers re-checked every declaration of this export and accepted it: `leanchecker` of Lean 4.35.0, which runs Lean's own kernel on its own, and `nanoda`, a separate implementation of Lean's type checker, set to reject any axiom other than the three above.

- *Axioms and admissions.* The axiom audit of the build prints the axioms of 437 named theorems, among them every final one, and checks the axioms of all 4609 declarations of the development, private and generated ones included. Only `propext`, `Classical.choice` and `Quot.sound` occur. For the final theorems it prints

  ```
  'FNE.cor_eight' depends on axioms: [propext, Classical.choice, Quot.sound]
  'FNE.cor_eight_undecidable_statement' depends on axioms: [propext, Classical.choice, Quot.sound]
  'FNE.cor_eight_undecidable' depends on axioms: [propext, Classical.choice, Quot.sound]
  'FNE.nonergodic_not_re' depends on axioms: [propext, Classical.choice, Quot.sound]
  'FNE.halts_not_computable' depends on axioms: [propext, Classical.choice, Quot.sound]
  ```

  Apart from the two open proofs of the challenge, no proof uses `sorry`, and no file declares an axiom or uses `native_decide`, `implemented_by` or `extern`.

- *Dependencies.* The 8624 Lean sources of Mathlib's library agree byte for byte with the official commit above. The build reused prebuilt Mathlib objects; the two kernel checkers re-check every declaration of Mathlib that the proofs use.

The natural rule on $\mathbb{Z}^6$

The statements. A second development formalizes the natural Wang rule of Section 2. Its statement file, `ForgettingNaturalRule/Statements.lean`, defines the cardinal checker, the five provider boxes, the quorum, the neighbourhood and the noise with the version 2.2.0

constants, in particular $\varepsilon_* = (2(18q)^2)^{-200}$. The present $\varepsilon_* = 10^{-24}$ is outside the formal statement. The alphabet is an option type on R, identified with a finite numbered alphabet; this relabelling does not change uniform noise or ergodicity. Tiles, finite lists of tiles and encoded machines are those of the existing machine-family core. The process, probability laws, ergodicity and translation invariance are those of the noise-diagram core [12]. Translation invariance means invariance under every translation of $\mathbb{Z}^6$, exactly as in Section 2.

There are five release targets. Each has no named hypothesis and is proved without `sorry`.

- `FNR.natural_classification` proves both sentences of Theorem D on the version 2.2.0 interval $0 < \varepsilon \le (2(18q)^2)^{-200}$: ergodicity if and only if R does not tile, and two distinct translation-invariant stationary laws when it does. This is Theorem 2.1 without the coalescence bound (9).
- `FNR.theorem_A` proves Theorem A on $\mathbb{Z}^6$ as printed in version 2.2.0, for $0 < \varepsilon \le (2(18q)^2)^{-200}$, the second sentence of Corollary 3.1 on that interval.
- `FNR.machine_corollary` proves all three sentences of Corollary 3.1 on the version 2.2.0 interval $0 < \varepsilon \le (2(18q)^2)^{-200}$: the computable reduction, the ergodicity equivalence for the same map, and undecidability at each fixed rate in that interval.
- `FNR.domino_reduction` proves the domino reduction, the first sentence of Corollary 3.1, for the encoded machine model of the existing core.
- `FNR.domino_undecidable` proves that the domino problem is undecidable, the classical theorem of Berger and Robinson [3, 43].

An algorithm here is a computable function in Mathlib's sense; tile sets are supplied as finite lists. The statement file was read blind against the printed paper by a reader who did not write it. That read found one mismatch: a conditional form of Theorem A had been given the name of the unconditional printed theorem. Before the file was frozen, `FNR.TheoremA` was corrected to the unconditional ergodicity reduction. The conditional forms remain in the development, outside the five release targets. The corrected statement file was then frozen.

What is not formalized. The flag-position lemmas and constants of version 3.0.0 are not formalized. The coalescence bound (9) and its all-input coefficient (7) are outside this core. So are Lemmas 2.8 to 2.10 and the printed quantitative estimates used only for that bound: the sparse-root and marked-path bounds are proved only in the weaker or qualitative forms described below, and uniqueness and attraction use the qualitative route. Nilpotency (Corollary 3.2) and every other consequence of Section 3, including Gibbs uniqueness, robustness, complexity completeness, replacement values, forgetting times, the uncountable-tiling theorem and the $\mathbb{Z}^7$ stack, are not formalized. Nor does this development extend the scope of the existing machine-family core in Sections 4 to 11. The open problems, the account of the work and the appendices are not formalized.

Proof route. The statements agree with the printed ones within that scope. The proof has the following differences, recorded as F1–F9 in the development's scope account.

F1. The non-tiling direction follows the qualitative complete-history route of Section 2, using the noise-diagram core's Lemma 8.5 for ergodicity [12]. Coherent letters of an unmarked live complete history directly give a tiling; the quantitative forbidden-square route is not used.

F2. The sparse-root estimate is weakened to $\mathbb{P}(v \in \mathcal{M}) \le \sum_{f \ge 1} (K_c \varepsilon)^f \le 1/4$, rather than the printed (15). The resulting memory gap is $1/2$, rather than $1 - 2p_*$, which suffices for nonergodicity and the two laws.

F3. Coherence is proved under the uniform age condition $s + d < t$ for displacements $0, e_1, e_2$, and applied to complete trajectories, where any finite age is available.

F4. The marked field is literally the raw marked set (11) of the noise-diagram core. Its contour, packing and counting lemmas are reused with $K_c \varepsilon \le 1/8$; only the qualitative conclusion that almost surely there is no infinite marked path is formalized.

F5. The memory branch runs both starts with two-sided noise from time zero, and identifies that process with the one-sided process.

F6. Memory between two Dirac starts gives nonergodicity. The existence of the translation-invariant stationary laws is proved as a separate clause.

F7. For those laws, temporal Cesàro limits come first, retaining the one-site bounds at every site. Each stationary law is then averaged along the six coordinate directions in turn. This reverses the order of the paper's averages; one-dimensional telescoping sums replace averaging over planar squares, the tiling space and its boundary count.

F8. The domino reduction uses this paper's computable family R_M, through the existing core's realization of Theorem C, rather than Berger's or Robinson's construction. The statement of the reduction does not name a construction.

F9. For this family, the implication from refutability to non-tiling uses Theorem B at the machine core's rate ε_0: halting makes the quorum rule ergodic, whereas a tiling set makes it nonergodic. It does not use a separate combinatorial lemma. This is the argument after Corollary 3.1.

Reused cores. The proof reuses three precedent developments on the same pins: this paper's version 1.4.0 machine-family core [10], the noise-diagram core of [12], published as the `lean/` folder of its version 1.2.0 [13], and the hard-model core of version 1.2.0 of [14]. The hard-model core is reached through the noise-diagram core. The latter supplies the PCA and marked-field layers; the former machine-family core supplies the computable family R_M, its tiling and refutation results, and the noncomputability of halting. The package rebuilds the imported sources: 61 machine-family modules, 23 noise-diagram modules and one hard-model module, together with the 16 modules of the natural-rule development and the three package modules `Challenge`, `Solution` and `Audit`, 104 reachable modules in all.

Pins and checks. The pins are Lean 4.35.0 and the same Lean and Mathlib commits printed above. The five targets use only `propext`, `Classical.choice` and `Quot.sound`. The fresh serial source replay, lint, admission census and strict axiom audit passed. The challenge imports only Mathlib and contains the frozen statements, including the reused vocabulary, with five intended proof holes. The sandboxed comparator accepted all five targets with its default kernel and no supplied exports. An export of their proof closure passed both `leanchecker` and `nanoda`; the latter was restricted to the three standard axioms. A negative control omitting `Classical.choice` was rejected. Dependency-source checks and comparisons with the precedent cores also passed. These are the recorded package checks; no Lean process was run for this text pass.

Code and reproduction. The sources and check records are in `lean-natural/` of this paper's release repository, beside `lean/` for the existing $\mathbb{Z}^8$ core:

`https://github.com/Apsiape/fixed-noise-ergodicity`.

The file `README-ARXIV.md` gives the reproduction procedure. The `vendor/` directory is fetched, not shipped: obtain the pinned dependency packages and the three precedent source cores. The serial `scripts/replay.py` rebuilds the reachable modules; the lint and census scripts check the sources; `lake comparator --config comparator.json` performs the sandboxed comparison. The documented export and kernel commands replay the five targets with `leanchecker` and `nanoda`, including the negative control. Mathlib objects are reused rather than rebuilt; the two standalone kernels check the exported proof closure, including the Mathlib declarations it uses. The noise-diagram core is public as the `lean/` folder of version 1.2.0 of that paper [13]; its Lean sources are byte-identical to the ones used in these checks, so the package can be reproduced from public sources.

The natural rule on $\mathbb{Z}^4$

The statements. A third development formalizes the qualitative natural raw-Wang classification of Theorem 12.5, stated as Theorem E in the introduction, and its direct undecidability corollaries. The statement of record is the frozen `ForgettingLower/Statements.lean`. It reuses the earlier cores' tiles, list-coded inputs, encoded machines, uniform whole-symbol replacement, probability laws, stationarity, ergodicity and weak convergence from every initial probability law. Its constants are the three boxes at $0, 160e_1, 160e_2$, $K = 256$, $q = 768$, $h = 512$ and $e_0 = 382205952^{-33720}$ of version 2.2.0; the present value 10^{-30} is proved on paper with exact arithmetic. The common neighbourhood is literally $\Gamma \times S$. The checker reads the five cardinal letters and checks all four shared edges. The alphabet is R plus blank, relabelled by a finite numbered alphabet.

There are three release targets, each proved with no named hypothesis and without `sorry`.

- `FLD.z4_classification`, of type `FLD.Z4Classification`, proves every qualitative sentence of Theorem 12.5 on the version 2.2.0 interval: the common neighbourhood, ergodicity exactly for non-tiling inputs throughout $0 < e \le 382205952^{-33720}$, two distinct invariant probability laws on tiling inputs, and finite-window all-start coupling from the past and attraction from every initial law on non-tiling inputs. It adds no spatial translation-invariance condition to the two laws.
- `FLD.z4_machine_corollary`, of type `FLD.Z4MachineCorollary`, proves one computable map from machines to tile lists, its domino specification, positive rates and ergodicity exactly for halting machines at every rate allowed in version 2.2.0, and the nonexistence of a computable ergodicity decider on that machine family.
- `FLD.z4_raw_undecidable`, of type `FLD.Z4RawUndecidable`, proves fixed-rate undecidability on the raw family $R \mapsto G_R^{(4)}$ for every rate allowed in version 2.2.0, in particular at 382205952^{-33720}, the version 2.2.0 form of the statement at the end of the proof of Theorem 12.5.

Choices and fidelity. The frozen statement file passed a blind comparison with the version 2.2.0 statement of Theorem 12.5, with verdict *FAITHFUL* and no mismatch or vacuity. Coupling from the past places all integer observation times and finite windows inside one almost-sure quantifier, uniformly over starts and sufficiently long horizons; countability and

time translation make this equivalent to the printed form. The machine proposition strengthens the version 2.2.0 fixed-rate sentence to every rate allowed in version 2.2.0 and adds the machine family and positive rates. Before the freeze, the literal raw-family proposition was added, two draft proof stubs were removed, and encoding choices were marked. The final statement file has SHA-256

`58cb7388400f05f2986d2b2cc9154a99a554eed4b384f10c8809b706d3b244c4`.

Proof route and scope. The proofs formalize the three-box contour and all-path density machinery, amortized same-site coherence, separate planar edge matching and extraction, uniqueness of complete histories, coupling from the past, and the memory construction. Tiling is extracted directly from coherent readings of a complete history. Sparse roots are bounded by $1/4$ and the memory gap by $1/2$; the printed intermediate bound $< 2e$ and gap $1 - 4e$ are not claimed. These weaker estimates suffice for the frozen qualitative classification. Temporal Cesàro limits give the two invariant laws, with no spatial averaging asserted in this target. The machine corollary reuses the proved, unconditional `FNR.domino_reduction`.

The $\mathbb{Z}^3$ family of Theorem 12.11, Kari's theorem, the rate and effective prefactor of Theorem 12.12, the new all-Gibbs corollary Corollary 12.14, the saturation theorem and every other result outside the three frozen targets are not formalized by this package. In particular this core does not establish the quantitative LIVE bounds used in the proof of Corollary 12.14. The earlier natural-rule and machine-family cores retain their stated scopes.

Pins and checks. The recorded package uses Lean 4.35.0 and the same Lean and Mathlib commits printed above. Its fresh serial replay accepted 133 modules: 31 lower-dimensional modules, 99 precedent-core modules and three package modules. The precedent split is 61 machine-family, 23 noise-diagram, 14 natural-rule and one hard-model module. A reachable-constant comparison agreed on 116 challenge and development constants, including kinds, type hashes and definition-value hashes. The bubblewrap-sandboxed comparator accepted the three targets with its default kernel and no supplied export. Their exported proof closure passed `leanchecker` and `nanoda`, the latter restricted to `propext`, `Classical.choice` and `Quot.sound`; the negative control omitting `Classical.choice` was rejected. The fresh axiom audit covered 206 strict roots, 464 lower-dimensional declarations and 7009 imported core declarations, including generated and private declarations, with only those three axioms. The admission census and source audit found no holes in the 31 original files; the three challenge holes were intentional. Dependency-source checks and byte comparisons with the four public precedent trees passed. These are recorded package checks; no Lean process was run for this writing revision.

Code and reproduction. The package will ship in `lean-lower/` beside `lean-natural/` and `lean/` in this paper's release repository. The recorded verified package commit is `a471ef3`; the final release tag or commit for the standard Lake layout is the tag `v2.2.0` of this release repository. The package's `README-ARXIV.md`, `REPORT.md`, `SCOPE.md` and `STATEMENT-READ.md` give reproduction, check receipts, exact scope and the blind statement read. The four precedent cores are public source dependencies; the checked imported modules were rebuilt from source. The nine pinned dependency packages use prebuilt objects whose source identities were checked, and the two standalone kernels replay the exported proof closure. The final layout conversion is separate from the verification recorded here; no new replay of that final layout is claimed in this writing revision.

D Proof of the totality source

Proof of Lemma 14.2. Input-erasure and output-erasure trajectories are respectively the trajectories of the nondeterministic automata $\hat{F}$ and $\tilde{F}$ attached to a CA F in [55, Section 2]; simulations and their rigidity are [55, Definitions 7–8]. Besides the statements named in C_T, the construction uses from [55] composition of rigidity (Lemma 11), sparse simulation and its parents and sparsity properties (Lemmas 14–19), the resource hypotheses of Lemma 24, the amplifier of Theorem 30, chain rigidity (Lemma 31), Claims 33–34 in Lemma 32, and the parent and covering statements of Lemma 35 and Corollary 36, which are printed without proof. The parameter and density arguments below supply the needed versions of that construction and of the argument surrounding Theorem 37.

Run $\varphi_e(0), \varphi_e(1), \dots$ sequentially and report after each completed run. Put

$$N_e = \{2^t : \text{a report occurs at step } t\}.$$

Membership is decidable in time polynomial in binary input length: reject nonpowers of two and otherwise simulate only $\log_2 n$ steps. The set is infinite exactly on totality. Apply the sparse-amplifier construction with effective levels $n \in N_e$.

Use actual colony length Q_n and work period U_n in the local program; the colony-length minus one in the printed Algorithm 7 is replaced by Q_n. A field with inclusive upper bound M has $\lceil \log_2(M+1) \rceil$ bits. Normalize invalid encodings only in that layer's own bookkeeping fields; every simulated-data bit remains legal data. Undefined agent transitions produce Die. Compile these tests and cleaned neighbour reads into each template before self-reference. Reserve the all-zero enum code for blank or no-kind, so an all-blank neighbourhood cannot satisfy a live kind test. In every universal template execute the final age increment only if Live is still true. Birth clears all fields before leaving blank or making a live birth. Every output is a valid own-field word, and every nonblank two-sided symbol is live. These changes preserve honest simulation and demanding paths. A changed output is blank, which is available under input erasure by erasing all inputs; the published hat-trajectory implications therefore apply with the condition stated above.

For $h(n) = \lceil \log_2(n+2) \rceil$ choose

$$Q_n = 2^{H+Bh(n)}, \qquad U_n = 2^{H+(B+6)h(n)}.$$

Symbolic polynomial composition of the guarded programs gives effective K, k, independent of the numerical H, B, bounding both interpreter requirements by $K(1+H+B)^k(n+2)^k$. Padding is confined to inert read-only parameters. Choose $B \geq \max(k, 1)$ and then $H \geq 4$ by finite search with $2^H \geq K(1+H+B)^k$. Exponential domination makes the search terminate. Then Q_n bounds the interpreter space and time, $Q_n \geq 32$, $4 \mid Q_n$, $U_n/Q_n \geq 64$ and $U_n \geq 6Q_n +$ required time. [55, Lemma 24] requires that Q_n bound the interpreter space and that $U_n \geq 6Q_n$ plus the interpreter time; the conditions $Q_n \geq 32$, $4 \mid Q_n$ and $U_n \geq 64Q_n$ are used in the overlap argument below. No totality test enters these choices. Choose the separate final universal wrapper with the same conventions and sufficiently large powers of two.

Write B_n, W_n for composed simulation widths and periods, and $Q'_n = Q_n M_n$, $U'_n = U_n M_n$, where $M_n = 2n+1$ at effective levels and 1 otherwise. The needed absolute overlap statement is: two S_n macrocells in one output-erasure trajectory with anchor differences at most $2B_n/3$ and $2W_n/3$ in absolute value have the same anchor. Here is the absolute form of the crossing argument. Order the macrocells by time. In the earlier one choose the vertical demanding column at $3Q/4$ or $Q/4$, according as the later one is to its right or left. The later returning

sweep runs from $((Q-1)M, QM)$ to $(0, 2QM)$ relative to its anchor; its sparse starting coordinate is $(Q-1)M$. The hypotheses $Q \geq 32$, $U \geq 64Q$ give strict horizontal crossing margins and keep the sweep inside the earlier column's safe time interval, including a delay $2U/3$.

Induct on the composed level. At levels above one, Claim 34 supplies a previous-row white macrocell under each black path vertex. At a polygonal crossing of their lower grids, a black and white anchor have absolute normalized offsets at most $1/2$ on each axis. Indeed compare the two crossing steps' bottom times. If their gap is at most $1/2$, intersecting intervals of length zero or one have an endpoint pair within $1/2$. If the gap exceeds $1/2$, use the earlier step's upper and the later step's lower black endpoints; both gaps are at most one minus that time gap. Absolute induction aligns the lower grids. Crossing black paths are then weakly connected, since a diagonal swap has a vertical cross-edge. For counter c, the conserved sparse phases are

$$t - c \pmod{M}, \qquad \begin{cases} i - c & \text{Right}, \\ i + c & \text{Left}, \\ i + c + 1 & \text{Returning} \end{cases} \pmod{M}.$$

Each nonblank output shares these phases with every nonblank actual predecessor, by the branches of the sparse rule. At the composed induction step M is M_{n-1}, the current sparse factor. Phase conservation forces divisibility by that factor; the absolute universal Corollary 26 then identifies the parent anchors. At the bottom level lower-grid divisibility is automatic. This proves the absolute overlap assertion used in the following packing.

Fix an effective $n \geq 1$ and a bottom-column time slab of length $W_{n+1} = W_n U'_n$. Sparse ancestry takes at most $2n$ steps. Its parent offset lies in $[-2n, 2n]$, correcting the printed narrower window. Universal live coverage, followed by the sparse connecting lift, assigns each nonblank lower S_n macrocell to an actual higher S_{n+1} parent in the same decoded output-erasure trajectory. Its relative anchor is in

$$[-3Q'_n, 2Q'_n] \times [-3U'_n, 0].$$

Bottom coverage uses lower anchors in

$$[-4B_n, 3B_n] \times [\ell W_{n+1} - 4W_n, (\ell+1)W_{n+1} - 1].$$

The higher parents lie in a box of width $12B_{n+1}$ and height $8W_{n+1}$. Absolute overlap allows at most one parent per half-size bin, so at most $25 \cdot 17 = 425 < 1024$ parents occur. Group lower macrocells by their actual parent. They lie on its grid and, by the constructed paths, are connected to its same source base in that one decoded trajectory. Grid congruence alone would not imply this connection.

In a fixed sparse weak component, phase conservation gives, on a fixed spatial column, at most $3r/M_n + 3 \leq 3r/n + 3$ cells in r successive source times. The lower-anchor box has at most eight coarse columns and $U'_n + 5$ coarse rows, so one parent contributes at most

$$8\big[3(U'_n + 5)/n + 3\big] \leq 168U'_n/n$$

lower cells. Here $U'_n \geq n + 5$ follows from the chosen periods. Each lower anchor covers at most $5W_n$ bottom times. The nonblank fraction in the original slab is therefore at most

$$\frac{5 \cdot 168 \cdot 1024}{n} < \frac{2^{20}}{n}.$$

This bound is uniform over trajectories, slabs and columns; it is used in place of the printed Claim 39. On totality the effective n are unbounded. Tile long column intervals by these slabs and discard a bounded remainder. Every two-sided output-erasure trajectory has zero temporal upper nonblank density in each column. For a TI law, bounded convergence therefore gives zero nonblank probability.

Put one strongly rigid universal wrapper around the bottom amplifier. By strong rigidity, any input-erasure wrapper trajectory, on each fixed grid offset, decodes to an output-erasure amplifier trajectory. Its decoded TI law is blank on totality. Unit translation invariance gives zero activity at every anchor offset. Lemma 29 covers each live wrapper cell by a nonblank macrocell at one of finitely many nearby anchors. The finite union has probability zero. The live guard and valid-word convention identify all nonblank symbols with live ones, proving (i).

If N_e is finite, choose a level above its last effective member. All higher links are universal. A completely nonblank row at arbitrarily high level K lifts honestly to the fixed level m with every cell live for $\prod_{j=m}^{K-1} U_j$ steps; a colony dies only at a work-period end if the next simulated row becomes blank. These horizons diverge. Compactness gives an all-live forward level-m trajectory. The finite simulations down to the bottom and through the final wrapper place a nonblank base in every fixed $B \times W$ block. Average over growing space–time boxes whose time windows recede from the initial boundary. A subsequential limit is a fully TI two-sided ordinary trajectory law of activity at least $1/(BW) > 0$. This proves (ii). □

References

[1] A. Atserias, A. Bulatov and V. Dalmau. On the power of k-consistency. In *Automata, Languages and Programming (ICALP 2007)*, LNCS 4596, Springer (2007), 279–290.

[2] L. Barto and M. Kozik. Constraint satisfaction problems solvable by local consistency methods. *J. ACM* 61 (2014), no. 1, Article 3. doi:10.1145/2556646.

[3] R. Berger. The undecidability of the domino problem. *Mem. Amer. Math. Soc.* 66 (1966). doi:10.1090/memo/0066.

[4] P. Berman and J. Simon. Investigations of fault-tolerant networks of computers. In *Proceedings of STOC 1988*, ACM (1988), 66–77. doi:10.1145/62212.62219.

[5] A. Bušić, J. Mairesse and I. Marcovici. Probabilistic cellular automata, invariant measures, and perfect sampling. *Adv. Appl. Probab.* 45 (2013), 960–980.

[6] G. Choquet. Existence et unicité des représentations intégrales au moyen des points extrémaux dans les cônes convexes. *Séminaire Bourbaki*, exposé 139 (1956–1958), 33–47.

[7] K. Čulik II, J. Pachl and S. Yu. On the limit sets of cellular automata. *SIAM J. Comput.* 18 (1989), 831–842. doi:10.1137/0218057.

[8] V. Dalmau, Ph. G. Kolaitis and M. Y. Vardi. Constraint satisfaction, bounded treewidth, and finite-variable logics. In *Principles and Practice of Constraint Programming (CP 2002)*, LNCS 2470, Springer (2002), 310–326.

[9] S. Douglas and N. Mghirbi. Exponential ergodicity and the space–time Gibbs states of one-dimensional probabilistic cellular automata. Preprint, version 1.2.0, Zenodo (2026). doi:10.5281/zenodo.24122307. Numbers of results refer to this version.

[10] S. Douglas and N. Mghirbi. Forgetting under noise is undecidable. Preprint, version 1.4.0, Zenodo (2026). doi:10.5281/zenodo.24126025. Machine-family Lean core.

[11] S. Douglas and N. Mghirbi. A nilpotent cellular automaton that remembers under noise. Preprint, version 1.2.0, Zenodo (2026). doi:10.5281/zenodo.24120279.

[12] S. Douglas and N. Mghirbi. One noisy cellular automaton with an uncomputable noise diagram. Preprint, version 1.1.0, Zenodo (2026). doi:10.5281/zenodo.24152233. All numbered locators refer to this version.

[13] S. Douglas and N. Mghirbi. One noisy cellular automaton with an uncomputable noise diagram. Preprint, version 1.2.0, with its Lean formalization, Zenodo (2026). doi:10.5281/zenodo.24174445.

[14] S. Douglas and N. Mghirbi. An uncomputable phase diagram at positive temperature for one classical lattice model. Preprint, version 1.2.0, Zenodo (2026). doi:10.5281/zenodo.24144192.

[15] R. L. Dobrushin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory Probab. Appl.* 13 (1968), 197–224.

[16] R. L. Dobrushin and S. B. Shlosman. Constructive criterion for the uniqueness of Gibbs field. In *Statistical Physics and Dynamical Systems*, Progress in Physics 10, Birkhäuser (1985), 347–370.

[17] T. Feder and M. Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory. *SIAM J. Comput.* 28 (1998), 57–104. doi:10.1137/S0097539794266766.

[18] P. Gács. Reliable cellular automata with self-organization. *J. Stat. Phys.* 103 (2001), 45–267; revised version arXiv:math/0003117v2 (2024).

[19] P. Gács. A new version of Toom's proof. Preprint, arXiv:2105.05968 (2021).

[20] P. Gács and J. Reif. A simple three-dimensional real-time reliable cellular array. *J. Comput. System Sci.* 36 (1988), 125–147. doi:10.1016/0022-0000(88)90024-4.

[21] S. Goldstein, R. Kuik, J. L. Lebowitz and C. Maes. From PCA's to equilibrium systems and back. *Comm. Math. Phys.* 125 (1989), 71–79.

[22] A. Halawi. *Asymptotic behaviour of probabilistic cellular automata*. MS thesis, American University of Beirut (2022). `https://scholarworks.aub.edu.lb/handle/10938/23549`. Theorem 4.2.3 and Proposition 4.2.4.

[23] E. Helly. Über Mengen konvexer Körper mit gemeinschaftlichen Punkten. *Jahresber. Deutsch. Math.-Verein.* 32 (1923), 175–176.

[24] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* 58 (1963), 13–30. doi:10.1080/01621459.1963.10500830.

[25] P. Jeavons, D. Cohen and M. Gyssens. Closure properties of constraints. *J. ACM* 44 (1997), 527–548.

[26] J. Kari. The nilpotency problem of one-dimensional cellular automata. *SIAM J. Comput.* 21 (1992), 571–586. doi:10.1137/0221036.

[27] A. S. Kechris. *Classical Descriptive Set Theory*. Graduate Texts in Mathematics 156, Springer (1995).

[28] Ph. G. Kolaitis and M. Y. Vardi. A game-theoretic approach to constraint satisfaction. In *Proceedings of AAAI 2000*, AAAI Press (2000), 175–181.

[29] G. L. Kurdyumov. An algorithm-theoretic method for the study of homogeneous random networks. In *Locally Interacting Systems and Their Application in Biology*, Lecture Notes in Math. 653, Springer (1978), 37–55.

[30] L. de Moura and S. Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, LNCS 12699, Springer (2021), 625–635.

[31] *lean-wang*: a Lean formalization of results on Wang tilings, GitHub repository `https://github.com/edemaine/lean-wang`, directory `LeanWang/Robinson/Closed104`.

[32] J. L. Lebowitz, C. Maes and E. R. Speer. Statistical mechanics of probabilistic cellular automata. *J. Stat. Phys.* 59 (1990), 117–170.

[33] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, ACM (2020), 367–381.

[34] I. Marcovici, M. Sablik and S. Taati. Ergodicity of some classes of cellular automata subject to noise. *Electron. J. Probab.* 24 (2019), paper no. 41, 1–44. doi:10.1214/19-EJP297.

[35] I. Marcovici and S. Taati. Positive-rate PCA and IPS with stationary Bernoulli measures are rapidly forgetful. Preprint, arXiv:2605.16904 (2026).

[36] H. Marsan, M. Sablik and I. Törmä. A perturbed cellular automaton with two phase transitions for the ergodicity. *J. Stat. Phys.* 193 (2026), 6. doi:10.1007/s10955-025-03564-0.

[37] J. Miękisz and S. Taati. Quasicrystalline Gibbs states in 4-dimensional lattice-gas models with finite-range interactions. *J. Stat. Phys.* 193 (2026), 166. doi:10.1007/s10955-026-03727-7.

[38] B. Mossé. Puissances de mots et reconnaissabilité des points fixes d'une substitution. *Theoret. Comput. Sci.* 99 (1992), 327–334.

[39] S. Mozes. Tilings, substitution systems and dynamical systems generated by them. *J. Analyse Math.* 53 (1989), 139–186.

[40] N. Ollinger. Two-by-two substitution systems and the undecidability of the domino problem. In *Logic and Theory of Algorithms (CiE 2008)*, LNCS 5028, Springer (2008), 476–485.

[41] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* 9 (1996), 223–252.

[42] J. Radon. Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten. *Math. Ann.* 83 (1921), 113–115. doi:10.1007/BF01464231.

[43] R. M. Robinson. Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.* 12 (1971), 177–209.

[44] J. Slawny. Ergodic properties of equilibrium states. *Comm. Math. Phys.* 80 (1981), 477–483.

[45] B. Solomyak. Nonperiodicity implies unique composition for self-similar translationally finite tilings. *Discrete Comput. Geom.* 20 (1998), 265–279.

[46] J. M. Swart, R. Szabó and C. Toninelli. Peierls bounds from Toom contours. *J. Theor. Probab.* (2026). doi:10.1007/s10959-026-01508-2; arXiv:2202.10999.

[47] A. L. Toom. Stable and attractive trajectories in multicomponent systems. In *Multicomponent Random Systems*, R. L. Dobrushin and Ya. G. Sinai (eds.), Adv. Probab. Related Topics 6, Dekker, New York (1980), 549–575.

[48] A. Toom. Algorithmic unsolvability of the ergodicity problem for binary cellular automata. *Markov Process. Related Fields* 6 (2000), 569–577.

[49] J. van den Berg and J. E. Steif. On the existence and nonexistence of finitary codings for a class of random fields. *Ann. Probab.* 27 (1999), 1501–1522. Theorem numbers as in the preprint version.

[50] H. Wang. Proving theorems by pattern recognition II. *Bell System Tech. J.* 40 (1961), 1–41.

[51] J. Kari. Theory of cellular automata: A survey. *Theoret. Comput. Sci.* 334 (2005), 3–33. doi:10.1016/j.tcs.2004.11.021.

[52] S. Douglas and N. Mghirbi. A Rice theorem for stationary states under fixed noise. Preprint, version 1.0.0, Zenodo (2026). doi:10.5281/zenodo.24241164. Lemma 6.1; Theorem C, $r = 1$.

[53] L. Gayral, M. Sablik and S. Taati. Characterisation of the set of ground states of uniformly chaotic finite-range lattice models. *Nonlinearity* 38 (2025), 075011. doi:10.1088/1361-6544/addbb8; arXiv:2302.07326v2. Propositions 14–15.

[54] H. Sandhu. Finite Certificates and Undecidable Gibbs Uniqueness. Zenodo preprint, version 1.0, 6 November 2026. doi:10.5281/zenodo.23983514.

[55] I. Törmä. A uniquely ergodic cellular automaton. *J. Comput. System Sci.* 81 (2015), 415–442. doi:10.1016/j.jcss.2014.10.001; arXiv:1310.0670v2. Numbered locators refer to that author manuscript.