ERGODIC THEOREMS FOR THE SHIFT ACTION AND POINTWISE VERSIONS OF THE ABÉRT–WEISS THEOREM*

BY

ANTON BERNSHTEYN

*Department of Mathematics, University of Illinois at Urbana–Champaign
Urbana, IL 61801, USA*

and

*Department of Mathematical Sciences, Carnegie Mellon University
Pittsburgh, PA 15213, USA
e-mail: bernsht2@illinois.edu, abernsht@math.cmu.edu*

ABSTRACT

Let Γ be a countably infinite group. A common theme in ergodic theory is to start with a probability measure-preserving (p.m.p.) action $\Gamma \curvearrowright (X, \mu)$ and a map $f \in L^1(X, \mu)$, and to compare the global average $\int f \, d\mu$ of f to the pointwise averages $|D|^{-1} \sum_{\delta \in D} f(\delta \cdot x)$, where $x \in X$ and D is a nonempty finite subset of Γ. The basic hope is that, when D runs over a suitably chosen infinite sequence, these pointwise averages should converge to the global value for μ-almost all x.

In this paper we prove several results that refine the above basic paradigm by uniformly controlling the averages over specific sets D rather than considering their limit as $|D| \to \infty$. Our results include ergodic theorems for the Bernoulli shift action $\Gamma \curvearrowright ([0;1]^\Gamma, \lambda^\Gamma)$ and strengthenings of the theorem of Abért and Weiss that the shift is weakly contained in every free p.m.p. action of Γ. In particular, we establish a purely Borel version of the Abért–Weiss theorem for finitely generated groups of subexponential growth. The central role in our arguments is played by the recently introduced measurable versions of the Lovász Local Lemma, due to the current author and to Csóka, Grabowski, Máthé, Pikhurko, and Tyros.

1. Introduction

The Lovász Local Lemma (the LLL for short) is a powerful tool in probabilistic combinatorics, introduced by Erdős and Lovász [16]. The LLL is mostly used to obtain existence results, and it is particularly well-suited for showing that a given structure X admits a coloring satisfying some "local" constraints. Roughly speaking, in order for the LLL to apply in this context, two requirements must be met: First, a random coloring should be "likely" to fulfill each individual constraint; second, the constraints must not interact with each other "too much." For the precise statement, see §6.A.

It has been a matter of interest to determine if the LLL can be used to derive conclusions that are, in some sense, "constructive" (as opposed to pure existence results). A decisive breakthrough was made by Moser and Tardos [26], who developed an algorithmic approach to the LLL. (The work of Moser and Tardos was preceded by a line of earlier results, starting with Beck's paper [5]; for more details, see the references in [26].) The Moser–Tardos method proved quite versatile and was adapted to establish "constructive" analogs of the LLL in a variety of different contexts. For example, Rumyantsev and Shen [27] proved a computable version of the LLL. Here we will be focused on the measurable versions of the LLL that were studied in [6] by the current author and in [14] by Csóka, Grabowski, Máthé, Pikhurko and Tyros (see also [23] for related work by Kun).

Measurable analogs of the LLL are designed to apply in the following framework. Let (X, μ) be a standard probability space and let C be a set of colors (we will only consider the case when C is finite). Suppose we are looking for a coloring $f: X \to C$ that fulfills a family $\mathscr{B}$ of constraints. Under suitable assumptions, the ordinary LLL implies that such a coloring f exists; however, this f need not behave well with respect to the measurable structure on (X, μ). In contrast to that, measurable versions of the LLL can provide a μ-measurable (or sometimes even Borel) function $f: X \to C$ that satisfies the constraints $\mathscr{B}$, or at least does so on a "large" subset of X. Such results appear to be particularly relevant in ergodic theory, since many concepts pertaining to measure-preserving group actions are phrased in terms of measurable partitions of the underlying probability space—which can naturally be thought of as measurable colorings. Some ergodic-theoretic applications of the LLL can be found

in [6, 7]. Here we present further consequences of the LLL in measurable dynamics, specifically in the study of ergodic averages and of weak containment of measure-preserving group actions.

Our arguments employ a general approach that is standard in combinatorics, in particular in graph coloring theory (see, e.g., the book [25] for many examples). The first step is to use concentration of measure to obtain strong upper bounds on probabilities of certain "bad" random events; the LLL is then invoked to eliminate all the "bad" events. Nontrivial results can also be derived by combining the concentration of measure bounds with more classical tools, such as the Borel–Cantelli lemma (Theorem 2.1 below is an example). Roughly speaking, using the LLL instead of the Borel–Cantelli lemma results in replacing pointwise convergence with approximation in the ∞-norm.

2. Statements of results

Throughout, Γ denotes a countably infinite group with identity element $\mathbf{1}$. We study **probability measure-preserving (p.m.p.) actions** of Γ, i.e., actions of the form $\alpha\colon \Gamma \curvearrowright (X, \mu)$, where (X, μ) is a standard probability space and the measure μ is α-invariant. We also consider, more generally, **Borel actions** $\alpha\colon \Gamma \curvearrowright X$, i.e., actions of Γ on a standard Borel space X by Borel automorphisms.

Given a set A, the **shift action** $\sigma_A\colon \Gamma \curvearrowright A^\Gamma$ on the set of all maps $x\colon \Gamma \to A$ is defined by

$$(\gamma \cdot x)(\delta) \coloneqq x(\delta\gamma) \quad \text{for all } x \in A^\Gamma \text{ and } \gamma,\, \delta \in \Gamma.$$

We are particularly interested in the case when A is the unit interval $[0; 1]$ equipped with the Lebesgue probability measure λ. (Owing to the measure isomorphism theorem [19, Theorem 17.41], any other atomless standard probability space could be used instead.) To unclutter the notation, set

$$(\Omega, \boldsymbol{\lambda}) \coloneqq ([0; 1]^\Gamma, \lambda^\Gamma)$$

and $\sigma \coloneqq \sigma_{[0;1]}$. Note that the action $\sigma\colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$ is measure-preserving.

2.A. ERGODIC THEOREMS FOR THE SHIFT ACTION. Let $\alpha \colon \Gamma \curvearrowright (X, \mu)$ be a p.m.p. action. Given $f \in L^1(X, \mu)$, we can compute its **global average**:

$$\mathbb{E}_\mu f \coloneqq \int_X f \, \mathrm{d}\mu,$$

and compare it to the **pointwise averages** of the form

$$\mathbb{E}_D f(x) \coloneqq \frac{1}{|D|} \sum_{\delta \in D} f(\delta \cdot x),$$

where $x \in X$ and D is a nonempty finite subset of Γ.   Note that $\mathbb{E}_D \colon L^1(X, \mu) \to L^1(X, \mu)$ is a linear operator of norm 1: The lower bound on $\|\mathbb{E}_D\|_{\mathrm{op}}$ is witnessed by the constant 1 function, while the upper bound follows from the fact that, since μ is α-invariant, we have $\mathbb{E}_\mu \mathbb{E}_D f = \mathbb{E}_\mu f$, and hence

$$\|\mathbb{E}_D f\|_1 = \mathbb{E}_\mu |\mathbb{E}_D f| \leqslant \mathbb{E}_\mu \mathbb{E}_D |f| = \mathbb{E}_\mu |f| = \|f\|_1.$$

Assuming the action α is ergodic, one hopes to show that the pointwise averages $\mathbb{E}_D f$ converge, in a suitable sense, to $\mathbb{E}_\mu f$, as D ranges over a given infinite family of finite subsets of Γ. Results of this kind are usually referred to as ergodic theorems (often with adjectives indicating the mode of convergence, such as "pointwise"). Two prototypical examples are von Neumann's [31] and Birkhoff's [8] ergodic theorems. Both of these classical results apply when $\Gamma = \mathbb{Z}$ and D ranges over the sets of the form $\{0, 1, \ldots, n-1\}$ with $n \in \mathbb{N}^+$. Von Neumann's theorem yields convergence in the 2-norm (assuming $f \in L^2(X, \mu)$ to begin with), while Birkhoff's result ensures pointwise convergence almost everywhere.   An extension of Birkhoff's pointwise ergodic theorem to all amenable Γ was obtained by Lindenstrauss [24]; there D ranges over a tempered Følner sequence (the special case of Lindenstrauss's result for $f \in L^2(X, \mu)$ follows from the earlier work of Shulman, see [29, §5.6]). Generalizing ergodic theorems beyond the realm of amenable groups is a major challenge; for further background, see, e.g., [4, 10, 9] and the references therein.

Here we work with an arbitrary group Γ; moreover, the only condition on the sequence $(D_n)_{n \in \mathbb{N}}$ of averaging sets is that $|D_n|$ grows sufficiently quickly with n. On the other hand, instead of studying arbitrary ergodic actions, we focus our attention on the shift action $\sigma \colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$ in the hope of exploiting its mixing properties. Our first result is a pointwise ergodic theorem for continuous functions $f \colon \Omega \to \mathbb{C}$:

THEOREM 2.1 (Pointwise ergodic theorem for continous maps on the shift): *Let* $(D_n)_{n \in \mathbb{N}}$ *be a sequence of finite subsets of* Γ *such that* $|D_n|/\log n \to \infty$. *Then, for all continuous* $f \colon \Omega \to \mathbb{C}$,

$$\lim_{n \to \infty} \mathbb{E}_{D_n} f(x) = \mathbb{E}_{\boldsymbol{\lambda}} f, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega.$$

Since the set of all continuous functions is dense in $L^1(\Omega, \boldsymbol{\lambda})$ and $\|\mathbb{E}_D\|_{\mathrm{op}} = 1$ for all nonempty finite $D \subset \Gamma$, Theorem 2.1 has the following immediate corollary:

COROLLARY 2.2 (Mean ergodic theorem for the shift): *Let* $(D_n)_{n \in \mathbb{N}}$ *be a sequence of finite subsets of* Γ *such that* $|D_n|/\log n \to \infty$. *Then, for all* $f \in L^1(\Omega, \boldsymbol{\lambda})$, *we have*

$$\lim_{n \to \infty} \mathbb{E}_{D_n} f = \mathbb{E}_{\boldsymbol{\lambda}} f \quad \text{in } L^1(\Omega, \boldsymbol{\lambda}).$$

It is natural to ask whether Theorem 2.1 can be extended to all $f \in L^1(\Omega, \boldsymbol{\lambda})$. The answer turns out to be negative even if the lower bound on the growth rate of the averaging sets is raised, as the constructions of Akcoglu and del Junco [2] and del Junco and Rosenblatt [15] (with minor modifications) demonstrate:

THEOREM 2.3 (ess. Akcoglu–del Junco [2] and del Junco–Rosenblatt [15]): *Suppose that* $\Gamma = \mathbb{Z}$ *and let* $h \colon \mathbb{N} \to \mathbb{N}$ *be an arbitrary function. There exists a sequence* $(D_n)_{n \in \mathbb{N}}$ *of finite subsets of* $\mathbb{Z}$ *with the following properties:*

– *each* D_n *is an interval, i.e., a set of the form* $\{s, s+1, \ldots, s+\ell-1\}$ *for* $s \in \mathbb{Z}$ *and* $\ell \in \mathbb{N}^+$;
– $|D_n| \geqslant h(n)$ *for all* $n \in \mathbb{N}$;
– *for every free p.m.p. action* $\mathbb{Z} \curvearrowright (X, \mu)$, *there is a Borel set* $A \subseteq X$ *such that*

$$\liminf_{n \to \infty} \mathbb{E}_{D_n} \mathbb{1}_A(x) = 0 \quad and \quad \limsup_{n \to \infty} \mathbb{E}_{D_n} \mathbb{1}_A(x) = 1, \quad \text{for } \mu\text{-a.e. } x \in X,$$

where $\mathbb{1}_A \colon X \to \{0,1\}$ *is the indicator function of* A. *Moreover, the family of such sets* A *is comeager in the measure algebra* $\mathrm{MAlg}(X, \mu)$.

For completeness, we sketch a proof of Theorem 2.3 using Rokhlin's lemma in the appendix.

As mentioned in the introduction, Theorem 2.1 follows by combining a concentration of measure inequality with the Borel–Cantelli lemma. We now turn to further results that can be obtained if the Borel–Cantelli lemma is replaced by the LLL.

For a p.m.p. action $\alpha\colon\Gamma\curvearrowright(X,\mu)$, $f\in L^1(X,\mu)$, and a nonempty finite set $D\subset\Gamma$, define the **discrepancy norm** of f with respect to D by the formula

$$\|f\|_D^{\mathrm{disc}}:=\|\mathbb{E}_Df-\mathbb{E}_\mu f\|_\infty.$$

(Here $\|\cdot\|_\infty$ is the ∞-norm in the sense of $L^\infty(X,\mu)$.) Even if $f\colon\Omega\to\mathbb{C}$ is continuous, its discrepancy norm may be separated from 0. For instance, consider the continuous map

$$f\colon\Omega\to[-1;1]\colon x\mapsto -1+2\cdot x(\mathbf{1}).$$

Then $\mathbb{E}_{\boldsymbol{\lambda}}f=0$ yet $\|\mathbb{E}_Df\|_\infty=1$, and hence $\|f\|_D^{\mathrm{disc}}=1$, for all nonempty finite $D\subset\Gamma$. However, we show that any $f\in L^1(\Omega,\boldsymbol{\lambda})$ can be written as a sum of two functions $g,h\in L^1(\Omega,\boldsymbol{\lambda})$, where g is small in the discrepancy norm, while h is small in the 1-norm:

THEOREM 2.4 (L^∞-ergodic theorem for the shift): *For all $f\in L^1(\Omega,\boldsymbol{\lambda})$ and $\varepsilon>0$, there exists $C>0$ with the following property:*

Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of finite subsets of Γ with $|D_n|\geqslant C\log(n+2)$ for all $n\in\mathbb{N}$. Then there exist $g,h\in L^1(\Omega,\boldsymbol{\lambda})$ such that $f=g+h$, $\|h\|_1\leqslant\varepsilon$, and $\|g\|_{D_n}^{\mathrm{disc}}\leqslant\varepsilon$ for all $n\in\mathbb{N}$.

Note that Theorem 2.4 also yields Corollary 2.2.

Our ultimate goal in this subsection is to sharpen Theorem 2.4 by considering other statistical properties of the function f, beside its average $\mathbb{E}_{\boldsymbol{\lambda}}f$. This is made precise by the following formalism. Let K be a compact metric space. We use $\mathrm{Prob}(K)$ to denote the space of all probability Borel measures on K equipped with the usual weak-$*$ topology (see, e.g., [19, §17.E]). Given a p.m.p. action $\alpha\colon\Gamma\curvearrowright(X,\mu)$ and a Borel function $f\colon X\to K$, define

$$\mathbb{M}_\mu f:=f_*(\mu),$$

where $f_*\colon\mathrm{Prob}(X)\to\mathrm{Prob}(K)$ is the pushforward map. (This notation is intended to be reminiscent of $\mathbb{E}_\mu f$, while the letter "$\mathbb{M}$" emphasizes that $\mathbb{M}_\mu f$ is a measure.) For $x\in X$ and a nonempty finite set $D\subset\Gamma$, let $\upsilon_{x,D}$ be the probability measure on X with (finite) support $D\cdot x$ given by

$$\upsilon_{x,D}(\{y\}):=\frac{1}{|D|}\cdot|\{\delta\in D:\delta\cdot x=y\}|\quad\text{for all }y\in D\cdot x.$$

If the α-stabilizer of x is trivial, then $\upsilon_{x,D}$ is simply the uniform probability measure on $D \cdot x$. Let

$$\mathbb{M}_D f(x) := f_*(\upsilon_{x,D}).$$

The measures $\mathbb{M}_\mu f$ and $\mathbb{M}_D f(x)$ are points in $\mathrm{Prob}(K)$ that encode the global and the pointwise statistics of f, respectively. In particular, when K is a subset of $\mathbb{C}$, $\mathbb{M}_\mu f$ and $\mathbb{M}_D f(x)$ contain the information about $\mathbb{E}_\mu f$ and $\mathbb{E}_D f(x)$; explicitly,

$$(2.5) \qquad \mathbb{E}_\mu f = \int_K z \, \mathrm{d}(\mathbb{M}_\mu f)(z) \quad \text{and} \quad \mathbb{E}_D f(x) = \int_K z \, \mathrm{d}(\mathbb{M}_D f(x))(z).$$

We wish to also take into account more detailed information about the interaction of f with the action α. Toward that end, let $\pi_f \colon X \to K^\Gamma$ denote the equivariant map given by

$$\pi_f(x)(\gamma) := f(\gamma \cdot x) \quad \text{for all } x \in X \text{ and } \gamma \in \Gamma.$$

The map π_f is called the **symbolic representation**, or the **coding map**, of the dynamical system (X, Γ, α, f). Notice that the projection function $\mathsf{p} \colon K^\Gamma \to K \colon \kappa \mapsto \kappa(\mathbf{1})$ satisfies $f = \mathsf{p} \circ \pi_f$ and gives rise to a continuous map $\mathsf{p}_* \colon \mathrm{Prob}(K^\Gamma) \to \mathrm{Prob}(K)$ such that $f_* = \mathsf{p}_* \circ (\pi_f)_*$. This observation shows that, by considering π_f, we achieve greater generality than just by working with f itself.

Given a standard probability space (X, μ) and a compact metric space $(K, \mathfrak{d})$, let $\mathfrak{B}(X, K)$ denote the set of all Borel functions from X to K. We equip $\mathfrak{B}(X, K)$ with a psedometric $\mathfrak{d}_\mu$ given by

$$\mathfrak{d}_\mu(f, g) := \int_X \mathfrak{d}(f(x), g(x)) \, \mathrm{d}\mu(x).$$

If K is a subset of $\mathbb{C}$ equipped with the metric $\mathfrak{d}(z_1, z_2) = |z_1 - z_2|$, then $\mathfrak{d}_\mu(f, g) = \|f - g\|_1$.

THEOREM 2.6 (Pushforward-ergodic theorem for the shift): *Let* $(K, \mathfrak{d})$ *be a compact metric space and let* $f \colon \Omega \to K$ *be a Borel function. For any* $\varepsilon > 0$ *and an open neighborhood* U *of the measure* $\mathbb{M}_{\boldsymbol{\lambda}} \pi_f$, *there exists* $C > 0$ *with the following property:*

Let $(D_n)_{n \in \mathbb{N}}$ *be a sequence of finite subsets of* Γ *with* $|D_n| \geqslant C \log(n+2)$ *for all* $n \in \mathbb{N}$. *Then there is a Borel map* $g \colon \Omega \to K$ *such that* $\mathfrak{d}_{\boldsymbol{\lambda}}(f, g) \leqslant \varepsilon$ *and*

$$\mathbb{M}_{D_n} \pi_g(x) \in U, \quad \text{for all } n \in \mathbb{N} \text{ and for } \boldsymbol{\lambda}\text{-a.e. } x \in X.$$

In the light of (2.5), it is clear that Theorem 2.4 is a special case of Theorem 2.6.

We end this subsection with a simple application of Theorem 2.6. Recall that a group Γ is called **residually finite** if the intersection of all its subgroups of finite index is trivial. The following is an easy observation:

PROPOSITION 2.7: *A countable group Γ is residually finite if and only if every open neighborhood $U \subseteq \mathrm{Prob}(\Omega)$ of $\boldsymbol{\lambda}$ contains a finitely supported measure ν that is shift-invariant.*

Proof. Let $(\nu_n)_{n \in \mathbb{N}}$ be a sequence of finitely supported shift-invariant measures on Ω that converges to $\boldsymbol{\lambda}$. This gives us a sequence of actions of Γ on the finite sets $X_n \coloneqq \mathrm{supp}(\nu_n)$ and, since $\nu_n \to \boldsymbol{\lambda}$, each nonidentity group element $\gamma \in \Gamma$ acts on X_n nontrivially for all large enough n. This shows that Γ is residually finite.

Conversely, suppose that Γ is residually finite and let $(\Delta_n)_{n \in \mathbb{N}}$ be a decreasing sequence of finite index subgroups of Γ with trivial intersection. For $k \in \mathbb{N}^+$, let

$$Q_k \coloneqq \{0, 1/k, \ldots, (k-1)/k\} \subset [0; 1].$$

Let $P(k, n)$ denote the set of all maps $x \colon \Gamma \to Q_k$ that are constant on the right cosets of Δ_n. Then the set $P(k, n)$ is finite and shift-invariant, and, letting $\nu_{k,n}$ be the uniform probability measure on $P(k, n)$, we see that $\nu_{k,n} \to \boldsymbol{\lambda}$ as $k, n \to \infty$. ∎

Motivated by Proposition 2.7, we say that a group Γ is **approximately residually finite** if for every open neighborhood U of $\boldsymbol{\lambda}$, there is a finitely supported measure ν such that $\gamma \cdot \nu \in U$ for all $\gamma \in \Gamma$. Proposition 2.7 implies that every residually finite group is approximately residually finite, so our terminology is consistent. An intuitive way of thinking about approximate residual finiteness is as follows: To show that a group Γ is approximately residually finite, we have to find finite subsets $X \subset \Omega$ that are "almost uniformly distributed" over the space $(\Omega, \boldsymbol{\lambda})$ and also remain such when shifted by any $\gamma \in \Gamma$. We remark that a random finite set X fails to have this property: For any $n \in \mathbb{N}^+$, the product action $\sigma^n \colon \Gamma \curvearrowright (\Omega^n, \boldsymbol{\lambda}^n)$ is ergodic, and hence if $x_1, \ldots, x_n \in \Omega$ are chosen randomly and independently from each other, then, with probability 1, for every open $V \subseteq \Omega$ there is some $\gamma \in \Gamma$ such that $\gamma \cdot x_1, \ldots, \gamma \cdot x_n \in V$. Nevertheless, we have the following:

COROLLARY 2.8 (to Theorem 2.6): *Every countable group is approximately residually finite.*

Proof. Let U be an open neighborhood of $\boldsymbol{\lambda}$. It suffices to exhibit a finitely supported measure $\nu \in U$ such that $\nu \cdot \gamma \in U$ for all $\gamma \in \Gamma$, where the **right shift action** $\Omega \curvearrowleft \Gamma$ is given by

$$(x \cdot \gamma)(\delta) \coloneqq x(\gamma\delta) \quad \text{for all } x \in \Omega \text{ and } \gamma, \delta \in \Gamma.$$

Applying Theorem 2.6 with $K = [0; 1]$ and $f = (x \mapsto x(\mathbf{1}))$, we obtain a nonempty finite set $D \subset \Gamma$ and a Borel map $g \colon \Omega \to [0; 1]$ such that $\mathbb{M}_D \pi_g(x) \in U$ for $\boldsymbol{\lambda}$-a.e. $x \in \Omega$. Since Γ is countable and the measure $\boldsymbol{\lambda}$ is right-shift-invariant, there is $x \in \Omega$ such that $\mathbb{M}_D \pi_g(x \cdot \gamma) \in U$ for all $\gamma \in \Gamma$. Set $\nu \coloneqq \mathbb{M}_D \pi_g(x)$. Then ν is finitely supported; furthermore, it is straightforward to verify, using the (left-)equivariance of π_g and the fact that the left and the right shift actions of Γ on Ω commute with each other, that $\nu \cdot \gamma = \mathbb{M}_D \pi_g(x \cdot \gamma)$ for all $\gamma \in \Gamma$. Hence, ν is as desired. ∎

Since the above argument only involves the properties of g on a countable subset of Ω, Corollary 2.8 can also be derived directly from the classical LLL, without using its measurable analogs.

2.B. POINTWISE VERSIONS OF THE ABÉRT–WEISS THEOREM. So far we have considered the action $\sigma \colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$ on its own. Now we would like to discuss the relationship between σ and other actions of Γ.

The concepts of weak containment and weak equivalence of p.m.p. actions were introduced by Kechris in [20, §10(C)]. They are inspired by the analogous notions for unitary representations and are closely related to the so-called local-global convergence in the theory of graph limits [18]. The relation of weak equivalence is much coarser than the isomorphism relation, which makes it relatively well-behaved. On the other hand, several interesting parameters associated with p.m.p. actions—such as their cost, type, etc.—turn out to be invariants of weak equivalence. Due to these favorable properties, the relations of weak containment and weak equivalence have attracted a considerable amount of attention in recent years. For a survey of the topic, see [12].

Roughly speaking, a p.m.p. action $\alpha \colon \Gamma \curvearrowright (X, \mu)$ is weakly contained in another p.m.p. action $\beta \colon \Gamma \curvearrowright (Y, \nu)$ if for every compact metric space K and for any Borel map $f \colon X \to K$, the interaction of f with α can be arbitrarily well "simulated" by a Borel map $g \colon Y \to K$ interacting with β. Here is a precise definition:

Definition 2.9 (Weak containment; [12, §2.2(2)]): Let $\alpha \colon \Gamma \curvearrowright (X, \mu)$ and
$\beta \colon \Gamma \curvearrowright (Y, \nu)$ be p.m.p. actions of Γ. We say that α is **weakly contained**
in β, in symbols $\alpha \preccurlyeq \beta$, if for any compact metric space K, a Borel function
$f \colon X \to K$, and an open neighborhood U of the measure $\mathbb{M}_\mu \pi_f$, there exists a
Borel map $g \colon Y \to K$ such that $\mathbb{M}_\nu \pi_g \in U$. If both $\alpha \preccurlyeq \beta$ and $\beta \preccurlyeq \alpha$, then α
and β are said to be **weakly equivalent**, in symbols $\alpha \simeq \beta$.

Weak containment can be defined in a number of equivalent ways, several of
which can be found in [12, §§2.1, 2.2]. The characterization given above is due
to Abért and Weiss [1, Lemma 8] (see also [30, Proposition 3.5]). We sometimes
write $(\alpha, \mu) \preccurlyeq (\beta, \nu)$ instead of $\alpha \preccurlyeq \beta$ in order to emphasize the dependence of
weak containment on the invariant measures μ and ν.

Burton [11, Corollary 4.2] (see [12, Theorem 3.3]) proved that there exist
continuum many distinct weak equivalence classes of (not necessarily ergodic)
p.m.p. actions of Γ. Glasner, Thouvenot and Weiss [17] and independently
Greg Hjorth (unpublished) proved that the pre-order of weak containment has
a maximum element (see also [12, Theorem 3.1]). A complementary result of
Abért and Weiss [1, Theorem 1] (see also [12, Theorem 3.5]) asserts that the shift
action $\sigma \colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$ is minimum among all p.m.p. actions $\alpha \colon \Gamma \curvearrowright (X, \mu)$ that
are (**almost everywhere**) **free**, i.e., such that the α-stabilizer of μ-a.e. $x \in X$
is trivial:

THEOREM 2.10 (Abért–Weiss [1, Theorem 1]): *Let* $\alpha \colon \Gamma \curvearrowright (X, \mu)$ *be an almost
everywhere free p.m.p. action of* Γ. *Then* $(\sigma, \boldsymbol{\lambda}) \preccurlyeq (\alpha, \mu)$; *or, explicitly, the
following statement holds:*

Let K *be a compact metric space and let* $f \colon \Omega \to K$ *be a Borel function.
Then, for any open neighborhood* U *of the measure* $\mathbb{M}_{\boldsymbol{\lambda}} \pi_f$, *there is a Borel map*
$g \colon X \to K$ *such that* $\mathbb{M}_\mu \pi_g \in U$.

We strengthen Theorem 2.10 by replacing the measure $\mathbb{M}_\mu \pi_g$ by its pointwise
analogs of the form $\mathbb{M}_D \pi_g(x)$. Moreover, our result applies to actions that are
not necessarily free but only "close enough" to being free. Specifically, for a
set $S \subseteq \Gamma$, we say that an action $\alpha \colon \Gamma \curvearrowright X$ is S-**free** if for all $\gamma, \delta \in S$ and
$x \in X$, $\gamma \cdot x = \delta \cdot x$ implies $\gamma = \delta$. (Thus, "free" is the same as "Γ-free.") Given
a sequence of sets $S_1, \ldots, S_n \subseteq \Gamma$, we say that α is $(S_1, \ldots, S_n)$-**free** if α is
S_i-free for each $1 \leqslant i \leqslant n$.

THEOREM 2.11 (Pointwise Abért–Weiss): *Let K be a compact metric space and let $f\colon \Omega \to K$ be a Borel function. For any open neighborhood U of the measure $\mathbb{M}_{\boldsymbol{\lambda}}\pi_f$, there exist $C > 0$ and a finite set $S \subset \Gamma$ with the following property:*

Let D be a finite subset of Γ with $|D| \geqslant C$ and let $\alpha\colon \Gamma \curvearrowright X$ be an (S, D)-free Borel action of Γ. Then, for any $\mu \in \mathrm{Prob}(X)$ and $\delta > 0$, there is a Borel map $g\colon X \to K$ such that

$$\mu(\{x \in X : \mathbb{M}_D\pi_g(x) \in U\}) \geqslant 1 - \delta.$$

Remarks 2.12: Let us make a few comments about the statement of Theorem 2.11.

(i) To see that Theorem 2.11 is a strengthening of the Abért–Weiss theorem, let $\alpha\colon \Gamma \curvearrowright (X, \mu)$ be a free p.m.p. action. Given a compact metric space K and a Borel function $f\colon \Omega \to K$, we can apply Theorem 2.11 to obtain a finite set $D \subset \Gamma$ and a Borel map $g\colon X \to K$ such that the pushforward measure $\mathbb{M}_D\pi_g(x)$ is arbitrarily close to $\mathbb{M}_{\boldsymbol{\lambda}}\pi_f$, for all points $x \in X$ away from a set of arbitrarily small measure. The α-invariance of μ yields

$$\mu = \int_X v_{x,D} \, \mathrm{d}\mu(x), \quad \text{hence} \quad \mathbb{M}_\mu\pi_g = \int_X \mathbb{M}_D\pi_g(x) \, \mathrm{d}\mu(x),$$

and thus $\mathbb{M}_\mu\pi_g$ is also close to $\mathbb{M}_{\boldsymbol{\lambda}}\pi_f$, as desired.

(ii) The measure μ in Theorem 2.11 is not required to be α-invariant (or even α-quasi-invariant) and is only used to bound the set of all $x \in X$ with $\mathbb{M}_D\pi_g(x) \notin U$.

(iii) We emphasize that the averaging set D in Theorem 2.11 is independent of the choice of $\delta > 0$; that is what makes this result particularly interesting. It is possible that the conclusion of Theorem 2.11 also holds with $\delta = 0$, but we do not know how to prove (or disprove) that in general; see Problem 8.2 in Section 8. (However, we can make δ be zero under some additional assumptions—see (iv) and Theorem 2.14 below.)

(iv) In contrast to the Abért–Weiss theorem, the conclusion of Theorem 2.11 is nontrivial even if $(X, \mu) = (\Omega, \boldsymbol{\lambda})$ and $\alpha = \sigma$. This case, however, is already covered by the ergodic Theorem 2.6, in fact even with $\delta = 0$.

(v) For actions α that are free and measure-preserving, Theorem 2.11 follows relatively straightforwardly by combining Theorem 2.6 with the usual Abért–Weiss theorem. We sketch the argument here. Let $\alpha\colon \Gamma \curvearrowright (X, \mu)$ be

a free p.m.p. action. Let K be a compact metric space and let $f \colon \Omega \to K$ be a Borel function. Fix an open neighborhood U of the measure $\mathbb{M}_{\boldsymbol{\lambda}} \pi_f$. By Theorem 2.6, for any sufficiently large finite set $D \subset \Gamma$, there is a Borel map $h \colon \Omega \to K$ with

$$(2.13) \qquad\qquad \mathbb{M}_D \pi_h(x) \in U, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega.$$

The equivariance of π_h yields $\mathbb{M}_D \pi_h(x) = (\pi_h)_*(v_{x,D}) = v_{\pi_h(x),D}$, and hence (2.13) is equivalent to

$$v_{\kappa,D} \in U, \quad \text{for } \mathbb{M}_{\boldsymbol{\lambda}} \pi_h\text{-a.e. } \kappa \in K^\Gamma.$$

Now we can use the Abért–Weiss theorem to obtain a Borel map $g \colon X \to K$ for which the pushforward measure $\mathbb{M}_\mu \pi_g$ is so close to $\mathbb{M}_{\boldsymbol{\lambda}} \pi_h$ that

$$\mu(\{x \in X : \mathbb{M}_D \pi_g(x) \in U\}) = \mathbb{M}_\mu \pi_g(\{\kappa \in K^\Gamma : v_{\kappa,D} \in U\}) \geqslant 1 - \delta,$$

for any given $\delta > 0$, as desired. For non-free actions α, a different, more direct proof is necessary.

(vi) The results of §2.A apply to an infinite sequence of averaging sets $(D_n)_{n \in \mathbb{N}}$, while in Theorem 2.11 we only consider a single set D. Our approach can be routinely adapted to extend Theorem 2.11 to the case of finitely many averaging sets; however, when the family of averaging sets is infinite, our methods are not applicable—see Remark 6.11.

Notice that the pointwise operator $\mathbb{M}_D$ is well-defined for an arbitrary Borel action $\alpha \colon \Gamma \curvearrowright X$ and does not require fixing a probability measure μ on X. Therefore, it makes sense to ask for a purely Borel version of the Abért–Weiss theorem, with the last line of Theorem 2.11 replaced by

$$\mathbb{M}_D \pi_g(x) \in U, \quad \text{for all } x \in X.$$

Here we establish such a version for finitely generated groups of subexponential growth and, more generally, for uniformly subexponential Borel actions. Let $\alpha \colon \Gamma \curvearrowright X$ be a Borel action of Γ. We say that α is **uniformly subexponential** if for every finite set $S \subset \Gamma$ and for all $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$ and for all $x \in X$, $|S^n \cdot x| \leqslant (1 + \varepsilon)^n$, where

$$S^n := \{\gamma_1 \cdots \gamma_n : \gamma_i \in S \text{ for all } 1 \leqslant i \leqslant n\}.$$

For example, if Γ is a finitely generated group of subexponential growth, then every action of Γ is uniformly subexponential.

THEOREM 2.14 (Borel Abért–Weiss for uniformly subexponential actions):   *Let K be a compact metric space and let $f\colon \Omega \to K$ be a Borel function. For any open neighborhood U of the measure $\mathbb{M}_{\boldsymbol{\lambda}}\pi_f$, there exist $C > 0$ and a finite set $S \subset \Gamma$ with the following property:*

Let D be a finite subset of Γ with $|D| \geqslant C$ and let $\alpha\colon \Gamma \curvearrowright X$ be a uniformly subexponential (S, D)-free Borel action of Γ. Then there is a Borel map $g\colon X \to K$ such that

$$\mathbb{M}_D \pi_g(x) \in U, \quad \text{for all } x \in X.$$

Note that, even though groups of subexponential growth are amenable, the averaging set D in the statement of Theorem 2.14 is not assumed to be a Følner set.

2.C. OUTLINE OF THE REMAINDER OF THE PAPER. This paper is organized as follows. Section 3 contains a few definitions and some preliminary results concerning the continuity of various basic operations, such as $f \mapsto f_*$. We commence the proofs of Theorems 2.1, 2.6, 2.11, and 2.14 in Section 4, where they are reduced to their special cases with a more "combinatorial" flavor. Then, in Section 5, we state and prove a certain concentration of measure inequality. At this point, we already have all the tools needed to derive Theorem 2.1, which is done in §5.B. In Section 6 we review the LLL and its measurable analogs, and in Section 7 we complete the proofs of Theorems 2.6, 2.11 and 2.14. It turns out that in order to prove Theorem 2.6, it is not enough to simply apply a known measurable version of the LLL—we actually have to go through the *proof* of one of them to obtain some additional information; this is done in §7.B. We conclude the paper with some open problems in Section 8. The appendix contains a proof of Theorem 2.3.

3. Preliminaries

3.A. FURTHER NOTATION.

Integers. We use $\mathbb{N}$ to denote the set of all nonnegative integers and identify each $k \in \mathbb{N}$ with the k-element set $\{i \in \mathbb{N} : i < k\}$. Let $\mathbb{N}^+ := \mathbb{N} \setminus \{0\}$. All finite sets (including each $k \in \mathbb{N}$) are assumed to carry discrete topologies.

Sets and functions. Each function f is identified with its graph, i.e., the set $\{(x,y) : y = f(x)\}$. This enables the use of set-theoretic notation, such as $\subseteq$, $|\cdot|$, etc., for functions. For a function f and a set S of its domain, $f|_S$ denotes the restriction of f to S. For sets A and B,

- $[B]^{<\infty}$ denotes the set of all finite subsets of B;
- $[B \to A]$ denotes the set of all partial functions $\varphi\colon B \rightharpoonup A$;
- $[B \to A]^{<\infty}$ denotes the set of all partial functions $\varphi\colon B \rightharpoonup A$ with $\mathrm{dom}(\varphi) \in [B]^{<\infty}$.

The identity function $X \to X$ on a set X is denoted by id_X.

Symbolic dynamics. Let A be a set and let $\alpha\colon \Gamma \curvearrowright X$ be an action of Γ. We extend the definition of the coding map to partial functions $f\colon X \rightharpoonup A$ by letting $\pi_f(x)\colon \Gamma \rightharpoonup A$ be given by

$$\pi_f(x)(\gamma) := \begin{cases} f(\gamma \cdot x) & \text{if } \gamma \cdot x \in \mathrm{dom}(f); \\ \text{undefined} & \text{otherwise,} \end{cases} \quad \text{for all } x \in X \text{ and } \gamma \in \Gamma.$$

We similarly extend the shift action $\sigma_A\colon \Gamma \curvearrowright A^\Gamma$ to an action $\Gamma \curvearrowright [\Gamma \to A]$ in the obvious way.

The free part of an action. For an action $\alpha\colon\Gamma\curvearrowright X$ of Γ, let $\mathrm{Free}(X)\subseteq X$ denote the set of all $x \in X$ whose α-stabilizer is trivial and let $\mathrm{Free}(\alpha)\colon \Gamma \curvearrowright \mathrm{Free}(X)$ denote the induced action of Γ on $\mathrm{Free}(X)$; we call $\mathrm{Free}(\alpha)$ the **free part** of α.

Miscellaneous. For a metric space $(K,\mathfrak{d})$, $a, b \in K$, and $\varepsilon > 0$, we write $a \approx_\varepsilon b$ to mean $\mathfrak{d}(a,b) < \varepsilon$.

3.B. Topological preliminaries.

Continuity of the coding map. Fix an arbitrary enumeration $\{\gamma_n\}_{n\in\mathbb{N}}$ of the elements of Γ. If $(K,\mathfrak{d})$ is a compact metric space, then the product topology on K^Γ is induced by the metric $\hat{\mathfrak{d}}$:

$$\hat{\mathfrak{d}}(\kappa,\eta) := \sum_{n=0}^{\infty} \frac{\mathfrak{d}(\kappa(\gamma_n),\eta(\gamma_n))}{2^{n+1}}.$$

Recall that if X is a standard Borel space and $\mu \in \mathrm{Prob}(X)$, then the space $\mathfrak{B}(X,K)$ is endowed with the pseudometric $\mathfrak{d}_\mu$. Additionally, we shall consider

the uniform metric $\mathfrak{d}_{\mathsf{uni}}$ given by

$$\mathfrak{d}_{\mathsf{uni}}(f, g) := \sup_{x \in X} \mathfrak{d}(f(x), g(x)).$$

LEMMA 3.1: *Let* $(K, \mathfrak{d})$ *be a compact metric space and let* $\alpha \colon \Gamma \curvearrowright X$ *be a Borel action of* Γ.

   (a) *If* $\mu \in \mathrm{Prob}(X)$ *is* α-*invariant, then the function*

$$(\mathfrak{B}(X, K), \mathfrak{d}_\mu) \to (\mathfrak{B}(X, K^\Gamma), \hat{\mathfrak{d}}_\mu) \colon f \mapsto \pi_f$$

   is distance-preserving, hence continuous.
   (b) *The function*

$$(\mathfrak{B}(X, K), \mathfrak{d}_{\mathsf{uni}}) \to (\mathfrak{B}(X, K^\Gamma), \hat{\mathfrak{d}}_{\mathsf{uni}}) \colon f \mapsto \pi_f$$

   is 1-Lipschitz, hence continuous.

Proof. (a) For all $f, g \in \mathfrak{B}(X, K)$, we have

$$\hat{\mathfrak{d}}_\mu(\pi_f, \pi_g) = \int_X \hat{\mathfrak{d}}(\pi_f(x), \pi_g(x)) \, \mathrm{d}\mu(x) = \int_X \sum_{n=0}^\infty \frac{\mathfrak{d}(f(\gamma_n \cdot x), g(\gamma_n \cdot x))}{2^{n+1}} \, \mathrm{d}\mu(x).$$

Switching the order of integration and summation, we rewrite the last expression as

$$\sum_{n=0}^\infty \frac{1}{2^{n+1}} \int_X \mathfrak{d}(f(\gamma_n \cdot x), g(\gamma_n \cdot x)) \, \mathrm{d}\mu(x).$$

Since μ is α-invariant, this is equal to

$$\sum_{n=0}^\infty \frac{1}{2^{n+1}} \int_X \mathfrak{d}(f(x), g(x)) \, \mathrm{d}\mu(x) = \sum_{n=0}^\infty \frac{\mathfrak{d}_\mu(f, g)}{2^{n+1}} = \mathfrak{d}_\mu(f, g).$$

   (b) For all $f, g \in \mathfrak{B}(X, K)$ and $x \in X$, we have

$$\hat{\mathfrak{d}}(\pi_f(x), \pi_g(x)) = \sum_{n=0}^\infty \frac{\mathfrak{d}(f(\gamma_n \cdot x), g(\gamma_n \cdot x))}{2^{n+1}} \leqslant \sum_{n=0}^\infty \frac{\mathfrak{d}_{\mathsf{uni}}(f, g)}{2^{n+1}} = \mathfrak{d}_{\mathsf{uni}}(f, g),$$

and the desired conclusion follows. ∎

Continuity of the pushforward operator. For a Polish space X, let $C_{\mathsf{b}}(X)$ denote the set of all bounded continuous real-valued functions on X. By definition, the weak-$*$ topology on $\mathrm{Prob}(X)$ is generated by the maps

$$\mathrm{Prob}(X) \to \mathbb{R} \colon \mu \mapsto \int \xi \, \mathrm{d}\mu,$$

where $\xi \in C_{\mathsf{b}}(X)$.

LEMMA 3.2: *Let X and K be Polish spaces and let $f\colon X \to K$ be continuous. Then $f_*\colon \operatorname{Prob}(X) \to \operatorname{Prob}(K)$ is also continuous.* ∎

Now we turn to the continuity properties of the mapping $f \mapsto f_*$.

LEMMA 3.3: *Let $(K, \mathfrak{d})$ be a compact metric space and let (X, μ) be a standard probability space. Then the function $(\mathfrak{B}(X, K), \mathfrak{d}_\mu) \to \operatorname{Prob}(K)\colon f \mapsto f_*(\mu)$ is continuous.*

Proof. Let $f, f_0, f_1, \ldots \in \mathfrak{B}(X, K)$ be such that $f_n \to f$ in $(\mathfrak{B}(X, K), \mathfrak{d}_\mu)$. To demonstrate that $(f_n)_*(\mu) \to f_*(\mu)$, let $\xi \in C_{\mathsf{b}}(X)$; we have to show that

$$(3.4) \qquad \int_K \xi \, \mathrm{d}(f_n)_*(\mu) \to \int_K \xi \, \mathrm{d}f_*(\mu).$$

We may scale ξ if necessary to make it bounded in absolute value by 1. Take any $\varepsilon > 0$. Since K is compact, ξ is uniformly continuous, so we can let $\delta > 0$ be such that $\xi(a) \approx_\varepsilon \xi(b)$ whenever $a \approx_\delta b$. For $n \in \mathbb{N}$, let X_n denote the set of all $x \in X$ with $\mathfrak{d}(f_n(x), f(x)) < \delta$. Since $\mathfrak{d}_\mu(f_n, f) \to 0$, we have $\mu(X_n) \to 1$, and hence, for all large enough $n \in \mathbb{N}$,

$$\left| \int_K \xi \, \mathrm{d}(f_n)_*(\mu) - \int_K \xi \, \mathrm{d}f_*(\mu) \right| \leqslant \int_X |\xi \circ f_n - \xi \circ f| \, \mathrm{d}\mu$$
$$\leqslant \varepsilon \mu(X_n) + 2(1 - \mu(X_n)) \leqslant 2\varepsilon.$$

Since ε was chosen arbitrarily, (3.4) follows. ∎

If K is a compact metric space, then the space $C_{\mathsf{b}}(K)$, equipped with the uniform norm, is separable. Therefore, there exists a countable set $\{\xi_n\}_{n \in \mathbb{N}}$ of continuous real-valued functions on K bounded in absolute value by 1 such that $\{a\xi_n : a \in \mathbb{R}, \ n \in \mathbb{N}\}$ is a dense subset of $C_{\mathsf{b}}(K)$. With this choice of $\{\xi_n\}_{n \in \mathbb{N}}$, the topology on $\operatorname{Prob}(K)$ is induced by the metric Δ^K:

$$\Delta^K(\mu, \nu) := \sum_{n=0}^{\infty} \frac{\left| \int_K \xi_n \, \mathrm{d}\mu - \int_K \xi_n \, \mathrm{d}\nu \right|}{2^{n+1}}.$$

LEMMA 3.5: *Let X be a Polish space and let $(K, \mathfrak{d})$ be a compact metric space. Then the map $(\mathfrak{B}(X, K), \mathfrak{d}_{\mathsf{uni}}) \to (\mathfrak{B}(\operatorname{Prob}(X), \operatorname{Prob}(K)), \Delta^K_{\mathsf{uni}})\colon f \mapsto f_*$ is continuous.*

Proof. Let $\{\xi_n\}_{n \in \mathbb{N}}$ be the set of functions used to define Δ^K. Take any $N \in \mathbb{N}^+$ and $\varepsilon > 0$. Since K is compact, each ξ_n is uniformly continuous, hence we can choose $\delta > 0$ so that $\xi_n(a) \approx_\varepsilon \xi_n(b)$ for all $n \leqslant N$, whenever $a \approx_\delta b$.

Let $f, g \in \mathfrak{B}(X, K)$ and suppose that $\mathfrak{d}_{\mathsf{uni}}(f, g) < \delta$. Then, for any $\mu \in \mathrm{Prob}(K)$, we have

$$
\Delta^K(f_*(\mu), g_*(\mu)) = \sum_{n=0}^{\infty} \frac{|\int_K \xi_n \, \mathrm{d}f_*(\mu) - \int_K \xi_n \, \mathrm{d}g_*(\mu)|}{2^{n+1}}
$$

$$
\leqslant \sum_{n=0}^{N} \frac{1}{2^{n+1}} \int_X |\xi_n \circ f - \xi_n \circ g| \, \mathrm{d}\mu + \frac{1}{2^{N-1}} < \varepsilon + \frac{1}{2^{N-1}}.
$$

Hence, $\Delta_{\mathsf{uni}}^K(f_*, g_*) < \varepsilon + 2^{-N+1}$. Since ε and N are arbitrary, this completes the proof. ∎

Density of continuous functions. Recall that a topological space X is **zero-dimensional** if it has a basis consisting of clopen sets.

LEMMA 3.6: *Let X be a zero-dimensional Polish space and let $(K, \mathfrak{d})$ be a compact metric space. If $\mu \in \mathrm{Prob}(X)$, then the set of all continuous maps $f \colon X \to K$ is dense in $(\mathfrak{B}(X, K), \mathfrak{d}_\mu)$.*

Proof. Without loss of generality, assume that the metric $\mathfrak{d}$ is bounded by 1. Let $f \in \mathfrak{B}(X, K)$ and $\varepsilon > 0$. Since K is compact, it contains a finite ε-net $Z = \{z_0, \ldots, z_{n-1}\} \subseteq K$. Let $g \colon X \to Z$ be the map that sends each $x \in X$ to the point $z \in Z$ that is closest to $f(x)$ (ties may be broken arbitrarily). Then $\mathfrak{d}_{\mathsf{uni}}(f, g) < \varepsilon$ by construction. Since the measure μ is regular [19, Theorem 17.10] and the space X is zero-dimensional, for each $0 \leqslant i < n$, there is a clopen set $U_i \subseteq X$ such that $\mu(U_i \triangle g^{-1}(z_i)) < \varepsilon/n$. For every $x \in X$, set $h(x) \coloneqq z_i$ if $x \in V_i \coloneqq U_i \setminus (U_0 \cup \cdots \cup U_{i-1})$ for some $0 \leqslant i < n$, and $h(x) \coloneqq z_0$ if $x \in V \coloneqq X \setminus (U_0 \cup \cdots \cup U_{n-1})$. Since the sets $V_0, \ldots, V_{n-1}, V$ are clopen, the map h is continuous. If $h(x) \neq g(x)$, then either $x \in V_i \setminus g^{-1}(z_i)$ for some $0 \leqslant i < n$, in which case $x \in U_i \setminus g^{-1}(z_i)$; or else, $x \in V \setminus g^{-1}(z_0)$, in which case $x \in g^{-1}(z_i) \setminus U_i$ for some $1 \leqslant i < n$. Since the metric $\mathfrak{d}$ is bounded by 1, we conclude that

$$
\mathfrak{d}_\mu(g, h) \leqslant \sum_{i=0}^{n-1} \mu(U_i \setminus g^{-1}(z_i)) + \sum_{i=1}^{n-1} \mu(g^{-1}(z_i) \setminus U_i) \leqslant \sum_{i=0}^{n-1} \mu(U_i \triangle g^{-1}(z_i)) < \varepsilon.
$$

Therefore, we have found a continuous function $h \colon X \to K$ with $\mathfrak{d}_\mu(f, h) < 2\varepsilon$. As ε is arbitrary, the proof is complete. ∎

4. Combinatorial reductions

For $k \in \mathbb{N}^+$, let u_k denote the uniform probability measure on k, i.e., let $u_k(\{i\}) \coloneqq 1/k$ for all $i < k$. Set $\Omega_k \coloneqq k^\Gamma$ and $\boldsymbol{u}_k \coloneqq u_k^\Gamma$. Recall that the space $(\Omega_k, \boldsymbol{u}_k)$ is equipped with the shift action σ_k.

Given $\varphi \in [\Gamma \to k]^{<\infty}$ and a partial map $c \colon \Gamma \rightharpoonup k$, we say that $\gamma \in \Gamma$ is an **occurrence** of φ in c if $\gamma \cdot c \supseteq \varphi$. The set of all occurrences of φ in c is denoted by $\mathcal{O}_\varphi(c)$. By definition, if $\gamma \in \mathcal{O}_\varphi(c)$, then, in particular, $\mathrm{dom}(\varphi)\gamma \subseteq \mathrm{dom}(c)$. Define

$$\Omega_k(\varphi) \coloneqq \{c \in \Omega_k : \mathbf{1} \in \mathcal{O}_\varphi(c)\} = \{c \in \Omega_k : \varphi \subset c\}.$$

Note that $\boldsymbol{u}_k(\Omega_k(\varphi)) = k^{-|\varphi|}$ (where $|\varphi|$ is the cardinality of the domain of φ). The family of sets $\{\Omega_k(\varphi) : \varphi \in [\Gamma \to k]^{<\infty}\}$ forms a basis for the topology on Ω_k consisting of clopen sets. From this fact and [19, Theorem 17.20], we obtain the following:

LEMMA 4.1: *Let $k \in \mathbb{N}^+$ and μ, μ_0, $\mu_1, \ldots \in \mathrm{Prob}(\Omega_k)$. Then $\lim_{n \to \infty} \mu_n = \mu$ if and only if, for all $\varphi \in [\Gamma \to k]^{<\infty}$, we have $\lim_{n \to \infty} \mu_n(\Omega_k(\varphi)) = \mu(\Omega_k(\varphi))$.* ∎

We also consider the space $\tilde{\Omega}_k \coloneqq (k^\mathbb{N})^\Gamma$, equipped with the product measure $\tilde{\boldsymbol{u}}_k \coloneqq (u_k^\mathbb{N})^\Gamma$ and the shift action $\sigma_{k^\mathbb{N}}$ of Γ. To simplify the notation, given $x \in \tilde{\Omega}_k$, $\gamma \in \Gamma$, and $n \in \mathbb{N}$, we write $x(\gamma, n)$ to mean $x(\gamma)(n)$ (however, $x(\gamma)$ still denotes the corresponding element of $k^\mathbb{N}$). If $k \geqslant 2$, then, by the measure isomorphism theorem [19, Theorem 17.41], the standard probability spaces $([0; 1], \lambda)$ and $(k^\mathbb{N}, u_k^\mathbb{N})$ are Borel isomorphic, which allows us to replace $\sigma \colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$ by $\sigma_{k^\mathbb{N}} \colon \Gamma \curvearrowright (\tilde{\Omega}_k, \tilde{\boldsymbol{u}}_k)$ in the statements of Theorems 2.6, 2.11, and 2.14. This gives us two main advantages. First, the space $\tilde{\Omega}_k$ is zero-dimensional; in particular, Lemma 3.6 applies to it. Second, the structure of $\tilde{\Omega}_k$ will be explicitly used in the proof of Theorem 2.6 presented in §7.B.

4.A. REDUCTION FOR THEOREM 2.1. In this subsection we reduce Theorem 2.1 to the following statement:

THEOREM 2.1′: *Let $k \in \mathbb{N}^+$ and let $(D_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of Γ with $|D_n|/\log n \to \infty$. Then, for all $S \in [\Gamma]^{<\infty}$ and $\varphi \colon S \to k$, we have*

$$\lim_{n \to \infty} \frac{|D_n \cap \mathcal{O}_\varphi(c)|}{|D_n|} = \frac{1}{k^{|S|}}, \quad \text{for } \boldsymbol{u}_k\text{-a.e. } c \in \Omega_k.$$

LEMMA 4.2: *Theorem 2.1′ implies Theorem 2.1.*

Proof. Assume Theorem 2.1′. Fix a sequence $(D_n)_{n\in\mathbb{N}}$ of nonempty finite subsets of Γ such that $|D_n|/\log n \to \infty$. Notice that Theorem 2.1 is equivalent to the following assertion:

$$(4.3) \qquad \lim_{n\to\infty} v_{x,D_n} = \boldsymbol{\lambda}, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega.$$

On the other hand, by Lemma 4.1, the conclusion of Theorem 2.1′ is equivalent to

$$(4.4) \qquad \lim_{n\to\infty} v_{c,D_n} = \boldsymbol{u}_k, \quad \text{for } \boldsymbol{u}_k\text{-a.e. } c \in \Omega_k.$$

CLAIM 4.2.1: *If* $\pi\colon (\Omega, \boldsymbol{\lambda}) \to (\Omega_k, \boldsymbol{u}_k)$ *is a factor map, then*

$$\lim_{n\to\infty} \mathbb{M}_{D_n}\pi(x) = \mathbb{M}_{\boldsymbol{\lambda}}\pi = \boldsymbol{u}_k, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega.$$

Proof. From the equivariance of π, it follows that for all $x \in \Omega$ and $D \in [\Gamma]^{<\infty} \setminus \{\varnothing\}$,

$$\mathbb{M}_D\pi(x) = \pi_*(v_{x,D}) = v_{\pi(x),D}.$$

Using (4.4) and the fact that, since π is a factor map, $\mathbb{M}_{\boldsymbol{\lambda}}(\pi) = \pi_*(\boldsymbol{\lambda}) = \boldsymbol{u}_k$, we conclude that

$$\mathbb{M}_{D_n}\pi(x) = v_{\pi(x),D_n} \xrightarrow[n\to\infty]{} \boldsymbol{u}_k, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega. \qquad \dashv$$

Define a function $\mathsf{p}\colon \Omega \to [0;1]$ by

$$\mathsf{p}(x) \coloneqq x(\mathbf{1}),$$

Notice that $\pi_{\mathsf{p}} = \mathrm{id}_\Omega$. For each $k \in \mathbb{N}^+$, let $Q_k \subset [0;1]$ be the set of all fractions of the form i/k, $0 \leqslant i < k$, and define $f_k\colon \Omega \to Q_k$ by

$$f_k(x) \coloneqq \max\{q \in Q_k : q \leqslant \mathsf{p}(x)\}.$$

Let $\pi_k \coloneqq \pi_{f_k}$. By definition, $f_k(x) \approx_{1/k} \mathsf{p}(x)$ for all $x \in \Omega$; in other words, the sequence $(f_k)_{k\in\mathbb{N}}$ converges to p uniformly. By Lemmas 3.1(b) and 3.5, this implies that

$$\pi_k \to \mathrm{id}_\Omega \quad \text{and} \quad (\pi_k)_* \to \mathrm{id}_{\mathrm{Prob}(\Omega)} \quad \text{uniformly.}$$

By construction, $(f_k)_*(\boldsymbol{\lambda})$ is the uniform probability measure on Q_k, and $(\pi_k)_*(\boldsymbol{\lambda})$ is the corresponding product measure on Q_k^Γ. Thus, we may apply Claim 4.2.1 to π_k and conclude that

$$\lim_{n\to\infty} \mathbb{M}_{D_n}\pi_k(x) = \mathbb{M}_{\boldsymbol{\lambda}}\pi_k, \quad \text{for } \boldsymbol{\lambda}\text{-a.e. } x \in \Omega.$$

We can put all of these facts together as follows:

$$
\begin{array}{ccc}
\mathbb{M}_{D_n}\pi_k(x) & \xrightarrow[\text{uniformly in } n]{k \to \infty} & \upsilon_{x,D_n} \\[2ex]
\Big\downarrow{\scriptstyle\approx}{\scriptstyle\infty} & & \\[2ex]
\mathbb{M}_{\boldsymbol{\lambda}}\pi_k & \xrightarrow{\hspace{2cm} k \to \infty \hspace{2cm}} & \boldsymbol{\lambda}
\end{array}
$$

It is clear from the above diagram that υ_{x,D_n} converges to $\boldsymbol{\lambda}$ as $n \to \infty$, proving (4.3). ∎

4.B. REDUCTIONS FOR THEOREMS 2.6, 2.11, AND 2.14. Theorem 2.6 reduces to the following statement:

THEOREM 2.6′: *For all $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$, there is $C > 0$ with the following property:*

Let $(D_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of Γ with $|D_n| \geqslant C \log(n+2)$ for all $n \in \mathbb{N}$. Then there exists a Borel map $g \colon \tilde{\Omega}_k \to k$ such that

$$
\tilde{\boldsymbol{u}}_k(\{x \in \tilde{\Omega}_k : g(x) \neq x(\boldsymbol{1}, 0)\}) \leqslant \varepsilon,
$$

and, for all $\varphi \colon S \to k$, we have

$$
\frac{|D_n \cap \mathcal{O}_\varphi(\pi_g(x))|}{|D_n|} \approx_\varepsilon \frac{1}{k^{|S|}}, \quad \text{for all } n \in \mathbb{N} \text{ and for } \tilde{\boldsymbol{u}}_k\text{-a.e. } x \in \tilde{\Omega}_k.
$$

LEMMA 4.5: *Theorem 2.6′ implies Theorem 2.6.*

Proof. Assume Theorem 2.6′. Taking advantage of the measure isomorphism theorem, we will prove the statement of Theorem 2.6 with $(\tilde{\Omega}_2, \tilde{\boldsymbol{u}}_2)$ in place of $(\Omega, \boldsymbol{\lambda})$. We equip the Cantor space $2^{\mathbb{N}}$ with the metric $\mathfrak{m}$ given by

$$
\mathfrak{m}(a, b) := \sum_{n=0}^{\infty} \frac{\mathbb{1}_{a(n) \neq b(n)}}{2^{n+1}}.
$$

Define $\mathsf{p} \colon \tilde{\Omega}_2 \to 2^{\mathbb{N}}$ by $\mathsf{p}(x) := x(\boldsymbol{1})$. Note that $\pi_\mathsf{p} = \mathrm{id}_{\tilde{\Omega}_2}$ and $\mathbb{M}_{\tilde{\boldsymbol{u}}_2}\pi_\mathsf{p} = \tilde{\boldsymbol{u}}_2$.

CLAIM 4.5.1: *It suffices to prove Theorem 2.6 with $K = 2^{\mathbb{N}}$ and $f = \mathsf{p}$.*

Proof. Let $(K, \mathfrak{d})$ be a compact metric space. Without loss of generality, assume that the metric $\mathfrak{d}$ is bounded by 1. Fix a Borel function $f \colon \tilde{\Omega}_2 \to K$, $\varepsilon > 0$, and an open neighborhood U of $\mathbb{M}_{\tilde{\boldsymbol{u}}_2} \pi_f$. The space $\tilde{\Omega}_2$ is zero-dimensional, so Lemmas 3.6, 3.1(a), and 3.3 allow us to assume that f is continuous (after replacing ε by, say, $\varepsilon/2$). Since $\tilde{\Omega}_2$ is compact, f is uniformly continuous, so we can pick $\delta \in (0; \varepsilon/2)$ such that $f(x) \approx_{\varepsilon/2} f(y)$ whenever $x \approx_\delta y$. By Lemma 3.2, the set $U' \coloneqq (\pi_f)_*^{-1}(U)$ is an open neighborhood of $\tilde{\boldsymbol{u}}_2$.

Let $\mathsf{q} \colon \tilde{\Omega}_2 \to 2^{\mathbb{N}}$ be a Borel map and consider the function $g \coloneqq f \circ \mathsf{q}$. Note that if $\mathfrak{m}_{\tilde{\boldsymbol{u}}_2}(\mathsf{p}, \mathsf{q}) \leqslant \delta^2$, then $\mathfrak{d}_{\tilde{\boldsymbol{u}}_2}(f, g) \leqslant \varepsilon$. Indeed, if $\mathfrak{m}_{\tilde{\boldsymbol{u}}_2}(\mathsf{p}, \mathsf{q}) \leqslant \delta^2$, then, by Markov's inequality,

$$\tilde{\boldsymbol{u}}_2(\{x \in \tilde{\Omega}_2 : \mathsf{p}(x) \napprox_\delta \mathsf{q}(x)\}) \leqslant \delta,$$

and, by the choice of δ and since $\mathfrak{d}$ is bounded by 1, we have

$$\mathfrak{d}_{\tilde{\boldsymbol{u}}_2}(f, g) \leqslant \varepsilon/2 + \delta < \varepsilon.$$

Additionally, if $D \in [\Gamma]^{<\infty} \setminus \{\varnothing\}$ and $x \in \tilde{\Omega}_2$ satisfy $\mathbb{M}_D \pi_{\mathsf{q}}(x) \in U'$, then

$$\mathbb{M}_D \pi_g(x) = (\pi_f)_*(\mathbb{M}_D \pi_{\mathsf{q}}(x)) \in U.$$

Therefore, if Theorem 2.6 holds for p, δ^2, and U', then it also holds for f, ε, and U, as desired. ⊣

The remainder of the argument is similar to the last part of the proof of Lemma 4.2. For each $n \in \mathbb{N}^+$, let Q_n be the set of all $a \in 2^{\mathbb{N}}$ such that $a(i) = 0$ for all $i \geqslant n$, and define $\mathsf{p}_n \colon \tilde{\Omega}_2 \to Q_n$ by

$$\mathsf{p}_n(x)(i) \coloneqq \begin{cases} x(\boldsymbol{1}, i) & \text{if } i < n; \\ 0 & \text{if } i \geqslant n. \end{cases}$$

Then $\mathsf{p}_n \to \mathsf{p}$ uniformly, so to prove Theorem 2.6 for p, it is enough to prove it for each p_n. Due to Lemma 4.1, Theorem 2.6 for p_1 is equivalent to Theorem 2.6′ applied with $k = 2$. For larger n, consider the mapping $\vartheta_n \colon 2^{\mathbb{N}} \to (2^n)^{\mathbb{N}}$ given by

$$\vartheta(a)(i) \coloneqq (a(in), \ a(in+1), \dots, a(in+n-1)) \quad \text{for all } a \in 2^{\mathbb{N}} \text{ and } i \in \mathbb{N},$$

where we identify the natural numbers less than 2^n with the n-tuples of zeros and ones. This mapping induces an equivariant isomorphism between $(\tilde{\Omega}_2, \tilde{\boldsymbol{u}}_2)$ and $(\tilde{\Omega}_{2^n}, \tilde{\boldsymbol{u}}_{2^n})$ and shows that Theorem 2.6 for p_n is equivalent to Theorem 2.6′ applied with $k = 2^n$. ∎

Similarly, Theorems 2.11 and 2.14 reduce to the following statements:

THEOREM 2.11′: *For all $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$, there is $C > 0$ with the following property:*

Let D be a finite subset of Γ with $|D| \geqslant C$ and let $\alpha \colon \Gamma \curvearrowright X$ be an (S, D)-free Borel action of Γ. Then, for any $\mu \in \operatorname{Prob}(X)$ and $\delta > 0$, there is a Borel map $g \colon X \to k$ such that, for all $\varphi \colon S \to k$,

$$\mu\left(\left\{x \in X : \frac{|D \cap \mathcal{O}_\varphi(\pi_g(x))|}{|D|} \approx_\varepsilon \frac{1}{k^{|S|}}\right\}\right) \geqslant 1 - \delta.$$

THEOREM 2.14′: *For all $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$, there is $C > 0$ with the following property:*

Let D be a finite subset of Γ with $|D| \geqslant C$ and let $\alpha \colon \Gamma \curvearrowright X$ be a uniformly subexponential (S, D)-free Borel action of Γ. Then there is a Borel map $g \colon X \to k$ such that, for all $\varphi \colon S \to k$,

$$\frac{|D \cap \mathcal{O}_\varphi(\pi_g(x))|}{|D|} \approx_\varepsilon \frac{1}{k^{|S|}} \quad \text{for all } x \in X.$$

The proof of the following lemma is essentially the same as of Lemma 4.5, and we omit it.

LEMMA 4.6: *Theorem 2.11′ implies Theorem 2.11, while Theorem 2.14′ implies Theorem 2.14.*   ∎

5. Using concentration of measure

5.A. THE MAIN PROBABILISTIC BOUND. The following inequality is the main probabilistic input for our arguments:

LEMMA 5.1: *Let $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$. Let D be a nonempty finite subset of Γ and let $\alpha \colon \Gamma \curvearrowright X$ be an (S, D)-free action of Γ. Take any $x \in X$ and pick a function $c \colon (SD \cdot x) \to k$ uniformly at random. Then, for all $\varphi \colon S \to k$,*

$$\mathbb{P}\left[\frac{|D \cap \mathcal{O}_\varphi(\pi_c(x))|}{|D|} \not\approx_\varepsilon \frac{1}{k^{|S|}}\right] \leqslant 2 \exp\left(-\varepsilon^2 \frac{|D|}{2|S|^3}\right).$$

Proof. We shall apply the following concentration of measure result, which is a consequence of Azuma's inequality for Doob martingales [3, §7.4]:

THEOREM 5.2 (Simple Concentration Bound; see [25, 79]): *Let ξ be a random variable determined by s independent trials such that changing the outcome of any one trial can affect ξ at most by b. Then*

$$\mathbb{P}[\xi \not\approx_t \mathbb{E}\xi] \leqslant 2\exp\left(-\frac{t^2}{2b^2 s}\right).$$

Let k, S, ε, D, α, and c be as in the statement of Lemma 5.1. Since the action α is S-free, for all $\varphi\colon S \to k$, we have

$$\mathbb{E}[|D \cap \mathcal{O}_\varphi(\pi_c(x))|] = \sum_{\delta \in D} \mathbb{P}[\delta \in \mathcal{O}_\varphi(\pi_c(x))] = \frac{|D|}{k^{|S|}}.$$

Consider any $y \in SD \cdot x$ and let c_1, $c_2\colon (SD \cdot x) \to k$ be two maps that agree on $(SD \cdot x) \setminus \{y\}$. Let $\varphi\colon S \to k$ and suppose that some $\delta \in D$ belongs to $\mathcal{O}_\varphi(\pi_{c_1}(x)) \bigtriangleup \mathcal{O}_\varphi(\pi_{c_2}(x))$. Then $y \in S \cdot (\delta \cdot x)$, i.e., $\delta \cdot x \in S^{-1} \cdot y$. Since α is D-free, there are at most $|S^{-1} \cdot y| = |S|$ possible values for δ. Thus, we may apply the Simple Concentration Bound with parameters

$$s := |SD \cdot x| \leqslant |S||D|, \quad b := |S|, \quad \text{and} \quad t := \varepsilon|D|,$$

to obtain

$$\mathbb{P}\left[|D \cap \mathcal{O}_\varphi(\pi_c(x))| \not\approx_{\varepsilon|D|} \frac{|D|}{k^{|S|}}\right] \leqslant 2\exp(-\varepsilon^2 \frac{|D|}{2|S|^3}),$$

as desired. ∎

5.B. PROOF OF THEOREM 2.1. We are now ready to prove Theorem 2.1 (or rather Theorem 2.1'). Let $k \in \mathbb{N}^+$ and let $(D_n)_{n\in\mathbb{N}}$ be a sequence of finite subsets of Γ such that $|D_n|/\log n \to \infty$. Take any $S \in [\Gamma]^{<\infty}$, $\varphi\colon S \to k$, and $\varepsilon > 0$. We will show that for $\boldsymbol{u}_k$-a.e. $c \in \Omega_k$ and for all sufficiently large $n \in \mathbb{N}$,

$$(5.3) \qquad \frac{|D_n \cap \mathcal{O}_\varphi(c)|}{|D_n|} \approx_\varepsilon \frac{1}{k^{|S|}},$$

which will imply the conclusion of Theorem 2.1'. For each $n \in \mathbb{N}$, let X_n denote the set of all $c \in \Omega_k$ for which (5.3) fails. By Lemma 5.1, we have

$$\sum_{n\in\mathbb{N}} \boldsymbol{u}_k(X_n) \leqslant \sum_{n\in\mathbb{N}} 2\exp\left(-\varepsilon^2 \frac{|D_n|}{2|S|^3}\right) < \infty,$$

since $\varepsilon^2|D_n|/(2|S|^3) > 2\log n$ for all sufficiently large n. An application of the Borel–Cantelli lemma completes the proof.

6. The Lovász Local Lemma and its measurable versions

6.A. THE CLASSICAL LLL. The reader is referred to [3, Chapter 5] and [25] for background on the LLL and its applications in combinatorics. The presentation below follows, with slight modifications, [6, Section 1.2].

Let X be a set and let $k \in \mathbb{N}^+$. A **bad** (k-)**event** over X is a nonempty subset $B \subseteq [X \to k]^{<\infty}$ such that for all $\varphi, \varphi' \in B$, $\mathrm{dom}(\varphi) = \mathrm{dom}(\varphi')$. If a bad event B is nonempty, then its **domain** is the set $\mathrm{dom}(B) := \mathrm{dom}(\varphi)$ for any (hence all) $\varphi \in B$; the domain of the empty bad event is, by definition, the empty set. The **probability** of a bad k-event B with domain F is defined to be

$$\mathbb{P}[B] := \frac{|B|}{k^{|F|}}.$$

We say that a map $f \colon X \to k$ **avoids** a bad k-event B if there is no $\varphi \in B$ such that $\varphi \subseteq f$. Note that if X is finite and $f \colon X \to k$ is chosen uniformly at random, then $\mathbb{P}[B]$ is the probability that f does not avoid B.

A (k-)**instance** (of the LLL) over a set X is an arbitrary set $\mathscr{B}$ of bad k-events. A **solution** to a k-instance $\mathscr{B}$ is a function $f \colon X \to k$ that avoids all $B \in \mathscr{B}$. For an instance $\mathscr{B}$ and a bad event $B \in \mathscr{B}$, the **neighborhood** of B in $\mathscr{B}$ is the set

$$\mathrm{N}_{\mathscr{B}}(B) := \{B' \in \mathscr{B} \setminus \{B\} : \mathrm{dom}(B') \cap \mathrm{dom}(B) \neq \varnothing\}.$$

The **degree** of B in $\mathscr{B}$ is defined to be

$$\deg_{\mathscr{B}}(B) := |\mathrm{N}_{\mathscr{B}}(B)|.$$

Let

$$p(\mathscr{B}) := \sup_{B \in \mathscr{B}} \mathbb{P}[B] \quad \text{and} \quad d(\mathscr{B}) := \sup_{B \in \mathscr{B}} \deg_{\mathscr{B}}(B).$$

An instance $\mathscr{B}$ is **correct for the Symmetric LLL** (the **SLLL** for short) if

$$e \cdot p(\mathscr{B}) \cdot (d(\mathscr{B}) + 1) < 1,$$

where $e = 2.71\ldots$ denotes the base of the natural logarithm. Note that if $\mathscr{B}$ is correct for the SLLL, then, in particular, $\deg_{\mathscr{B}}(B) < \infty$ for all $B \in \mathscr{B}$ (instances $\mathscr{B}$ with this property are called **locally finite** in [6]).

THEOREM 6.1 (Erdős–Lovász [16]; Symmetric Lovász Local Lemma): *Let* $k \in \mathbb{N}^+$ *and let* $\mathscr{B}$ *be a k-instance of the LLL over a set X. If $\mathscr{B}$ is correct for the SLLL, then $\mathscr{B}$ has a solution.*

The Symmetric LLL was introduced by Erdős and Lovász (with 4 in place of e) in their seminal paper [16]; the constant was subsequently improved by Lovász (the sharpened version first appeared in [28]). Theorem 6.1 is a special case of the SLLL in the so-called variable framework (the name is due to Kolipaka and Szegedy [22]), which encompasses most typical applications. For the full statement of the SLLL, see [3, Corollary 5.1.2]. Deducing Theorem 6.1 from [3, Corollary 5.1.2] is routine when X is finite (see, e.g., [25, 41]); the case of infinite X then follows by compactness. A more general version of Theorem 6.1 for infinite X, with k replaced by an arbitrary standard probability space, was proved by Kun [23, Lemma 13].

Theorem 6.1 can be extended to instances $\mathscr{B}$ with $d(\mathscr{B}) = \infty$, provided that the probability $\mathbb{P}[B]$ of a bad event $B \in \mathscr{B}$ decays sufficiently quickly as $|\mathrm{dom}(B)|$ increases. An instance $\mathscr{B}$ is **correct for the General LLL** (the **GLLL** for short) if $\mathrm{N}_{\mathscr{B}}(B)$ is countable for every $B \in \mathscr{B}$, and there is a function $\omega\colon \mathscr{B} \to [0;1)$, called a **witness** to the correctness of $\mathscr{B}$, such that for all $B \in \mathscr{B}$,

$$\mathbb{P}[B] \leqslant \omega(B) \prod_{B' \in \mathrm{N}_{\mathscr{B}}(B)} (1 - \omega(B')).$$

THEOREM 6.2 (General Lovász Local Lemma; [3, Lemma 5.1.1]): *Let $k \in \mathbb{N}^+$ and let $\mathscr{B}$ be a k-instance of the LLL over a set X. If $\mathscr{B}$ is correct for the GLLL, then $\mathscr{B}$ has a solution.*

A standard calculation (see [3, proof of Corollary 5.1.2]) shows that if an instance $\mathscr{B}$ is correct for the SLLL, then it is also correct for the GLLL (hence the name "General LLL").

6.B. MEASURABLE VERSIONS OF THE LLL. Let X be a standard Borel space and let $k \in \mathbb{N}^+$. Then the set of all bad k-events is also naturally equipped with the structure of a standard Borel space (indeed, each bad event is a finite set, so the set of all bad k-events is a Borel subset of the space $[[X \to k]^{<\infty}]^{<\infty}$). Thus, it makes sense to talk about **Borel** instances of the LLL, i.e., Borel sets of bad events.

Given a Borel k-instance $\mathscr{B}$ over X that is correct for the SLLL, it is natural to wonder if it has a Borel solution $f\colon X \to k$. Although the answer is negative in general (see [13, Theorem 1.6]), Csóka, Grabowski, Máthé, Pikhurko and Tyros [14] answered the question in the affirmative for uniformly subexponential

instances. Given an instance $\mathscr{B}$ over a set X, an element $x \in X$, and an integer $n \in \mathbb{N}$, let $R_{\mathscr{B}}^n(x)$ denote the set of all $y \in X$ such that either $y = x$, or there exists a sequence $B_1, \ldots, B_m \in \mathscr{B}$ with $m \leqslant n$ satisfying

$$x \in \mathrm{dom}(B_1), \quad \mathrm{dom}(B_i) \cap \mathrm{dom}(B_{i+1}) \neq \varnothing \text{ for all } 1 \leqslant i < m,$$
$$\text{and} \quad y \in \mathrm{dom}(B_m).$$

The instance $\mathscr{B}$ is **uniformly subexponential** if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$ and for all $x \in X$,

$$|R_{\mathscr{B}}^n(x)| < (1 + \varepsilon)^n.$$

THEOREM 6.3 (Csóka–Grabowski–Máthé–Pikhurko–Tyros [14, Theorem 1.3], Borel SLLL for uniformly subexponential instances): *Let $k \in \mathbb{N}^+$ and let $\mathscr{B}$ be a Borel k-instance of the LLL over a standard Borel space X. If $\mathscr{B}$ is correct for the SLLL and uniformly subexponential, then $\mathscr{B}$ has a Borel solution $f \colon X \to k$.*

For a k-instance $\mathscr{B}$ over a set X and a map $f \colon X \to k$, we define the **defect** $\mathrm{Def}(f; \mathscr{B})$ of f with respect to $\mathscr{B}$ by

$$(6.4) \quad \mathrm{Def}(f, \mathscr{B}) := \{x \in X : x \in \mathrm{dom}(\varphi) \text{ for some } \varphi \in B \in \mathscr{B} \text{ with } \varphi \subseteq f\}.$$

Evidently, f is a solution to $\mathscr{B}$ if and only if $\mathrm{Def}(f, \mathscr{B}) = \varnothing$. Thus, in the absence of a Borel solution to $\mathscr{B}$, it is natural to seek a Borel map $f \colon X \to k$ whose defect is "small" in some sense. The next result was proved by the current author in [6]:

THEOREM 6.5 ([6, Theorem 5.1], approximate SLLL): *Let $k \in \mathbb{N}^+$ and let $\mathscr{B}$ be a Borel k-instance of the LLL over a standard Borel space X. If $\mathscr{B}$ is correct for the SLLL, then for any $\mu \in \mathrm{Prob}(X)$ and $\delta > 0$, there is a Borel function $f \colon X \to k$ with*

$$\mu(\mathrm{Def}(f, \mathscr{B})) \leqslant \delta.$$

It is an open question whether the conclusion of Theorem 6.5 holds with $\delta = 0$; see Problem 8.1 in Section 8. Also, Theorem 6.5 fails for instances that are correct for the GLLL instead of the SLLL (see [6, Theorem 7.1] and Remark 6.11 below). However, when the underlying structure is in a certain sense induced by the shift action σ, even instances that are only correct for the GLLL can be solved with a null defect—see Theorem 6.10 in the next subsection.

6.C. USING THE LLL OVER GROUP ACTIONS. Now we describe a convenient set-up for applying the LLL to problems in ergodic theory.

Let $\alpha\colon \Gamma \curvearrowright X$ be an action of Γ and let $\Phi \subseteq [\Gamma \to k]^{<\infty}$ be a bad k-event over Γ with domain $F \in [\Gamma]^{<\infty}$. For each $x \in X$, define a bad k-event $B_x(\Phi, \alpha)$ over X via

$$B_x(\Phi, \alpha) \coloneqq \{\varphi\colon (F \cdot x) \to k : \pi_\varphi(x)|_F \in \Phi\}.$$

Note that if $B_x(\Phi, \alpha) \neq \varnothing$, then $\mathrm{dom}(B_x(\Phi, \alpha)) = F \cdot x$. (If α is not F-free, then $B_x(\Phi, \alpha)$ may be empty even if Φ is not.) By construction, a function $f\colon X \to k$ avoids $B_x(\Phi, \alpha)$ precisely when $\pi_f(x)$ avoids Φ. Define an instance $\mathscr{B}(\Phi, \alpha)$ of the LLL over X as follows:

$$\mathscr{B}(\Phi, \alpha) \coloneqq \{B_x(\Phi, \alpha) : x \in X\}.$$

Clearly, if X is a standard Borel space and $\alpha\colon \Gamma \curvearrowright X$ is a Borel action, then the instance $\mathscr{B}(\Phi, \alpha)$ is Borel. A function $f\colon X \to k$ is a solution to $\mathscr{B}(\Phi, \alpha)$ if and only if $\pi_f(x)$ avoids Φ for all $x \in X$. Hence, it is somewhat more convenient to define the **defect** of a map $f\colon X \to k$ as the set of all $x \in X$ such that $\pi_f(x)$ does not avoid Φ:

$$\mathrm{Def}(f, \Phi, \alpha) \coloneqq \{x \in X : \pi_f(x)|_F \in \Phi\}.$$

There is a straightforward relationship between this definition and the one in (6.4), namely,

(6.6) $$\mathrm{Def}(f, \mathscr{B}(\Phi, \alpha)) = F \cdot \mathrm{Def}(f, \Phi, \alpha).$$

Using the above notation, we can formulate the following corollaries of Theorems 6.3 and 6.5:

COROLLARY 6.7 (to Theorem 6.3): *Let* $\alpha\colon \Gamma \curvearrowright X$ *be a uniformly subexponential Borel action of* Γ *and let* $k \in \mathbb{N}^+$. *Let* Φ *be a bad* k-*event over* Γ *and suppose that the instance* $\mathscr{B}(\Phi, \alpha)$ *is correct for the SLLL. Then* $\mathscr{B}(\Phi, \alpha)$ *has a Borel solution* $f\colon X \to k$.

COROLLARY 6.8 (to Theorem 6.5): *Let* $\alpha\colon \Gamma \curvearrowright X$ *be a Borel action of* Γ *and let* $k \in \mathbb{N}^+$. *Let* Φ *be a bad* k-*event over* Γ *and suppose that the instance* $\mathscr{B}(\Phi, \alpha)$ *is correct for the SLLL. Then, for any* $\mu \in \mathrm{Prob}(X)$ *and* $\delta > 0$, *there is a Borel function* $f\colon X \to k$ *with*

$$\mu(\mathrm{Def}(f, \Phi, \alpha)) < \delta.$$

Remark 6.9: In the statement of Corollary 6.8, the measure μ is not assumed to be α-invariant. Because of that, to derive Corollary 6.8, one has to apply Theorem 6.5 not to μ itself, but to the measure obtained by shifting μ by one of the elements of $\mathrm{dom}(\Phi)$, and then use (6.6).

More generally, let $(\Phi_n)_{n \in \mathbb{N}}$ be a sequence of bad k-events over Γ. For an action $\alpha \colon \Gamma \curvearrowright X$ and a map $f \colon X \to k$, define

$$\mathscr{B}((\Phi_n)_{n \in \mathbb{N}}, \alpha) := \bigcup_{n=0}^{\infty} \mathscr{B}(\Phi_n, \alpha)$$

and

$$\mathrm{Def}(f, (\Phi_n)_{n \in \mathbb{N}}, \alpha) := \bigcup_{n=0}^{\infty} \mathrm{Def}(f, \Phi_n, \alpha).$$

When $\alpha = \sigma$, we have the following strengthening of Corollary 6.8:

THEOREM 6.10 ([6, Corollary 6.7], measurable GLLL over the shift): *Let* $k \in \mathbb{N}^+$ *and let* $(\Phi_n)_{n \in \mathbb{N}}$ *be a sequence of bad k-events over* Γ. *Suppose that the instance* $\mathscr{B}((\Phi_n)_{n \in \mathbb{N}}, \mathrm{Free}(\sigma))$ *is correct for the GLLL. Then there is a Borel function* $f \colon \Omega \to k$ *with*

$$\boldsymbol{\lambda}(\mathrm{Def}(f, (\Phi_n)_{n \in \mathbb{N}}, \sigma)) = 0.$$

Remark 6.11: Theorem 6.10 can fail for actions other than σ: According to [6, Theorem 7.1], if Γ is amenable, then the analog of Theorem 6.10 holds for a free ergodic p.m.p. action $\alpha \colon \Gamma \curvearrowright (X, \mu)$ if and only if there is a factor map $\pi \colon (X, \mu) \to (\Omega, \boldsymbol{\lambda})$.

Theorem 6.10 is a special case of [6, Theorem 6.6], whose full statement is rather technical and will not be needed here. Roughly speaking, [6, Theorem 6.6] asserts that any combinatorial argument proceeding via a series of iterative applications of the GLLL can be performed in a measurable fashion over the shift action $\sigma \colon \Gamma \curvearrowright (\Omega, \boldsymbol{\lambda})$.

7. Proofs of Theorems 2.6, 2.11, and 2.14

7.A. PROOFS OF THEOREMS 2.11 AND 2.14. We first establish Theorems 2.11 and 2.14, as their proofs are somewhat more straightforward than that of Theorem 2.6 (for instance, they only use the Symmetric LLL rather than the more technical General LLL).

Let $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$. For a nonempty finite subset $D \subset \Gamma$, let $\Phi(k, S, \varepsilon, D)$ denote the bad k-event over Γ with domain SD consisting of all maps $c \colon SD \to k$ such that

$$\frac{|D \cap \mathcal{O}_\varphi(c)|}{|D|} \not\approx_\varepsilon \frac{1}{k^{|S|}} \quad \text{for some } \varphi \colon S \to k.$$

By definition, if $\alpha \colon \Gamma \curvearrowright X$ is a Borel action of Γ and $g \colon X \to k$ is a Borel map, then we have

$$x \in \mathrm{Def}(g, \Phi(k, S, \varepsilon, D), \alpha)$$

(7.1)
$$\iff \frac{|D \cap \mathcal{O}_\varphi(\pi_g(x))|}{|D|} \not\approx_\varepsilon \frac{1}{k^{|S|}} \text{ for some } \varphi \colon S \to k.$$

LEMMA 7.2: *Let $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$. There exists $C > 0$ such that for all $D \in [\Gamma]^{<\infty}$ with $|D| > C$ and for every (S, D)-free action $\alpha \colon \Gamma \curvearrowright X$, the instance $\mathscr{B}(\Phi(k, S, \varepsilon, D), \alpha)$ is correct for the SLLL.*

Proof. Let $D \in [\Gamma]^{<\infty} \setminus \{\varnothing\}$ and let $\alpha \colon \Gamma \curvearrowright X$ be (S, D)-free. Set

$$\Phi := \Phi(k, S, \varepsilon, D), \quad \mathscr{B} := \mathscr{B}(\Phi, \alpha), \quad \text{and} \quad B_x := B_x(\Phi, \alpha) \text{ for all } x \in X.$$

Due to Lemma 5.1, we have

$$p(\mathscr{B}) \leqslant 2k^{|S|} \exp\left(-\varepsilon^2 \frac{|D|}{2|S|^3} \right).$$

To upper bound $d(\mathscr{B})$, note that for each $x \in X$,

$$\mathrm{N}_{\mathscr{B}}(B_x) = \{ B_y \in \mathscr{B} \setminus \{B_x\} : (SD \cdot y) \cap (SD \cdot x) \neq \varnothing \}.$$

Since $(SD \cdot y) \cap (SD \cdot x) \neq \varnothing$ if and only if $y \in (SD)^{-1} SD \cdot x$, we obtain

$$\deg_{\mathscr{B}}(B_x) \leqslant |(SD)^{-1} SD| - 1 \leqslant |S|^2 |D|^2 - 1.$$

(We subtracted 1 since y cannot be equal to x.) Hence, $d(\mathscr{B}) \leqslant |S|^2 |D|^2 - 1$, and $\mathscr{B}$ is correct for the SLLL as long as

$$e \cdot 2k^{|S|} \exp\left(-\varepsilon^2 \frac{|D|}{2|S|^3} \right) \cdot |S|^2 |D|^2 < 1,$$

which holds whenever $|D|$ is sufficiently large. ∎

Theorems 2.11′ and 2.14′ now follow immediately by combining (7.1) and Lemma 7.2 with Corollaries 6.8 and 6.7 respectively.

7.B. PROOF OF THEOREM 2.6. For the purposes of proving Theorem 2.6, the role of Lemma 7.2 is played by the following fact:

LEMMA 7.3: *Let $k \in \mathbb{N}^+$, $S \in [\Gamma]^{<\infty}$, and $\varepsilon > 0$. There exists $C > 0$ with the following property:*

Let $(D_n)_{n \in \mathbb{N}}$ be a sequence of finite subsets of Γ with $|D_n| \geqslant C \log(n + 2)$ for all $n \in \mathbb{N}$ and let $\alpha \colon \Gamma \curvearrowright X$ be a free action of Γ. Set

$$\Phi_n := \Phi(k, S, \varepsilon, D_n) \text{ for all } n \in \mathbb{N},$$

$$\mathscr{B} := \mathscr{B}((\Phi_n)_{n \in \mathbb{N}}, \alpha), \quad \text{and} \quad B_{n,x} := B_x(\Phi_n, \alpha) \text{ for all } n \in \mathbb{N} \text{ and } x \in X.$$

Then the instance $\mathscr{B}$ is correct for the GLLL. Moreover, there is a function $\omega \colon \mathbb{N} \to [0; 1)$ such that

$$(7.4) \qquad \sum_{n=0}^{\infty} |SD_n| \cdot \frac{\omega(n)}{1 - \omega(n)} < \varepsilon,$$

and the mapping $\tilde{\omega} \colon \mathscr{B} \to [0; 1) \colon B_{n,x} \mapsto \omega(n)$ is a witness to the correctness of $\mathscr{B}$.

Proof. Fix any $0 < a < \varepsilon^2/(2|S|^3)$. We claim that if C is large enough, then the function

$$\omega(n) := \exp(-a|D_n|)$$

has the desired properties. To begin with, we are going to assume that C is so large that

$$\exp(-a \cdot C \log(2)) < 1/2,$$

and that the function $\xi \mapsto \xi \exp(-a\xi)$ is decreasing for all $\xi \geqslant C \log 2$. For any such C, we have

$$\sum_{n=0}^{\infty} |SD_n| \cdot \frac{\omega(n)}{1 - \omega(n)} \leqslant |S| \sum_{n=0}^{\infty} |D_n| \cdot \frac{\exp(-a|D_n|)}{1 - \exp(-a|D_n|)}$$

$$\leqslant 2|S| \sum_{n=0}^{\infty} |D_n| \exp(-a|D_n|)$$

$$\leqslant 2|S|C \sum_{n=0}^{\infty} \frac{\log(n + 2)}{(n + 2)^{Ca}}.$$

The last expression approaches 0 as $C \to \infty$, so we can guarantee (7.4).

Consider any $n \in \mathbb{N}$ and $x \in X$. By Lemma 5.1, we have

$$\mathbb{P}[B_{n,x}] \leqslant 2k^{|S|} \exp\Big(- \varepsilon^2 \frac{|D_n|}{2|S|^3} \Big).$$

If $\mathrm{dom}(B_{n,x}) \cap \mathrm{dom}(B_{m,y}) \neq \varnothing$ for some $m \in \mathbb{N}$ and $y \in X$, then

$$y \in (SD_m)^{-1} SD_n \cdot x,$$

and hence for any particular $m \in \mathbb{N}$, there are at most $|S|^2 |D_m||D_n|$ choices of such y. Therefore, the mapping $\tilde{\omega} \colon \mathscr{B} \to [0;1)$ is a witness to the correctness of $\mathscr{B}$ as long as we have

$$(7.5) \qquad 2k^{|S|} \exp\Big(- \varepsilon^2 \frac{|D_n|}{2|S|^3} \Big) \leqslant \omega(n) \prod_{m=0}^{\infty} (1 - \omega(m))^{|S|^2 |D_m||D_n|},$$

for all $n \in \mathbb{N}$. Using the definition of ω and then taking the logarithm of both sides of (7.5) and dividing them by $(-|D_n|)$, we rewrite (7.5) as

$$(7.6) \qquad -\frac{\log(2k^{|S|})}{|D_n|} + \frac{\varepsilon^2}{2|S|^3} \geqslant a - |S|^2 \sum_{m=0}^{\infty} |D_m| \log(1 - \exp(-a|D_m|)).$$

Let us first look at the left-hand side of (7.6). We have

$$-\frac{\log(2k^{|S|})}{|D_n|} + \frac{\varepsilon^2}{2|S|^3} \geqslant -\frac{\log(2k^{|S|})}{C \log 2} + \frac{\varepsilon^2}{2|S|^3} \xrightarrow[C\to\infty]{} \frac{\varepsilon^2}{2|S|^3}.$$

As for the right-hand side of (7.6), note that $-\log(1-\xi) < 2\xi$ for all $0 < \xi < 1/2$, so

$$a - |S|^2 \sum_{m=0}^{\infty} |D_m| \log(1 - \exp(-a|D_m|))$$

$$< a + 2|S|^2 \sum_{m=0}^{\infty} |D_m| \exp(-a|D_m|)$$

$$\leqslant a + 2|S|^2 C \sum_{m=0}^{\infty} \frac{\log(m+2)}{(m+2)^{Ca}} \xrightarrow[C\to\infty]{} a.$$

Since a was chosen to be less than $\varepsilon^2/(2|S|^3)$, we conclude that (7.6) holds for all large C. ∎

From (7.1), Lemma 7.3, and Theorem 6.10, we can derive most of Theorem 2.6′. The only part that is missing is that the map $g \colon \tilde{\Omega}_k \to k$ can be chosen so that

$$\tilde{\boldsymbol{u}}_k(\{x \in \tilde{\Omega}_k : g(x) \neq x(\mathbf{1}, 0)\}) \leqslant \varepsilon.$$

To argue this, we have to review the proof of Theorem 6.10. As mentioned in the introduction, the tool used to prove Theorem 6.10 is the Moser–Tardos algorithm, developed by Moser and Tardos in [26]. Here we outline only the most relevant elements of the Moser–Tardos theory when applied to our current situation. For further details, see [26] and [6, §3].

For the rest of this subsection, fix $k \in \mathbb{N}^+$ and a sequence $(\Phi_n)_{n \in \mathbb{N}}$ of bad k-events over Γ. For each $n \in \mathbb{N}$, set $F_n := \mathrm{dom}(\Phi_n)$. Define

$$\mathscr{B} := \mathscr{B}((\Phi_n)_{n \in \mathbb{N}}, \mathrm{Free}(\sigma_{k^{\mathbb{N}}})),$$

and

$$B_{n,x} := B_x(\Phi_n, \sigma_{k^{\mathbb{N}}}) \text{ for all } n \in \mathbb{N} \text{ and } x \in \tilde{\Omega}_k.$$

Consider the following inductive construction:

Set $t_0(x) := 0$ for all $x \in \tilde{\Omega}_k$.

STEP $i \in \mathbb{N}$: Define

$$g_i(x) := x(\mathbf{1}, t_i(x)) \quad \text{for all } x \in \tilde{\Omega}_k;$$

$$A_i' := \{(n, x) \in \mathbb{N} \times \tilde{\Omega}_k : g_i \text{ does not avoid } B_{n,x}\}.$$

Choose $A_i \subseteq A_i'$ to be an arbitrary Borel maximal subset of A_i' with the property that

$$(F_n \cdot x) \cap (F_m \cdot y) = \varnothing \quad \text{for all distinct pairs } (n, x), \ (m, y) \in A_i.$$

(Such A_i exists by, e.g., [21, Lemma 7.3].) Let

$$T_i := \bigcup_{(n,x) \in A_i} (F_n \cdot x) \quad \text{and} \quad t_{i+1}(x) := \begin{cases} t_i(x) + 1 & \text{if } x \in T_i; \\ t_i(x) & \text{otherwise.} \end{cases}$$

By definition, $g_0(x) = x(\mathbf{1}, 0)$ for all $x \in \tilde{\Omega}_k$. We call a sequence $\mathcal{A} := (A_i)_{i=0}^{\infty}$ obtained via the above procedure a **Borel Moser–Tardos process**. Note that there is not a unique Borel Moser–Tardos process, as there is some freedom in the choice of the Borel maximal subset $A_i \subseteq A_i'$.

Let $\mathcal{A} = (A_i)_{i=0}^{\infty}$ be a Borel Moser–Tardos process. For $x \in \tilde{\Omega}_k$, define $t(x) \in \mathbb{N} \cup \{\infty\}$ by

$$t(x) := \lim_{i \to \infty} t_i(x).$$

We say that x is $\mathcal{A}$-**stable** if $t(x) < \infty$, i.e., if the corresponding sequence $t_0(x)$, $t_1(x), \ldots$ is eventually constant. Let $\mathrm{St}(\mathcal{A}) \subseteq \tilde{\Omega}_k$ denote the set of all $\mathcal{A}$-stable

elements. For $x \in \mathrm{St}(\mathcal{A})$, we can define

(7.7)
$$g(x) := x(\mathbf{1}, t(x)).$$

It is easy to verify (see [6, Proposition 3.3]) that if $F_n \cdot x \subseteq \mathrm{St}(\mathcal{A})$, then $x \notin \mathrm{Def}(g, \Phi_n, \sigma_{k^{\mathbb{N}}})$. The **index** $\mathrm{Ind}(n, x, \mathcal{A}) \in \mathbb{N} \cup \{\infty\}$ of a pair $(n, x) \in \mathbb{N} \times \tilde{\Omega}_k$ in $\mathcal{A}$ is defined by the formula

$$\mathrm{Ind}(n, x, \mathcal{A}) := |\{i \in \mathbb{N} : (n, x) \in A_i\}|.$$

Note that for all $x \in \mathrm{Free}(\tilde{\Omega}_k)$, we have

(7.8)
$$t(x) = \sum_{n=0}^{\infty} \sum_{\delta \in F_n} \mathrm{Ind}(n, \delta^{-1} \cdot x, \mathcal{A}),$$

and hence such x is $\mathcal{A}$-stable if and only if the expression on the right-hand side of (7.8) is finite. The following theorem is the central result of the Moser–Tardos theory:

THEOREM 7.9 (Moser–Tardos [26]; see also [6, Theorem 3.5]): *Let* $\omega \colon \mathbb{N} \to [0;1)$ *be a function such that the mapping* $\tilde{\omega} \colon \mathscr{B} \to [0;1) \colon B_{n,x} \mapsto \omega(n)$ *is a witness to the correctness of* $\mathscr{B}$*. Then, for any Borel Moser–Tardos process* $\mathcal{A}$ *and for all* $n \in \mathbb{N}$*, we have*

$$\int_{\tilde{\Omega}_k} \mathrm{Ind}(n, x, \mathcal{A}) \, d\tilde{\boldsymbol{u}}_k(x) \leqslant \frac{\omega(n)}{1 - \omega(n)}.$$

COROLLARY 7.10 (to Theorem 7.9): *Let* $\omega \colon \mathbb{N} \to [0;1)$ *be such that* $\tilde{\omega} \colon \mathscr{B} \to [0;1) \colon B_{n,x} \mapsto \omega(n)$ *is a witness to the correctness of* $\mathscr{B}$*. Then there is a Borel function* $g \colon \tilde{\Omega}_k \to k$ *such that*

$$\tilde{\boldsymbol{u}}_k(\mathrm{Def}(g, (\Phi_n)_{n \in \mathbb{N}}, \sigma_{k^{\mathbb{N}}})) = 0$$

and

$$\tilde{\boldsymbol{u}}_k(\{x \in \tilde{\Omega}_k : g(x) \neq x(\mathbf{1}, 0)\}) \leqslant \sum_{n=0}^{\infty} |F_n| \cdot \frac{\omega(n)}{1 - \omega(n)}.$$

Proof. First we show that the sum

$$S := \sum_{n=0}^{\infty} |F_n| \cdot \frac{\omega(n)}{1 - \omega(n)}$$

is finite. Without loss of generality, assume that $\Phi_0 \neq \varnothing$. Consider any $x \in \mathrm{Free}(\tilde{\Omega}_k)$. Since $\tilde{\omega}$ is a witness to the correctness of $\mathscr{B}$, we have

$$\mathbb{P}[B_{0,x}] \leqslant \omega(0) < 1,$$

so $F_0 \neq \varnothing$. Hence, for every $n \in \mathbb{N}^+$, there exist at least $|F_n|$ distinct y with $B_{n,y} \in \mathrm{N}_{\mathscr{B}}(B_{0,x})$. Therefore,

$$\prod_{n=1}^{\infty} (1 - \omega(n))^{|F_n|} \geqslant \mathbb{P}[B_{0,x}] > 0,$$

which implies that $\sum_{n=0}^{\infty} |F_n|\omega(n)$ is finite. In particular, for all sufficiently large n we have $\omega(n) \leqslant 1/2$ and $\omega(n)/(1 - \omega(n)) \leqslant 2\omega(n)$, and hence S is also finite.

Let $\mathcal{A} = (A_i)_{i=0}^{\infty}$ be an arbitrary Borel Moser–Tardos process and let g be given by (7.7). From (7.8) and the Moser–Tardos theorem, we get

$$\int_{\tilde{\Omega}_k} t(x)\, \mathrm{d}\tilde{\boldsymbol{u}}_k(x) = \sum_{n=0}^{\infty} \sum_{\delta \in F_n} \int_{\tilde{\Omega}_k} \mathrm{Ind}(n, \delta^{-1} \cdot x, \mathcal{A})\, \mathrm{d}\tilde{\boldsymbol{u}}_k(x)$$

$$[\tilde{\boldsymbol{u}}_k \text{ is shift-invariant}] \quad = \sum_{n=0}^{\infty} |F| \cdot \int_{\tilde{\Omega}_k} \mathrm{Ind}(n, x, \mathcal{A})\, \mathrm{d}\tilde{\boldsymbol{u}}_k(x) \leqslant S < \infty.$$

In particular, $t(x) < \infty$ for $\tilde{\boldsymbol{u}}_k$-a.e. $x \in \tilde{\Omega}_k$, i.e., $\tilde{\boldsymbol{u}}_k(\mathrm{St}(\mathcal{A})) = 1$, so

$$\tilde{\boldsymbol{u}}_k(\mathrm{Def}(g, (\Phi_n)_{n \in \mathbb{N}}, \sigma_{k^{\mathbb{N}}})) = 0.$$

Furthermore, if $x \in \mathrm{St}(\mathcal{A})$ and $g(x) \neq x(\boldsymbol{1}, 0) = g_0(x)$, then $t(x) \geqslant 1$; thus,

$$\tilde{\boldsymbol{u}}_k(\{x \in \tilde{\Omega}_k : g(x) \neq x(\boldsymbol{1}, 0)\}) \leqslant \tilde{\boldsymbol{u}}_k(\{x \in \tilde{\Omega}_k : t(x) \geqslant 1\}) \leqslant \int_{\tilde{\Omega}_k} t(x)\, \mathrm{d}\tilde{\boldsymbol{u}}_k(x) \leqslant S,$$

as desired. ∎

Since the domain of $\Phi(k, S, \varepsilon, D)$ is, by definition, SD, (7.4) in the statement of Lemma 7.3 and Corollary 7.10 yield the remaining part of Theorem 2.6′.

8. Open problems

The following is perhaps the central open question regarding the behavior of the LLL in the measurable setting:

Problem 8.1: Does the SLLL hold measurably with a null defect? In other words, can one replace $\mu(\mathrm{Def}(f, \mathscr{B})) \leqslant \delta$ by $\mu(\mathrm{Def}(f, \mathscr{B})) = 0$ in the conclusion of Theorem 6.5?

A positive solution to Problem 8.1 would allow one to strengthen Theorem 2.11 by taking $\delta = 0$. For now, we leave this potential strengthening as an open problem.

Problem 8.2: Does Theorem 2.11 hold with $\delta = 0$?

As mentioned in §6.B, the SLLL fails in the purely Borel context [13, Theorem 1.6]. However, it is still conceivable that a purely Borel pointwise version of the Abért–Weiss theorem, similar to Theorem 2.14, holds in full generality, in which case a different proof approach might be needed to establish it. We state it here as another open question.

Problem 8.3: Let K be a compact metric space and let $f\colon \Omega \to K$ be a Borel function. Fix an open neighborhood U of the measure $\mathbb{M}_{\boldsymbol{\lambda}}\pi_f$. Does there always exist a nonempty finite set $D \subset \Gamma$ such that the following statement holds?

Let $\alpha\colon \Gamma \curvearrowright X$ be a free Borel action of Γ. Then there is a Borel map $g\colon X \to K$ such that

$$\mathbb{M}_D \pi_g(x) \in U, \quad \text{for all } x \in X.$$

Appendix A. Proof of Theorem 2.3

Let $\alpha\colon \mathbb{Z} \curvearrowright (X, \mu)$ be a free p.m.p. action of $\mathbb{Z}$. For a Borel set $A \subseteq X$, let $[A]$ denote the class of all Borel sets $B \subseteq X$ with $\mu(A \triangle B) = 0$. The measure algebra $\mathrm{MAlg}(X, \mu)$ is the space of all classes $[A]$ with the metric

$$\mathfrak{d}([A], [B]) := \mu(A \triangle B).$$

Note that the space $\mathrm{MAlg}(X, \mu)$ is Polish. For a sequence $(D_n)_{n\in\mathbb{N}}$ of nonempty finite subsets of Γ, let

$$\mathfrak{L}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$$
$$:= \{[A] \in \mathrm{MAlg}(X, \mu) : \liminf_{n\to\infty} \mathbb{E}_{D_n} \mathbb{1}_A(x) = 0 \text{ for } \mu\text{-a.e. } x \in X\};$$
$$\mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$$
$$:= \{[A] \in \mathrm{MAlg}(X, \mu) : \limsup_{n\to\infty} \mathbb{E}_{D_n} \mathbb{1}_A(x) = 1 \text{ for } \mu\text{-a.e. } x \in X\}.$$

It is straightforward to check that the sets

$$\mathfrak{L}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}}) \quad \text{and} \quad \mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$$

are G_δ in $\mathrm{MAlg}(X, \mu)$. Therefore, to establish the conclusion of Theorem 2.3, it is enough to ensure that both these sets are dense. Below we only give the argument that shows that $\mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$ is dense; the proof for $\mathfrak{L}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$ is the same, mutatis mutandis.

LEMMA A.1: *Let $h\colon \mathbb{N} \to \mathbb{N}$ be an arbitrary function and let $\varepsilon > 0$. There exists a finite sequence $D_0, \ldots, D_{N-1}$ of finite subsets of $\mathbb{Z}$ with the following properties:*

- *each D_n is an interval;*
- *$|D_n| \geqslant h(n)$ for all $0 \leqslant n < N$;*
- *for every free p.m.p. action $\mathbb{Z} \curvearrowright (X, \mu)$, there is a Borel set $A \subseteq X$ with $\mu(A) \leqslant \varepsilon$ such that*

$$\mu(\{x \in X : D_n \cdot x \subseteq A \text{ for some } 0 \leqslant n < N\}) \geqslant 1 - \varepsilon.$$

Proof. Take any N so large that

$$\frac{2}{N+1} < \varepsilon \quad \text{and} \quad (1 - \varepsilon/2)\frac{N}{N+1} > 1 - \varepsilon.$$

Let $\ell := \max_{n=0}^{N-1} h(n)$ and, for each $0 \leqslant n < N$, define

$$D_n := \{n\ell, \, n\ell + 1, \, \ldots, \, n\ell + \ell - 1\}.$$

We claim that this sequence of intervals works. Let $\alpha\colon \mathbb{Z} \curvearrowright (X, \mu)$ be a free p.m.p. action of $\mathbb{Z}$ induced by a measure-preserving transformation $T\colon X \to X$. By Rokhlin's lemma, there exists a Borel set $R \subseteq X$ such that its translates R, $TR, \ldots, T^{(N+1)\ell-1}R$ are pairwise disjoint and their union has measure at least $1 - \varepsilon/2$. Let

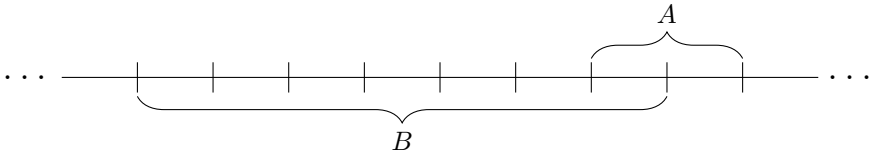$$A := \bigcup_{i=(N-1)\ell}^{(N+1)\ell-1} T^i R \quad \text{and} \quad B := \bigcup_{i=0}^{N\ell-1} T^i R.$$

Figure 1. A cartoon of the sets A and B.

Then

$$\mu(A) = 2\ell\mu(R) \leqslant 2/(N+1) < \varepsilon$$

and

$$\mu(B) = N\ell\mu(R) \geqslant (1 - \varepsilon/2)N/(N+1) > 1 - \varepsilon,$$

and for each $x \in B$, there is some $0 \leqslant n < N$ with $D_n \cdot x \subseteq A$, as desired (see Fig. 1). ∎

Let $h\colon \mathbb{N} \to \mathbb{N}$ be any function. Applying Lemma A.1 repeatedly, we construct an increasing sequence of natural numbers $(N_i)_{i\in\mathbb{N}}$ starting with $N_0 := 0$ and a sequence of finite intervals $(D_n)_{n\in\mathbb{N}}$ with $|D_n| \geqslant h(n)$, such that for every free p.m.p. action $\mathbb{Z} \curvearrowright (X,\mu)$, there exists a sequence of Borel sets $(A_i)_{i\in\mathbb{N}}$ with $\mu(A_i) \leqslant 2^{-i-1}$ and $\mu(B_i) \geqslant 1 - 2^{-i-1}$, where

$$B_i := \{x \in X : D_n \cdot x \subseteq A_i \text{ for some } N_i \leqslant n < N_{i+1}\}.$$

For $k \in \mathbb{N}$, let $A_{\geqslant k} := \bigcup_{i=k}^{\infty} A_i$. We claim that $A_{\geqslant k} \in \mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$. Indeed,

$$\limsup_{n\to\infty} \mathbb{E}_{D_n} \mathbb{1}_{A_{\geqslant k}}(x) = 1 \quad \text{for all } x \in \limsup_{i\to\infty} B_i,$$

and, by Fatou's lemma,

$$\mu(\limsup_{i\to\infty} B_i) \geqslant \limsup_{i\to\infty} \mu(B_i) = 1.$$

Now if $[S] \in \mathrm{MAlg}(X,\mu)$, then $[S \cup A_{\geqslant k}] \in \mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$ as well, and

$$\mathfrak{d}([S], [S \cup A_{\geqslant k}]) \leqslant \mu(A_{\geqslant k}) \leqslant \sum_{i=k}^{\infty} \mu(A_i) \leqslant 2^{-k}.$$

Since k is arbitrary, this shows that $\mathfrak{U}(\alpha, X, \mu, (D_n)_{n\in\mathbb{N}})$ is dense in $\mathrm{MAlg}(X,\mu)$, as desired.

References

[1] M. Abért and B. Weiss, *Bernoulli actions are weakly contained in any free action*, Ergodic Theory and Dynamical Systems **33** (2013), 323–333.

[2] M. A. Akcoglu and A. del Junco, *Convergence of averages of point transformations*, Proceedings of the American Mathematical Society **49** (1975), 265–266.

[3] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[4] C. Anantharaman, J.-P. Anker, M. Babillot, A. Bonami, B. Demange, S. Grellier, F. Havard, P. Jaming, E. Lesigne, P. Maheux, J.-P. Otal, B. Schapira and J.-P. Schreiber, *Théorèmes ergodiques pour les actions de groupes*, Monographies de L'Enseignement Mathématique, Vol. 41, L'Enseignement Mathématique, Geneva, 2010.

[5] J. Beck, *An algorithmic approach to the Lovász local lemma*, Random Structures & Algorithms **2** (1991), 343–365.

[6] A. Bernshteyn, *Measurable versions of the Lovász local lemma and measurable graph colorings*, Advances in Mathematics **353** (2019), 153–223.

[7] A. Bernshteyn, *Building large free subshifts using the local lemma*, Groups, Geometry, and Dynamics, to appear, https://arxiv.org/abs/1802.07123.

[8] G. D. Birkhoff, *Proof of the ergodic theorem*, Proceedings of the National Acadamey of Sciences of the United States of America **17** (1931), 656–660.

[9] L. Bowen and A. Nevo, *Pointwise ergodic theorems beyond amenable groups*, Ergodic Theory and Dynamical Systems **33** (2013), 777–820.

[10] A. Bufetov and A. Klimenko, *On markov operators and ergodic theorems for group actions*, European Journal of Combinatorics **33** (2012), 1427–1443.

[11] P. J. Burton, *Topology and convexity in the space of actions modulo weak equivalence*, Ergodic Theory and Dynamical Systems **38** (2018), 2508–2536.

[12] P. J. Burton and A. S. Kechris, *Weak containment of measure preserving group actions*, Ergodic Theory and Dynamical Systems, to appear, `https://doi.org/10.1017/etds.2019.26`.

[13] C. Conley, S. Jackson, A. Marks, B. Seward and R. Tucker-Drob, *Hyperfiniteness and Borel combinatorics*, Journal of the European Mathematical Society, to appear, `https://arxiv.org/abs/1611.02204`.

[14] E. Csóka, L. Grabowski, A. Máthé, O. Pikhurko and K. Tyros, *Borel version of the local lemma*, `https://arxiv.org/abs/1605.04877`.

[15] A. del Junco and J. Rosenblatt, *Counterexamples in ergodic theory and number theory*, Mathematische Annalen **245** (1979), 185–197.

[16] P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, in *Infinite and Finite Sets*, Colloquia Mathematica Societatis János Bolyai, Vol. 10, North-Holland, Amsterdam, 1975, pp. 609–627.

[17] E. Glasner, J.-P. Thouvenot and B. Weiss, *Every countable group has the weak Rohlin property*, Bulletin of the London Mathematical Society **138** (2006), 932–936.

[18] H. Hatami, L. Lovász and B. Szegedy, *Limits of locally-globally convergent graph sequences*, Geometric and Functional Analysis **24** (2014), 269–296.

[19] A. S. Kechris, *Classical Descriptive Set Theory*, Graduate Texts in Mathematics, Vol. 156, Springer, New York, 1995.

[20] A. S. Kechris, *Global Aspects of Ergodic Group Actions*, Mathematical Surveys and Monographs, Vol. 160, American Mathematical Society, Providence, RI, 2010.

[21] A. S. Kechris and B. D. Miller, *Topics in Orbit Equivalence*, Lecture Notes in Mathematics, Vol. 1852, Springer, Berlin, 2004.

[22] K. Kolipaka and M. Szegedy, *Moser and Tardos meet Lovász*, in *STOC'11—Proceedings of the 43rd ACM Symposium on Theory of Computing* ACM, New York, 2011, pp. 235–244.

[23] G. Kun, *Expanders have a spanning Lipschitz subgraph with large girth*, `https://arxiv.org/abs/1303.4982`.

[24] E. Lindenstrauss, *Pointwise theorems for amenable groups*, Inventiones Mathematicae **146** (2001), 259–295.

[25] M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method*, Algorithms and Combinatorics, Vol. 23, Springer, Berlin, 2002.

[26] R. Moser and G. Tardos, *A constructive proof of the general Lovász local lemma*, Journal of the ACM **57** (2010), Article no. 11.

[27] A. Rumyantsev and A. Shen, *Probabilistic constructions of computable objects and a computable version of Lovász local lemma*, Fundamenta Informaticae **132** (2014), 1–14.

[28] J. H. Spencer, *Asymptotic lower bounds for Ramsey functions*, Discrete Mathematics **20** (1977), 69–76.

[29] A. Tempelman, *Ergodic Theorems for Group Actions*, Mathematics and its Applications, Vol. 78, Kluwer Academic Publishers, Dordrecht, 1992.

[30] R. D. Tucker-Drob, *Weak equivalence and non-classifiability of measure preserving actions*, Ergodic Theory and Dynamical Systems **35** (2015), 293–336.

[31] J. von Neumann, *Proof of the quasi-ergodic hypothesis*, Proceedings of the National Academy of Sciences of the United States of America **18** (1932), 70–82.