The Annals of Probability
2017, Vol. 45, No. 5, 2867–2898
DOI: 10.1214/16-AOP1127

FINITARY COLORING

BY ALEXANDER E. HOLROYD, ODED SCHRAMM AND DAVID B. WILSON

Microsoft Research

Suppose that the vertices of $\mathbb{Z}^d$ are assigned random colors via a finitary factor of independent identically distributed (i.i.d.) vertex-labels. That is, the color of vertex v is determined by a rule that examines the labels within a finite (but random and perhaps unbounded) distance R of v, and the same rule applies at all vertices. We investigate the tail behavior of R if the coloring is required to be proper (i.e., if adjacent vertices must receive different colors). When $d \geq 2$, the optimal tail is given by a power law for 3 colors, and a tower (iterated exponential) function for 4 or more colors (and also for 3 or more colors when $d = 1$). If proper coloring is replaced with any shift of finite type in dimension 1, then, apart from trivial cases, tower function behavior also applies.

1. Introduction. A *q-coloring* of $\mathbb{Z}^d$ is a random element $X=(X_v)_{v\in\mathbb{Z}^d}$ of $\{1,\dots,q\}^{\mathbb{Z}^d}$ that assigns distinct colors to neighboring sites; that is, almost surely $X_u \neq X_v$ whenever $|u - v| = 1$, where $|\cdot|$ is the 1-norm on $\mathbb{Z}^d$. We say that X is a *factor* of an i.i.d. process if it can be expressed as $X = F(Y)$ for some family of i.i.d. random variables $Y = (Y_v)_{v\in\mathbb{Z}^d}$ and some measurable map F that is translation-equivariant (i.e., that commutes with the action of every translation of $\mathbb{Z}^d$). We say that X is a *finitary* factor of an i.i.d. process, or simply that X is *ffiid*, if furthermore, for almost every y (with respect to the law of Y) there exists $r < \infty$ such that whenever y' agrees with y on the ball $B(r) := \{v \in \mathbb{Z}^d : |v| \leq r\}$, the resulting values assigned to the origin $0 \in \mathbb{Z}^d$ agree, that is, $F(y')_0 = F(y)_0$. In that case, we write $R(y)$ for the minimum such r, and we call the random variable $R = R(Y)$ the *coding radius* of the factor. In other words, in an ffiid coloring, the color at the origin can be determined by examining the i.i.d. variables within distance given by the coding radius (which is a finite but perhaps unbounded random variable).

We focus on the questions: for which q and d does an ffiid q-coloring of $\mathbb{Z}^d$ exist, and what can be said about the tail behavior of its coding radius? As a motivating example before stating our main results, we briefly describe a simple construction of an ffiid 4-coloring of $\mathbb{Z}^2$ whose coding radius has exponential tail decay; see Figure 1 for an illustration. Let $(B_v)_{v\in\mathbb{Z}^2}$ be i.i.d. labels taking values $+$ and $-$ with equal probabilities. Since the critical probability for site percolation

Received June 2015; revised May 2016.
MSC2010 subject classifications. 60G10, 05C15, 37A50.
Key words and phrases. Coloring, finitary factor, tower function, shift of finite type.

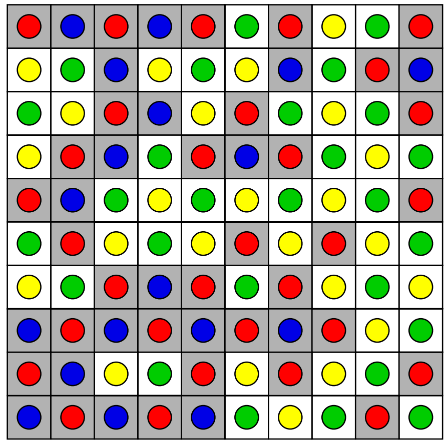

FIG. 1. *An ffiid* 4*-coloring of* $\mathbb{Z}^2$ *whose coding radius has exponential tails. Each* (*subcritical*) *site percolation cluster is assigned a checkerboard coloring.*

is greater than $\frac{1}{2}$, almost surely all $(+)$-clusters and $(-)$-clusters are finite. Next, we color each $(+)$-cluster with colors 1 and 2 in a checkerboard pattern. To ensure translation-equivariance, the phase of the checkerboard must be chosen locally. Here is one way to do this. Assign color 1 to the lexicographically largest site w in the $(+)$-cluster, and also to all other sites v in the cluster for which the sum of the coordinates of $w - v$ is even; assign the remaining sites in the cluster color 2. Checkerboard the $(-)$-clusters with colors 3 and 4 in the same manner. The resulting 4-coloring is ffiid. To determine the color of the origin, we must examine the labels B_v in its cluster and its boundary. Since the radius of the cluster has exponential tails, so does the coding radius.

In fact, much faster decay than exponential is possible in many cases, while only a power law is possible in others. For a nonnegative integer r, define the *tower function* by $\operatorname{tower}(r) := \exp^r 1 = \exp \cdots \exp 1$, where the exponential is iterated r times. For convenience, we also write $\operatorname{tower}(r) := \operatorname{tower}\lfloor r \rfloor$ for $r \in \mathbb{R}^+$. Here are our main results.

THEOREM 1 (Tower function coloring). *Let $d = 1$ and $q \ge 3$, or let $d \ge 2$ and $q \ge 4$. There exist positive constants c and C depending on q and d such that the following hold.*

(i) *There exists an ffiid q-coloring of $\mathbb{Z}^d$ whose coding radius R satisfies*

$$\mathbb{P}(R > r) < 1/\operatorname{tower}(cr), \qquad \forall r \ge 0.$$

(ii) *Every ffiid q-coloring of $\mathbb{Z}^d$ satisfies*

$$\mathbb{P}(R > r) > 1/\operatorname{tower}(Cr), \qquad \forall r \ge 0.$$

THEOREM 2 (Power law 3-coloring). *Let $d \geq 2$.*

(i) *There exists a positive constant α (depending on d) and an ffiid 3-coloring of $\mathbb{Z}^d$ whose coding radius satisfies*

$$\mathbb{P}(R > r) < r^{-\alpha}, \qquad \forall r \geq 0.$$

(ii) *Every ffiid 3-coloring of $\mathbb{Z}^d$ satisfies*

$$\mathbb{E}(R^2) = \infty.$$

Since it is easy to see that no ffiid 2-coloring of $\mathbb{Z}^d$ is possible for any $d \geq 1$, Theorems 1 and 2 determine the functional form (up to the various constants) of the optimal tail decay of the coding radius for all q and d. Our proofs in principle yield explicit bounds on the constants c, C and α, but c and C are very far apart in most cases, while α is much smaller than 2.

Isometry equivariance. We will prove that the colorings in the (i) parts of both theorems can be chosen to have the stronger property that the map F from the i.i.d. variables to the coloring is equivariant under all *isometries* of $\mathbb{Z}^d$. To motivate this distinction, note that the percolation-based construction of the 4-coloring described above is not isometry-equivariant, because using the lexicographic ordering of $\mathbb{Z}^2$ breaks rotation and reflection symmetry. However, the construction can be modified as follows. Take $(U_v)_{v\in\mathbb{Z}^2}$ i.i.d. uniform on $[0, 1]$ and independent of $(B_v)_{v\in\mathbb{Z}^2}$, and assign color 1 to the site w in a $(+)$-cluster with the largest U_w, and to all other sites of the same parity in the cluster [and similarly for $(-)$-clusters]. The resulting process is an isometry-equivariant factor of the i.i.d. variables $Y_v := (B_v, U_v)$, with the same coding radius as before.

Shifts of finite type. Next, we consider some generalizations, focusing on the case $d = 1$. Coloring is a special case of the more general notion of a shift of finite type, in which the requirement that adjacent colors differ is replaced with arbitrary local constraints. Write $[q] := \{1, \dots, q\}$. We call elements of $[q]^{\mathbb{Z}^d}$ *configurations*. Let $d = 1$. A *shift of finite type* is a (deterministic) set of configurations S characterized by an integer k and a set $W \subseteq [q]^k$ of allowed local patterns as follows:

$$S = S(q, k, W) := \{x \in [q]^{\mathbb{Z}} : (x_{i+1}, \dots, x_{i+k}) \in W \ \forall i \in \mathbb{Z}\}.$$

We want to exclude a certain uninteresting case. For $w \in W$, let $T(w)$ be the set of times at which the pattern w can recur, that is, the set of $t \geq 1$ for which there exists $x \in S$ with $(x_1, \dots, x_k)$ and $(x_{t+1}, \dots, x_{t+k})$ both equal to w. We call the shift of finite type *nonlattice* if there exists $w \in W$ for which $T(w)$ has greatest common divisor 1 (and otherwise it is *lattice*). If S is nonlattice, then necessarily $S \neq \varnothing$. For example, the set of all deterministic q-colorings of $\mathbb{Z}$ is a shift of finite type, and is nonlattice if and only if $q \geq 3$.

THEOREM 3 (Shifts of finite type). *Let S be a shift of finite type on $\mathbb{Z}$.*

(i) *If S is nonlattice then there exists an ffiid process X such that $X \in S$ a.s., with coding radius R satisfying*

$$\mathbb{P}(R > r) \le 1/\operatorname{tower}(cr), \qquad \forall r > 0.$$

(ii) *If S contains no constant configuration $(\cdots aaa \cdots)$, then for any ffiid process X such that $X \in S$ a.s., the coding radius satisfies*

$$\mathbb{P}(R > r) \ge 1/\operatorname{tower}(Cr), \qquad \forall r > 0.$$

Here, c, C are constants in $(0, \infty)$ that depend on S.

It is easily seen that for any *lattice* shift of finite type S, no ffiid process belongs a.s. to S. Indeed, no *mixing* process belongs to S (see Proposition 19). On the other hand, a constant configuration is trivially an ffiid process with $R \equiv 0$. Together with these observations, Theorem 3 thus covers all cases for $d = 1$.

The concept of a shift of finite type extends in the obvious way to $\mathbb{Z}^d$ (by requiring that the configuration restricted to every ball of radius k lies in some fixed set W). For $d \ge 2$, we do not know what possible restrictions on the coding radius can be imposed by the requirement that an ffiid process belong to a given shift of finite type, besides the possibilities already seen: tower functions (e.g., 4-coloring), power laws (e.g., 3-coloring), and the two trivial cases of constant sequences and lattice shifts of finite type.

Finite dependence. Closely related to ffiid processes is the notion of k-dependence. A process $X = (X_v)_{v \in \mathbb{Z}^d}$ on $\mathbb{Z}^d$ is called *k-dependent* if $(X_v)_{v \in A}$ is independent of $(X_v)_{v \in B}$ for any subsets $A, B \subseteq \mathbb{Z}^d$ that satisfy $|u - v| > k$ for all $u \in A$ and $v \in B$. A process is *finitely dependent* if it is k-dependent for some k. A process X is *stationary* if $(X_{v+u})_{v \in \mathbb{Z}^d}$ and $(X_v)_{v \in \mathbb{Z}^d}$ are equal in law for all u. On the other hand, X is a *block factor* (of an i.i.d. process) if it is an ffiid process with *bounded* coding radius. When $d = 1$, we say that X is a *k-block factor* if there exist i.i.d. Y and a fixed measurable function g of k variables such that $X_i = g(Y_{i+1}, \dots, Y_{i+k})$ a.s. for all $i \in \mathbb{Z}$.

Clearly, any k-block factor on $\mathbb{Z}$ is stationary and $(k-1)$-dependent. Much less obviously, the converse is false. This was an open question for some time (see, e.g., [17]); the first counterexample appeared in [1]. Furthermore, there exist 1-dependent stationary processes that are not k-block factors for any k; see [6]. See, for example, [14] for more on the history of this question, which apparently has its origins in [16].

Since Theorem 1(ii) implies that no k-block factor q-coloring exists for any k and q, it is natural to ask whether there is a stationary k-dependent q-coloring. It is easily seen that the answer is no if $k = 0$ or if $q = 2$. We also establish a negative answer in the first nontrivial case: $k = 1$ and $q = 3$.

THEOREM 4. *There is no stationary* 1*-dependent* 3*-coloring of* $\mathbb{Z}$.

Surprisingly, it has recently been proved [14] that there exist both a stationary 1-dependent 4-coloring and a stationary 2-dependent 3-coloring of $\mathbb{Z}$. Thus, the above question is answered for all k and q. Moreover, coloring therefore provides a very clean and natural proof of the nonequivalence of finitely dependent processes and block factors. (Previous counterexamples have tended to be somewhat contrived.)

By combining the 1-dependent 4-coloring of [14] with results of the current article, it is also proved in [14] that for all $d \geq 2$ there exists a stationary k-dependent 4-coloring of $\mathbb{Z}^d$, for some $k = k(d)$, and also that for any nonlattice shift of finite type S on $\mathbb{Z}$ there exists a stationary k-dependent process that lies in S a.s., for some $k = k(S)$.

Combined with our Theorem 3(ii), this last result provides an even more striking illustration of the difference between finitely dependent processes and block factors: *any* nonlattice shift of finite type with no constant sequence serves to distinguish between them.

The argument we use to prove Theorem 2(ii) will also show (Corollary 25) that no stationary k-dependent 3-coloring of $\mathbb{Z}^d$ exists for any k and $d \geq 2$. See [13, 15] for further recent work on k-dependent coloring.

Outline of proofs. The existence of an ffiid q-coloring of $\mathbb{Z}^d$ satisfying a tower function bound with *some* number of colors $q = q(d)$ depends on a known method that was originally motivated by applications in distributed computing. The method appeared first in [7], and was developed further in [11, 19] and many subsequent articles. The version that we use is essentially that of [19].

Translated to our setting and terminology, the method mentioned above implies the existence of a *block* factor of an i.i.d. process that is "almost" a coloring, in the sense that the probability of a violation (i.e., of two given neighbors having the same color) is extremely small as a function of the block radius. Such processes can be constructed by starting with a discrete i.i.d. process and iteratively applying an appropriate radius-1 block factor that reduces the number of colors by a logarithmic function without producing new violations.

In order to obtain an ffiid coloring, we next proceed to "stitch together" an infinite family of the processes described above, with different block radii and violation probabilities. This can be done even on a general graph of bounded degree. In fact, the resulting factor satisfies a much stronger property than automorphism equivariance: to determine the color at a vertex, we do not need to know the graph structure, except within the coding radius.

The most elaborate and novel part of the proof of Theorem 1(i) involves reduction of the number of colors to 4 in all dimensions $d \geq 2$. This is done by applying carefully constructed block factors to colorings with more colors, in order to obtain a 2-valued process with *bounded* clusters. After this, we conclude by checkerboarding the clusters with two pairs of colors in the manner mentioned earlier. Many of the techniques in this proof are quite general, and have wider applicability. (One application appears in [14].)

The tower function lower bound Theorem 1(ii) is also a consequence of a known result from distributed computing, which was proved in [20], building on earlier work in [18]. We provide a proof that is arguably simpler and more direct than the original proof.

Turning to Theorem 2, the proof of the existence of an ffiid 3-coloring with power law coding radius is considerably simpler when $d = 2$. The construction in this case is based on critical bond percolation and its dual, on a square lattice rotated by 45 degrees. We assign colors to individual clusters based on their locations in a tree structure arising from surrounding circuits. The power law bound is a consequence of a Russo–Seymour–Welsh estimate.

The proof of Theorem 2(i) for general $d \geq 2$ is broadly similar but more involved. Instead of percolation clusters, we use a partition of $\mathbb{Z}^d$ that we construct via an iterative scheme. The sets of the partition are not themselves independent sets, but contain pairs of neighbors. Therefore, each set is assigned a checkerboard coloring using 2 of the available 3 colors, and this necessitates a more complicated tree argument. The method is quite general, and can be extended to other graphs.

The second moment bound Theorem 2(ii) is a consequence of the existence of a *height function* for 3-colorings of $\mathbb{Z}^2$. The total height change around a large contour must be zero, otherwise it is impossible to extend the 3-coloring to the interior. However, if the coding radius has finite second moment, the height changes along distant parts of the contour are asymptotically uncorrelated, leading to a contradiction.

Finally, the result on shifts of finite type is again obtained from the result on tower function coloring by the use of appropriate block factors, while the impossibility of 1-dependent 3-coloring is proved by a conditioning argument.

2. Tower function lower bound.

In this section, we prove Theorem 1(ii). The following is the key fact. An essentially equivalent result was proved in [20], building on earlier work of [18]. We give a simple direct proof. Another exposition and applications appear in [2]. Recall that $[q] := \{1, \ldots, q\}$.

PROPOSITION 5. *Let $(U_i)_{i\in\mathbb{Z}}$ be i.i.d. random variables taking values in an arbitrary set B, and let r and q be positive integers. For any measurable function $f : B^r \to [q]$,*

$$\mathbb{P}\big[f(U_1, \ldots, U_r) = f(U_2, \ldots, U_{r+1})\big] \geq \frac{1}{2^{2^{\cdot^{\cdot^{\cdot^{2^{4q}}}}}}},$$

where there are $r - 1$ exponentiation operations in the tower.

If the U_i's have a continuous distribution, then

$$\mathbb{P}\big[(U_1, \dots, U_r) = (U_2, \dots, U_{r+1})\big] = 0,$$

so it is not obvious *a priori* that the probability in Proposition 5 must be positive. If the U_i's have a discrete distribution, the probability is positive, but it is not clear *a priori* that there is a positive lower bound depending only on r and q that holds for all such distributions. The results of Section 3 below show that the tower function bound is essentially tight.

PROOF OF PROPOSITION 5. We will use induction on r. Let $\delta(r, q)$ be the largest number for which

$$\mathbb{P}\big[f(U_1, \dots, U_r) = f(U_2, \dots, U_{r+1})\big] \geq \delta(r, q)$$

for all choices of B, f, and the law of the U_i. When $r = 1$, it is elementary that

$$\mathbb{P}\big[f(U_1) = f(U_2)\big] = \sum_{a=1}^{q} \mathbb{P}\big[f(U_1) = a\big]^2 \geq \frac{1}{q},$$

so $\delta(1, q) = 1/q \geq 1/(4q)$, proving the result when $r = 1$.

Now suppose $r \geq 2$. Let $\varepsilon := \delta(r-1, 2^q)/(2q)$, and define for $u_1, \dots, u_{r-1} \in B$:

$$S(u_1, \dots, u_{r-1}) := \big\{a \in [q] : \mathbb{P}\big[f(u_1, \dots, u_{r-1}, U_r) = a\big] \geq \varepsilon\big\}.$$

This is the set of values that f assumes with probability $\geq \varepsilon$ given the first $r-1$ arguments. Since S is a function on B^{r-1} taking at most 2^q possible values (the subsets of $[q]$), by the definition of δ we have

$$\mathbb{P}\big[S(U_1, \dots, U_{r-1}) = S(U_2, \dots, U_r)\big] \geq \delta\big(r-1, 2^q\big).$$

But the definition of S implies

$$\mathbb{P}\big[f(U_1, \dots, U_r) \notin S(U_1, \dots, U_{r-1})\big] \leq q\varepsilon,$$

so we deduce

$$\mathbb{P}\big[f(U_1, \dots, U_r) \in S(U_2, \dots, U_r)\big] \geq \delta\big(r-1, 2^q\big) - q\varepsilon. \tag{1}$$

By the definition of S again, conditional on $U_2, \dots, U_r$, each element of $S(U_2, \dots, U_r)$ has probability at least ε as a possible value for the random variable $f(U_2, \dots, U_{r+1})$, and this remains true if we condition also on U_1 (since U_1 and U_{r+1} are conditionally independent given $U_2, \dots, U_r$). Therefore, almost surely

$$\begin{aligned} &\mathbb{P}\big[f(U_2, \dots, U_{r+1}) = f(U_1, \dots, U_r) | U_1, \dots, U_r\big] \\ &\quad \geq \mathbf{1}\big[f(U_1, \dots, U_r) \in S(U_2, \dots, U_r)\big] \times \varepsilon. \end{aligned}$$

Taking the expectation and using (1) gives

$$\mathbb{P}[f(U_1,\dots,U_r)=f(U_2,\dots,U_{r+1})] \ge [\delta(r-1,2^q)-q\varepsilon]\varepsilon = \frac{\delta(r-1,2^q)^2}{4q}.$$

Thus

$$\delta(r,q) \ge \frac{\delta(r-1,2^q)^2}{4q}. \tag{2}$$

All that remains is to use (2) to check the claimed bound on δ. For $r=2$, we obtain

$$\delta(2,q) \ge \frac{1}{2^{2q}}\frac{1}{4q} \ge \frac{1}{2^{4q}}$$

as required. We now use induction on r with base case $r=2$. Since obviously $\delta(r,1)=1$, we assume $q\ge 2$. Suppose $\delta(r,q)\ge 1/2^{2^{\cdots 2^{4q}}}$ where there are $r-1$ exponentiation operations in the tower. Then (2) gives

$$\delta(r+1,q) \ge \frac{1}{\left(2^{2^{\cdot^{\cdot^{\cdot^{2^{4\times 2^q}}}}}}\right)^2 4q} \ge \frac{1}{\left(2^{2^{\cdot^{\cdot^{\cdot^{2^{4\times 2^q}}}}}}\right)^4} = \frac{1}{16^{2^{\cdot^{\cdot^{\cdot^{2^{4\times 2^q}}}}}}}.$$

Observe that when $x\ge\frac{2}{3}$ we have $16^{2^x}=2^{2^{x+2}}\le 2^{2^{4x}}=2^{16^x}$, so

$$\delta(r+1,q) \ge \frac{1}{2^{2^{\cdot^{\cdot^{\cdot^{16^{4\times 2^q}}}}}}}.$$

But $16^{4\times 2^q}=2^{2^{q+4}}\le 2^{2^{4q}}$ for $q\ge 2$, which completes the induction. $\square$

The following notation will be useful. Suppose X is an ffiid process with underlying i.i.d. process Y and coding radius R, and recall that $R=R(Y)$ where R is a map from configurations $y=(y_v)_{v\in\mathbb{Z}^d}$ to $\mathbb{Z}$. For $v\in\mathbb{Z}^d$, define the coding radius at v to be the random variable

$$R_v := R(\theta_{-v}Y),$$

where θ_{-v} denotes translation by $-v$, defined by $(\theta_{-v}y)(u):=y_{u+v}$. Thus, R_v is the radius around v up to which we need to examine the Y variables in order to determine X_v. Note that $R=R_0$, and that the random variables $(R_v)_{v\in\mathbb{Z}^d}$ are identically distributed.

PROOF OF THEOREM 1(II). Let X be an ffiid q-coloring of $\mathbb{Z}^d$. Suppose first that $d=1$. Fix $r>0$ and define a modified process X' by

$$X'_v := \begin{cases} X_v, & R_v\le r;\\ \infty, & R_v>r.\end{cases}$$

Then X' is an ffiid process with coding radius bounded above by r, that is, X' is a $(2r+1)$-block-factor. Since X is a coloring,

$$\mathbb{P}(X'_0 = X'_1) = \mathbb{P}(X'_0 = X'_1 = \infty) = \mathbb{P}(R_0, R_1 > r) \le \mathbb{P}(R > r).$$

On the other hand, Proposition 5 gives

$$\mathbb{P}(X'_0 = X'_1) \ge \frac{1}{2^{2^{\cdot^{\cdot^{\cdot^{2^{4(q+1)}}}}}}},$$

with $2r$ exponentiations in the tower. This is at least $1/\operatorname{tower}(Cr)$ for some C depending only on q, as required.

Now suppose $d \ge 2$. The restriction of the coloring X to the axis $\mathbb{Z} \times \{0\}^{d-1}$ is itself an ffiid q-coloring of $\mathbb{Z}$, with underlying i.i.d. process $(Z_i)_{i \in \mathbb{Z}}$ given by the slices $Z_i := (Y_{(i,w)})_{w \in \mathbb{Z}^{d-1}}$ (where Y is the underlying i.i.d. process for X). Furthermore, the coding radius of the 1-dimensional process is at most the coding radius of X, so the required bound follows from the 1-dimensional case proved above. □

3. Tower coloring on general graphs. In preparation for the proof of Theorem 1(i), in this section we prove that on any graph of maximum degree Δ, there is an ffiid $(\Delta+1)$-coloring whose coding radius has tower function tails, and that is an *automorphism*-equivariant factor of the underlying i.i.d. process. In particular, on $\mathbb{Z}^d$ this gives an isometry-equivariant $(2d+1)$-coloring—we will improve this to 4 colors for all $d \ge 2$ in the next section.

In fact, we will construct a coloring with a much stronger property than automorphism-equivariance: the color at a vertex can be determined locally without knowledge of the graph itself—we need only examine the i.i.d. labels and the graph structure within the coding radius, and the construction is invariant even under graph-automorphisms of this local structure. We now make this precise.

Let $G = (V, E)$ be a simple undirected graph. We write $u \sim v$ if $\langle u, v\rangle \in E$. A *configuration* on G is an element $z = (z_v)_{v \in V}$ of $\mathbb{R}^V$ that assigns labels to the vertices. A *labeled rooted graph* is a triple (G, o, z) consisting of a simple graph $G = (V, E)$, a root $o \in V$, and a configuration z on G. We call two labeled rooted graphs *isomorphic* if there is a graph isomorphism between them that preserves the root and the labels. We call two labeled rooted graphs *isomorphic to distance* r if the labeled rooted subgraphs induced by the respective sets of vertices within graph-distance r of their roots are isomorphic. A *local graph function* is a function f from labeled rooted graphs to $\mathbb{R}$, such that for every (G, o, z) there exists $r \le \infty$ such that $f(G, o, z) = f(G', o', z')$ whenever (G', o', z') and (G, o, z) are isomorphic to distance r. Let $R = R(f, G, o, z)$ be the minimum such r.

A local graph function f induces a map F between configurations on graphs as follows. Let G be a graph and let z be a configuration on G. Define the configuration $F(z)$ by $(F(z))_v := f(G, v, z)$. We call F a *graph-factor map*. A *process* on G is a random configuration $Z = (Z_v)_{v \in V}$, and it is A*-valued* if each Z_v takes values in a set $A \subseteq \mathbb{R}$. If Z is a process on G and $X = F(Z)$ then we say that the process X is a graph-factor of Z, and for $v \in V$ we call $R_v := R(f, G, v, Z)$ the *coding radius* at v. If $R_v < \infty$ a.s. for all v then it is a *finitary* graph-factor, and if $R_v \leq r$ a.s. for all v and some deterministic $r < \infty$ then it is a *block* graph-factor. We call X *graph-ffiid* if it is a finitary graph-factor of some i.i.d. process. Recall that $[q] := \{1, \dots, q\}$. A process X on a graph G is a q*-coloring* if it is $[q]$-valued, and a.s. $X_u \neq X_v$ whenever $u \sim v$.

THEOREM 6 (Tower coloring on graphs). *Let $\Delta \geq 1$ be an integer. There exists $C = C(\Delta) > 0$ such that for every graph G of maximum degree Δ, there is a graph-ffiid $(\Delta + 1)$-coloring of G such that for every vertex v, the coding radius R_v satisfies*

$$\mathbb{P}(R_v > r) < 1/\operatorname{tower}(Cr), \qquad r > 0.$$

The proof will actually give an even stronger fact: the same local graph function may be used for all graphs of maximum degree Δ. The proof will proceed by combining in a suitable way a family of block graph-factors that are almost colorings in the sense that the probability that neighbors share a color decays very rapidly as a function of the block coding radius. As remarked earlier, the existence of such block-factor processes is essentially equivalent to known results in the distributed computing literature. However, the different focus in the latter field makes it difficult to translate the results directly into mathematical ones of the form we need. For the reader's convenience, we therefore provide a complete proof, which is quite straightforward.

We will make extensive use of the fact that if F and G are block graph-factor maps with coding radii at most r and s then the composition $F \circ G$ is a block graph-factor map with coding radius at most $r + s$. We also need the following simple result on set systems. Refinements and generalizations appear in [9].

LEMMA 7 (Set systems). *For each positive integer d, there exists $c = c(d) > 0$ such that, provided $n \leq e^{ck}$, there exists a family of n sets $S_1, \dots, S_n \subseteq [k]$ satisfying*

$$S_{i_0} \nsubseteq S_{i_1} \cup \cdots \cup S_{i_d}$$

for all distinct $i_0, \dots, i_d \in [n]$.

PROOF. Let $S_1, \dots, S_n$ be i.i.d. uniformly random subsets of $[k]$. The probability that $S_{d+1} \subseteq S_1 \cup \cdots \cup S_d$ is $(1 - 2^{-d})^k = C^k$, say, where $C = C(d) \in (0, 1)$. Therefore, the expected number of vectors $(i_0, \dots, i_d)$ of distinct entries such that $S_{i_0} \subseteq S_{i_1} \cup \cdots \cup S_{i_d}$ is at most $n^{d+1} C^k$. This is strictly less than 1 provided

$n < (C^{-1/(d+1)})^k$, which implies that there exist families of sets for which there are no such vectors. $\square$

We next prove the existence of "almost colorings" as mentioned above. Fix $\Delta \geq 1$. Let $c = c(\Delta)$ be as in Lemma 7, and define a sequence $n_1 < n_2 < \cdots$ as follows. Let n_1 be the smallest positive integer such that $\lfloor e^{cn_1} \rfloor > n_1$, and define inductively for $i \geq 1$:

$$n_{i+1} := \lfloor e^{cn_i} \rfloor.$$

It is easy to check that $n_i \geq \text{tower}(c'i)$ for all i and some $c' = c'(\Delta) > 0$. The following is a variant of a result of [19].

PROPOSITION 8 (Almost colorings). *Let $G = (V, E)$ be a graph of maximum degree Δ, and define $(n_k)_{k \geq 1}$ as above. For each $k \geq 1$ there exists an $[n_1] \cup \{\infty\}$-valued block graph-ffiid process $Y = Y^k$, with coding radius bounded above by k for every vertex, and with the following properties. For adjacent vertices $u \sim v$, we have either $Y_u \neq Y_v$ or $Y_u = \infty = Y_v$. For any vertex v, we have $\mathbb{P}(Y_v = \infty) \leq \Delta/n_k$.*

PROOF. We will construct a sequence of processes $Z^k, \dots, Z^1$, each a radius-1 block graph-factor of the previous one, ending with the required process $Y = Z^1$. (The reverse indexing is a notational convenience.) The process Z^i will be $[n_i] \cup \{\infty\}$-valued. Let $(Z_v)_{v \in V}$ be i.i.d. random variables, each uniform on $[n_k]$. Define the first process Z^k by setting $Z_v^k := \infty$ if $Z_v = Z_u$ for some $u \sim v$, and otherwise setting $Z_v^k := Z_v$.

Now suppose that $Z^k, \dots, Z^{i+1}$ have been defined. We will construct Z^i from Z^{i+1}. Fix a family of n_{i+1} subsets $(S_j)_{j \in [n_{i+1}]}$ of $[n_i]$ so that none is contained in the union of any Δ others; Lemma 7 and the definition of n_i ensure that this is possible. For a vertex v, write $S(v) := S_{Z_v^{i+1}}$ for the corresponding set, where we take $S_\infty := \varnothing$. Now define

$$Z_v^i := \min\Bigl(S(v) \setminus \bigcup_{u \sim v} S(u)\Bigr), \tag{3}$$

where $\min \varnothing := \infty$.

We claim that for adjacent vertices $u \sim v$, and any i, either $Z_u^i \neq Z_v^i$, or both are ∞, and moreover, for any v we have $Z_v^i = \infty$ if and only if $Z_v^k = \infty$. This follows easily by induction on i. It certainly holds for $i = k$. By (3) and Lemma 7, if $Z_v^i = \infty$ then either $Z_v^{i+1} = \infty$ or $Z_v^{i+1} = Z_u^{i+1}$ for some $u \sim v$. Moreover, for $u \sim v$, if $Z_u^i \neq \infty \neq Z_v^i$ then $Z_v^i \in S(v) \setminus S(u)$ and $Z_u^i \in S(u) \setminus S(v)$, so $Z_u^i \neq Z_v^i$.

Finally, we set $Y = Z^1$. It is evident from the construction that Y is a block graph-ffiid process with coding radius at most k. We have

$$\mathbb{P}(Y_v = \infty) = \mathbb{P}(Z_v^k = \infty) = \mathbb{P}(Z_v = Z_u \text{ for some } u \sim v) \leq \Delta/n_k. \qquad \square$$

In addition to the above result, we will use the following simple procedure for eliminating colors, which has other applications also. Let $\mathbb{Z}^+$ denote the positive integers. Suppose that X is a $\mathbb{Z}^+ \cup \{\infty\}$-valued process on a graph $G = (V, E)$. Let $a \in \mathbb{Z}^+$. We define a new process $\mathcal{E}_a X$ by

$$(\mathcal{E}_a X)_v := \begin{cases} \min(\mathbb{Z}^+ \setminus \{X_u : u \sim v\}), & X_v = a; \\ X_v, & X_v \neq a. \end{cases}$$

Thus, the map $\mathcal{E}_a$ replaces color a with the smallest color that is absent from the neighbors of the vertex. This replacement color is in $[\Delta + 1]$ if G has maximum degree Δ. Neighboring vertices have distinct colors in $\mathcal{E}_a X$ provided they do in X. Note that $\mathcal{E}_a$ is a radius-1 block graph-factor map.

A simple application of the map defined above is that if X is a q-coloring of a graph of maximum degree Δ, then $\mathcal{E}_{\Delta+2}\mathcal{E}_{\Delta+3}\cdots\mathcal{E}_q X$ is a $(\Delta + 1)$-coloring. We use this idea in a more subtle way in the next proof.

PROOF OF THEOREM 6. Let $G = (V, E)$ be a graph of maximum degree Δ. Let $(n_i)_{i\geq 1}$ be defined as above, and let $Y^1, Y^2, \dots$ be the processes of Proposition 8, each constructed from the same i.i.d. family $(U_v)_{v\in V}$ (say by taking $Z_v = \lceil n_k U_v \rceil$ at the beginning of the proof of Proposition 8, where U_v is uniform on [0, 1]). Recall that each Y^k is $[n_1] \cup \{\infty\}$-valued, and is a coloring except at the vertices that are labeled ∞ (and that the probability of label ∞ decreases rapidly with k).

We now construct a sequence of $[\Delta + 1] \cup \{\infty\}$-valued processes $X^0, X^1, X^2, \dots$. The desired coloring will be formed by taking their limit. First, let $X^0_v := \infty$ for all v. Assuming $X^0, \dots, X^{k-1}$ have been defined, we next construct X^k from X^{k-1} and Y^k. To do this, we first define an auxiliary $[\Delta + 1 + n_1] \cup \{\infty\}$-valued process W^k via

$$W^k_v := X^{k-1}_v \wedge (Y^k_v + \Delta + 1).$$

In other words, we construct W^k from X^{k-1} by replacing occurrences of ∞ with the process Y^k from the previous lemma, with the colors increased by $\Delta + 1$ so that they are distinct from the existing ones (of course, we take $\infty + \Delta + 1 := \infty$). We now obtain X^k from W^k by eliminating these extra colors:

$$X^k := \mathcal{E}_{\Delta+2}\mathcal{E}_{\Delta+3}\cdots\mathcal{E}_{\Delta+1+n_1} W^k.$$

Note that for any vertex v, if $X^k_v \neq \infty$ for some k then X^j_v is constant for all $j \geq k$. We therefore define $X_v := \lim_{k\to\infty} X^k_v$. By Proposition 8, for all k,

$$\mathbb{P}(X_v = \infty) \leq \mathbb{P}(X^k_v = \infty) \leq \mathbb{P}(Y^k_v = \infty) \leq \Delta/n_k \xrightarrow{k\to\infty} 0,$$

and it follows that X is a $(\Delta + 1)$-coloring of G. Now, for any block graph-ffiid process Z, write $r(Z)$ for the smallest constant r such that the coding radius at every vertex is bounded above by r. Then

$$r(X^k) \leq n_1 + r(W^k) \leq n_1 + [r(X^{k-1}) \vee r(Y^k)] = n_1 + [r(X^{k-1}) \vee k].$$

Hence, we have $r(X^k) \le n_1 k + 1$ for all k. It follows that X is graph-ffiid with coding radius R_v satisfying

$$\mathbb{P}(R_v > n_1 k + 1) \le \mathbb{P}(X_v^k = \infty) \le \Delta / n_k$$

for every v. As remarked earlier, we have $n_i \le \operatorname{tower}(c' i)$ for some $c' = c'(\Delta) > 0$, so the claimed bound on $\mathbb{P}(R_v > r)$ follows. $\square$

4. Tower 4-coloring. In this section, we prove Theorem 1(i). Theorem 6 in the last section already gives an isometry-equivariant ffiid $(2d+1)$-coloring of $\mathbb{Z}^d$ for all $d \ge 1$, thus proving the $d = 1$ case. For $d \ge 2$, the idea will be to use Theorem 6 to obtain a coloring of a spread-out lattice, and then apply carefully constructed block factors. We start by proving some more general results that have applications elsewhere also.

We shift our focus back to processes on $\mathbb{Z}^d$. A *factor map* is a measurable map $F : \mathbb{R}^{\mathbb{Z}^d} \to \mathbb{R}^{\mathbb{Z}^d}$ between configurations that is translation-equivariant, that is, that commutes with the action of every translation of $\mathbb{Z}^d$. Isometry-equivariance is defined analogously. If $X = F(Y)$ for a factor map F then we say that X is a factor of Y. Finitary factors and coding radius are defined as in the Introduction. A *block factor* map is a finitary factor map whose coding radius is bounded above, that is, $R \le k$ a.s. for some deterministic $k < \infty$. Recall that $R_v := R \circ \theta^{-v}$ denotes the coding radius at vertex $v \in \mathbb{Z}^d$.

We say that a nonnegative random variable R has *tower tails* if it satisfies $\mathbb{P}(R > r) < 1/\operatorname{tower}(cr)$ for all $r > 0$ and some $c \in (0, \infty)$. We call a process *tower ffiid* if it is ffiid and its coding radius has tower tails. The following simple fact will be used extensively.

LEMMA 9 (Block factors). *If X is a tower ffiid process on $\mathbb{Z}^d$ then any block factor of X is tower ffiid.*

PROOF. Let X be a tower factor of the i.i.d. process Y, and let W be a block factor of X, with coding radius bounded above by k. Clearly, W is a factor of Y. Write R for the coding radius of X, and as usual let R_v be the coding radius at $v \in \mathbb{Z}^d$. If R' denotes the coding radius of W viewed as a factor of Y, then

$$\mathbb{P}(R' > r) \le \mathbb{P}\Bigg[\bigcup_{v \in B(k)} \{R_v > r - k\}\Bigg] \le \frac{c_1 k^d}{\operatorname{tower}(c_2(r-k))} \le \frac{1}{\operatorname{tower}(c_3 r)},$$

for some constants $c_i = c_i(k, d) \in (0, \infty)$. $\square$

Let $\|\cdot\|_p$ denote the p-norm on $\mathbb{Z}^d$, and recall that we usually work with the 1-norm $|\cdot| = \|\cdot\|_1$. For most purposes, the distinction is unimportant, because the norms are equivalent and we are not concerned with exact constants. However, our construction of a 4-coloring will use both the 1- and ∞-norms.

A process $(X_v)_{v \in \mathbb{Z}^d}$ is a *range-m q-coloring* with respect to the p-norm if it is $[q]$-valued, and almost surely $X_u \ne X_v$ whenever $0 < \|u - v\|_p \le m$.

COROLLARY 10 (Long-range coloring). *Fix integers $d, m \geq 1$ and a choice of norm $\|\cdot\|_p$. There exists a tower ffiid range-m q-coloring of $\mathbb{Z}^d$ with respect to $\|\cdot\|_p$, for some number of colors $q = q(d, m, p)$. Moreover, the factor may be chosen to be isometry-equivariant.*

PROOF. This is a special case of Theorem 6, applied to the graph $\mathbb{Z}^d_{(m)}$ with vertex set $\mathbb{Z}^d$ and with an edge between distinct $u, v \in \mathbb{Z}^d$ whenever $\|u - v\|_p \leq m$. We can take $q := |\{v \in \mathbb{Z}^d : \|v\|_p \leq m\}|$. $\square$

Let $m \geq 1$ be an integer. A $\{0, 1\}$-valued process $J = (J_v)_{v \in \mathbb{Z}^d}$ is an *m-net* with respect to the p-norm if a.s. for every vertex u there exists v with $\|u - v\|_p \leq m$ and $J(v) = 1$, but there do not exist distinct vertices u, v with $\|u - v\|_p \leq m$ and $J(u) = J(v) = 1$. In other words, the support of J is a maximal independent set in the graph $\mathbb{Z}^d_{(m)}$ defined in the above proof. In dimension $d = 1$, the distance between any two consecutive 1's of an m-net lies in the interval $[m + 1, 2m + 1]$.

COROLLARY 11 (Nets). *Fix integers $d, m \geq 1$ and a choice of norm $\|\cdot\|_p$. There exists a tower ffiid m-net on $\mathbb{Z}^d$. Moreover, the factor may be chosen to be isometry-equivariant.*

PROOF. By Corollary 10, let X be a tower-ffiid range-m q-coloring. Let $\mathcal{E}_a$ be the color-elimination map defined in Section 3, for the graph $\mathbb{Z}^d_{(m)}$ defined in the last proof. Recall that $\mathcal{E}_a$ attempts to eliminate color a by replacing it with the smallest color that is absent from the range-m neighborhood of a vertex. Now we attempt to eliminate all colors:

$$Y := \mathcal{E}_1 \mathcal{E}_2 \cdots \mathcal{E}_q X.$$

The resulting process Y is a coloring, and it is tower ffiid by Lemma 9 (since $\mathcal{E}_a$ is a block-factor map). We claim that $J_v := \mathbf{1}[Y_v = 1]$ yields the required m-net J. Indeed, Y has no two 1's within distance m, while, if $X_v = a$ say, when we apply the map $\mathcal{E}_a$, the color at v becomes 1 provided there is currently no other 1 within distance m (and 1's remain 1's at subsequent steps). $\square$

In preparation for the proof of Theorem 1(i), we record the following simple geometric fact.

LEMMA 12. *Fix a norm. Let $d \geq 1$ and let $c > 0$ be a real constant. For any $m \geq 1$ and any m-net J, the number of 1's of J within distance cm of any fixed $u \in \mathbb{Z}^d$ is at most C, where C is a constant depending only on c, d and the norm (not on m).*

PROOF. The balls of radius $m/2$ centered at different 1's are disjoint; consider their volumes. $\square$

The next lemma enables a 4-coloring of $\mathbb{Z}^d$ to be constructed from a 2-valued process with bounded clusters (via the checkerboard construction mentioned in the Introduction). As is customary, we denote by $\mathbb{Z}^d$ the graph having vertex set $\mathbb{Z}^d$ and an edge between u and v whenever $\|u-v\|_1 = 1$. If X is a process on $\mathbb{Z}^d$ then an *a-cluster* of X is the vertex set of a connected component of the subgraph of $\mathbb{Z}^d$ induced by the (random) set of all v with $X_v = a$. The *diameter* (with respect to the ∞-norm) of a set $A \subseteq \mathbb{Z}^d$ is $\sup\{\|u-v\|_\infty : u, v \in A\}$.

LEMMA 13 (Checkerboarding). *Fix integers $d, b \geq 1$. Suppose Y is a $[2]$-valued process on $\mathbb{Z}^d$ in which each cluster has diameter at most b a.s. There exists a 4-coloring of $\mathbb{Z}^d$ that is a block-factor of Y. Moreover, if $(U_v)_{v\in\mathbb{Z}^d}$ are i.i.d., uniform on $[0, 1]$ and independent of Y, there exists a 4-coloring that is an isometry-equivariant block-factor of the joint process (Y, U).*

PROOF. We checkerboard each 1-cluster with 1's and 3's, and each 2-cluster with 2's and 4's. More formally, for $v \in \mathbb{Z}^d$, let $w = w(v) \in \mathbb{Z}^d$ be the lexicographically largest vertex in the same (1- or 2-)cluster as v. (Or, for the isometry-equivariant version, let w be the vertex in the cluster for which U_w is largest.) Let $X_v := Y_v + 1 + (-1)^{\|v-w\|_1}$. Then X is a block-factor of Y because the clusters are bounded. $\square$

Finally, our proof of Theorem 1(i) will require the following technical lemma. A *slab* is a set of edges of $\mathbb{Z}^d$ that is an image under some isometry of $\mathbb{Z}^d$ of the set

$$\{\langle x, x+e_1\rangle : x \in \{0\} \times \{1, \ldots, L\}^{d-1}\},$$

for some $L > 0$, where $e_1 = (1, 0, \ldots, 0)$ is the 1st coordinate vector. The slab has *direction* $j \in \{1, \ldots, d\}$ if coordinate j is the image of coordinate 1 under the isometry. By the distance between two sets of edges, we mean the distance between their respective sets of incident sites.

LEMMA 14 (Slabs). *Suppose that H is a subgraph of $\mathbb{Z}^d$ whose edge set is the union of a collection of slabs, such that no two slabs of a given direction are within $\|\cdot\|_\infty$-distance 2. Each connected component of H has $\|\cdot\|_\infty$-diameter at most 1.*

PROOF. Consider the component of 0, and first consider edges in direction 1. The given condition implies that for either $s = 0$ or $s = 1$, all of the edges

$$\{\langle x, x+e_1\rangle : x \in \{s, s-2\} \times \{-1, 0, 1\}^{d-1}\}$$

are absent from H. Since similar statements hold for each coordinate, we deduce that for some cube of $\|\cdot\|_\infty$-diameter 1 containing 0, all the edges on the exterior boundary are absent from H. $\square$

By the *box* of radius $r \in \mathbb{Z}$ centered at $v \in \mathbb{Z}^d$, we mean the ∞-norm ball $\{u \in \mathbb{Z}^d : \|u - v\|_\infty \le r\}$. The *boundary* of a subset A of $\mathbb{Z}^d$ is the set of edges incident to a site in A and a site in A^C. The boundary of a box is a union of a set of $2d$ slabs; we call them the *faces* of the box.

PROOF OF THEOREM 1(I). As remarked earlier, the case $d = 1$ and $q = 3$ already follows as a special case of Theorem 6, therefore, we need to construct a 4-coloring of $\mathbb{Z}^d$ for $d \ge 2$. By Lemmas 9 and 13, it suffices to construct a tower ffiid [2]-valued process Z with bounded clusters.

Let $M = M(d)$ be a (large) positive integer to be fixed later. By Corollary 11, let J be an M-net on $\mathbb{Z}^d$ with respect to $\|\cdot\|_\infty$, and let $S := \{v \in \mathbb{Z}^d : J(v) = 1\}$ be its support. Also, by Corollary 10, let Y be a range-$(4M + 3)$ q-coloring of $\mathbb{Z}^d$ with respect to $\|\cdot\|_\infty$ (where we allow q to be chosen as a function of M). Take J and Y to be finitary factors of the same i.i.d. process. We will construct a process Z with bounded clusters as a block factor of (J, Y). The coloring Y will appear in the construction only in the form of its restriction to S. (In fact, an alternative variant of the proof would be to instead use a coloring of the random graph with vertex set S and with an edge between elements at distance at most $4M + 3$, using Theorem 6.)

We wish to assign an integer $r(s) \in [M, 2M)$ to each element s of S in such a way that, if we place a box of radius $r(s)$ centered at each $s \in S$, then no two faces of a given direction are within $\|\cdot\|_\infty$-distance 2 of each other. (So that we can apply Lemma 14.) This will be done iteratively in the order given by the coloring Y.

Assuming radii have been chosen for all s of colors $Y_s < j$ (which is vacuously true when $j = 1$), we will simultaneously choose a radius $r(s)$ for each $s \in S$ of color $Y_s = j$ in such a way that no faces of the box of radius $r(s)$ centered at s come within distance 2 of those faces already chosen. By Lemma 12, there are at most C elements of S within $\|\cdot\|_\infty$-distance $4M + 2$ of s, where C is a constant that depends only on d (not on M, j, q or s). Any face of an existing box centered at one of these elements prohibits at most 7 possible values for $r(s)$ in $[M, 2M)$. Therefore, at most $C' := 14dC$ possible values for $r(s)$ are prohibited by the condition on faces (in particular, this C' depends only on d). Also, since all radii are less than $2M$ but Y is a range-$(4M + 3)$ coloring, the radii $r(s)$ for all those $s \in S$ with color j can be chosen simultaneously without interfering with each other (i.e., without two of them violating the face condition). Therefore, if we choose $M = C' + 1$ then these radii can indeed be chosen for each $j = 1, \dots, q$ in turn. For definiteness and to ensure isometry-equivariance, choose each $r(s)$ to be the smallest allowable value in $[M, 2M)$ at the appropriate step.

Now construct a $\{+1, -1\}$-valued process Z as follows. Any vertex v is covered by at least one of the boxes chosen above (since J is an M-net), but by only finitely many. Let $s = s(v) \in S$ be the center of the one that has the lowest numbered color in Y. Let $Z_v := (-1)^{\|s-v\|_1}$. In other words, each box is labeled checkerboard-fashion, with the parity determined by the position of its center, and with lower-colored boxes taking priority over higher ones. (We are not using Lemma 13 here, despite the similarity of the construction.)

Let G be the (random) subgraph of $\mathbb{Z}^d$ in which two adjacent vertices u, v are connected by an edge if and only if $Z_u = Z_v$. By the construction of the boxes, G is a subgraph of a graph H satisfying the conditions of Lemma 14, so each cluster of Z has $\|\cdot\|_\infty$-diameter at most 1, as required.

In each step $1, \dots, q$ of the above procedure, a site $s \in S$ only needed to examine S, Y and the earlier choices of radii within a neighborhood of radius $4M + 2$ in order to determine its radius $r(s)$. Thus, the entire procedure constitutes a block-factor map from (J, Y) to Z (and indeed it is an isometry-equivariant map). Therefore, Lemma 9 gives that Z is tower ffiid. $\square$

We note that the above argument actually gives the following fact. This has been used in [14] to prove the existence of k-dependent 4-colorings of $\mathbb{Z}^d$ for all $d \geq 2$.

COROLLARY 15. *Let $d \geq 1$. There exists m such that for any q there exists a block factor map F with the following property. If X is a range-m q-coloring of $\mathbb{Z}^d$ then $F(X)$ is a 4-coloring of $\mathbb{Z}^d$. In addition, if $(U_v)_{v \in \mathbb{Z}^d}$ are i.i.d. uniform on* $[0, 1]$ *and independent of X then similarly there exists an isometry-equivariant block factor map F' such that $F'((X, U))$ is a 4-coloring of $\mathbb{Z}^d$.*

PROOF. We take $m = 4M + 3$ in the proof of Theorem 1(i) above. Since a range-$(4M + 3)$ coloring is also a range-M coloring, the construction in the proof of Corollary 11 gives us an M-net J as a block factor of X, and we also take $Y = X$. $\square$

5. Shifts of finite type. In this section, we prove Theorem 3, for which we will use the following construction. Let $S = S(q, k, W)$ be a shift of finite type on $\mathbb{Z}$. Let $G = G_S$ be the directed graph with vertex set W, and with a directed edge from $u = (u_1, \dots, u_k)$ to $v = (v_1, \dots, v_k)$ if and only if $(u_2, \dots, u_k) = (v_1, \dots, v_{k-1})$. For any $x \in [q]^{\mathbb{Z}}$, clearly we have $x \in S$ if and only if the sequence $((x_{i+1}, \dots, x_{i+k}))_{i \in \mathbb{Z}}$ forms a directed (bi-infinite) path in G.

PROPOSITION 16 (Shifts of finite type from nets). *Let S be a nonlattice shift of finite type on $\mathbb{Z}$. There exist an integer $m \geq 1$ and a block-factor map F such that if J is an m-net then $F(J)$ belongs to S a.s.*

PROOF. Let $S = S(q,k,W)$ and let $G = G_S$ be the directed graph defined above. For $w \in W$, the set of recurrence times $T(w)$ is precisely the set of positive integers t for which there exists a (not necessarily self-avoiding) directed cycle of length t in G that contains the vertex w. Suppose that the greatest common divisor of $T(w)$ is 1. Since $T(w)$ is closed under addition, it is a standard fact of number theory that there exists some m such that $T(w)$ contains all integers greater than m.

Therefore, for each integer $t \in [m+1, 2m+1]$, we can fix a directed cycle of G of length t containing w. Let $w = y_0^t, t_1^t, \dots, y_t^t = w$ be its vertices in order. Let J be an m-net. Construct a W-valued process Z from J as follows. For each $i \in \mathbb{Z}$ with $J_i = 1$, let $Z_i = w$. If $i < j$ are the locations of two consecutive 1's in J, let $t = j - i \in [m+1, 2m+1]$, and let $(Z_i, \dots, Z_j) = (y_0^t, t_1^t, \dots, y_t^t)$. Finally, define a process X by letting X_i be the first entry of the k-vector Z_i for each $i \in \mathbb{Z}$. Clearly, $X \in S$, and X is a block factor of J because the intervals between 1's of J have bounded lengths. □

PROOF OF THEOREM 3(I). This follows immediately from Corollary 11 and Proposition 16. □

We note that our argument yields the following, which is used in [14].

COROLLARY 17. *Let S be a nonlattice shift of finite type on $\mathbb{Z}$. There exist m such that for any q, there exists a block-factor map F such that if X is a range-m q-coloring of $\mathbb{Z}$ then $F(J)$ belongs to S.*

PROOF. This follows from Proposition 16 and the proof of Corollary 11. □

PROOF OF THEOREM 3(II). Suppose that $S = S(q,k,W)$. If S contains no constant sequence, then the graph $G = G_S$ has no self-loops. Suppose X is an ffiid process that belongs to S a.s. Then the block process $W = (W_i)_{i\in\mathbb{Z}}$ given by $W_i := (X_{i+1}, \dots, X_{i+k})$ is a q^k-coloring of $\mathbb{Z}$, and it is clearly a block factor of X. Let R be the coding radius of X, and let R' be the coding radius of W viewed as a factor of the i.i.d. process underlying X. Theorem 1(ii) implies $\mathbb{P}(R' > r) \geq 1/\operatorname{tower}(Cr)$ for all r and some c, while as in the proof of Lemma 9, $\mathbb{P}(R' > r) \leq k\mathbb{P}(R > r-k)$. Hence, $\mathbb{P}(R > r) \geq 1/\operatorname{tower}(C'r)$ for some $C' = C'(C,k)$. □

Finally, in this section, we show that a *lattice* shift of finite type admits no ffiid process, as mentioned in the Introduction. In fact, we prove a stronger statement. A process X on $\mathbb{Z}$ is called *mixing* if for any events A and B in the σ-field generated by X we have $\mathbb{P}(A \cap \theta^n B) \to \mathbb{P}(A)\mathbb{P}(B)$ as $n \to \infty$. [Here, if A is the event $\{X \in \mathcal{A}\}$ then $\theta^n A$ is the translated event $\{(X_{i+n})_{i\in\mathbb{Z}} \in \mathcal{A}\}$.] The following is a standard fact.

LEMMA 18. *Any factor of an i.i.d. process on $\mathbb{Z}$ is mixing.*

PROOF. Suppose X is a factor of the i.i.d. process Y. Fix events $A, B \in \sigma(X)$ and any $\varepsilon > 0$. There exist cylinder events $A_\varepsilon, B_\varepsilon$ of Y such that $\mathbb{P}(A \triangle A_\varepsilon)$, $\mathbb{P}(B \triangle B_\varepsilon) < \varepsilon$, and by translation-equivariance, $\mathbb{P}(\theta^n B \triangle \theta^n B_\varepsilon) < \varepsilon$. For n sufficiently large, A_ε and $\theta^n B_\varepsilon$ are independent, and hence $|\mathbb{P}(A \cap \theta^n B) - \mathbb{P}(A)\mathbb{P}(B)| < 4\varepsilon$. $\square$

PROPOSITION 19. *Let S be a lattice shift of finite type on $\mathbb{Z}$. There is no mixing stationary process X for which $X \in S$ a.s.*

PROOF. Suppose that such an X does exist. Since X is mixing, it is ergodic. Hence, there exists some $w \in W$ that a.s. appears infinitely often in the process W given by $W_i := (X_{i+1}, \dots, X_{i+k})$. Fix such a w, and let t be the greatest common divisor of the recurrence set $T(w)$. Then a.s. the random set $\{i \in \mathbb{Z} : W_i = w\}$ lies in $L + t\mathbb{Z}$ for some random L in $[t]$. Since the set is a.s. nonempty, L is measurable with respect to $\sigma(X)$, and by stationarity L must be uniformly distributed over $[t]$. Therefore, letting A be the event that $L = t$, we have $\mathbb{P}(A \cap \theta^n A) = \mathbf{1}[t \text{ divides } n]/t$, which does not converge as $n \to \infty$, contradicting the fact that X is mixing. $\square$

6. Power law coloring. In this section, we construct ffiid 3-colorings of $\mathbb{Z}^d$ for $d \geq 2$ with power law tails, proving Theorem 2(i). A simpler version of the argument is available when $d = 2$; we give this first.

PROOF OF THEOREM 2(I), CASE $d = 2$. First, construct a random graph H with vertex set $\mathbb{Z}^2$ by choosing, for each unit square of $\mathbb{Z}^2$, exactly one of the two diagonals to be an edge of H, with each diagonal having probability $1/2$, and where the choices are independent for different squares. It is of course trivial to do this as a translation-equivariant block factor of an i.i.d. process indexed by the vertices. For an *isometry*-equivariant construction, one can proceed as follows. Let $(U_v)_{v \in \mathbb{Z}^2}$ be i.i.d. uniform on $[0, 1]$ and let $(B_v)_{v \in \mathbb{Z}^2}$ be i.i.d. uniform on $\{\pm 1\}$, independent of each other. For a unit square s, define $B'_s := \prod_{i=1}^4 B_{s_i}$, where $s_1, \dots, s_4$ are the vertices (in counterclockwise order, say). Then $(B'_s)_s$ is an i.i.d. uniform ± 1-valued family indexed by unit squares (as can be seen by considering in lexicographic order the unit squares that make up an n by n square, and noting that each is independent of those preceding it). Now place an edge between s_1, s_3 if $(U_{s_1} + U_{s_3} - U_{s_2} - U_{s_4})B_s > 0$, and otherwise place it between s_2, s_4.

Observe that H is precisely a critical bond percolation model on the even sublattice of $\mathbb{Z}^2$ (interpreted as a copy of $\mathbb{Z}^2$ rotated by $\pi/4$ and enlarged by $\sqrt{2}$) together with its planar dual on the odd sublattice. See Figure 2. Note that for the purpose of constructing an ffiid process, it is important that we treat the even and odd sublattices identically.

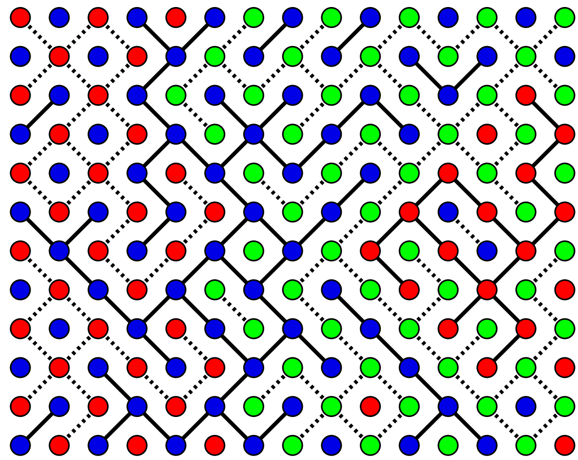

FIG. 2. *Random diagonals, the resulting bond percolation process (solid lines), its planar dual (dashed lines) and a corresponding 3-coloring.*

Call the connected components of H *clusters*, and call two clusters *adjacent* if some vertex of one is adjacent in $\mathbb{Z}^2$ to some vertex of the other. (Adjacent clusters belong to sublattices of opposite parity, of course.) We will assign one of the 3 colors to each cluster. This will result in a coloring of $\mathbb{Z}^2$ provided adjacent clusters receive distinct colors, as in Figure 2.

All clusters are finite a.s. (since there is no percolation at the critical point $1/2$ of bond percolation on $\mathbb{Z}^2$; see, e.g., [12]). For each cluster K, there is precisely one cluster $\pi(K)$ that surrounds K (i.e., intersects every infinite path from K) and is adjacent to K (see, e.g., [12]). We call $\pi(K)$ the *parent* of K, and K a *child* of $\pi(K)$. Any two adjacent clusters are parent and child in exactly one direction. If $K' = \pi^m(K)$ for some $m \ge 0$ (where π^m denotes the mth iterate of π) then we say that K' is an *ancestor* of K and that K is a *descendant* of K'. Note that each cluster has infinitely many ancestors but only finitely many descendants.

Next, we assign a label Y_K to each cluster K, in such a way that conditional on H the labels are i.i.d. and uniform on $\{\pm 1\}$. To do this, take $(V_v)_{v\in\mathbb{Z}^2}$ i.i.d. uniform on $[0, 1]$ and $(W_v)_{v\in\mathbb{Z}^2}$ i.i.d. uniform on $\{\pm 1\}$, and let $Y_K = W_u$ where u is the vertex of K for which V_u is greatest. Call a cluster K *special* if $Y_K = 1$ but $Y_{\pi(K)} = -1$. Now we define the coloring. Assign color 1 to each special cluster. For a nonspecial cluster K, let $m \ge 1$ be the smallest positive integer for which the ancestor $\pi^m(K)$ is special, and assign K color 2 or 3 according to whether m is odd or even, respectively.

The above clearly gives a coloring. To check that it is ffiid and bound the coding radius, note that to determine the color of the origin, it suffices to examine the various i.i.d. labels of the parent of the most recent special ancestor of the cluster of the origin, together with those of all its descendants, and the vertices of $\mathbb{Z}^2$ within distance 2 of these clusters. The coding radius R is at most the radius around 0 of this set of vertices. To bound R, define a family of nested annuli $A_n := \{x \in \mathbb{Z}^2 : 2^n \leq |x| < 2^{n+1}\}$ centered at the origin. By the Russo–Seymour–Welsh theorem, the probability that H contains a circuit in the even sublattice that lies in A_n and surrounds the origin is bounded strictly away from 0 as $n \to \infty$, and similarly for the odd sublattice; see, for example, [12]. Take $p > 0$ and $N \geq 1$ such that both probabilities are at least p for all $n > N$. Now let E_m be the event that the following all hold: A_{4m} and A_{4m+2} each contain such a circuit in the even sublattice, while A_{4m+1} and A_{4m+3} each contain one in the odd sublattice, and moreover, the cluster that contains the outermost such circuit in A_{4m+1} is special. Now the events $(E_m)_{m\geq 1}$ are independent, and $\mathbb{P}(E_m) > p^4/4$ if $4m > N$. If E_m occurs, then the cluster of the origin has a special ancestor whose parent lies within the ball $B(2^{4m+4})$. Therefore, $\mathbb{P}(R > 2^{4m+4} + 2) \leq (1 - p^4/4)^m$ for $4m > N$, which gives the claimed power law tail bound. $\square$

Unfortunately, the above method gives only a very small positive power α in the bound $\mathbb{P}(R > r) < cr^{-\alpha}$. The best available lower bound for the Russo–Seymour–Welsh circuit probability p is roughly 2^{-36}. And, even with more elaborate bookkeeping, the best that can be obtained from the above argument is $\mathbb{P}(R > 2^m) \leq (1 - p/2)^m$, giving $\alpha \approx p/(2\log 2)$. It would be of interest to obtain a more reasonable power (either for this 3-coloring of $\mathbb{Z}^2$ or another one).

We now move on to the case of general $d \geq 2$. The strategy will be broadly similar to that for $d = 2$ above, but with the following main differences. We can no longer use critical percolation together with its planar dual; instead, we use an iterative procedure to construct a partition of $\mathbb{Z}^d$ with a similar tree structure. However, unlike the percolation clusters, individual sets of this partition will themselves contain pairs of neighboring vertices. Therefore, rather than a single color, each set will be assigned a checkerboard 2-coloring comprising 2 of the 3 available colors. This in turn will necessitate a more subtle version of the family tree coloring procedure. The method of proof is quite general, and can be applied to other graphs (with an appropriate number of colors that depends on the graph).

The first part of the construction is deterministic, and can be done on any graph. (In fact, it can be generalized to metric spaces.) Let $G = (V, E)$ be a simple undirected graph, and let δ denote graph-distance on V. Denote the closed ball $B(u, r) := \{v \in V : \delta(u, v) \leq r\}$. As usual, the diameter of a set $S \subseteq V$ is $\operatorname{diam}(S) := \sup\{\delta(u, v) : u, v \in S\}$, the radius around a point $u \in S$ is $\operatorname{rad}_u(S) := \sup\{\delta(u, v) : v \in S\}$, and the (graph) distance between two sets $S, T \subseteq V$ is $\delta(S, T) := \inf\{\delta(s, t) : s \in S, t \in T\}$.

Here is the construction. Define

$$r_j := 13^j, \qquad j \geq 1$$

and suppose we are given a family of sets $V_1, V_2, \ldots \subseteq V$. (In our application below, the sets will be chosen randomly, in such a way that no two elements of V_j

are within distance $4r_j$ of each other.) We call elements of V_j *j-centers*. Call the ball of radius r_j centered at any j-center a *j-ball*. To each j-ball, we will associate a subset of V, called a *j-tile*. The collection of all tiles will be our partition.

The 1-tiles are precisely the 1-balls. Now assume that j-tiles have been defined for all $j \le n$, and let $\mathcal{T}_n$ denote the set of all such tiles. Let $\mathcal{G}_n$ be the graph with vertex set $\mathcal{T}_n$ in which two tiles are neighbors in $\mathcal{G}_n$ if the distance between them is at most 2. Define an *n-clump* to be the union of the tiles that correspond to a connected component of $\mathcal{G}_n$. By the n-clump of a tile, we mean the n-clump containing that tile.

Now let B be an $(n+1)$-ball. Let S_B denote the union of B and all the n-clumps that are within distance at most 2 from B. Define the $(n+1)$-tile T_B to be the set of all $v \in V$ that are within distance at most 1 from S_B but are not in $\bigcup \mathcal{T}_n$. The $(n+1)$-tiles are all such T_B.

At the same time as defining tiles, we impose a family tree structure on them. Every tile T' of $\mathcal{T}_n$ that is a subset of S_B is declared a *child* of T_B, provided T' was not already declared a child of some other tile at some earlier stage. If T' is a child of T, then T is a *parent* of T'. *A priori* a tile might have no parents, or more than one, but we will see next that for our choice of V_j's the parent is unique.

Each tile has a *center*, defined to be the center of the ball B used to define the tile T_B. (The center is not necessarily an element of the tile.)

LEMMA 20 (Tiling). *Let $G = (V, E)$ be an infinite connected graph, let $V_1, V_2, \ldots \subseteq V$ be sets of centers, and construct tiles as described above. Suppose that every $v \in V$ lies in some ball, and that no two j-centers are within distance $4r_j$ (for each $j \ge 1$). Then the set of all tiles is a partition of V. Each tile is nonempty, and has exactly one parent. If T, T' are distinct tiles neither of which is a child of the other then $\delta(T, T') > 1$. If there is a j-tile centered at v, then the tile and its associated j-clump are subsets of the ball $B(v, \frac{3}{2}r_j)$, and are functions of $V_1, \ldots, V_j$ restricted to this ball.*

PROOF. The key step is to prove by induction that the diameter of a j-clump is at most $3r_j$. This certainly holds for $j = 1$. Assume that it holds for $j = n$. Let B be an $(n+1)$-ball with center u. Recalling the definition of the associated tile T_B, we can bound its radius:

$$\mathrm{rad}_u(T_B) \le r_{n+1} + 2 + 3r_n + 1.$$

Let $\widehat{T}_B$ be the union of T_B with all the n-clumps that are within distance at most 2 from T_B. Then

$$\mathrm{rad}_u(\widehat{T}_B) \le \mathrm{rad}_u(T_B) + 2 + 3r_n \le r_{n+1} + 6r_n + 5 < \frac{3}{2} r_{n+1},$$

by our choice of r_j. For distinct $(n+1)$-balls B_1, B_2, the centers are at distance at least $4r_{n+1}$, therefore, $\delta(\widehat{T}_{B_1}, \widehat{T}_{B_2}) > (4 - 2 \cdot \frac{3}{2}) r_{n+1} = r_{n+1} > 2$. It follows that the

$(n+1)$-clump of T_B is $\widehat{T}_B$, and hence that this clump has diameter at most $3r_{n+1}$. This completes the induction.

From the above inequality, in fact the radius of T_B's clump $\widehat{T}_B$ is at most $\frac{3}{2}r_{n+1}$, and by the construction of T_B, the tile and the clump are functions of $V_1, \dots, V_j$ restricted to the ball of radius $\frac{3}{2}r_{n+1}$ centered at u, as claimed.

Now, if v lies in an n-ball B then either v lies in T_B, or it lies in some tile of $\mathcal{T}_{n-1}$. Thus, every v lies in some tile. On the other hand, we showed above that any two n-tiles are disjoint (and in fact are at distance greater than 2), while by the construction, an n-tile is disjoint from $\bigcup \mathcal{T}_{n-1}$. Thus, the tiles partition V.

To see that the n-tile T_B is nonempty, recall that $T_B \supseteq B \setminus \bigcup \mathcal{T}_{n-1}$. But we cannot have $\bigcup \mathcal{T}_{n-1} \supseteq B$, because B is connected, while each component of $\bigcup \mathcal{T}_{n-1}$ lies in an $(n-1)$-clump, and thus has strictly smaller diameter than B.

Let T, T' be distinct tiles neither of which is a child of the other. As remarked above, if both are n-tiles then $\delta(T, T') > 2 > 1$. On the other hand, if $T = T_B$ is an n-tile and $T' \in \mathcal{T}_{n-1}$ then, by the definition of S_B, either $T' \subseteq S_B$ or $\delta(T', S_B) > 2$. In the former case, T' was already assigned a parent before stage n, and thus all vertices of V that are within distance 1 of T' lie in $\bigcup \mathcal{T}_{n-1}$, so $\delta(T, T') > 1$. In the latter case, the definition of T_B implies that $\delta(T, T') > 1$ also.

If B is an n-ball, then S_B is contained in the n-clump of T_B, but we showed above that for distinct n-balls B_1 and B_2, the clumps of T_{B_1} and T_{B_2} are disjoint. Thus, any tile has at most one parent. It remains to show that an n-tile T has at least one parent. Since G is infinite and connected but the n-clump of T is bounded, there exists $w \in V$ that is at distance 1 from the clump but not in $\bigcup \mathcal{T}_n$. This w lies in some ball B, which must be a m-ball for some $m > n$ (otherwise w would lie in $\bigcup \mathcal{T}_n$), and thus S_B contains T. Hence, either T_B is the parent of T, or another tile was declared the parent of T before T_B was constructed. $\square$

As before, we write $\pi(T)$ for the parent of a tile T. If $T' = \pi^m(T)$ for some $m \ge 0$ then we call T' an *ancestor* of T, and T a *descendant* of T'. Let $\mathcal{F}$ denote the graph whose vertices are the tiles, and where two tiles are adjacent if they are at distance 1. Thus, $\mathcal{F}$ is a forest with exactly one end per component. (In our application below, $\mathcal{F}$ will actually be a tree.)

In order to bound the coding radius of our coloring, we need the following additional property.

LEMMA 21. *Assume the conditions of Lemma 20. If the ball B contains the vertex v then some descendant of tile T_B contains the vertex v.*

PROOF. Suppose B is an n-ball. Either v lies in T_B itself, or it lies in some previously constructed tile T that is a subset of S_B. In the latter case, either T is a child of T_B, or T was earlier declared a child of some other tile $\pi(T) = T_{B'}$, say, where B' is an n'-ball and $n' < n$. In that case, $S_{B'}$ contains T (by the definition of

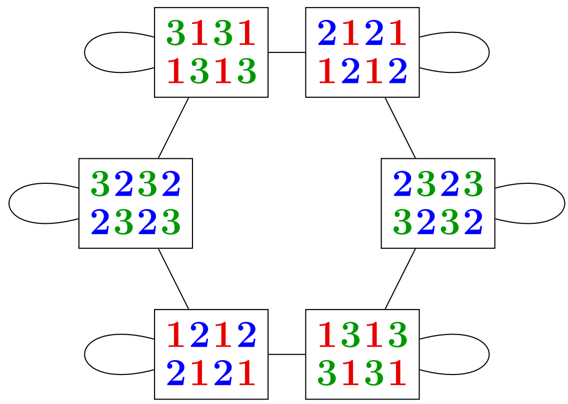

FIG. 3. *The graph $\mathcal{Q}$ of checkerboard colorings of $\mathbb{Z}^2$. (A small part of each coloring is shown.)*

child). Since the n'-clump of T is a subset of the n-clump of T (by the definition of clump), we have that $S_{B'} \subseteq S_B$ and therefore $\pi(T) \subseteq S_B$. Now we iterate this argument: the parent $\pi^2(T)$ of $\pi(T)$ is either T_B, or it is some other tile constructed after $\pi(T)$ but before T_B, in which case $\pi^2(T) \subseteq S_B$, and so on. Eventually, we conclude that T_B is an ancestor of T. $\square$

PROOF OF THEOREM 2(i) FOR GENERAL $d \geq 2$. We first construct a random tiling of $\mathbb{Z}^d$. Define $r_j = 13^j$ as above. For $j \geq 1$, let W_j be a random subset of $\mathbb{Z}^d$ in which each vertex is included with probability r_j^{-d}, independently for different vertices. Let V_j be the set of elements of W_j that have no other element of W_j within distance $4r_j$. Let the sets $(V_j)_{j\geq 1}$ be independent of each other. Construct tiles using the sets of centers $(V_j)_{j\geq 1}$ as described above. Note that the probability that $v \in \mathbb{Z}^d$ lies in some j-ball is at least η for some $\eta = \eta(d) > 0$ that does not depend on j. Therefore, every v lies in some ball, so Lemma 20 applies.

Let Q be the set of all deterministic colorings of $\mathbb{Z}^d$ that use any 2 colors from $\{1, 2, 3\}$. Then Q has 6 elements, since there are $\binom{3}{2}$ choices of 2 colors, and 2 possible checkerboard phases. Consider the graph $\mathcal{Q}$ with vertex set Q, and with an edge between two colorings if one can be obtained from the other by exchanging one color for the unused color, together with a self-loop at each vertex. Thus, $\mathcal{Q}$ is a hexagon with self-loops, and hence has diameter $D := 3$. See Figure 3. We will assign a coloring in Q to each tile, and this will result in a coloring of $\mathbb{Z}^d$ provided adjacent tiles receive colorings that are adjacent in $\mathcal{Q}$. For every pair of vertices of $\mathcal{Q}$, fix a canonical shortest path between them.

Conditional on the tiling, flip an independent fair coin for each tile (e.g., by flipping a coin for every vertex of $\mathbb{Z}^d$ and using the coin at the center of the tile). Call a tile T *special* if its coin is Heads but no ancestor $\pi^m(T)$ with $1 \leq m < D$ has Heads. Let A be the random set of special tiles. For any tile T, let $a(T)$ be its most recent special *strict* ancestor, that is, the tile $\pi^m(T)$ where $m \geq 1$ is the smallest positive integer for which this tile is special. (Such an m exists a.s.)

To each special tile T, assign a uniformly random element $h(T)$ of Q (again, this can be done via the center). The idea will be that $h(T)$ will be used to color certain descendants of T. However, the phase must be chosen locally. Therefore, let $h'(T)$ denote the 2-coloring $h(T)$ translated by u, where u is the center of T. Thus, $h'(T)$ is either $h(T)$ or the coloring that results from exchanging the 2 colors, according to the parity of u.

We now construct a new function g from the tiles to Q. The idea is that a special tile T tries to force its descendants to use $h'(T)$, succeeding if they are at least D levels below, but any special descendants get to take over this task.

To make this precise, for any tile T, we choose a shortest path in $\mathcal{Q}$ from $h'(a(a(T)))$ to $h'(a(T))$. Here, we again need to be careful with phase: let u be the center of $a(a(T))$, and first consider the canonical path between $h(a(a(T))$ and the translation of $h'(a(T))$ by $-u$, then translate all the colorings of this path by u to obtain a new path. Let $h'(a(a(T))) = z_0, z_1, \dots, z_\ell = h'(a(T))$ denote this path. Now, if the distance from T to $a(T)$ in $\mathcal{F}$ is j, let $g(T) = z_{\min(j,\ell)}$. We claim that g is a graph homomorphism from $\mathcal{F}$ to $\mathcal{Q}$. Indeed, consider the parent $T' := \pi(T)$ of T. If T' is not special, then $a(T') = a(T)$, so by the path construction, $g(T)$ and $g(T')$ are neighbors in $\mathcal{Q}$. On the other hand, if T' is special, then $a(T')$ is at distance at least D from T' in $\mathcal{F}$, so $g(T') = h'(a(T'))$; but T is at distance 1 from $a(T) = T'$ in $\mathcal{F}$, so $g(T)$ is a neighbor of $h'(a(a(T)) = h'(a(T'))$, so again $g(T)$ and $g(T')$ are neighbors in $\mathcal{Q}$.

Now we define X by assigning the checkerboard coloring $g(T) \in Q$ to all the vertices of the tile T. By Lemma 20, each edge of the lattice either connects two vertices in the same tile, or connects a vertex in one tile to a vertex in that tile's parent. Since the colorings $g(T)$ and $g(\pi(T))$ are neighbors in $\mathcal{Q}$, they are compatible, so X is in fact a 3-coloring.

It is immediate from the construction that X is an automorphism-equivariant factor of the various i.i.d. labels. To check that it is ffiid and bound the coding radius, note that the color X_0 can be determined by examining the tile $a(\pi^D(T))$ and its descendants, where T is the tile containing 0. For $m \geq 1$, let E_m be the event that there is a j-ball containing 0 for each of $j = 2Dm, 2Dm + 1, \dots, 2D(m + 1) - 1$, and the coins associated to the corresponding tiles T_j are Heads for T_{2Dm+D} and Tails for $T_{2Dm+D+1}, \dots, T_{2D(m+1)-1}$. On E_m, tile T_{2Dm+D} is special while tiles $T_{2Dm+1}, \dots, T_{2Dm+D-1}$ tiles are not, and this is enough to determine the coloring of tile T_{2Dm}. By Lemma 21, tile T_{2Dm} has a descendent containing 0. But all the descendants of a tile are in its clump, so from Lemma 20 it follows that the coding radius is at most $(\frac{3}{2} + 4)r_{2D(m+1)-1}$. (The 4 comes from the construction of V_j from W_j). On the other hand, the events $(E_m)_{m\geq 1}$ are independent, and their probabilities are bounded below by some $p = p(d) > 0$. Thus, $\mathbb{P}(R > \frac{11}{2} \times 13^{2D(m+1)}) \leq (1 - p)^m$, giving the required power law bound. $\square$

7. Second moment bound. In this section, we prove Theorem 2(ii), which will follow from a lower bound on spatial correlations that holds for any stationary 3-coloring of $\mathbb{Z}^2$. The key to the proof is that there is a height function associated to the 3-colorings. If correlations were to decay too fast, then the height changes around a large contour would not cancel.

We need the following simple lemma, the proof of which is deferred to the end of the section. A process Z on $\mathbb{Z}$ is called *right-tail-trivial* if every event in $\mathcal{T}_+ := \bigcap_{n\in\mathbb{Z}} \sigma(Z_n, Z_{n+1}, \dots)$ has probability zero or one.

LEMMA 22. *If $(Z_i)_{i\in\mathbb{Z}}$ is a ± 1-valued stationary right-tail-trivial process, either it is a.s. deterministic or* $\limsup_{n\to\infty} \operatorname{Var} \sum_{i=1}^n Z_i = \infty$.

Let X be a 3-coloring of $\mathbb{Z}^2$. We will prove a lower bound on spatial correlations involving pairs of edges. Let $u, v \in \mathbb{Z}^2$ be neighboring vertices. Since X is a coloring, $X_v - X_u \equiv \pm 1 \pmod 3$. Therefore, define $h(u, v) \in \{-1, +1\}$ by $h(u, v) \equiv X_v - X_u \pmod 3$. Now define

$$\rho(r) := \sup\bigl\{\operatorname{Cov}\bigl[h(u_1, v_1), h(u_2, v_2)\bigr] : \\ \|u_1 - v_1\|_1 = \|u_2 - v_2\|_1 = 1, \|u_1 + v_1 - u_2 - v_2\|_1 \geq 2r\bigr\}.$$

Note that ρ is nonnegative (since interchanging u_1 and v_1 reverses the sign of the covariance), and nonincreasing.

PROPOSITION 23 (Correlations). *Let X be a stationary 3-coloring of $\mathbb{Z}^2$, and suppose that its restriction $(X_{(i,0)})_{i\in\mathbb{Z}}$ to the axis is right-tail-trivial. Then with ρ defined as above*,

$$\sum_{r=1}^{\infty} r\rho(r) = \infty.$$

The key point is that the function h defined above can be interpreted as the difference along an edge of a height function. (See, e.g., [4, 10] for background.) Indeed, suppose $w_0, \dots, w_3$ are the vertices of a unit square of $\mathbb{Z}^2$ in counterclockwise order, and write $w_4 = w_0$. Then $\sum_{j=0}^{3} h(w_j, w_{j+1}) = 0$ (since the sum lies in $\{0, \pm 2, \pm 4\}$ but equals 0 modulo 3). Therefore, for arbitrary vertices $u, v \in \mathbb{Z}^2$ we can define $h(u, v) := \sum_{j=0}^{m-1} h(w_j, w_{j+1})$ where $u = w_0, w_1, \dots, w_m = v$ is any path from u to v; it follows from the above observation that this sum does not depend on the choice of path.

PROOF OF PROPOSITION 23. Let X be a 3-coloring with the given properties, and suppose for a contradiction that $\sum_{r=1}^{\infty} r\rho(r) = C < \infty$.

Write $v_j := (j, 0)$, and let $n \geq 1$. We will bound the variance of $h(v_0, v_n)$ by expressing it in two different ways. Summing along the axis gives

$$h(v_0, v_n) = \sum_{j=1}^{n} h(v_{j-1}, v_j),$$

while by summing around three sides of a square:

$$\begin{aligned} h(v_0, v_n) = & \sum_{j=1}^{n} h\big((0, j-1), (0, j)\big) + \sum_{j=1}^{n} h\big((j-1, n), (j, n)\big) \\ & + \sum_{j=1}^{n} h\big((n, n-j+1), (n, n-j)\big). \end{aligned}$$

Thus, we may compute $\operatorname{Var} h(v_0, v_n) = \operatorname{Cov}[h(v_0, v_n), h(v_0, v_n)]$ as the covariance of the two representations. This gives

$$\operatorname{Var} h(v_0, v_n) \leq 2 \sum_{r=1}^{2n} r\rho(r) + \sum_{r=n}^{2n} 2n\rho(r) \leq 4C. \tag{4}$$

Now let $Z_i := h(v_i, v_{i+1})$. The assumption on the coloring X implies that the process $Z = (Z_i)_{i \in \mathbb{Z}}$ is right-tail-trivial, so using (4) and applying Lemma 22 shows that Z is deterministic, which is to say that either $Z_i = 1$ for all i a.s. or $Z_i = -1$ for all i a.s. Without loss of generality, consider the former case. Then the coloring $(X_{(i,0)})_{i \in \mathbb{Z}}$ restricted to the axis is supported on the set of three 3-periodic colorings of the form $\cdots 123123 \cdots$. Stationarity implies that these three colorings must each have probability $1/3$, but the resulting process is not right-tail-trivial. $\square$

PROOF OF THEOREM 2(ii). Let X be an ffiid 3-coloring of $\mathbb{Z}^d$ with $d \geq 2$, and suppose for a contradiction that the coding radius satisfies $\mathbb{E}R^2 < \infty$. Similar to the proof of Theorem 1(ii), we may assume without loss of generality that $d = 2$, since restricting an ffiid process to the plane $\mathbb{Z}^2 \times \{0\}^{d-2}$ gives another ffiid process, and does not increase the coding radius. We claim also that X restricted to $\mathbb{Z} \times \{0\}$ is right-tail-trivial as required for Proposition 23. This is a consequence of a fact from ergodic theory: any process that is a factor of an i.i.d. process and takes values in $A^{\mathbb{Z}}$, where A is a finite set, is right-tail-trivial; see, for example, [8], Theorem 1 on page 283, Exercise 1 on page 280 and Definition 3 on page 181. Alternatively, an elementary argument shows that an ffiid process with $\mathbb{E}R^2 < \infty$ satisfies a stronger tail-triviality condition; we explain this at the end of the section—specifically we use Lemma 24 with $d = 2$.

We now bound $\rho(r)$. Let $\|u_1 + v_1 - u_2 - v_2\|_1 \geq 2r$ and $\|u_i - v_i\|_1 = 1$ for $i = 1, 2$. Write $H_i := h(u_i, v_i) - \mathbb{E}h(u_i, v_i)$, so that $\mathbb{E}H_i = 0$ and $|H_i| \leq 2$. Recall that R_v denotes the coding radius at vertex v, and define the event $E_i := \{R_{u_i} \vee$

$R_{v_i} > r/2 - 1\}$. Thus, the random variables $H_1 \mathbf{1}_{\overline{E}_1}$ and $H_2 \mathbf{1}_{\overline{E}_2}$ are independent, since they are functions of disjoint sets of i.i.d. variables. Writing $\varepsilon = \varepsilon(r) := \mathbb{P}(R > r/2 - 1)$, note that $\mathbb{P}(E_i) \le 2\varepsilon$, and also $\mathbb{E}(H_i \mathbf{1}_{\overline{E}_i}) = -\mathbb{E}(H_i \mathbf{1}_{E_i}) \le 4\varepsilon$. Therefore,

$$\begin{aligned} \mathbb{E}(H_1 H_2) &= \mathbb{E}(H_1 H_2 \mathbf{1}_{E_1 \cup E_2}) + \mathbb{E}(H_1 H_2 \mathbf{1}_{\overline{E}_1} \mathbf{1}_{\overline{E}_2}) \\ &\le 4\mathbb{P}(E_1 \cup E_2) + \mathbb{E}(H_1 \mathbf{1}_{\overline{E}_1}) \mathbb{E}(H_2 \mathbf{1}_{\overline{E}_2}) \\ &\le 16\varepsilon + 16\varepsilon^2 \le 32\varepsilon, \end{aligned}$$

and thus $\rho(r) \le 32\mathbb{P}(R > r/2 - 1)$. Proposition 23 gives $\sum_r r\rho(r) = \infty$, so $\sum_r r\mathbb{P}(R > r/2 - 1) = \infty$, which implies $\mathbb{E}R^2 = \infty$. $\square$

We conclude the section by giving the proof of Lemma 22, and also the elementary argument for tail-triviality mentioned above.

PROOF OF LEMMA 22. Let $(Z_i)_{i \in \mathbb{Z}}$ be stationary and ± 1-valued, and suppose $\operatorname{Var} \sum_{i=1}^n Z_i \le C^2$ for all n. We will deduce that Z is deterministic.

Let $\mathcal{F}_j$ be the σ-field generated by $Z_j, Z_{j+1}, \dots$, and consider the space of random variables $L^2(\mathcal{F}_j)$, with the norm $\|X\|_2 := (\mathbb{E}X^2)^{1/2}$. Write $\mu := \mathbb{E}Z_0$ and $S_j^k := \sum_{i=j}^{k-1} (Z_i - \mu)$, so that $\|S_j^k\|_2^2 \le C^2$ and in particular $S_j^k \in L^2(\mathcal{F}_j)$. Now define $\phi_j : L^2(\mathcal{F}_j) \to [0, \infty)$ by

$$\phi_j(X) := \limsup_{n \to \infty} \mathbb{E}\big(X + S_j^n\big)^2.$$

We will prove that ϕ_j has a unique global minimizer in $L^2(\mathcal{F}_j)$.

First note that by the Cauchy–Schwarz inequality and the uniform bound on $\|S_j^k\|_2^2$, the function ϕ_j satisfies the bounds

$$\|X\|_2^2 - 2C\|X\|_2 \le \phi_j(X) \le \|X\|_2^2 + 2C\|X\|_2 + C^2,$$

and, therefore, $\phi_j(X) < \infty$ for all $X \in L^2(\mathcal{F}_j)$, while $\phi_j(X) \to \infty$ as $\|X\|_2 \to \infty$.

We next claim that ϕ_j is continuous and strictly convex. To check continuity, let $X, Y \in L^2(\mathcal{F}_j)$ satisfy $\|X - Y\|_2 = \varepsilon$. Writing $Y + S_j^n = (X + S_j^n) + (Y - X)$ and using Cauchy–Schwarz again,

$$\phi_j(Y) \le \phi_j(X) + 2\varepsilon\phi_j(X)^{1/2} + \varepsilon^2. \tag{5}$$

Applying (5) in both directions, using (5) again to bound $\phi_j(Y)$ in terms of $\phi_j(X)$ on the right-hand side, and simplifying, we obtain

$$\big|\phi_j(Y) - \phi_j(X)\big| \le 2\varepsilon\phi_j(X)^{1/2} + 3\varepsilon^2,$$

from which continuity follows.

To check that ϕ_j is strictly convex, observe that for $X, Y \in L^2(\mathcal{F}_j)$,

$$\begin{aligned}\phi_j\left(\frac{X+Y}{2}\right) &= \limsup_{n\to\infty} \mathbb{E}\left[\left(\frac{(X+S_j^n)+(Y+S_j^n)}{2}\right)^2\right] \\ &= \limsup_{n\to\infty} \mathbb{E}\left[\frac{(X+S_j^n)^2+(Y+S_j^n)^2}{2} - \frac{(X-Y)^2}{4}\right] \\ &\le \frac{\phi_j(X)+\phi_j(Y)}{2} - \frac{\|X-Y\|_2^2}{4}.\end{aligned}$$

It now follows (see, e.g., [3], Theorem 2.11, Remarks 2.12, 2.13) that ϕ_j has a unique mimimizer. Let $X_j \in L^2(\mathcal{F}_j)$ minimize ϕ_j. For $j < k < n$ we have $S_j^n = S_j^k + S_k^n$, and hence $\phi_k(X + S_j^k) = \phi_j(X)$. Therefore,

$$X_j + S_j^k = X_k.$$

By construction, $(X_i)_{i\in\mathbb{Z}}$ is stationary. Since $X_j = X_{j+1} - S_j^{j+1} = X_{j+1} - Z_j + \mu$ we have

$$X_j \equiv X_{j+1} + 1 + \mu \mod 2. \tag{6}$$

Thus, $X_j \bmod 2 \in L^2(\mathcal{F}_{j+1})$, and by iterating we see that $X_j \bmod 2$ is in the right tail of $(Z_i)_{i\in\mathbb{Z}}$. Therefore, $X_j \bmod 2$ is an a.s. constant for each j. Since $(X_i)_{i\in\mathbb{Z}}$ is stationary, we have $X_0 \equiv X_1 \bmod 2$ a.s. Hence, (6) gives $\mu \equiv 1 \bmod 2$, that is, $\mu \in \{-1, +1\}$, so Z_0 is a.s. deterministic. $\square$

A process X on $\mathbb{Z}^d$ is called *fully tail-trivial* if every event in $\mathcal{T}(X) := \bigcap_{r\ge 0} \sigma(X_v : v \notin B(r))$ has probability zero or one. Of course, the restriction of a fully tail-trivial process to the axis is also fully tail-trivial and, therefore, right-tail-trivial. Hence, the following lemma suffices for our needs in the proof of Theorem 2(ii) above.

LEMMA 24. *Suppose X is an ffiid process on $\mathbb{Z}^d$. If the coding radius R satisfies $\mathbb{E}R^d < \infty$ then X is fully tail-trivial.*

PROOF. Let X be a finitary factor of the i.i.d. process Y with coding radius satisfying $\mathbb{E}R^d < \infty$. For $u, v \in \mathbb{Z}^d$ we write $u \hookrightarrow v$ for the event $\{|u - v| \le R_v\}$, that is, the event that u is within the ball that must be examined to determine X_v.

For positive integers $n < N$, define

$$E_{n,N} := \{\exists u \in B(n) \text{ and } v \notin B(N) \text{ s.t. } u \hookrightarrow v\}.$$

We claim that for any n we have $\mathbb{P}(E_{n,N}) \to 0$ as $N \to \infty$. Indeed, by translation-invariance and the assumption on R,

$$\sum_{v\in\mathbb{Z}^d} \mathbb{P}(0 \hookrightarrow v) = \sum_{v\in\mathbb{Z}^d} \mathbb{P}(-v \hookrightarrow 0) = \sum_{v\in\mathbb{Z}^d} \mathbb{P}(v \hookrightarrow 0) = \mathbb{E}|B(R)| < \infty.$$

(This is an instance of the "mass-transport principle"—see [5] for background.) Hence for any n,

$$\sum_{u\in B(n),v\in\mathbb{Z}^d} \mathbb{P}(u \hookrightarrow v) < \infty,$$

and thus

$$\mathbb{P}(E_{n,N}) \le \sum_{u\in B(n),v\notin B(N)} \mathbb{P}(u \hookrightarrow v) \to 0 \qquad \text{as } N \to \infty,$$

as claimed.

Now, fix $\varepsilon > 0$ and a tail event $A \in \mathcal{T}(X)$. Since X is a function of Y, we have $A \in \sigma(Y)$, so we can find an approximating cylinder event: there exist n and $A' \in \sigma(Y_v : v \in B(n))$ such that $\mathbb{P}(A\triangle A') < \varepsilon$. Let $A'' = A \setminus E_{n,N}$. By the above claim, for N large enough we have $\mathbb{P}(A\triangle A'') < \varepsilon$. On the other hand, since A is a tail event, $A \in \sigma(X_v : v \notin B(N))$, and so by the definition of $E_{n,N}$ we deduce $A'' \in \sigma(Y_v : v \notin B(n))$. Thus, A'' and A' are independent. Since ε was arbitrary this implies that A is independent of itself, that is, $\mathbb{P}(A) \in \{0, 1\}$. □

8. Finitely dependent coloring.

In this section, we prove two results on k-dependent coloring. See [14] for more on this topic.

PROOF OF THEOREM 4. Suppose $(X_i)_{i\in\mathbb{Z}}$ is a translation-invariant 1-dependent coloring. Let $\mathcal{F}_i := \sigma(\dots, X_{i-1}, X_i)$, and define the random variable

$$Y_i := \mathbb{P}(X_{i+1} = 1 \mid \mathcal{F}_i).$$

Let $m = \operatorname{ess\,sup} Y_1$; then since $X_2 \ne 1$ on the event $X_1 = 1$ we have a.s.

$$Y_1 \le m\mathbf{1}[X_1 \ne 1].$$

Since X_2 is independent of $\mathcal{F}_0$, we deduce

$$\begin{aligned}\mathbb{P}(X_2 = 1) = \mathbb{P}(X_2 = 1 \mid \mathcal{F}_0) &= \mathbb{E}(Y_1 \mid \mathcal{F}_0)\\ &\le m\mathbb{P}(X_1 \ne 1 \mid \mathcal{F}_0) = m(1 - Y_0),\end{aligned}$$

and since this holds a.s. we deduce

$$\mathbb{P}(X_2 = 1) \le m(1 - m) \le 1/4.$$

Similarly, we have $\mathbb{P}(X_2 = k) \le 1/4$ for each color $k = 1, \dots, q$, so $q \ge 4$. □

Finally, we note the following consequence of the results of the previous section.

COROLLARY 25. *Let $d \ge 2$ and $k \ge 1$. There exists no stationary k-dependent 3-coloring of $\mathbb{Z}^d$.*

PROOF. By restricting to a plane, it is enough to prove the $d = 2$ case. We use Proposition 23. It is elementary to check that the restriction of a k-dependent process to the axis is right-tail-trivial, and that the correlation function $\rho(r)$ is zero for $r > k + 2$. □

In contrast with Corollary 25, in dimension $d = 1$, a stationary 2-dependent 3-coloring was constructed in [14].

Open problems.

(i) What is the largest α for which there exists an ffiid 3-coloring of $\mathbb{Z}^d$ whose coding radius has finite α-moment, for each $d \geq 2$? (Our results show that it is at most 2, and at least some small positive number.)
(ii) Does there exist, for some $d \geq 2$, a shift of finite type S on $\mathbb{Z}^d$ that contains no constant configuration, but that admits some ffiid process X with $X \in S$ a.s. whose coding radius tail decays strictly faster than a tower function?
(iii) Does there exist, for some $d \geq 2$, a shift of finite type S that admits an ffiid X with all moments of the coding radius finite, but that admits no X with tower function decay? (For example, can the optimal tail decay be exponential?)
(iv) Does there exist, for some $d \geq 2$, a shift of finite type S that admits an ffiid X, but such that all moments of the coding radius are infinite for every such X?

Our results imply negative answers to (ii), (iii), (iv) in dimension $d = 1$.

Acknowledgments. We thank Itai Benjamini, Jeff Steif and Benjamin Weiss for valuable discussions. We thank an anonymous referee for helpful suggestions.

REFERENCES

[1] AARONSON, J., GILAT, D., KEANE, M. and DE VALK, V. (1989). An algebraic construction of a class of one-dependent processes. *Ann. Probab.* **17** 128–143. MR0972778

[2] ALON, N. and FELDHEIM, O. N. (2014). A note on general sliding window processes. *Electron. Commun. Probab.* **19** no. 66, 7. MR3262072

[3] BARBU, V. and PRECUPANU, T. (2012). *Convexity and Optimization in Banach Spaces*, 4th ed. Springer, Dordrecht. MR3025420

[4] BAXTER, R. J. (1982). *Exactly Solved Models in Statistical Mechanics*. Academic Press, New York.

[5] BENJAMINI, I., LYONS, R., PERES, Y. and SCHRAMM, O. (1999). Group-invariant percolation on graphs. *Geom. Funct. Anal.* **9** 29–66. MR1675890

[6] BURTON, R. M., GOULET, M. and MEESTER, R. (1993). On 1-dependent processes and k-block factors. *Ann. Probab.* **21** 2157–2168. MR1245304

[7] COLE, R. and VISHKIN, U. (1986). Deterministic coin tossing with applications to optimal parallel list ranking. *Inf. Control* **70** 32–53. MR0853994

[8] CORNFELD, I. P., FOMIN, S. V. and SINAĬ, Y. G. (1982). *Ergodic Theory*. *Grundlehren der Mathematischen Wissenschaften* **245**. Springer, New York. MR0832433

[9] ERDŐS, P., FRANKL, P. and FÜREDI, Z. (1985). Families of finite sets in which no set is covered by the union of r others. *Israel J. Math.* **51** 79–89. MR0804477

[10] GALVIN, D. (2003). On homomorphisms from the Hamming cube to **Z**. *Israel J. Math.* **138** 189–213. MR2031957

[11] GOLDBERG, A. V., PLOTKIN, S. A. and SHANNON, G. E. (1988). Parallel symmetry-breaking in sparse graphs. *SIAM J. Discrete Math.* **1** 434–446. MR0968851

[12] GRIMMETT, G. (1999). *Percolation*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* **321**. Springer, Berlin. MR1707339

[13] HOLROYD, A. E. (2017). One-dependent coloring by finitary factors. *Ann. Inst. H. Poincaré Probab. Statist.* **53** 753–765. MR3634273

[14] HOLROYD, A. E. and LIGGETT, T. M. (2016). Finitely dependent coloring. *Forum of Math., Pi* **4** e9. MR3570073

[15] HOLROYD, A. E. and LIGGETT, T. M. (2015). Symmetric 1-dependent colorings of the integers. *Electron. Commun. Probab.* **20** no. 31, 8. MR3327870

[16] IBRAGIMOV, I. A. and LINNIK, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, Groningen. MR0322926

[17] JANSON, S. (1984). Runs in m-dependent sequences. *Ann. Probab.* **12** 805–818. MR0744235

[18] LINIAL, N. (1987). Distributive graph algorithms—global solutions from local data. In *28th Annual Symposium on Foundations of Computer Science* 331–335. IEEE, New York.

[19] LINIAL, N. (1992). Locality in distributed graph algorithms. *SIAM J. Comput.* **21** 193–201. MR1148825

[20] NAOR, M. (1991). A lower bound on probabilistic algorithms for distributive ring coloring. *SIAM J. Discrete Math.* **4** 409–412. MR1105946

MICROSOFT RESEARCH
ONE MICROSOFT WAY
REDMOND, WASHINGTON 98052
USA
E-MAIL: holroyd@microsoft.com
dbwilson@microsoft.com