SINAI FACTORS OF NONSINGULAR SYSTEMS: BERNOULLI SHIFTS AND ANOSOV FLOWS

ZEMER KOSLOFF AND TERRY SOO

Abstract. We show that a totally dissipative system has all nonsingular systems as factors, but that this is no longer true when the factor maps are required to be finitary. In particular, if a nonsingular Bernoulli shift has a limiting marginal distribution p, then it cannot have, as a finitary factor, an independent and identically distributed (iid) system of entropy larger than $H(p)$; on the other hand, we show that iid systems with entropy strictly lower than $H(p)$ can be obtained as finitary factors of these Bernoulli shifts, extending Keane and Smorodinsky's finitary version of Sinai's factor theorem to the nonsingular setting. As a consequence of our results we also obtain that every transitive twice continuously differentiable Anosov diffeomorphism on a compact manifold, endowed with volume measure, has iid factors, and if the factor is required to be finitary, then the iid factor cannot have Kolmogorov-Sinai entropy greater than the measure-theoretic entropy of a Sinai-Ruelle-Bowen measure associated with the Anosov diffeomorphism.

1. Introduction

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space. We will often be concerned with the case where μ is a probability measure, so that $\mu(\Omega) = 1$. Let $T : \Omega \to \Omega$. The map T is ***ergodic*** if $\mu(E \triangle T^{-1}(E)) = 0$ implies that $0 \in \{\mu(E), \mu(\Omega/E)\}$. We say that T is ***nonsingular*** if $\mu \circ T^{-1} \sim \mu$, that is, the measures have the same null sets; in the case that $\mu \circ T^{-1} = \mu$ we say that T is ***measure-preserving***. We refer to $(\Omega, \mathcal{F}, \mu, T)$ as a ***nonsingular dynamical system***. In the case where T is ergodic and probability-preserving, Kolmogorov-Sinai entropy [26, 34, 62], a single nonnegative real number, is assigned to the dynamical system and measures the amount of randomness contained in the system, all of which can be accounted for by a Bernoulli subsystem in the following way.

Let A be a finite set of symbols, $(\rho_i)_{i \in \mathbb{Z}}$ be a sequence of probability measures on A, and $\nu = \bigotimes_{i \in \mathbb{Z}} \rho_i$ be the product measure on the sequence space $A^{\mathbb{Z}}$ endowed with the usual product topology and the Borel product sigma-algebra $\mathcal{B} = \mathcal{B}(A^{\mathbb{Z}})$. When the product probability space $(A^{\mathbb{Z}}, \mathcal{B}, \nu)$

2020 *Mathematics Subject Classification.* 37A40, 37A35, 37C05, 60G07, 94A24.

Key words and phrases. Bernoulli shifts, finitary factors, nonsingular dynamics, dissipative systems, Anosov diffeomorphisms, entropy.

ZK is partially supported by Israel Science Foundation grant 1180/22.

is endowed with the ***left-shift*** $S : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ given by $(Sx)_i = x_{i+1}$, we refer to the dynamical system as a ***Bernoulli shift***; if all the measures $\rho_i \equiv p$ are identical, then we say that system is ***independent and identically distributed (iid)***; sometimes we will also refer to iid systems as ***stationary*** Bernoulli shifts. Thus in the nonsingular setting, a Bernoulli shift is a sequence of independent, but not necessarily identical, random variables. Sometimes we will simply refer to a system by its corresponding measure or endowed mapping. The Kolmogorov-Sinai entropy of iid systems is given by the usual Shannon entropy [61]: $H(p) := -\sum_{a \in A} p(a) \log p(a)$. We say that a nonsingular system $(\Omega, \mathcal{F}, \mu, T)$ has an ***iid factor*** of entropy h' if there exists a ***factor map*** $\phi : \Omega \to A^{\mathbb{Z}}$ such that the mapping is equivariant, $\phi \circ T = S \circ \phi$, and $\mu \circ \phi^{-1} \sim q^{\mathbb{Z}}$, where q is a probability measure on A with $H(q) = h'$; we emphasize that in the nonsingular case, we allow the possibility that the push-forward of μ under ϕ is *not* exactly $q^{\mathbb{Z}}$, only that it is a measure that is equivalent to $q^{\mathbb{Z}}$. The Sinai factor theorem [63] gives that an ergodic probability-preserving system with entropy h has all iid factors of entropies less than or equal to h.

A common notion of *chaos* for nonsingular dynamical systems is that the system can be used to simulate a bi-infinite collection of iid coin tosses and the Sinai factor theorem shows that the chaotic probability-preserving systems are precisely the systems with positive Kolmogorov-Sinai entropy [50]. In the absence of a similarly robust notion of entropy and entropy theory, it is unclear which nonsingular systems are chaotic and how to define entropy. We show that all totally dissipative systems are chaotic and that the analogue of Keane and Smorodinsky's finitary version of Sinai's factor theorem [30] remains true for a large class of nonsingular Bernoulli shifts.

1.1. Dissipative systems are universal.

Recently, we showed that a large class of nonsingular Bernoulli shifts have stationary Bernoulli shift factors [36, 37]. A curious aspect of our results is that it also applies to dissipative Bernoulli shifts, and it turns out there was a deeper reason why we never had to assume conservativity. We say that a nonsingular system is ***factor-universal*** if it has every nonsingular system as a factor. Recall a nonsingular system is dissipative if it has a wandering set of positive measure, is conservative if there are no wandering sets of positive measure, and is totally dissipative if the measurable union of the wandering sets is the whole space; see Section 2 for more precise definitions.

Theorem 1.1. *A totally dissipative system on a non-atomic measure space is factor-universal.*

We prove Theorem 1.1 in the slightly more general setting of a countable group action; see Theorem 2.2. We will also consider a version of Theorem 1.1 in the setting of a flow in Theorem 2.5; the case of flows is more subtle, and it is not true that every totally dissipative flow is factor-universal.

Theorem 1.1 may come as a surprise, since a totally dissipative transformation is, by a result of Hopf, isomorphic in the nonsingular category to the simple shift transformation $x \mapsto x + 1$ equipped with Lebesgue measure on the real line, and the latter is a rather predictable and seemingly uninteresting system; see Proposition 2.1. Although the proof is surprisingly simple, our result has several interesting applications in bridging several definitions of chaos in dynamical systems.

In the context of Ilya Prigogine's theory of dissipative structures [53], Theorem 1.1 provides an abstract mathematical explanation of how dissipative structures are systems which can possess both chaotic behavior and elliptic islands, since under this mathematical framework *any* system is a subsystem of a dissipative system.

Theorem 1.1 together with the works of Sinai and Livsic [42] and Gurevic and Oseledec [23], implies that *every* transitive C^2 Anosov diffeomorphism has stationary Bernoulli factors, and thus all transitive C^2 Anosov diffeomorphisms can be used to simulate an iid sequence, in an equivariant way.

Theorem 1.2. *A transitive C^2 Anosov diffeomorphism on a compact manifold, endowed with the natural volume measure, has iid factors, which may be chosen to have infinite entropy in the case the diffeomorphism is dissipative.*

See Section 2.3 for definitions and more details; see also Section 2.3.2 and Theorem 2.7 for the case of an Anosov flow.

We note that in Theorem 1.1 there is a key assumption that the system is totally dissipative, and that there is no regularity assumption on the associated factor mapping, other than measurabililty. If the factor map is required to be continuous almost-surely, then we have the following upper bound on the entropy of an iid factor.

Theorem 1.3 (Upper bound for Anosov diffeomorphisms)**.** *An iid system that is obtained as a continuous almost-surely factor of a transitive C^2 Anosov diffeomorphism on a compact manifold, endowed with the natural volume measure, will have Kolmogorov-Sinai entropy that is bounded by the measure-theoretic entropy of a Sinai-Ruelle-Bowen measure associated with the Anosov diffeomorphism.*

See Section 4.3 for details. In this context, a factor map that is continuous almost-surely is sometimes referred to as being ***finitary***. In some special cases, such as the measure-preserving case of a hyperbolic toral automorphism [4, 32], we know that the upper bound in Theorem 1.3 is achieved; these results rely on symbolic dynamics and the finitary constructions of Keane and Smorodinsky [30, 31] for stationary Bernoulli shifts, which we will extend to the nonsingular setting.

1.2. **Finitary factors of nonsingular Bernoulli shifts.** When we turn our attention to the case of nonsingular Bernoulli shifts, and consider factor maps that are finitary [60], so that they are continuous almost surely, we have results that are consistent with those of entropy theory and Keane and Smorodinsky [30]. We note that there are both dissipative and conservative nonsingular Bernoulli shifts that do not have an absolutely continuous invariant probability measure, with the first examples of the latter given by Krengel [39] and Hamachi [24]. It follows from Theorem 1.1 that many Bernoulli shifts have infinite entropy iid factors. Specifically, all dissipative nonsingular Bernoulli shifts are totally dissipative and thus shift universal; see also Example 1.8 and Proposition 1.9. It what follows, we address together, dissipative and conservative nonsingular Bernoulli shifts, with the added finitary regularity assumption on the factor mapping.

A product measure $\rho = \bigotimes_{i\in\mathbb{Z}} \rho_i$ on $A^{\mathbb{Z}}$ satisfies the ***Doeblin condition*** if there exists $\delta > 0$ such that for all $a \in A$ and $i \in \mathbb{Z}$, we have $\rho_i(a) > \delta$. We say that the ***limiting marginal measure*** is p if $\rho_{|i|}(a) \to p(a)$ for all $a \in A$. Motivated by Sinai's original factor theorem, we say that a family of iid factor maps of this system have ***near optimal*** entropy if they can produce systems of entropy $H(p) - \varepsilon$ for every $\varepsilon > 0$. An iid factor has ***optimal*** entropy if it has entropy $H(p)$ and has ***super optimal*** entropy if it has entropy greater than $H(p)$.

Let A and B be finite sets. Consider a factor map $\phi : A^{\mathbb{Z}} \to B^{\mathbb{Z}}$, where $A^{\mathbb{Z}}$ is equipped with a nonsingular measure μ. Consider the zeroth coordinate projection $\bar{\phi}$ given by $\bar{\phi}(x) = \phi(x)(0)$. We say that ϕ is ***finitary*** if for every $b \in B$ there exists C_b which is a countable union of cylinder subsets of $A^{\mathbb{Z}}$ such that $\mu(\bar{\phi}^{-1}(b)\triangle C_b) = 0$. This condition is equivalent to the map being continuous almost surely with respect to μ and also equivalent to ϕ having an almost surely finite coding radius, see for example [52, page 281]. This condition was first introduced by Weiss [71] and studied by Denker and Keane [15], see also [60].

Theorem 1.4 (Upper bound for finitary factors)**.** *A nonsingular Bernoulli shift that has a limiting measure, does not have a finitary iid factor with super optimal entropy.*

We will prove Theorem 1.4 as a consequence of a more general result, Theorem 4.1, which by using symbolic dynamics [64] and some finer results regarding Anosov diffeomorphisms [12], will also allow us to infer Theorem 1.3.

Turning our attention to positive results, we recently proved the following theorem regarding finitary factors in [37, Theorem 1].

Theorem 1.5 (Low entropy factors [37])**.** *Every nonsingular Bernoulli shift which satisfies the Doeblin condition has a finitary iid factor.*

In Theorem 1.5, it is obvious from our construction that the finitary factor is non-optimal with respect to entropy. Keane and Smorodinsky [30, 31] defined finitary factor isomorphisms between any two measure-preserving Bernoulli shifts of finite entropy using a *marker-filler* method; our proof of Theorem 1.6 will also make use of these ideas adapted to the nonsingular setting.

Theorem 1.6 (Near optimal entropy factors)**.** *If a nonsingular Bernoulli shift satisfies the Doeblin condition, and has a limiting measure, then it has near optimal entropy finitary iid factors.*

By an elementary argument, we obtain the following extension, where we remove the Doeblin condition and allow for the possibility of a countable number of symbols.

Corollary 1.7. *If nonsingular Bernoulli shift on a countable (or finite) number of symbols has a limiting measure, then it has near optimal entropy finitary iid factors.*

In Theorem 1.7, in the case that the Bernoulli shift has infinite entropy, the finitary iid factors can be taken to have arbitrarily large finite entropy; see Section 3.5 for details.

We remark that there are many conservative Bernoulli shifts which do not have an absolutely continuous invariant measure, and understanding these shifts is an active area of research; see for example, [9, 24, 35, 14]. Specifically, Theorem 1.6 does not follow from Theorem 1.1 or Keane and Smorodinsky's version of the Sinai factor theorem.

Example 1.8. We illustrate our theory with following example which was studied in [69], where the theory provides simple examples of nonsingular conservative Bernoulli shifts that do not admit an absolutely continuous invariant measure. Consider $A = \{0, 1\}$, and the family of product measures ρ^c with marginals

$$\rho^c_n(0) = \frac{1}{2} + \frac{c}{\sqrt{n}} \cdot \mathbf{1}[n \geq 1, c/\sqrt{n} < 1/2], \tag{1}$$

where $c > 0$ is a parameter. It is easy to check using Kakutani's theorem (see Section 3.2) that ρ^c is nonsingular and is not equivalent to the product measure $(\frac{1}{2}, \frac{1}{2})^{\mathbb{Z}}$, but by Theorem 1.6, it still has a finitary iid factor of entropy almost $H(\frac{1}{2}, \frac{1}{2}) = \log 2$ and by Theorem 1.4 finitary iid factors must have entropy at most $\log 2$.

We proved that there exists a critical $c^* \in (1/6, \infty)$ such that the shift is totally dissipative for $c > c^*$ and for $c < c^*$ the shift is conservative [37, Theorem 3]. Thus when $c > c^*$, by Theorem 1.1, the shift has any system (even a circle rotation) as a (non-finitary) factor. ♢

With regards to nonsingular Bernoulli shifts, we do not know if finitary optimal entropy factors must exist, and we cannot exclude the possibility of super optimal entropy non-finitary factors, even when the shift is conservative.

A nonsingular dissipative Bernoulli shift that does *not* satisfy the Doeblin condition and has a trivial limiting measure may have the following varied behaviour.

Proposition 1.9. *There exists a nonsingular Bernoulli shift which has all iid factors, but no finitary iid factor.*

Our proof of Proposition 1.9 will follow from Theorem 1.1 and Theorem 1.4. See Section 5.

1.3. **Some remarks about random number generation and Theorem 1.6.** The practical problem of random number generation has a long history [40, 70] and continues to be a topic of current research. A central goal of (true) random number generation is to generate iid fair coin tosses from an iid biased source, in an efficient way, see for example [51]. A somewhat more realistic, and less studied assumption is to assume the source is *noisy*, that is given by independent, but not identical random variables [73]; the goal in this setting is often to efficiently generate independent bits that are *approximately* fair. In the language of random number generation, we are interested in generating, in an equivariant way, from a noisy source, a sequence that is statistically indistinguishable from an iid sequence of near optimal entropy. The equivariance requirement means that if the source input is given on a ticker tape, then the same procedure is applied everywhere on the ticker tape, and positions on the ticker tape do not need to be additionally labelled. The factor map given by Theorem 1.6 is finitary, so that our procedure, in principle, can be implemented by a machine.

1.4. **Some tools used in our proofs.** Our proof of Theorem 1.1 for countable group actions is simple and given in the next section. The corresponding version for flows is more involved, and as a result the version of Theorem 1.2 for Anosov flows is a bit harder and requires understanding of their ergodic decompositions.

In Section 3, our proof of Theorem 1.6 draws from a combination of ideas of Keane and Smorodinsky [30, 31], Harvey, Holroyd, Peres, and Romik [25], Karen Ball [8], some tools from information theory, and our last paper on factors in the nonsingular setting [37]. In [25], the authors used marker-filler methods in combination of with a version of Elias' [18] construction of an unbiased (random length) iid binary sequence from a stationary source to produce *source universal* finitary factors. Ball also employed marker-filler methods in her construction of *monotone* finitary factors; see also [54]. We will prove a disintegration of a non-stationary product measure that will have enough regularity which will allow us to recover enough stationarity.

In Section 4, our proof of Theorem 1.4 will make use of some ideas from Kalikow's [29, Chapter 4.1, Theorem 365] beautiful elementary proof of Kolmogorov's theorem [33, 34] that the three-shift is not a factor of the two-shift; whereas Kolmogorov's proof was an easy consequence that entropy cannot increase under factor maps, Kalikow's proof has the advantage that it uses "essentially nothing." This proof can be modified and extended to our nonsingular setting, at the cost of restricting to finitary factors, leaving the possibility that the analogue of Kolmogorov's theorem in the nonsingular setting does not hold, even when the shift is conservative. We will prove an abstraction of Theorem 1.4 which will also apply to other symbolic systems. In order to obtain the bound for Anosov diffeomorphisms, we use symbolic dynamics, and carry out entropy calculations that simultaneously involve two measures; the first volume lemma of Bowen and Ruelle [12], allows us to reconcile the volume and its Sinai-Ruelle-Bowen measure.

2. Dissipative actions are factor universal

Let $(\Omega, \mathcal{F}, \mu)$ be a non-atomic (standard) σ-finite measure space. An (nonsingular) ***automorphism*** of $(\Omega, \mathcal{F}, \mu)$ is an invertible measurable map $R : \Omega \to \Omega$ which satisfies $\mu \circ R^{-1} \sim \mu$. Denote the automorphism group of $(\Omega, \mathcal{F}, \mu)$ by $\mathrm{Aut}(\Omega, \mathcal{F}, \mu)$. A ***$G$-action*** of a locally compact group G on $(\Omega, \mathcal{F}, \mu)$, sometimes abbreviated as $G \curvearrowright (\Omega, \mathcal{F}, \mu)$, is a group homeomorphism $g \mapsto T_g$ from G to $\mathrm{Aut}(\Omega, \mathcal{F}, \mu)$.

The action $(\Omega', \mathcal{F}', \mu', (S_g)_{g\in G})$ is a ***factor*** of $(\Omega, \mathcal{F}, \mu, (T_g)_{g\in G})$ if there exists a measurable map $\pi : \Omega \to \Omega'$ such that π is equivariant, so that μ almost everywhere, for all $g \in G$, we have $\pi \circ T_g = S_g \circ \pi$, and in addition $\mu' \sim \mu \circ \pi^{-1}$; if π is invertible, modulo null sets, then we say that the two systems are ***isomorphic***. We say that the action $G \curvearrowright (\Omega, \mathcal{F}, \mu)$ is ***factor-universal*** if every G-action $G \curvearrowright (\Omega', \mathcal{F}', \mu')$ is a factor.

2.1. Discrete group actions.

Let G be a countable group. Given an action $G \curvearrowright (\Omega, \mathcal{F}, \mu)$, a set $W \in \mathcal{F}$ is ***wandering*** if $\{T_g W\}_{g\in G}$ are pairwise disjoint; the action is ***dissipative*** if there exists a wandering set with $\mu(W) > 0$ and it is ***totally dissipative*** if there is a wandering set W such that the union of its translates is the whole space modulo a null set: $\biguplus_{g\in G} T_g W = \Omega \bmod \mu$. The action is ***conservative*** if every wandering set is of measure zero. A set $A \in \mathcal{F}$ is ***G-invariant*** if for all $g \in G$, we have $T_g^{-1} A = A$.

The following action is the prototypical example of a totally dissipative action of a non-atomic measure space. Let $Z = [0, 1] \times G$ and ν be the product measure of the uniform measure on $[0, 1]$ and the counting measure on G. The ***translation action*** of G on Z is defined by $S_g(x, h) = (x, hg)$. The following result is due to Hopf, and closely related to the Hopf decomposition which partitions the space into a dissipative part and a conservative part; see [1, Propostion 1.1.2 and Exercise 1.2.1].

Proposition 2.1 (Hopf)**.** *If $G \curvearrowright (\Omega, \mathcal{F}, \mu)$ is a totally dissipative action of a non-atomic measure space, then is isomorphic to the translation action of G on $[0, 1] \times G$.*

Proof. By passing to an absolutely continuous probability we may assume that $\mu(\Omega) = 1$. Since the action is totally dissipative, let $W \in \mathcal{F}$ be a wandering set whose disjoint union of translates are the whole space. Recall we assume that the measure space is standard and non-atomic and $\mu(W) > 0$. By the Borel isomorphism theorem [67, Theorem 3.4.23], let $\beta : W \to [0, 1]$ be a bijective measurable map such that $\frac{d\mu\circ\beta}{d\mu} = \frac{1}{\mu(W)}$. Define $\pi : \Omega \to [0, 1] \times G$ by $\pi(x) = \left(\beta\left(T_g^{-1} x\right), g\right)$ where $g \in G$ is the unique group element such that $x \in T_g W$. Clearly, π is bijective and equivariant from which it follows that π is an isomorphism. $\square$

Theorem 2.2. *A totally dissipative action of countable group on a non-atomic measure space is factor universal.*

Proof. Let the first coordinate, $(\Omega_1, \mathcal{F}_1, \mu_1, (T_g)_{g\in G})$ be a totally dissipative G-action of a non-atomic measure space, and let the second coordinate, $(\Omega_2, \mathcal{F}_2, \mu_2, (S_g)_{g\in G})$ be any nonsingular G-action. Consider the direct product action $(T_g \otimes S_g)_{g\in G}$, which is a nonsingular G-action on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. Clearly the projection proj_2 from $\Omega_1 \times \Omega_2$ to the second coordinate is a factor map.

It is easy to verify that the product action inherits the totally dissipative property from the first coordinate. By Proposition 2.1, there exists an isomorphism $\psi : \Omega_1 \to \Omega_1 \times \Omega_2$ between the first action and the direct product action. We deduce that the composition $\mathrm{proj}_2 \circ \psi$ illustrated by

$$\Omega_1 \xrightarrow{\psi} (\Omega_1 \times \Omega_2) \xrightarrow{\mathrm{proj}_2} \Omega_2,$$

is the desired factor map from the first coordinate to the second. $\square$

The following shows that the latter are the only factor universal actions.

Proposition 2.3. *If $G \curvearrowright (\Omega, \mathcal{F}, \mu)$ has a totally dissipative factor, then it is totally dissipative.*

Proof. Towards a contradiction, suppose that the action is not totally dissipative and it has a totally dissipative factor $(\Omega', \mathcal{F}', \mu', (S_g)_{g\in G})$, via the factor map π. Since the conservative part in the Hopf decomposition is G-invariant, it follows that by restricting the factor map to the conservative part we may assume without loss of generality that $G \curvearrowright (\Omega, \mathcal{F}, \mu)$ is a conservative action.

Choose some $W \in \mathcal{F}'$ a wandering set for $(S_g)_{g\in G}$ of positive μ'-measure. It is easy to see that set $\pi^{-1}W \in \mathcal{F}$ is a wandering set of positive measure, contradicting the conservativity of $G \curvearrowright (\Omega, \mathcal{F}, \mu)$. $\square$

2.2. $\mathbb{R}$ **and** $\mathbb{R}^d$**-flows.** A nonsingular $\mathbb{R}^d$-action $\big(\Omega, \mathcal{F}, \mu, (\phi_t)_{t\in\mathbb{R}^d}\big)$ is a $\mathbb{R}^d$-**flow**, and the case $d = 1$ will simply be referred to as a ***flow***. An $\mathbb{R}^d$-flow is ***totally dissipative*** if for every $s > 0$, we have that $(\phi_{sn})_{n\in\mathbb{Z}^d}$ is a totally dissipative $s\mathbb{Z}^d$ action; this is equivalent to verifying the dissipativity for the single case of $s = 1$ [1, Corollary 1.6.5]. Sometimes we will denote Lebesgue measure on $\mathbb{R}^d$ by $\mathrm{Leb} = \mathrm{Leb}_{\mathbb{R}^d}$. Let $(\Gamma, \mathcal{C}, \nu)$ be a standard probability space [56]. A ***translation flow of*** $\mathbb{R}^d$ ***with respect to*** Γ is the dissipative action on $\Gamma \times \mathbb{R}^d$, equipped the product measure $\nu \otimes \mathrm{Leb}$, given by $\tau_s(z, t) = (z, t+s)$, for all $s, t \in \mathbb{R}^d$. When Γ is the unit interval and ν is Lebesgue measure, then we say that translation flow is ***canonical***. Note that when $\Gamma = \{*\}$ is a one point space, the translation flow is ergodic, and we may omit Γ; by the theorem below, all ergodic dissipative flows are isomorphic to this translation flow.

The following analogue of Proposition 2.1 was proved by Krengel [38, Satz 4.2] for flows and by Rosinski [57, Theorem 2.2] for $\mathbb{R}^d$-flows by constructing a relevant cross section.

Theorem 2.4 (Krengel and Rosinski)**.** *For every totally dissipative $\mathbb{R}^d$-flow on a nonatomic measure space, there exists a standard probability space, $(\Gamma, \mathcal{C}, \nu)$ such that the flow is isomorphic to the translation flow of $\mathbb{R}^d$ with respect to Γ.*

We note that in Rosinski [57], $(\Gamma, \mathcal{C}, \nu)$ is a σ-finite standard, measure space. By passing to an ν equivalent probability measure, we may assume that ν is a probability measure since the identity map gives the obvious isomorphism. In addition, we may substitute for $(\Gamma, \mathcal{C}, \nu)$, any other probability space that is isomorphic in the category of measure. The space Γ corresponds to the ergodic decomposition of the flow. In particular, the translation flow τ is ergodic if and only if Γ is a one point space. By Theorem 2.4, an ergodic dissipative flow of a non-atomic measure space is isomorphic to the translation flow on $\mathbb{R}^d$ with respect to the one point space $\Gamma = \{*\}$, which we refer to as ***the ergodic dissipative*** $\mathbb{R}^d$***-flow***.

Theorem 2.5 (Factors of totally dissipative flows)**.**

(a) The canonical translation flow of $\mathbb{R}^d$ is factor-universal.
(b) The ergodic dissipative $\mathbb{R}^d$-flow is a factor of any totally dissipative $\mathbb{R}^d$-flow on a non-atomic measure space.

Proof. For the proof of part (a), let $(\Omega, \mathcal{F}, \mu, (\phi_s)_{s\in\mathbb{R}^d})$ be a given nonsingular flow, which may not be dissipative. Let τ be the canonical translation flow. The product flow $\phi \otimes \tau = (\phi_s \otimes \tau_s)_{s\in\mathbb{R}^d}$ is a totally dissipative flow on a non-atomic measure space. By Theorem 2.4, there exists a standard probability space, $(\Gamma, \mathcal{C}, \nu)$ such that $\phi \otimes \tau$ is isomorphic to the translation flow of $\mathbb{R}^d$ with respect to Γ; let $\Theta : \Omega \times ([0, 1] \times \mathbb{R}) \to \Gamma \times \mathbb{R}^d$ be isomorphism.

By a routine variation of the Borel isomorphism theorem, there exists a homomorphism of probability spaces $g : [0, 1] \to \Gamma$ such that $\mathrm{Leb}_{[0,1]} \circ g^{-1} = \nu$. Let $\mathrm{proj}_\Omega : \Omega \times ([0, 1] \times \mathbb{R}^d) \to \Omega$ be the projection onto Ω. Then the following compositions of mappings given by the diagram

$$[0, 1] \times \mathbb{R}^d \xrightarrow{g\otimes \mathrm{id}_{\mathbb{R}^d}} \Gamma \times \mathbb{R} \xrightarrow{\Theta^{-1}} \Omega \times ([0, 1] \times \mathbb{R}^d) \xrightarrow{\mathrm{proj}_\Omega} \Omega,$$

gives a factor map from the canonical translation flow of $\mathbb{R}^d$ to the given flow.

For part (b), again by Theorem 2.4, every dissipative $\mathbb{R}^d$-flow is isomorphic to the translation flow of $\mathbb{R}^d$ with respect to some standard probability space Γ. Next, note that the projection $\mathrm{proj}_{\mathbb{R}^d} : \Gamma \times \mathbb{R}^d \to \mathbb{R}^d$ is a factor map from the dissipative translation flow of $\mathbb{R}^d$ with respect to Γ to the ergodic dissipative $\mathbb{R}^d$-flow. □

Remark 2.6. The case of flows differs from the countable group case since the translation flow on $\mathbb{R}^d$ given by $\tau_s(x) = x + s$ is ergodic. Consequently, not all totally dissipative flows on a non-atomic measure space are factor universal. We thank Jon Aaronson for pointing out a misinterpretation of Theorem 2.4 which led to an incorrect formulation of Theorem 2.5 in an earlier version of this manuscript. $\lozenge$

2.3. Applications to chaos in C^2 Anosov diffeomorphisms and flows.

Let M be a compact Riemannian manifold without boundary and vol_M be the volume measure on M. By the change of variables formula, for every diffeomorphism f, the system $(M, \mathcal{B}(M), \mathrm{vol}_M, f)$ is nonsingular. Let $\mathrm{Diff}^k(M)$ be the collection of C^k-diffeomorphisms. A diffeomorphisms is ***(topologically) transitive*** if there exists $x \in M$ such that $\{f^n(x) : \ n \in \mathbb{Z}\}$ is dense in M. Similarly a C^2***-flow*** is a homeomorphism $s \mapsto \phi_s$ from $\mathbb{R}$ to $\mathrm{Diff}^2(M)$ and the flow is transitive if it has a dense $\mathbb{R}$-orbit. Anosov systems are a central object of study in ergodic theory and dynamical systems [7, 11, 66].

2.3.1. *Applications to Anosov diffeomorphisms.* A diffeomorphism $f : M \to M$ is ***Anosov*** (uniformly hyperbolic) if for all $x \in M$, there exists a decomposition of the tangent bundle over x, given by $T_xM = E^{\mathrm{s}}_x \oplus E^{\mathrm{u}}_x$ such that:

- For all $x \in M$, we have $Df(x)E^{\mathrm{s}}_x = E^{\mathrm{s}}_{f(x)}$ and $Df(x)E^{\mathrm{u}}_x = E^{\mathrm{u}}_{f(x)}$.
- For any metric on TM, there exists $a > 0$ and $\lambda \in (0, 1)$ such that for all $v \in E^{\mathrm{s}}_x$, for all $n \in \mathbb{Z}^+$, we have
$$\|Df^n(x)v\| \leq a\lambda^n \|v\|,$$
and similarly for all $v \in E^{\mathrm{u}}_x$, for all $n \in \mathbb{Z}^+$, we have
$$\left\|Df^{-n}(x)v\right\| \leq a\lambda^n \|v\|.$$

Our proof of Theorem 1.2 is a consequence of Livsic-Sinai [42], when a Sinai-Ruelle-Bowen (SRB) measure [72] is available; on the other hand, Gurevic and Oseledec [23] showed that in the absence of a volume absolutely continuous invariant probability (a.c.i.p), the Anosov diffeomorphism is dissipative and thus Theorem 2.2 applies; see also Theorem 2.8. We remark that in the categorical sense *most* C^2 Anosov diffeomorphism do not have a volume a.c.i.p. [65]; see also [11, Corollary 4.15].

Proof of Theorem 1.2. Let f be a C^2 Anosov diffeomorphism. If there exists a volume absolutely continuous invariant probability (a.c.i.p) μ, then μ is an SRB measure and $(M, \mathcal{B}(M), \mu, f)$ has positive entropy [42]. By Sinai factor theorem $(M, \mathcal{B}(M), \mu, f)$ has an iid factor, and since $\mu \sim \mathrm{vol}_M$, this remains true when μ is replaced by vol_M.

In the absence of a volume a.c.i.p., $(M, \mathcal{B}(M), \mathrm{vol}_M, f)$ is a totally dissipative transformation of a non-atomic measure space [23], and it follows from Theorem 2.2 that every stationary Bernoulli shift is a factor. $\square$

2.3.2. *Applications to Anosov flows.* We will prove a version of Theorem 1.2 for the case of flows. Following Ornstein [46], we say that a probability-preserving flow $\left(\Omega, \mathcal{F}, \mu, (\phi_t)_{t\in\mathbb{R}^d}\right)$ is a ***Bernoulli flow*** if for every $s > 0$, the discrete-time system $(\Omega, \mathcal{F}, \mu, \phi_s)$ is isomorphic to a stationary Bernoulli shift. Bernoulli flows are the analogues of iid-systems for $\mathbb{R}$-flows and two Bernoulli flows whose time-one maps have equal Kolmogorov-Sinai entropy are isomorphic [45, 46]. Canonical examples of Bernoulli flows include the Totoki flow [22, 68, 46], and infinite entropy flows arising from Brownian motions and Poisson processes [47, 49].

A C^2 flow $(\phi_t)_{t\in\mathbb{R}}$ on compact Riemannian manifold M is an ***Anosov flow*** if for all $x \in M$, the decomposition $T_xM = E^{\mathrm{s}}_x \oplus E^{\mathrm{c}}_x \oplus E^{\mathrm{u}}_x$ is $D\phi_t$ equivariant, E^{c} is one dimensional and corresponds to the direction of the flow, and there exists $A > 0$ and $b > 0$ such that for all $x \in M$,

$$\|D\phi_t v\| \leq Ae^{-bt}\|v\|, \quad \text{for every } t > 0 \text{ and } v \in E^{\mathrm{s}}_x.$$
$$\|D\phi_{-t} v\| \leq Ae^{-bt}\|v\|, \quad \text{for every } t > 0 \text{ and } v \in E^{\mathrm{u}}_x.$$

Theorem 2.7. *A transitive C^2 Anosov flow, endowed with the natural volume measure, has a Bernoulli flow as a factor.*

Our proof of Theorem 2.7 proceeds as in Theorem 1.2. Sinai [65] showed that for every transitive C^2 Anosov flow there exists two probability measures μ^+ and μ^- such that for every continuous function $g : M \to \mathbb{R}$, for vol_M-almost every $x \in M$, we have

$$\lim_{N\to\infty} \frac{1}{N}\int_0^N g\circ\phi_t(x)dt = \int g d\mu^+, \quad \text{and}$$
$$\lim_{N\to\infty} \frac{1}{N}\int_{-N}^0 g\circ\phi_t(x)dt = \int g d\mu^-;$$

see also [12] for the more general setting of *Axiom A* flows. Again, the measures μ^+ and μ^- are called ***SRB measures***. The following was proved by Gurevic and Oseledec for C^2 Anosov diffeomorphisms, and we present the identical proof for flows for completeness.

Theorem 2.8 (Gurevic and Oseledec)**.** *Let $(M, (\phi_s)_{s\in\mathbb{R}})$ be a transitive C^2 Anosov flow. If there is no volume a.c.i.p., then $(M, \mathrm{vol}_M, (\phi_s)_{s\in\mathbb{R}})$ is totally dissipative.*

Proof. There exists a volume a.c.i.p. measure if and only if $\mu^+ = \mu^-$ [42], so we have $\mu^+ \neq \mu^-$. Let $T = \phi_1$ be the time-one map of the flow. Recall that it suffices to show that T is totally dissipative [1, Corollary 1.6.5]. Moreover, note that T is totally dissipative if and only if T^{-1} is totally dissipative.

Fix a continuous function on $g : M \to \mathbb{R}$ with $\int g d\mu^+ \neq \int g d\mu^-$. Set $\delta := \frac{1}{2}\left|\int g d\mu^+ - \int g d\mu^-\right| > 0$. Let $\epsilon > 0$. As μ^+ is a SRB measure, there exists $N > 0$ such that the set

$$A = \left\{x \in M :\ \forall n > N, \text{ we have } \left|\frac{1}{n}\int_0^n g\circ\phi_t(x)dt - \int g d\mu^+\right| < \delta\right\},$$

satisfies $\mu(A) > 1-\epsilon$. For all $n \in \mathbb{N}$, we have

$$\left(\int_0^n g \circ \phi_t dt\right) \circ T^{-n}(x) = \int_{-n}^0 g \circ \phi_t(x) dt.$$

As the latter tends to $\int g d\mu^-$ as $n \to \infty$ for almost every $x \in M$, for vol_M-almost every $x \in A$, we have

$$\#\{n \in \mathbb{N} : \ T^{-n}x \in A\} < \infty.$$

Since $\epsilon > 0$ is arbitrary, it follows from Halmos recurrence theorem [1, Theorem 1.1.1] that $(M, \mathrm{vol}_M, T^{-1})$ is totally dissipative. □

In order to use Theorem 2.5 (a), we will need the following lemma.

Lemma 2.9. *A transitive and totally dissipative C^2 flow $(\phi_s)_{s\in\mathbb{R}}$ on a compact manifold M is isomorphic to the canonical translation flow of $\mathbb{R}$.*

Proof. We already know from Theorem 2.4 that the flow is isomorphic to the translation flow of $\mathbb{R}$, with respect to some standard probability space $(\Gamma, \mathcal{C}, \nu)$; thus it suffices to show that Γ is non-atomic. Towards a contradiction, let $\pi : \Gamma \times \mathbb{R} \to M$ be the isomorphism and $g \in \Gamma$ be a ν-atom. Since π is an isomorphism, there exists $x \in M$ such that following its orbit on a manifold for one unit of time is the same as moving along the unit interval, so that

$$\begin{aligned}(\nu \otimes \mathrm{Leb}) \circ \pi^{-1}(\{\phi_s(x) : s \in [0,1]\}) &= (\nu \otimes \mathrm{Leb})(\{g\} \times [0,1]) \\ &= \nu(g)\mathrm{Leb}([0,1]) > 0.\end{aligned}$$

As $(\nu \otimes \mathrm{Leb}) \circ \pi^{-1}$ and vol_M are equivalent measures we have that

$$\mathrm{vol}_M \{\phi_s(m) : s \in [0,1]\} > 0.$$

We recall that the orbit $s \mapsto \phi_s(x)$ is smooth and is the solution to first order autonomous ordinary differential equation [66, page 795]. However, it is well-known that the image of a smooth curve on a manifold, with dimension two or higher, has no volume [59]. □

Proof of Theorem 2.7. If the flow has a volume a.c.i.p. μ, then μ is a Gibbs measure [42]. It follows from a result of Ornstein and Weiss [48] and Ratner [55] that flow endowed with $\mu \sim \mathrm{vol}_M$ is isomorphic to a Bernoulli flow.

If there is no volume a.c.i.p., then by Theorem 2.8 the flow endowed with volume measure is totally dissipative. By Lemma 2.9, the flow is isomorphic to the canonical translation of $\mathbb{R}$, and hence by Theorem 2.5 (a), *every* Bernoulli flow is a factor. □

3. Near optimal Sinai factors

3.1. **Markers and fillers.** Let $k_{\text{mark}} \in \mathbb{Z}^+$ be a large positive integer which we will specify later. Let A be a finite set containing the two distinct symbols a and b. Given $x \in A^{\mathbb{Z}}$, we call an integer interval $[n, n+2k_{\text{mark}}]$ a ***marker***, if $x_{n+i} = a$ for all $0 \leq i \leq 2k_{\text{mark}} - 1$ and $x_{n+2k_{\text{mark}}} = b$. Any integer that does not belong to a marker, belongs to a maximal ***filler***, so that markers and fillers partition the integers. Thus with markers and fillers, our task is to find an encoding of a finite non-stationary sequence into a finite stationary sequence that retain most of the entropy. A technical problem arises that we cannot control the size of the fillers, and for our construction, it will be convenient to have a version of fillers of a fixed size. The following idea is from Ball [8] and its presentation is adapted from [54]. We define a bi-infinite sequence of ***alternating*** intervals $I(x) = (I_i)_{i\in\mathbb{Z}}$ that partition $\mathbb{Z}$ into intervals of length k_{mark} and 1 in the following way. Locate all the markers of x. Any $n \in \mathbb{Z}$ that belongs to the right endpoint of a marker is an interval of length 1, following a marker will always be an interval of length k_{mark}, and if x restricted to the interval of length k_{mark} is not a string of k_{mark} consecutive a's, then the following interval will also be one of length k_{mark}, otherwise, the following intervals will all be of length 1, until the a stops occurring; the following interval will be one of length k_{mark}. We say that a ***switch*** occurs in an alternating interval if it is an interval of length k_{mark} and is a string of consecutive a's or if its an interval of length 1 and the symbol that is *not* a has appeared. For definiteness, we require that $0 \in I_0$, and $\sup I_i < \inf I_j$ if $i < j$. In what follows it will be more convenient to use the language of random variables.

Proposition 3.1 (Ball's alternating intervals)**.** *Let $X = (X_i)_{i\in\mathbb{Z}}$ be random variables corresponding to the law of a Bernoulli shift on A. Let k_{mark} be a positive integer. Conditioned on the alternating intervals $I(X) = (I_i)_{i\in\mathbb{Z}}$, the random sequence X has the following properties:*

- *The random variables $(X|_{I_i})_{i\in\mathbb{Z}}$ are independent.*
- *On each alternating interval I_i of size 1 not immediately left of an interval of size k_{mark}, we know that $X|_{I_i} = a$; otherwise $X|_{I_i} \neq a$, and a switch occurs.*
- *On each alternating interval I_i of size k_{mark} that is not immediately left of an interval of size 1, the law of $X|_{I_i}$, is the law of $X|_{I_i}$ conditioned not to be a string of a's; otherwise $X|_{I_i}$ is a string of a's, and a switch occurs.*

Proof. Let ρ be the law of X. Let $n \in \mathbb{Z}^+$. Consider the random variables $Y^n = (Y_i)_{i=-n}^{\infty}$ sampled using the procedure. We start by sampling k_{mark} elements from the measure $(\rho_{-n}, \ldots, \rho_{-n+k-1})$, and we continue to sample in blocks of k_{mark} until a switch occurs, in which case we sample one coordinate at a time, until a switch occurs, and then we go back to sampling k_{mark} elements at a time. Since switches are stopping times, it follows that Y^n

has the same law as $X^n = (X_i)_{i=-n}^{\infty}$, for all $n \in \mathbb{Z}^+$. The above properties clearly hold for the weak limit of Y^n and thus hold for X. □

3.2. **A Kakutani equivalent coding.** Recall that by Kakutani's theorem [28], we have that two infinite direct product measures μ and ν on $A^{\mathbb{Z}}$ satisfying the Doeblin condition are equivalent if and only if

$$\sum_{n \in \mathbb{Z}} \sum_{a \in A} (\mu_n(a) - \nu_n(a))^2 < \infty. \tag{2}$$

Note that Kakutani's theorem can be used to identify which Bernoulli shifts are nonsingular and which Bernoulli shifts are equivalent to stationary ones. One can imagine a coin flipper who progressively get better at flipping a coin, but does not get better so quickly that their flips are indistinguishable from iid ones.

We will say that a statement holds for a product measure μ modulo or up to Kakutani equivalence if there is an equivalent measure ν for which the corresponding statement holds. A key idea in our proof of Theorem 1.5 is that although a Bernoulli shift may not be given by independent and identical observations, non-identical observations can be combined in a way to yield identical observations, up to Kakutani equivalence. Specifically, we adapted von Neumann's observation that one can simulate a fair coin from a possibly biased coin, using what is now commonly referred to as rejection sampling [70]; flip the biased coin twice: if we get HT, we report this as H and if we get TH, we report this as a T, and we are repeat this procedure if we get either HH or TT. We remark that in the nonsingular setting, we are faced with the added difficulty of using different coins for each flip.

Given a product measure $\rho = \bigotimes_{i \in \mathbb{Z}} \rho_i$, we let $\rho_i^{\oplus k}$ be the probability measure on A^k distributed as $(\rho_i, \ldots, \rho_{i+k-1})$; we will use the notation $\rho_i^{\otimes k} = (\rho_i, \ldots, \rho_i)$ to denote the k-fold product of the measure ρ_i. An important fact we will exploit is that the sequence of measures $\rho_i^{\oplus k}$ and $\rho_i^{\otimes k}$ are equivalent, in the sense of Kakutani; see Lemma 3.5 and also [37, Theorem 20]. Some of the proofs and lemmas are concerned with demonstrating that we can substitute a more difficult statement involving $\rho_i^{\oplus k}$ with a simpler statement involving $\rho_i^{\otimes k}$. If Q is a probability measure and B is a measurable subset, then the probability measure given by

$$Q(\cdot|B) := \frac{Q(\cdot \cap B)}{Q(B)}$$

will sometimes be referred to as the ***conditional probability of*** Q ***given*** B.

Proposition 3.2 (Kakutani equivalent coding)**.** *Suppose ρ is a nonsingular Bernoulli product measure that satisfies the Doeblin condition and has a limiting marginal measure p on A. Let $\varepsilon > 0$. There exists $\alpha > 0$ and k_{mark} sufficiently large such that for all $k \geq k_{\mathrm{mark}}$ there exists a subset $\mathcal{G}_k \subset A^k$ omitting the sequence $a^k \in A^k$, a finite set B_k with the following properties:*

(a) $\#B_k \geq 2^{k(H(p)-\varepsilon)}$.
(b) For all but finitely many $i \in \mathbb{Z}$*, we have* $\rho_i^{\oplus k}(A^k \setminus \mathcal{G}_k) \leq e^{-\alpha k}$.
(c) There exists a single function $\psi : \mathcal{G}_k \to B_k$ *such that*

$$\sum_{i\in\mathbb{Z}} \sum_{c\in B_k} \left(\rho_i^{\oplus k}(\psi^{-1}(c)|\mathcal{G}_k) - \frac{1}{\#B_k} \right)^2 < \infty. \tag{3}$$

In our proof of Proposition 3.2, we make the following observation. For simplicity consider the binary case where $A = \{0,1\}$. Each $\rho_i^{\oplus k}$ can be viewed as a disintegration given by first sampling from a distribution that gives the total number of ones and then sampling from a distribution that places the locations of the ones and zeros in the positions $(i, \dots, i+k-1)$; in the case where ρ is given by an identical product measure, the second distribution is given by a uniform distribution. Since ρ has a limiting distribution, by a large deviations argument, when k is large, we can control the first distribution, so that we know up to an exponential error the number of ones that do occur, and then it turns out even when ρ is a nonsingular measure, the second distribution can be assumed to be uniform up to Kakutani equivalence. We remark that in the iid case of classical statistics, the first distribution corresponds to a *sufficient statistic* in the sense of Fisher [20], which gives all the necessary information for estimating the parameter given by probability of an occurrence of a single one; in contrast, we are more focused on the secondary uniform distribution, which contains no information about the parameter.

Before we give the details of the proof of Proposition 3.2, we show how it is used to prove Theorem 1.6.

3.3. **Using Proposition 3.2.** We say that an alternating interval of length k_{mark} is ***good*** if its values are in $\mathcal{G}_{k_{\text{mark}}}$. Thus Proposition 3.2 allows us to replace k_{mark} symbols that are asymptotically distributed as $p^{k_{\text{mark}}}$ with a single random variable that is uniformly distributed on a set of size at least $2^{k_{\text{mark}}(H(p)-\epsilon)}$, modulo Kakutani equivalence. In our proof of Theorem 1.6, we will use Keane and Smorodinsky's finitary factor [30] to generate a string of $k_{\text{mark}}+1$ symbols from the uniform random variables, so that at each good alternating interval of length k_{mark} we will have $k_{\text{mark}}+1$ symbols, where the one *extra* symbol can be possibly distributed to integers that do not belong to a good alternating interval; the following lemma will be used to distribute the extra symbol.

Lemma 3.3 (Matching)**.** *Consider a nonsingular Bernoulli shift that satisfies the Doeblin condition and has a limiting measure. Suppose each good alternating interval is assigned the colour green, an alternating interval of size* 1 *assigned the colour red, and an alternating interval of size* k_{mark} *that is not good assigned the colour maroon. For* k_{mark} *chosen sufficiently large, there is an equivariant (non-perfect) matching of green to red and maroon* *intervals, where each red interval has* 1 *green partner, and each maroon interval has* k_{mark} *green partners.*

As in our previous constructions of nonsingular factors [36, 37], we will build upon a variant of the Mešalkin's [43] explicit isomorphism of the stationary Bernoulli shifts $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$, which is adapted from Holroyd and Peres [27].

Proof of Lemma 3.3. The ***Mešalkin*** matching has the following inductive description. Let $W \in \{\text{red}, \text{maroon}, \text{green}\}^{\mathbb{Z}}$ be a random sequence of those colours; here think of W as a colouring of the indexed alternating intervals I, so that W_i is the colour of I_i. In the first instance, if W_n is red or maroon, and W_{n+1} is green, then we match n to $n+1$; that is, a red or maroon integer is matched to a green integer that is to its immediate right. Next, we remove all red and green integers that have 1 partner, and maroon integers that have k_{mark} partners, and repeat. The matching is ***successful*** if all red integers have a green integer partner, and all maroon integers have k_{mark} green partners; the green partners will always be to the right of their red or maroon partners. Note that by definition the resulting matching is equivariant with respect to the left-shift, and when applied to the coloured alternating intervals will also give an equivariant matching. It suffices to show, via an elementary random walk argument, that we can choose k_{mark} sufficiently large so that the excess of green integers compared to red and maroon integers will ensure that the Mešalkin matching is successful almost surely, when applied to the indexed alternating intervals.

Let ρ be a nonsingular Bernoulli shift on A that satisfies the Doeblin condition and has a limiting measure. Let X have law ρ. Without loss of generality assume that the symbols $g, \hat{g}, g' \notin A$. Let $I(X) = (I_m)_{m \in \mathbb{Z}}$ be the alternating intervals. If I_m is of size k_{mark}, then let $J_m = g$ if it is good, $J_m = \hat{g}$ if $X|_{I_m} = a^{k_{\text{mark}}}$, and $J_m = g'$ if it is otherwise not good; if I_m is of size 1, let $J_m = X|_{I_m}$. By Proposition 3.1, conditional on $I(X)$, the sequence J_m is a non-homogeneous Markov chain on $A \cup \{g, \hat{g}, g'\}$ with transitions:

$$\begin{aligned}
p_{gg}(t) &\geq \min_{i\in\mathbb{Z}} \rho_i^{\oplus k_{\mathrm{mark}}}(\mathcal{G}_{k_{\mathrm{mark}}}) := v_{k_{\mathrm{mark}}},\\
p_{g\hat{g}}(t) + p_{gg'}(t) &\leq 1 - v_{k_{\mathrm{mark}}},\\
\sum_{c\in A} p_{\hat{g}c}(t) &= 1,\\
p_{g'g} &\geq v_{k_{\mathrm{mark}}},\\
p_{g'\hat{g}} + p_{g'g'} &\leq 1 - v_{k_{\mathrm{mark}}},\\
p_{aa}(t) &\leq \max_{i\in\mathbb{Z}} \rho_i(a) < 1,\\
p_{ac}(t) &\geq 1 - \max_{i\in\mathbb{Z}} \rho_i(a) > 0 \text{ for all } c \in A\setminus\{a\},\\
p_{cg}(t) &\geq v_{k_{\mathrm{mark}}} \text{ for all } c \in A\setminus\{a\},\\
p_{cg'}(t) + p_{c\hat{g}} &\leq 1 - v_{k_{\mathrm{mark}}} \text{ for all } c \in A\setminus\{a\}.
\end{aligned}$$

Thus only the state g corresponds to a green interval, the other states all correspond to red or maroon intervals.

For all $j < n$, let

$$S_j^n := \sum_{m=j}^{n} \Big(\mathbf{1}[J_m = g] - k_{\mathrm{mark}} \cdot \mathbf{1}[J_m \in \{\hat{g}, g'\}] - \mathbf{1}[J_m \in A]\Big),$$

so that S_j^n is an excess of the difference between green intervals and red intervals with penalty 1, and maroon intervals, with penalty k_{mark}. Recall that by Proposition 3.2 (b), the probability that an alternating interval of size k_{mark} is not good can be made exponentially small and we can replace $v_{k_{\mathrm{mark}}}$, with a term $v'_{k_{\mathrm{mark}}}$, such that $k_{\mathrm{mark}}\big(1 - v'_{k_{\mathrm{mark}}}\big) \to 0$ as $k_{\mathrm{mark}} \to \infty$, at the cost of the inequalities above *failing* for finitely many times $t \in \mathbb{Z}$. A routine variation in a standard probabilistic argument in renewal theory [16] gives that for k_{mark} sufficiently large there exists $0 < C < 1$ such that for any $j \in \mathbb{Z}$, we have

$$\mathbb{P}\Big(\liminf_{n\to\infty} \frac{1}{n-j+1} S_j^n \geq C \mid I(X)\Big) = 1. \tag{4}$$

For each $j \in \mathbb{Z}$, if I_j is red (or maroon) let $Z_j = n$ if I_{n+j} is the (last k_{mark}) matched green interval under the Mešalkin matching; if I_j is green, set $Z_j = 0$, and if Z_j is red or maroon and there are not enough green partners, then set $Z_j = \infty$. Let

$$R_j = \inf\Big\{\ell \geq 1 : S_j^{j+\ell} > 0\Big\}.$$

From the definition of the Mešalkin matching,

$$\mathbb{P}(Z_j > n \mid I(X)) = \mathbb{P}(R_j > n \mid I(X))$$

and the right hand side tends to zero by (4). Hence Z_j is finite almost surely, and Mešalkin matching is successful almost surely. □

Proof of Theorem 1.6. Let $\varepsilon > 0$. Consider the following choice of parameter.

- Let $\varepsilon' = \varepsilon/3$.
- Choose k_{mark} sufficiently large as to satisfy Lemma 3.3 and Proposition 3.2, where in the notation of the proposition, ε is replaced by ε'.
- Furthermore, choose k_{mark} sufficiently large so that

$$H(p) \cdot \frac{k_{\text{mark}}}{k_{\text{mark}} + 1} > H(p) - \varepsilon'.$$

Let X be a nonsingular Bernoulli shift with law ρ. We define and identify markers, alternating intervals, and switches as in Section 3.1. Let $I(X)$ be the alternating intervals. Let $\mathfrak{I} \subset I(X)$ be the good alternating intervals. By Proposition 3.1, it follows that conditioned on $\mathfrak{I}$, the random variables $(X|_J)_{J \in \mathfrak{I}}$ are independent each with law $\rho_J(\cdot|\mathcal{G}_{k_{\text{mark}}})$, given by ρ restricted to J and conditioned to take values in $\mathcal{G}_{k_{\text{mark}}}$.

We apply Proposition 3.2 to associate to each good alternating intervals a single random variable that up to Kakutani equivalence, is uniformly distributed with entropy

$$h > k_{\text{mark}}(H(p) - \varepsilon').$$

Furthermore, we apply Keane and Smorodinsky's finitary factor [30] to replace each uniform random variable by a string of independent symbols from B, with law q, of length $k_{\text{mark}} + 1$, where

$$(k_{\text{mark}} + 1)H(q) > h - \varepsilon'.$$

Thus at each good alternating interval (of length k_{mark}) we have $k_{\text{mark}} + 1$ symbols; the *extra* symbol is distributed via the matching procedure given in Lemma 3.3, so that there is an independent symbol with law q for every integer. We disregard any remaining extra symbols that were not matched, and thus obtain an iid factor of entropy

$$H(q) > \frac{h - \varepsilon'}{k_{\text{mark}} + 1} > (H(p) - \varepsilon') \cdot \frac{k_{\text{mark}}}{k_{\text{mark}} + 1} - \varepsilon' > H(p) - \epsilon.$$

We note that the external components involved in our construction: the Mešalkin matching furnished by Lemma 3.3 and the Keane and Smorodinsky factor are finitary. Hence it follows, by definition that our construction is also finitary. □

3.4. **The proof of Proposition 3.2.** Generalizing our earlier discussion at the end of Section 3.2, we note that the sufficient statistic for a finite number iid observations of a categorical distribution is given the frequency counts of the types. In what follows, we will use the theory of types to help prove Proposition 3.2.

3.4.1. *Types.* The ***empirical probability measure of*** $x \in A^k$, denoted by $\mathrm{emp}(x)$, is the probability measure on A, given by

$$\mathrm{emp}(x)(a) = \frac{1}{k}\sum_{j=1}^{k} \mathbf{1}[x_j = a] = \frac{\#\{j : x_j = a\}}{k}.$$

Given $k \in \mathbb{Z}^+$ we say that $p \in \mathrm{Prob}(A)$ is of ***denominator*** k if $kp(a) \in \mathbb{N}$ for all $a \in A$. The collection of probability distributions of denominator k are precisely the ones which can arise as empirical probability measures for $x \in A^k$. For $p \in \mathrm{Prob}(A)$ of denominator k, let

$$\mathrm{type}_k(p) = \{x \in A^k : \mathrm{emp}(x) = p\} \subset A^k$$

be the k***-type class*** of p, which is all the sequences from A of length k which have an empirical measure that is equal to p.

We will use the following version of Proposition 3.2 where we use conditioning to impose strict control over the types that can occur, so that a large deviations argument is not required.

Proposition 3.4. *Suppose ρ is a nonsingular Bernoulli product measure that satisfies the Doeblin condition. Let $k \in \mathbb{Z}^+$. Let $p \in \mathrm{Prob}(A)$ be of denominator k and $V \subset \mathrm{type}_k(p)$. Recall that $\rho_i^{\oplus k}(\cdot|V)$ is the probability measure on A^k given by taking $(\rho_i, \dots, \rho_{i+k-1})$ conditioned to be on V. Then for all $F \subset V$, we have*

$$\sum_{i\in\mathbb{Z}} \left(\rho_i^{\oplus k}(F|V) - \frac{\#F}{\#V}\right)^2 < \infty. \tag{5}$$

Notice that if ρ is an identical product measure, then Proposition 3.4 is not difficult since each summand is identically zero, see (6). The following lemma, which we will use to prove Proposition 3.4 connects this elementary observation to our nonsingular setting.

Lemma 3.5. *Suppose ρ is a nonsingular Bernoulli product measure that satisfies the Doeblin condition. For all $c \in A^k$, we have*

$$\sum_{i\in\mathbb{Z}} \left(\rho_i^{\oplus k}(c) - \rho_i^{\otimes k}(c)\right)^2 < \infty.$$

Proof. For all $c = (c_0, \dots, c_{k-1}) \in A^k$, we have

$$\begin{aligned}\rho_i^{\oplus k}(c) - \rho_i^{\otimes k}(c) &= \rho_i(c_0)\left(\prod_{j=1}^{k-1}\rho_{i+j}(c_j) - \prod_{j=1}^{k-1}\rho_i(c_j)\right)\\ &= \rho_i(c_0)\sum_{j=1}^{k-1}\left(\frac{\prod_{\ell=1}^{j}\rho_{i+\ell}(c_\ell)}{\rho_{i+j}(c_j)}\right)(\rho_{i+j}(c_j) - \rho_i(c_j))\left(\frac{\prod_{\ell=j}^{k-1}\rho_i(c_\ell)}{\rho_i(c_j)}\right).\end{aligned}$$

As ρ satisfies the Doeblin condition, there exists $C > 0$ such that for all $1 \leq j \leq k-1$, we have

$$\rho_i(c_0)\left(\frac{\prod_{\ell=1}^{j}\rho_{i+\ell}(c_\ell)}{\rho_{i+j}(c_j)}\right)|\rho_{i+j}(c_j)-\rho_i(c_j)|\left(\frac{\prod_{\ell=j}^{k-1}\rho_i(c_\ell)}{\rho_i(c_j)}\right) \\ \leq C\,|\rho_{i+j}(c_j)-\rho_i(c_j)|\,.$$

Consequently for all $i \in \mathbb{Z}$ and $c \in A^k$, we have

$$\left(\rho_i^{\oplus k}(c)-\rho_i^{\otimes k}(c)\right)^2 \leq C^2(k-1)^2\sum_{j=1}^{k-1}\left(\rho_{i+j}(c_j)-\rho_i(c_j)\right)^2.$$

As ρ is nonsingular and satisfies the Doeblin condition, Kakutani's theorem implies that for all $m \geq 1$, we have

$$A_m := \sum_{i\in\mathbb{Z}}\sum_{j=1}^{m}\sum_{x\in A}\left(\rho_{i+j}(x)-\rho_i(x)\right)^2 < \infty.$$

Hence for all $c \in A^k$, we have

$$\sum_{i\in\mathbb{Z}}\left(\rho_i(c)-\rho_i^{\otimes k}(c)\right)^2 \leq C(k-1)^2\sum_{i\in\mathbb{Z}}\sum_{j=1}^{k-1}\left(\rho_{i+j}(c_j)-\rho_i(c_j)\right)^2 \\ \leq C(k-1)^2A_{k-1} < \infty. \qquad \square$$

Proof of Proposition 3.4. Fix $V \subset \mathrm{type}(p)$ and $F \subset V$. A nice observation that is used in [25] is that for all $c, d \in \mathrm{type}_k(p)$, we have

$$\rho_i^{\otimes k}(c) = \rho_i^{\otimes k}(d).$$

Consequently for all $i \in \mathbb{Z}$, and $c \in V$, we have

$$\rho_i^{\otimes k}(c|V) = \frac{\rho_i^{\otimes k}(c)}{\rho_i^{\otimes k}(V)} = \frac{1}{\#V}. \tag{6}$$

Since V is a finite set it follows from Lemma 3.5 that

$$\sum_{i\in\mathbb{Z}}\left(\rho_i^{\oplus k}(V)-\rho_i^{\otimes k}(V)\right)^2 = \sum_{i\in\mathbb{Z}}\left(\sum_{c\in V}\rho_i^{\oplus k}(c)-\rho_i^{\otimes k}(c)\right)^2 \\ \leq (\#V)^2\sum_{c\in V}\sum_{i\in\mathbb{Z}}\left(\rho_i^{\oplus k}(c)-\rho_i^{\otimes k}(c)\right)^2 < \infty. \tag{7}$$

Fix $c \in V$. By the Doeblin condition for ρ, there exists $D > 0$ such that for all $i \in \mathbb{Z}$, we have

$$\begin{aligned}\left[\rho_i^{\oplus k}(c|V) - \rho_i^{\otimes k}(c|V)\right]^2 &= \\ &\left(\rho_i^{\oplus k}(c)\left(\frac{1}{\rho_i^{\oplus}(V)} - \frac{1}{\rho_i^{\otimes k}(V)}\right) + \frac{\rho_i^{\oplus k}(c) - \rho_i^{\otimes k}(c)}{\rho_i^{\otimes k}(V)}\right)^2 \\ &\leq 4D\left[\left(\rho_i^{\oplus k}(V) - \rho_i^{\otimes k}(V)\right)^2 + \left(\rho_i^{\oplus k}(c) - \rho_i^{\otimes k}(c)\right)^2\right].\end{aligned}$$

Hence from (7) and (6) we have,

$$\sum_{i\in\mathbb{Z}}\left(\rho_i^{\oplus k}(c|V) - \frac{1}{\#V}\right)^2 < \infty,$$

from which the desired result is immediate:

$$\begin{aligned}\sum_{i\in\mathbb{Z}}\left(\rho_i^{\oplus k}(F|V) - \frac{\#F}{\#V}\right)^2 &= \sum_{i\in\mathbb{Z}}\left(\sum_{c\in F}\left(\rho_i^{\oplus k}(c|V) - \frac{1}{\#V}\right)\right)^2 \\ &\leq (\#V)^2\sum_{c\in F}\sum_{i\in\mathbb{Z}}\left(\rho_i^{\oplus k}(c|V) - \frac{1}{\#V}\right)^2 < \infty.\end{aligned}$$

□

3.4.2. *Sanov's theorem and the set $\mathcal{G}_k$ in Proposition 3.2.* In order to use Proposition 3.4, we will need to introduce some results from large deviations theory. Let $\delta > 0$. Recall that the ***total variation distance*** between $q_1, q_2 \in \mathrm{Prob}(A)$ is defined by

$$\mathrm{d_{TV}}(q_1, q_2) := \sum_{a\in A}|q_1(a) - q_2(a)|\,; \tag{8}$$

we will endow $\mathrm{Prob}(A)$ with this metric. Recall that $p = \lim_{|n|\to\infty}\rho_n$. For $k \in \mathbb{N}$, let

$$U(\delta) := \{q \in \mathrm{Prob}(A) : \mathrm{d_{TV}}(q, p) < \delta\} \tag{9}$$

and

$$\hat{\mathcal{G}}_{k,\delta} := \left\{x \in A^k : \mathrm{d_{TV}}(\mathrm{emp}(x), p) < \delta\right\} = \bigcup_{q\in U(\delta)}\mathrm{type}_k(q).$$

Recall that the ***Kullback-Leibler divergence*** between $p, q \in \mathrm{Prob}(A)$ is defined by

$$\mathrm{D_{KL}}(q||p) := \sum_{a\in A}p(a)\log\left(\frac{p(a)}{q(a)}\right).$$

The function $q \mapsto D(q||p)$ is a continuous function from

$$\{q \in \mathrm{Prob}(A) : q \ll p\} \text{ to } [0, \infty),$$

where $q \ll p$ means that q is absolutely continuous with respect to p.

The following is an adaptation of standard results on concentration of measure and well-known bounds of the method of types; see for example [13, Chapter 11].

Lemma 3.6. *For every $\varepsilon > 0$ there exists $\delta > 0$, and a positive integer $k_0 \in \mathbb{Z}^+$ and $\beta > 0$ such that for all $k > k_0$:*

(a) If $\mathrm{d_{TV}}(q,p) < \delta$ and q is of denominator k, then $\#\mathrm{type}_k(q) \geq 2^{\left(H(p)-\frac{\varepsilon}{2}\right)k}$.

(b) For all but finitely many $n \in \mathbb{Z}$, we have $\rho_n^{\oplus k}\left(\hat{\mathcal{G}}_{k,\delta}\right) \geq 1 - e^{-\beta k}$.

Proof. By [13, Theorem 11.1.3], for all $q \in \mathrm{Prob}(A)$ and $k \in \mathbb{Z}^+$, if q has denominator k, then

$$\#\mathrm{type}_k(q) \geq \frac{1}{(k+1)^{|A|}} 2^{kH(q)}.$$

Since the entropy map $q \mapsto H(q)$ is continuous, there exists $\delta > 0$ such that if $\mathrm{d_{TV}}(q,p) < \delta$, then $H(q) > H(p) - \frac{\varepsilon}{3}$. Let k_1 be such that if $k > k_1$, then $(k+1)^{|A|} < e^{\frac{\varepsilon}{6}k}$. Hence, if $k > k_1$ and $\mathrm{d_{TV}}(q,p) < \delta$, then

$$\#\mathrm{type}_k(q) \geq 2^{\left(H(p)-\frac{\varepsilon}{2}\right)k},$$

establishing part (a).

The set $K := \mathrm{Prob}(A) \setminus U(\delta)$ is compact so that

$$2\beta := \min_{q \in K} \mathrm{D_{KL}}(q||p) > 0.$$

By Sanov's theorem [58], for every $k \in \mathbb{Z}^+$, we have

$$p^{\otimes k}\left(A^k \setminus \hat{\mathcal{G}}_{k,\delta}\right) = p^{\otimes k}\left(x \in A^k : \ \mathrm{emp}(x) \in K\right) \leq (k+1)^{|A|} e^{-2k\beta}.$$

Choose a positive integer $k_0 \geq k_1$ such that for all $k \geq k_0$, we have

$$p^{\otimes k}\left(A^k \setminus \hat{\mathcal{G}}_{k,\delta}\right) \leq \frac{e^{-\beta k}}{2}. \tag{10}$$

Let $C := \min_{a \in A} p(a)$. For every $n \in \mathbb{Z}$, and every $k \geq k_0$ and $x \in A^k$, we have

$$\begin{aligned}\rho_n^{\oplus k}(x) = \prod_{j=n}^{n+k-1} \rho_j(x_j) &= \prod_{j=n}^{n+(k-1)} p(x_j)\left(1 + \frac{\rho_j(x_j) - p(x_j)}{p(x_j)}\right) \\ &\leq p^{\otimes k}(x) \prod_{j=n}^{n+k-1}\left(1 + \frac{\rho_j(x_j) - p(x_j)}{C}\right).\end{aligned}$$

Since

$$\lim_{|n|\to\infty} \max_{x \in A^k} \prod_{j=n}^{n+k-1}\left(1 + \frac{\rho_j(x_j) - p(x_j)}{C}\right) = 1,$$

for all but finitely many $n \in \mathbb{Z}$, we have

$$\max_{x \in A}\left(\rho_n^{\oplus k}(x)/p^{\otimes k}(x)\right) \leq 2. \tag{11}$$

Hence with (10), we obtain that for all but finitely many $n \in \mathbb{Z}$, we have

$$\rho_n^{\oplus k}\left(A^k \setminus \hat{\mathcal{G}}_{k,\delta}\right) \leq 2p^{\otimes k}\left(A^k \setminus \hat{\mathcal{G}}_{k,\delta}\right) \leq e^{-\beta k}.$$

as desired for part (b). □

We now combine Proposition 3.4 with Lemma 3.6 to obtain Proposition 3.2. Recall the role of Proposition 3.2 in the proof of Theorem 1.6 was to extract, using a single procedure, (up to Kakutani equivalence) an iid sequence of sufficiently high entropy (discrete uniform) random variables from the good alternating intervals. Lemma 3.6 gives that for k sufficiently large most k-type classes will be large enough so that the random variable corresponding to picking an element from such a type class will have sufficiently high entropy. Although there is a uniform lower bound on the size of the type classes, they vary. Part of our proof of Proposition 3.2 will involve *chopping* a type class up into equal sets of the desired size; exerting further control over the sizes helps to prove that the resulting random variables are stationary. See the proof below for details.

Proof of Proposition 3.2. Let $\varepsilon > 0$. Recall that p is the limiting measure. Choose δ and k_0 as in Lemma 3.6 and $k_{\text{mark}} := \max\left(k_0, \lceil \frac{2}{\varepsilon} \rceil\right)$. For all $k \geq k_{\text{mark}}$ set $B_k = \left\{1, \ldots, 2^{\lceil k(H(p)-\varepsilon)\rceil}\right\}$. We will first construct $\mathcal{G}_k$, and then define the mapping $\psi : \mathcal{G}_k \to B_k$. Set

$$U_k'(\delta) := \{q \in U(\delta) : q \text{ is of denominator } k \text{ and } \text{type}_k(q) \neq \emptyset\},$$

where $U(\delta)$ is as given in (9). Let $q \in U_k'(\delta) \subset \text{Prob}(A)$. By Lemma 3.6, since $k \geq \frac{2}{\epsilon}$, we have that $\#\text{type}_k(q) > \#B_k$. Set

$$m(q,k) := \left\lfloor \tfrac{\#\text{type}_k(q)}{\#B_k} \right\rfloor.$$

For each $q \in U_k'(\delta)$, let $F_k(q)$ be a fixed subset of $\text{type}_k(q)$ of cardinality $m(q,k)(\#B_k)$; here we can think of *chopping* up the set $\#\text{type}_k(q)$ into $m(q,k)$ portions, stacked upon each other, and discarding away the rest. Let

$$\mathcal{G}_k := \biguplus_{q \in U_k'(\delta)} F_k(q) \subset \hat{\mathcal{G}}_{k,\delta} \tag{12}$$

be given by a disjoint union.

In order to define $\psi : \mathcal{G}_k \to B_k$, note that since for all q in the disjoint union (12), the cardinality of $F_k(q)$ is an integer multiple of the cardinality of B_k, we choose a $m(q,k)$-to-1 mapping $\psi|_{F_k(q)} : F_k(q) \to B_k$ such that for all $c \in B_k$, we have

$$\frac{\#\left\{x \in F_k(q) : \ \psi(x) = c\right\}}{\#F_k(q)} = \frac{1}{\#B_k}; \tag{13}$$

putting together these choices, we obtain the desired map ψ.

We now verify the conditions (a),(b), and (c) of the proposition. Condition (a) holds by our choice of B_k.

Now we verify that we did not chop too much away. By construction, for all $q \in U'_k(\delta)$, by Lemma 3.6 (a), we have

$$\frac{p^{\otimes k}\left(\mathrm{type}_k(q) \setminus F_k(q)\right)}{p^{\otimes k}(\mathrm{type}_k(q))} = \frac{\#(\mathrm{type}_k(q) \setminus F_k(q))}{\#\mathrm{type}_k(q)} \\ \leq \frac{\#B_k}{\#\mathrm{type}_k(q)} \ \leq \ e^{-\frac{\varepsilon k}{2}}. \tag{14}$$

In addition, from the proof of Lemma 3.6, specifically inequality (11), for all but finitely many $n \in \mathbb{Z}$, for all $V \subset A^k$, we have

$$\frac{1}{2}p^{\otimes k}(V) \leq \rho_n^{\oplus k}(V) \leq 2p^{\otimes k}(V).$$

Consequently, for all but finitely many $n \in \mathbb{Z}$, we have

$$\frac{\rho_n^{\oplus k}\left(\hat{\mathcal{G}}_{k,\delta} \setminus \mathcal{G}_k\right)}{\rho_n^{\oplus k}\left(\hat{\mathcal{G}}_{k,\delta}\right)} \leq 4 \cdot \frac{p^{\otimes k}\left(\hat{\mathcal{G}}_{k,\delta} \setminus \mathcal{G}_k\right)}{p^{\otimes k}\left(\hat{\mathcal{G}}_{k,\delta}\right)}$$

$$= \frac{4}{p^{\otimes k}\left(\hat{\mathcal{G}}_{k,\delta}\right)} \sum_{\substack{q \in U(\delta) \\ \mathrm{type}_k(q) \neq \emptyset}} \frac{p^{\otimes k}\left(\mathrm{type}_k(q) \setminus F_k(q)\right)}{p^{\otimes k}(\mathrm{type}_k(q))} p^{\otimes k}(\mathrm{type}_k(q)),$$

and by (14), we obtain

$$\leq \frac{4}{p^{\otimes k}\left(\hat{\mathcal{G}}_{k,\delta}\right)} \sum_{\substack{q \in U(\delta) \\ \mathrm{type}_k(q) \neq \emptyset}} e^{-\frac{\varepsilon k}{2}} p^{\otimes k}(\mathrm{type}_k(q))$$

$$\leq 4e^{-\frac{\varepsilon k}{2}}.$$

Hence it follows from Lemma 3.6 (b) that there exists $\beta > 0$ such that for all but finitely many $n \in \mathbb{Z}$, we have

$$\rho_n^{\oplus k}\left(A^k \setminus \mathcal{G}_k\right) = \rho_n^{\oplus k}\left(A^k \setminus \hat{\mathcal{G}}_{k,\delta}\right) + \rho_n^{\oplus k}\left(\hat{\mathcal{G}}_{k,\delta} \setminus \mathcal{G}_k\right) \leq e^{-\beta k} + 4e^{-\frac{\varepsilon k}{2}}.$$

By enlarging k_{mark} if necessary, property (b) in Proposition 3.2 holds for any $\alpha < \min(\beta, \frac{\varepsilon}{2})$.

We will now prove that property (c) holds. Fix $c \in B_k$. Since B_k is a finite set, it suffices to show that

$$\sum_{i \in \mathbb{Z}} \left(\rho_i^{\oplus k}(\psi^{-1}(c) | \mathcal{G}_k) - \frac{1}{\#B_k}\right)^2 < \infty.$$

By Proposition 3.4 for all $q \in U'_k(\delta)$, we have with (13) that

$$\sum_{i \in \mathbb{Z}} \left[\rho_i^{\oplus k}\left(\psi^{-1}(c) \cap F_k(q) | F_k(q)\right) - \frac{1}{\#B_k}\right]^2 < \infty. \tag{15}$$

Recall by (12), by definition,

$$\sum_{q\in U_k'(\delta)} \frac{\rho_i^{\oplus k}\left(F_k(q)\right)}{\rho_i^{\oplus k}(\mathcal{G}_k)} = 1$$

and by [13, Theorem 11.1.1] or elementary counting, we have

$$\#U_k'(\delta) \le (k+1)^{|A|}.$$

Thus for all $c \in B_k$ and $i \in \mathbb{Z}$, we have

$$\begin{aligned}
&\left(\rho_i^{\oplus k}(\psi^{-1}(c)|\mathcal{G}_k) - \frac{1}{\#B_k}\right)^2 \\
&= \left(\sum_{q\in U_k'(\delta)} \frac{\rho_i^{\oplus k}\left(F_k(q)\right)}{\rho_i^{\oplus k}(\mathcal{G}_k)}\left[\rho_i^{\oplus k}\left(\psi^{-1}(\bar{c})\cap F_k(q)|F_k(q)\right) - \frac{1}{\#B_k}\right]\right)^2 \\
&\le (k+1)^{|A|}\sum_{q\in U_k'(\delta)}\left(\frac{\rho_i\left(F_k(q)\right)}{\rho_i(\mathcal{G}_k)}\right)^2\left[\rho_i^{\oplus k}\left(\psi^{-1}(\bar{c})\cap F_k(q)|F_k(q)\right) - \frac{1}{\#B_k}\right]^2 \\
&\le (k+1)^{|A|}\sum_{q\in U_k'(\delta)}\left[\rho_i^{\oplus k}\left(\psi^{-1}(\bar{c})\cap F_k(q)|F_k(q)\right) - \frac{1}{\#B_k}\right]^2.
\end{aligned}$$

Hence summing over both sides of the inequality in the index $i \in \mathbb{Z}$, and applying (15), we obtain condition (c). □

3.5. **The proof of Corollary 1.7.** Let A be a countable set. Recall that a product measure $\rho = \bigotimes_{n\in\mathbb{Z}} \rho_i$ on $A^{\mathbb{Z}}$ has a limiting measure p if ρ_i converges to p in the usual total variation distance; see (8).

Proof of Corollary 1.7. Let ρ be a nonsingular Bernoulli shift on a possibly countable set A that has a limiting measure p.

We start with removing the Doeblin assumption for the case of a finite set A.

$$\text{Assume that for all } a \in A, \text{ we have } p(a) > 0. \tag{16}$$

Since $\lim_{|i|\to\infty} \mathrm{d}_{\mathrm{TV}}(\rho_i, p) = 0$, there exists $N \in \mathbb{Z}^+$ such that for all $|i| \ge N$, we have

$$\rho_i(a) > \frac{1}{2}\min_{a\in A} p(a) = \delta'.$$

Since ρ is nonsingular, for all $i \in \mathbb{Z}$ and $a \in A$, we have $\rho_i(a) > 0$. Hence the Doeblin condition is satisfied with

$$\delta := \min\left(\left(p_i(a) : a \in A, |i| \le N\right), \delta'\right),$$

and Theorem 1.6 applies.

Otherwise set $B = \{a \in A :\ p(a) > 0\}$. If B is a singleton, then $H(p) = 0$ and the Theorem is vacuously true. Suppose that B has more than one element. Fix $b \in B$ and let $\pi : A^{\mathbb{Z}} \to B^{\mathbb{Z}}$ be given by

$$\pi(x)_j = \begin{cases} b, & \text{if } x_j \in \{b\} \biguplus (A \setminus B), \\ x_j, & \text{if } x_j \in B \setminus \{b\}. \end{cases}$$

Since the value of $\pi(x)_j$ just depends on the coordinate x_j, it is easy to see that π is a finitary factor map from ρ to $\nu := \rho \circ \pi^{-1}$; furthermore ν is a product measure on $B^{\mathbb{Z}}$ with limiting marginal $p|_B$, and $H(p|_B) = H(p)$. Thus condition (16) holds for ν, and we already know it has near optimal entropy finitary iid factors, from which it follows by composition with π that same holds for ρ.

Now with the Doeblin condition removed for the case of a finite set A, we consider the countable case, where for concreteness we take $A = \mathbb{N}$. The following cut-off functions will allow us to apply the result for the case of a finite set, established earlier. For each $n \in \mathbb{N}$, consider $\theta_n = \mathbb{N} \to \{0, \ldots, n\}$ given by

$$\theta_n(k) = \theta(k) = \min\{k, n\},$$

and $\Theta^n : \mathbb{N}^{\mathbb{Z}} \to \{0, \ldots, n\}^{\mathbb{Z}}$ given by

$$\Theta^n(x)_j = \Theta(x)_j = \theta(x_j) = \min\{x_j, n\}.$$

Clearly, Θ is a finitary factor map from ρ to $\nu := \rho \circ \Theta^{-1}$ and ν is a product measure on $\{0, \ldots, n\}^{\mathbb{Z}}$ with limiting measure $p \circ \theta_n^{-1}$; moreover, $H(p \circ \theta_n^{-1}) \to H(p)$, as $n \to \infty$. Thus choosing n finite and sufficiently large we again obtain near optimal entropy finitary iid factors by composing Θ with the finitary factor that we obtained in the finite case. □

4. Upper bounds on the entropy of a finitary factor

4.1. **Entropy rates for for bounding the entropy of symbolic factors.**

The next theorem is a mathematical abstraction of Theorem 1.4 that will allow for applications to Anosov diffeomorphisms and Bernoulli shifts.

Let A be a finite set. For $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}^+$, we write for $M < n$ the set,

$$[x]_M^n = \{y \in A^{\mathbb{Z}} : \; y_k = x_k \text{ for all } M \leq k \leq n\}$$

for the unique $[M, n]$-cylinder set containing x. Let μ be a Borel regular measure on $A^{\mathbb{Z}}$ which is nonsingular with respect to the left-shift. It will be convenient to use the language of random variables. For example, recall that if Z is a discrete random variable or vector, then its Shannon entropy is given by

$$H(Z) = -\sum_a \mathbb{P}(Z = a) \log \mathbb{P}(Z = a).$$

Let (n_k) be a subsequence of $\mathbb{Z}^+$ and $X \in A^Z$ be a random variable with law μ, so that $\mathbb{P}(X \in \cdot) = \mu(\cdot)$. The ***lower entropy rate of the measure*** μ ***with respect to*** (n_k) is given by

$$h(\mu, (n_k)) := \liminf_{k \to \infty} \frac{H(X_1, \ldots, X_{n_k})}{n_k}.$$

We recall that in the measure-preserving case, we may take $n_k = k$ and subadditivity ensures the actual limit exists and is the Kolmogorov-Sinai entropy. We will be able to use the lower entropy rate as a substitute for Kolmogorov-Sinai entropy in the nonsingular setting, provided the existence of certain physical proxies for μ, which are akin to SRB measures.

We say measure μ_{ph} on $A^{\mathbb{Z}}$ is ***a mean-physical measure for*** μ ***with respect to*** (n_k), if for every cylinder set $C \subset A^{\mathbb{Z}}$, we have

$$\lim_{k\to\infty} \frac{1}{n_k} \sum_{k=0}^{n_k-1} \mu\left(T^{-k}C\right) = \mu_{\mathrm{ph}}(C).$$

When $n_k = k$ we simply say that μ_{ph} is a ***mean-physical measure for*** μ. We say μ_{ph} is a ***physical measure for*** μ if for μ-almost every $x \in A^{\mathbb{Z}}$, we have

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^n x} \xrightarrow[n\to\infty]{} \mu_{\mathrm{ph}}, \text{ weakly,}$$

where δ_y is the usual point mass giving unit mass to a set containing the point y and zero mass otherwise. The portmanteau theorem [17, Theorem 2.4, page 87] implies that for every cylinder set $C \subset A^{\mathbb{Z}}$ and for μ-almost every $x \in A^{\mathbb{Z}}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_C \circ T^k(x) = \mu_{\mathrm{ph}}(C).$$

It is easy to see that every physical measure is a mean-physical measure, and every mean-physical measure is shift-invariant.

Theorem 4.1. *Let $\left(A^{\mathbb{Z}}, \mathcal{B}, \mu, T\right)$ be a nonsingular symbolic system. If there exists μ_{ph}, a mean-physical measure for μ with respect to a subsequence (n_k), then every iid system that is obtained as a finitary factor of μ has entropy no greater than $h(\mu, (n_k))$, the lower entropy rate of μ with respect to (n_k).*

The following lemma will allow us to transfer the finitary assumption into a form involving finite union of cylinder sets, which will be useful in the proof of Theorem 4.1.

Lemma 4.2. *Let $\left(A^{\mathbb{Z}}, \mathcal{B}, \mu, T\right)$ be a nonsingular symbolic dynamical system, $\pi : A^{\mathbb{Z}} \to B^{\mathbb{Z}}$ a finitary nonsingular factor map from μ to $q^{\mathbb{Z}}$ and $\{B_s : s \in B\}$ be the partition of $B^{\mathbb{Z}}$ according to the zeroth coordinate. For every continuous probability measure ν on $A^{\mathbb{Z}}$ and $\epsilon > 0$, there exists $\{C_s\}_{s\in B}$ and $C_\emptyset$ such that:*

(a) For each $s \in B$, the set $C_s \subset \mathcal{B}$ is a finite union of cylinder sets of $A^{\mathbb{Z}}$; the set $C_\emptyset$ is a complement of a finite union of cylinder sets.
(b) The sets $\{C_s\}_{s\in B\cup\{\emptyset\}}$ form a partition of $A^{\mathbb{Z}}$.
(c) For all $s \in B$, we have $\mu\left(C_s \setminus \pi^{-1}B_s\right) = 0$.
(d) $\mu\left(C_\emptyset \triangle \biguplus_{s\in B}\left(\pi^{-1}B_s \setminus C_s\right)\right) = 0$ and $\nu(C_\emptyset) < \epsilon$.

Proof. As a consequence of the finitary property of π, for every $s \in B$, there exists a sequence $\{D_n(s)\}_{n\in\mathbb{Z}^+}$ of pairwise disjoint cylinder sets such that

$$\mu\Big(\pi^{-1}B_s \triangle \biguplus_{n\in\mathbb{Z}^+} D_n(s)\Big) = 0.$$

Since $\left\{\pi^{-1}B_s\right\}_{s\in S}$ is a partition of $A^{\mathbb{Z}}$ modulo μ, we have

$$\biguplus_{s\in B}\biguplus_{n\in\mathbb{Z}^+} D_n(s) = A^{\mathbb{Z}} \mod \mu.$$

As ν is a continuous probability measure, for N sufficiently large, we have

$$\nu\Big(\biguplus_{s\in B}\biguplus_{n=N+1}^{\infty} D_n(s)\Big) < \epsilon.$$

Set $C_\emptyset := \biguplus_{s\in B}\biguplus_{n=N+1}^{\infty} D_n(s)$. For each $s \in B$, let $C_s := \biguplus_{n=1}^{N} D_n(s)$. The lemma is immediate. □

We will also require some elementary inequalities from information theory. Recall that Fano inequality [19] gives that if Z and Z' are finite-valued (A-valued) random variables, defined on the same probability space and $p_e = \mathbb{P}(Z \neq Z')$, then

$$H(Z|Z') \le H(p_e, 1-p_e) + p_e(\log(\#A) - 1). \tag{17}$$

In our proof of Theorem 4.1, we will use Fano's inequality to compare the entropies of two finite random strings, one of which is an approximation of the other.

Lemma 4.3. *Consider the nonsingular system $\left(A^{\mathbb{Z}}, \mathcal{B}, \mu, T\right)$ and μ_{ph} be a mean-physical measure for μ with respect to the subsequence (n_k). Let $\epsilon > 0$. Consider the set-up and notation of Lemma 4.2, take $\nu = \mu_{\mathrm{ph}}$ and obtain the set $C_\emptyset$ with $\mu_{\mathrm{ph}}(C_\emptyset) < \varepsilon$. Define $\beta : A^{\mathbb{Z}} \to (B \cup \{\emptyset\})^{\mathbb{Z}}$ via*

$$\beta(x)_n = s \text{ if and only if } T^n x \in C_s.$$

Let $X \in A^{\mathbb{Z}}$ be a random variable with law μ. Set

$$p_k := \mathbb{P}(\pi(X)_k \neq \beta(X)_k) = \mu \circ T^{-k}(C_\emptyset).$$

Then for all $n \in \mathbb{Z}^+$, we have

$$H(\pi(X)_1, \dots, \pi(X)_n) \le H(\beta(X)_1, \dots, \beta(X)_n) + \sum_{k=1}^{n} H(p_k, 1-p_k) + (\log(\#B+1) - 1)\sum_{k=1}^{n} p_k.$$

Proof. The proof follows from a routine application of the chain rule for entropy and Fano's inequality. □

Proof of Theorem 4.1. We will continue to use the notation of Lemmas 4.2 and 4.3. Let $\delta > 0$. It is elementary that we may choose $\epsilon > 0$ in Lemma 4.3 so that if $\limsup_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} p_k\right) < \epsilon$, then for all sufficiently large n, we have

$$\frac{1}{n}\sum_{k=1}^{n} H(p_k, 1-p_k) + \frac{(\log(\#B+1)-1)}{n}\sum_{k=1}^{n} p_k < \delta. \tag{18}$$

Since $C_\emptyset$ is a (disjoint) finite union of cylinder sets and μ_{ph} is a mean-physical measure for μ with respect to (n_k), we have

$$\lim_{k\to\infty}\frac{1}{n_k}\sum_{j=1}^{n_k} p_j = \mu_{\text{ph}}(C_\emptyset) < \epsilon.$$

By Lemma 4.3 and (18), for all sufficiently large k, we have

$$\frac{H(\pi(X)_1,\dots,\pi(X)_{n_k})}{n_k} \le \frac{H(\beta(X)_1,\dots,\beta(X)_{n_k})}{n_k} + \delta. \tag{19}$$

Since π is a factor map, and $\mu\circ\pi^{-1} \sim q^{\mathbb{Z}}$, it follows from a variation of the Shannon-McMillan-Breiman theorem [21, Theorem 8] that

$$\lim_{n\to\infty}\frac{H(\pi(X)_1,\dots,\pi(X)_n)}{n} = H(q). \tag{20}$$

Let M be sufficiently large so that each finite union of cylinder sets C_s for $s \in B\cup\{\emptyset\}$ is a finite union of cylinder sets defined on the coordinates $[-M, M]$. Consequently, $(\beta(X)_1,\dots\beta(X)_n)$ is a function of $X_{-M},\dots,X_{n+M}$ and

$$\frac{H(\beta(X)_1,\dots,\beta(X)_{n_k})}{n_k} \le \frac{2M\log(\#B+1)}{n_k} + \frac{H(X_1,\dots,X_{n_k})}{n_k},$$

where we have used the chain rule for entropy together with the fact that $\beta_i(X)$ take at most $\#B+1$ values. Hence

$$\liminf_{k\to\infty}\frac{H(\beta(X)_1,\dots,\beta(X)_{n_k})}{n_k} \le \liminf_{n\to\infty}\frac{H(X_1,\dots,X_{n_k})}{n_k} = h(\mu,(n_k));$$

together with (19) and (20), for an arbitrary $\delta > 0$, we have

$$H(q) = \lim_{k\to\infty}\frac{H(\pi(X)_1,\dots,\pi(X)_{n_k})}{n_k} \le \liminf_{k\to\infty}\frac{H(X_1,\dots,X_{n_k})}{n_k} + \delta. \quad \square$$

4.2. **Bernoulli shifts.** We will prove the following more general version of Theorem 1.4 from which the advertised result is immediate.

For a nonsingular Bernoulli measure ρ, write

$$h_+ = h := \liminf_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} H(\rho_k) \text{ and } h_- = \liminf_{n\to\infty}\frac{1}{n}\sum_{k=-n}^{-1} H(\rho_k).$$

Theorem 4.4. *If $q^{\mathbb{Z}}$ is a finitary factor of a nonsingular Bernoulli shift ρ, then*

$$H(q) \le \min(h_-, h_+).$$

Proof of Theorem 1.4. We have the additional assumption of a limiting measure p, which, with the continuity of H, implies $h_+ = h_- = H(p)$, from which the result is immediate from Theorem 4.4. □

Proof of Theorem 4.4. Let (n_k) be a subsequence such that

$$\lim_{k\to\infty} \frac{1}{n_k} \sum_{j=1}^{n_k} H(\rho_j) = h_+.$$

Consider the sequence of measures on $A^{\mathbb{Z}}$ given by $\zeta_k = \frac{1}{n_k}\sum_{j=1}^{n_k} \rho \circ T^j$. By the Banach-Alaoglu theorem [5] there exists a further subsequence $m_\ell = n_{k_\ell}$ and a probability measure ν on $A^{\mathbb{Z}}$ such that ζ_{m_ℓ} converges to ν weakly. Since cylinder sets are clopen, it follows that ν is a mean-physical measure for ρ with respect to (m_ℓ). By Theorem 4.1, we have

$$H(p) \leq h(\rho, (m_\ell)) = h_+.$$

Finally, we note that a mapping is equivariant with respect to the left-shift T if and only if it is also equivariant with respect to the right-shift T^{-1}. Interchanging T with T^{-1} in the argument gives $H(q) \leq h_-$ and thus $H(q) \leq \min(h_+, h_-)$ as desired. □

4.3. **Anosov diffeomorphisms.** We recall symbolic dynamics for Anosov diffeomorphisms will allow us to employ Theorem 4.1. Let V be a finite set. Given an ***adjacency*** matrix $\mathcal{A} \in \{0,1\}^{V\times V}$, the ***subshift of finite type*** corresponding to $\mathcal{A}$ is defined by

$$\Sigma_{\mathcal{A}} := \left\{ s \in V^{\mathbb{Z}} : \ \forall k \in \mathbb{Z},\ \mathcal{A}_{s_k, s_{k+1}} = 1 \right\},$$

where it is endowed with the usual left-shift T. For more information on symbolic dynamics and shifts of finite type see [41, 44].

Theorem 4.5 (Symbolic dynamics from Sinai [64])**.** *Let $f \in \mathrm{Diff}^2(M)$ be a transitive C^2 Anosov diffeomorphism on a compact manifold M and $\epsilon > 0$. Then there exist a finite set V, an adjacency matrix $\mathcal{A}$, and a covering $\mathcal{R} = \{R_v\}_{v\in V}$ of M by closed sets of diameter less than or equal to ϵ such that*

(a) For distinct $v, v' \in V$, the interiors of R_v and $R_{v'}$ have no intersection.
(b) The coding map $\pi : \Sigma_{\mathcal{A}} \to M$ given by

$$\pi(s) = \bigcap_{n\in\mathbb{Z}} R_{s_n}$$

is continuous, onto and finite-to-one map which is equivariant so that all $s \in \Sigma_{\mathcal{A}}$, we have $f \circ \pi(s) = \pi \circ T(s)$; that is, π is a finite-to-one semi-conjugacy of the topological dynamical systems.
(c) For every $y \notin \bigcup_{n\in\mathbb{Z}} \bigcup_{v\in V} f^n \partial R_v$, the inverse mapping $\pi^{-1}(y)$ is a single point in $\Sigma_{\mathcal{A}}$.

In the context of Theorem 4.5, the sets $\mathcal{R}$ are referred to as a *Markov partition* and the subshift a *topological Markov chain*; see also [3, 10] and [2]. By Theorem 4.5 if an iid system is an almost-surely continuous (finitary) factor of an Anosov diffeomorphism endowed with the natural volume measure, then it will also be a finitary factor of a symbolic system, and making it possible to apply Theorem 4.1. We will need some technical lemmas to deal with the boundary of the Markov partition.

Recall that f has in addition an SRB measure μ_f which is an ergodic f-invariant measure such for all continuous function $\varphi : M \to M$, we have for vol_M-almost every $y \in M$ that

$$\frac{1}{n}\sum_{k=0}^{n-1} \varphi \circ f^k(y) \xrightarrow[n\to\infty]{} \int \varphi d\mu_f.$$

This asymptotic condition has a useful reformulation in terms of weak convergence of measures. For $y \in M$, define a sequence of measures

$$\nu_n^y := \frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k y}.$$

The SRB property gives that vol_M-almost every $y \in M$, the measure ν_n^y converges weakly to μ_f as $n \to \infty$. A point $y \in M$ is a ***generic point*** for μ_f if ν_n^y converges weakly to μ_f. We will use this formulation in the proof of the following lemma.

Lemma 4.6. *Let f be a transitive, C^2 Anosov diffeomorphism on a compact manifold M. Consider symbolic dynamics for f as given in Theorem 4.5. Then* $\mathrm{vol}_M\big(\bigcup_{n\in\mathbb{Z}}\bigcup_{v\in V} f^n\partial R_v\big) = 0$.

We note that in Lemma 4.6, a volume a.c.i.p. may not exist, making the proof a bit harder. Our proof will involve revisiting some technical lemmas and arguments regarding Markov partitions that can be found in [11]. We will show that a point on the boundary fails to be generic with respect to an SRB measure. We thank Yuri Lima for his interest in the lemma and pointing out an important error in an earlier version of its proof.

Proof of Lemma 4.6. It suffices to show that $\partial R = \bigcup_{v\in V} \partial R_v$ has zero volume, since the volume measure is nonsingular, and the union in question is countable.

We recall that by [11, Lemma 3.11] we may express the closed boundary in question as the union $\partial R = \partial R^{\mathrm{u}} \cup \partial R^{\mathrm{s}}$, where these sets are often referred to as the ***unstable*** and ***stable*** boundaries; furthermore from [11, Proposition 3.15], we have containments,

$$f(\partial R^{\mathrm{s}}) \subseteq \partial R^{\mathrm{s}} \text{ and } f^{-1}(\partial R^{\mathrm{u}}) \subseteq \partial R^{\mathrm{u}}.$$

The stable containment implies that each point z in the stable boundary has a forward orbit-closure $\mathrm{orb}^+(z) := \overline{\{f^n(z) : n \in \mathbb{N}\}}$ that will have an empty intersection with some closed neighbourhood of each point that is not in the boundary. Fix a point $y_0 \notin \partial R$. By Urysohn's separation lemma [67, Proposition 2.1.18], for each $z \in \partial R^{\mathrm{s}}$ there exists $\Upsilon_z : M \to [0,1]$ such that $\Upsilon = 0$ on $\mathrm{orb}^+(z)$ and $\Upsilon = 1$ on a closed neighbourhood of y_0. Since the SRB measures are fully supported, we have that $\int \Upsilon_z d\mu_f > 0$. Similarly, the unstable containment implies an analogous statements for the backward orbit-closure for a point on the unstable boundary.

From [65], let μ_f and $\mu_{f^{-1}}$ be SRB measures so that on a subset M' of full volume, we have that for all $y \in M'$ the weak convergences:

$$\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k y} \xrightarrow[n\to\infty]{} \mu_f \quad \text{and} \quad \frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^{-k} y} \xrightarrow[n\to\infty]{} \mu_{f^{-1}}.$$

However by definition of Υ, for $z \in \partial R^{\mathrm{s}}$, for all $n \in \mathbb{Z}^+$ we have

$$\frac{1}{n}\sum_{k=0}^{n-1} \Upsilon_z \circ f^k(z) = 0 < \int \Upsilon_z d\mu_f,$$

and the forward weak convergence fails for z, and it is not a generic point for μ_f. Similarly, if z belongs to the unstable boundary, the backwards weak convergence fails. Hence if z is in the boundary, one of the weak convergences fails, so that $\partial R \subseteq M \setminus M'$, and must have zero volume. □

Remark 4.7. Lemma 4.6 gives that the symbolic coding π in Theorem 4.5 is a nonsingular continuous almost everywhere (finitary) isomorphism between the symbolic dynamics with the left-shift $(\Sigma_{\mathcal{A}}, \mathcal{B}, \eta, T)$ and the Anosov system $(M, \mathcal{B}(M), \mathrm{vol}_M, f)$, where $\eta = \mathrm{vol}_M \circ \pi$. Note that $\Sigma_{\mathcal{A}} \subseteq V^{\mathbb{Z}}$ has unit measure under η.

Notice that an SRB measure μ_f corresponds to a mean-physical measure for η in the symbolic space given by $\eta_{\mathrm{ph}} = \mu_f \circ \pi$. Thus Theorem 4.1 immediately gives that the entropy of any finitary factor is bounded by $h(\eta, (k))$. ◊

Lemma 4.8. *With the notation of Remark 4.7 in force, if the maximum diameter of the Markov partitions is sufficiently small, then*

$$h(\eta, (k)) \leq \min(h_{\mu_f}(f), h_{\mu_{f^{-1}}}(f)).$$

Proof of Theorem 1.3. In Theorem 4.5, one can choose the diameter of the Markov partition to be sufficiently small. Thus the proof is immediate from Remark 4.7 and Lemma 4.8. □

4.3.1. *The proof of Lemma 4.8.* We will execute entropy calculations with the notation of Remark 4.7 and also the Markov partition of Theorem 4.5. These entropy calculations are somewhat more difficult as they involve two measures simultaneously, only one of which is measure-preserving. In this subsection, we will typically use 'x' to denote an element of the symbolic space, and 'y' and 'z' to denote elements of the manifold.

Let $X \in \Sigma_{\mathcal{A}}$ be a random variable with law η. Given $k \in \mathbb{Z}^+$, define $I_k : \Sigma_{\mathcal{A}} \to (0, \infty)$ via

$$I_k(x) = -\log \mathbb{P}(X_0 = x_0 | X_{-1} = x_{-1}, \ldots, X_{-k} = x_{-k}).$$

Also, we recall the standard notation that for a probability space $(\Omega, \mathcal{F}, \zeta)$ and a partition α of Ω, where $\alpha(\omega)$ is the part to which ω belongs, we have

$$H_\zeta(\alpha) = -\sum_{a \in \alpha} \zeta(a) \log \zeta(a) \text{ and } I_\zeta[\alpha](\omega) = -\log \zeta(\alpha(\omega)),$$

so that

$$\int I_\zeta[\alpha] d\zeta = H_\zeta(\alpha).$$

More specifically, for $y \in M$, we have

$$I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}](y) = -\log\left(\frac{\mathrm{vol}\left(\vee_{i=0}^k f^{-i}\mathcal{R}(y)\right)}{\mathrm{vol}\left(\vee_{i=1}^k f^{-i}\mathcal{R}(y)\right)}\right). \tag{21}$$

Lemma 4.9. *Let $k \in \mathbb{Z}^+$. For η-almost every $x \in \Sigma_{\mathcal{A}}$, we have*

$$\frac{1}{n}\sum_{j=0}^{n-1} I_k \circ T^j(x) \xrightarrow[n\to\infty]{} \int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] d\mu_f.$$

Proof. By the definition of $\eta = \mathrm{vol}_M \circ \pi$ as the lift of the volume, for a subset of full η-measure, for all $x \in \Sigma'_{\mathcal{A}}$, we have

$$\Xi_n^x := \frac{1}{n}\sum_{j=0}^{n-1} \delta_{f^j \pi(x)} \xrightarrow[n\to\infty]{} \mu_f, \text{ weakly.}$$

Let $x \in \Sigma'_{\mathcal{A}}$. If $\pi(x) = y \in \bigcap_{i=0}^k f^{-i} R_{x_i}$, then

$$I_k(x) = I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}](\pi(x)) = -\log\left(\frac{\mathrm{vol}\left(\cap_{i=0}^k f^{-i} R_{x_i}\right)}{\mathrm{vol}\left(\cap_{i=1}^k f^{-i} R_{x_i}\right)}\right).$$

We have

$$\frac{1}{n}\sum_{j=0}^{n-1} I_k \circ T^j(x) = \frac{1}{n}\sum_{j=0}^{n-1} I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] \circ f^j(\pi(x)). \tag{22}$$

By Lemma 4.6, the function $I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}]$ is Riemann integrable, since it is discontinuous only on the boundaries of the Markov partition. By the portmanteau theorem,

$$\frac{1}{n}\sum_{j=0}^{n-1} I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] \circ f^j(\pi(x)) = \int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] d\Xi_n^x$$

$$\xrightarrow[n\to\infty]{} \int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] d\mu_f,$$

from which the desired result follows from (22). $\square$

Notice that in Lemma 4.9, we did not obtain convergence to $H_{\mu_f}(\mathcal{R}|\vee_{i=1}^k f^{-i}\mathcal{R})$, but instead ended up with an expression that contains both the volume measure and its SRB measure; in order to replace this expression with one involving only the SRB measure, we will also need to make use of the description of an SRB measure as a Gibbs measure of the geometric potential, and refer to results from Bowen and Ruelle [12]. Let $\varphi^{(\mathrm{u})} : M \to [0,\infty)$ be defined by $\varphi^{(\mathrm{u})}(y) = -\log \lambda(y)$, where $\lambda(y)$ is the Jacobian of the linear map, given by $D_f : E_y^{\mathrm{u}} \to E_{f(y)}^{\mathrm{u}}$.

By [12, Lemma 4.1], the map $\varphi^{(\mathrm{u})}$ is Hölder continuous and by [12, Proposition 4.4], μ_f is the unique equilibrium measure for $\varphi^{(\mathrm{u})}$ and

$$h_{\mu_f}(f) = -\int \varphi^{(\mathrm{u})} d\mu_f. \tag{23}$$

We also recall the *first volume lemma* [12, Lemma 4.2]. Fix a metric d on M. For $\epsilon > 0$, $n \in \mathbb{N}$, and $z \in M$, consider the Bowen ball given by

$$B_z(\epsilon, n) := \Big\{ y \in M : \max_{0 \le j \le n} d(f^j z, f^j y) \le \epsilon \Big\}.$$

Lemma 4.10 (First volume lemma [12]). *Fix a Riemannian metric d on M so that the volume measure* vol *is derived from d. For all small $\epsilon > 0$, there exists $C = C_\epsilon > 1$ such that for all $z \in M$ and $n \in \mathbb{N}$, we have*

$$\frac{1}{C} \exp\left(\sum_{j=0}^n \varphi^{(\mathrm{u})} \circ f^j(z) \right) \le \mathrm{vol}(B_z(\epsilon, n)) \le C \exp\left(\sum_{j=0}^n \varphi^{(\mathrm{u})} \circ f^j(z) \right).$$

In what follows, we will always assume that the maximum diameter of the atoms of the Markov partition, say ϵ, is small enough so that the Lemma 4.10 holds.

Lemma 4.11. *Under the assumption that the maximal diameter of the Markov partition is sufficiently small, we have*

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}] d\mu_f \le h_{\mu_f}(f).$$

Proof. For $k \in \mathbb{N}$, set

$$a_k := -\int \log \left(\frac{\mathrm{vol}\left(\vee_{i=0}^k f^{-i}\mathcal{R}(y)\right)}{\mu_f\left(\vee_{i=0}^k f^{-i}\mathcal{R}(y)\right)} \right) d\mu_f(y),$$

and the version of a_k, where we start the join of the partitions at $i = 1$:

$$b_k := -\int \log \left(\frac{\mathrm{vol}\left(\vee_{i=1}^k f^{-i}\mathcal{R}(y)\right)}{\mu_f\left(\vee_{i=1}^k f^{-i}\mathcal{R}(y)\right)} \right) d\mu_f(y).$$

A key observation is that since μ_f is f-preserving, for all $k \in \mathbb{Z}^+$, we have $b_k = a_{k-1}$; applying this key relation to (21), we have

$$\begin{aligned}\int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}]d\mu_f &= \int I_{\mu_f}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}]d\mu_f + a_k - a_{k-1} \\ &= H_{\mu_f}\left(\mathcal{R}\middle|\vee_{j=1}^k f^{-j}\mathcal{R}\right) + a_k - a_{k-1}.\end{aligned}$$

Consequently for all $n \in \mathbb{Z}^+$, we have

$$\frac{1}{n}\sum_{k=1}^n \int I_{\mathrm{vol}}[\mathcal{R}| \vee_{i=1}^k f^{-i}\mathcal{R}]d\mu_f = \frac{a_n - a_1}{n} + \frac{1}{n}\sum_{k=1}^n H_{\mu_f}\left(\mathcal{R}\middle|\vee_{j=1}^k f^{-j}\mathcal{R}\right).$$

Since μ_f is f-invariant and $\mathcal{R}$ is a generating partition,

$$\lim_{k\to\infty} H_{\mu_f}\left(\mathcal{R}\middle|\vee_{j=1}^k f^{-j}\mathcal{R}\right) = h_{\mu_f}(f).$$

Therefore the second term on the right hand side converges to $h_{\mu_f}(f)$ as $n \to \infty$. Hence it remains to show that

$$a := \liminf_{n\to\infty} \frac{a_n}{n} \le 0.$$

Let ϵ be the maximal diameter of the atoms in the Markov partition so that for all $y \in M$ and $n \in \mathbb{N}$, we have the inclusion

$$\vee_{i=0}^n f^{-i}\mathcal{R}(y) \subset B_y(\epsilon, n).$$

This inclusion together with Lemma 4.10 imply that

$$\begin{aligned}\frac{1}{n}\int \log \mathrm{vol}\left(\vee_{i=0}^n f^{-i}\mathcal{R}(y)\right) d\mu_f(y) &\le \frac{1}{n}\int \log \mathrm{vol}\left(B_y(\epsilon, n)\right) d\mu_f(y) \\ &\le \frac{\log C_\epsilon}{n} + \frac{1}{n}\int \left(\sum_{j=1}^n \varphi^{(\mathrm{u})} \circ f^j(y)\right) \mu_f(y) \\ &= \frac{\log C_\epsilon}{n} + \int \varphi^{(\mathrm{u})}\mu_f \\ &= \frac{\log C_\epsilon}{n} - h_{\mu_f}(f),\end{aligned}$$

where the first equality uses that μ_f is f-preserving, and the last equality is from (23).

Since $\mathcal{R}$ is a generating partition, we have

$$\begin{aligned}h_{\mu_f}(f) &= \lim_{n\to\infty} \frac{1}{n} H_{\mu_f}\left(\vee_{i=0}^n f^{-i}\mathcal{R}\right) \\ &= -\lim_{n\to\infty} \frac{1}{n}\int \log \mu_f\left(\vee_{i=0}^n f^{-i}\mathcal{R}(y)\right) d\mu_f(y).\end{aligned}$$

Hence

$$\begin{aligned}
a &= -\limsup_{n\to\infty}\frac{1}{n}\left(\int \log \mathrm{vol}\left(\vee_{i=0}^{n} f^{-i}\mathcal{R}(y)\right)d\mu_f + H_{\mu_f}\left(\vee_{i=0}^{n} f^{-i}\mathcal{R}\right)\right)\\
&\leq -\limsup_{n\to\infty}\left(\frac{\log C_\epsilon}{n} - h(\mu_f) + \frac{1}{n}H_{\mu_f}\left(\vee_{i=0}^{n} f^{-i}\mathcal{R}\right)\right)\\
&\leq 0. \qquad \square
\end{aligned}$$

Proof of Lemma 4.8. Let X be a random variable with law η. By the chain rule for entropy, for every $n \in \mathbb{Z}^+$, we have

$$\begin{aligned}
\frac{1}{n}H(X_1,\ldots,X_n) &= \frac{1}{n}H(X_1) + \frac{1}{n}\sum_{j=2}^{n} H(X_j|X_{j-1},\ldots X_1)\\
&\leq \frac{1}{n}\sum_{j=1}^{k} H(X_j) + \frac{1}{n}\sum_{j=k+1}^{n} H(X_j|X_{j-1},\ldots X_{j-k})\\
&\leq \frac{1}{n}k\log(\#V) + \frac{1}{n}\sum_{j=k+1}^{n} H(X_j|X_{j-1},\ldots X_{j-k}). \qquad (24)
\end{aligned}$$

Here, the second inequality uses that entropy can only decrease under further conditioning and for the last one we recall that X_j takes values on the set V values. For every $k < j \leq n$, we have

$$H(X_j|X_{j-1},\ldots X_{j-k}) = \int I_k \circ T^j d\eta,$$

and

$$\frac{1}{n}\sum_{j=k+1}^{n} H(X_j|X_{j-1},\ldots X_{j-k}) = \int\left(\frac{1}{n}\sum_{j=k+1}^{n} I_k \circ T^j(x)\right)d\eta(x).$$

The bounded convergence theorem and Lemma 4.9 give that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=k+1}^{n} H(X_j|X_{j-1},\ldots X_{j-k}) = \int I_{\mathrm{vol}}[\mathcal{R}|\vee_{i=1}^{k} f^{-i}\mathcal{R}]d\mu_f.$$

Hence from (24), for every $k \in \mathbb{Z}^+$, we have

$$\limsup_{n\to\infty}\frac{1}{n}H(X_1,\ldots,X_n) \leq \int I_{\mathrm{vol}}[\mathcal{R}|\vee_{i=1}^{k} f^{-i}\mathcal{R}]d\mu_f.$$

Moreover, summing over the index k in the inequality above, by Lemma 4.11 we have

$$\limsup_{n\to\infty}\frac{1}{n}H(X_1,\ldots,X_n) \leq \liminf_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\int I_{\mathrm{vol}}[\mathcal{R}|\vee_{i=1}^{k} f^{-i}\mathcal{R}]d\mu_f \leq h_{\mu_f}(f).$$

Thus we have $h(\eta,(k)) \leq h_{\mu_f}(f)$; applying a similar argument involving T^{-1} and f^{-1}, we also obtain that $h(\eta,(k)) \leq h_{\mu_{f^{-1}}}(f^{-1}) = h_{\mu_{f^{-1}}}(f)$. $\square$

5. A dissipative Bernoulli shift with no finitary factors

Recall that Kakutani's theorem [28] gives that two infinite direct product measures μ and ν on $A^{\mathbb{Z}}$ are either equivalent or mutually singular, and they are equivalent if and only if

$$\sum_{n\in\mathbb{Z}}\sum_{a\in A}(\sqrt{\mu_n(a)}-\sqrt{\nu_n(a)})^2<\infty.$$

Lemma 5.1. *Let* $\mu=\bigotimes_{n\in\mathbb{Z}}\mu_i$ *be the product measure on* $\{0,1\}^{\mathbb{Z}}$ *with marginals given by*

$$\mu_i(0)=\frac{10}{\sqrt{|i|+2}}=1-\mu_i(1).$$

The associated Bernoulli shift is nonsingular and totally dissipative action of a non-atomic measure space.

The proof will require some calculations; in particular, to show that the shift is dissipative we will apply a sufficient condition from [35, Lemma 2.2], which requires verifying that

$$\sum_{n\in\mathbb{Z}}\int\sqrt{\frac{d\mu\circ T^n}{d\mu}}d\mu<\infty. \tag{25}$$

Proof of Lemma 5.1. Since

$$\sum_{n\in\mathbb{Z}}\min(\mu_n(0),\mu_n(1))=\sum_{n\in\mathbb{Z}}\frac{1}{\sqrt{1+|n|}}=\infty,$$

it follows that the measure μ is non-atomic; see [6]. The shift is nonsingular because

$$\begin{aligned}&\sum_{n\in\mathbb{Z}}\left(\left(\sqrt{\mu_n(0)}-\sqrt{\mu_{n-1}(0)}\right)^2+\left(\sqrt{\mu_n(1)}-\sqrt{\mu_{n-1}(1)}\right)^2\right)\\ &\leq 20\sum_{n\in\mathbb{Z}}\left(\frac{\sqrt[4]{|n|+2}-\sqrt[4]{|n-1|+2}}{(\sqrt[4]{|n|+2})\cdot(\sqrt[4]{|n-1|+2})}\right)^2<\infty.\end{aligned}$$

It remains to show that the shift is dissipative. By Kakutani's theorem, for all $n\in\mathbb{Z}^+$, for μ almost every $x\in\{0,1\}^{\mathbb{Z}}$, we have

$$\frac{d\mu\circ T^n}{d\mu}(x)=\prod_{k\in\mathbb{Z}}\frac{\mu_{k-n}(x_k)}{\mu_k(x_k)}.$$

As in [69], since for all $0<a,b<1$, we have

$$\sqrt{ab}+\sqrt{(1-a)(1-b)}\leq 1-\frac{(b-a)^2}{2},$$

and as μ is a product measure, for all $n \in \mathbb{Z}^+$, we have

$$\begin{aligned}\int \sqrt{\frac{d\mu \circ T^n}{d\mu}} d\mu &= \prod_{k\in\mathbb{Z}} \int \sqrt{\frac{\mu_{k-n}(s)}{\mu_k(s)}} d\mu_k(s) \\ &= \prod_{k\in\mathbb{Z}} \left(\sqrt{\mu_k(0)\mu_{k-n}(0)} + \sqrt{\mu_k(1)\mu_{k-n}(1)} \right) \\ &\leq \prod_{k\in\mathbb{Z}} \left(1 - \frac{(\mu_k(0) - \mu_{k-n}(0))^2}{2} \right) \\ &\leq \exp(-\tfrac{1}{2} \sum_{k\in\mathbb{Z}} (\mu_k(0) - \mu_{k-n}(0))^2).\end{aligned}$$

For all sufficiently large $n \in \mathbb{Z}^+$ and $0 \leq k \leq n$, we have

$$\frac{n}{(\sqrt{2+k+n})(\sqrt{2+k+n} + \sqrt{2+k})} \geq \frac{1}{5}.$$

Consequently,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \left(\mu_k(0) - \mu_{k-n}(0)\right)^2 &\geq \sum_{k=n+1}^{2n} (\mu_k(0) - \mu_{k-n}(0))^2 \\ &= \sum_{k=0}^{n} \left(\frac{10}{\sqrt{k+2}} - \frac{10}{\sqrt{2+k+n}} \right)^2 \\ &= \sum_{k=1}^{n} \frac{1}{2+k} \left(\frac{10n}{(\sqrt{2+k+n})(\sqrt{2+k+n} + \sqrt{2+k})} \right)^2 \\ &\geq \sum_{k=1}^{n} \frac{4}{2+k} = 4(1+o(1)) \log(1+n).\end{aligned}$$

Hence we have for all n sufficiently large,

$$\int \sqrt{\frac{d\mu \circ T^n}{d\mu}} d\mu \leq \exp\left(-\tfrac{3}{2} \log(n+1)\right) = \frac{1}{(n+1)^{\frac{3}{2}}},$$

a summable p-series. □

Proof of Proposition 1.9. Immediate from Theorems 1.1 and 1.4, and Lemma 5.1. □

Acknowledgements

Terry Soo thanks Anthony Quas for teaching him Kalikow's argument that the 3-shift is not a factor of the 2-shift, over dinner, when he was his postdoctoral fellow in Victoria. We also thank Yuri Lima, Benjamin Weiss, Omri Sarig, and Jon Aaronson for many enlightening discussions, and Ronnie Kosloff for explaining the relation of Theorem 1.1 to Prigogine's work.

References

[1] J. Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[2] R. L. Adler, A. G. Konheim, and M. H. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.*, 114:309–319, 1965.

[3] R. L. Adler and B. Weiss. *Similarity of automorphisms of the torus*. Memoirs of the American Mathematical Society, No. 98. American Mathematical Society, Providence, R.I., 1970.

[4] M. A. Akcoglu, A. del Junco, and M. Rahe. Finitary codes between Markov processes. *Z. Wahrsch. Verw. Gebiete*, 47(3):305–314, 1979.

[5] L. Alaoglu. Weak topologies of normed linear spaces. *Ann. of Math. (2)*, 41:252–267, 1940.

[6] D. Aldous and J. Pitman. On the zero-one law for exchangeable events. *Ann. Probab.*, 7(4):704–723, 1979.

[7] D. V. Anosov. Geodesic flows on closed Riemannian manifolds of negative curvature. *Trudy Mat. Inst. Steklov.*, 90:209, 1967.

[8] K. Ball. Monotone factors of i.i.d. processes. *Israel J. Math.*, 150:205–227, 2005.

[9] M. Björklund, Z. Kosloff, and S. Vaes. Ergodicity and type of nonsingular Bernoulli actions. *Invent. Math.*, 224(2):573–625, 2021.

[10] R. Bowen. Markov partitions for Axiom A diffeomorphisms. *Amer. J. Math.*, 92:725–747, 1970.

[11] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, revised edition, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.

[12] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29(3):181–202, 1975.

[13] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[14] A. I. Danilenko and M. Lemańczyk. K-property for Maharam extensions of non-singular Bernoulli and Markov shifts. *Ergodic Theory Dynam. Systems*, 39(12):3292–3321, 2019.

[15] M. Denker and M. Keane. Almost topological dynamical systems. *Israel J. Math.*, 34(1-2):139–160 (1980), 1979.

[16] J. L. Doob. Renewal theory from the point of view of the theory of probability. *Trans. Amer. Math. Soc.*, 63:422–438, 1948.

[17] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, second edition, 1995.

[18] P. Elias. The Efficient Construction of an Unbiased Random Sequence. *Ann. Math. Statist.*, 43(3):865 – 870, 1972.

[19] R. M. Fano. *Transmission of information: A statistical theory of communications*. The M.I.T. Press, Cambridge, Mass.; John Wiley & Sons, Inc., New York-London, 1961.

[20] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222(602):309–368, 1922.

[21] R. M. Gray and J. C. Kieffer. Asymptotically mean stationary measures. *Ann. Probab.*, 8(5):962–973, 1980.

[22] B. M. Gurevič. A certain condition for the existence of a K-decomposition for a special flow. *Uspehi Mat. Nauk*, 24(5 (149)):233–234, 1969.

[23] B. M. Gurevič and V. I. Oseledec. Gibbs distributions, and the dissipativity of C-diffeomorphisms. *Dokl. Akad. Nauk SSSR*, 209:1021–1023, 1973.

[24] T. Hamachi. On a Bernoulli shift with nonidentical factor measures. *Ergodic Theory Dynam. Systems*, 1(3):273–283 (1982), 1981.

[25] N. Harvey, A. E. Holroyd, Y. Peres, and D. Romik. Universal finitary codes with exponential tails. *Proc. Lond. Math. Soc. (3)*, 94(2):475–496, 2007.
[26] H. Holden and R. Piene, editors. *The Abel Prize 2013–2017*. Springer, Cham, 2019.
[27] A. E. Holroyd and Y. Peres. Extra heads and invariant allocations. *Ann. Probab.*, 33(1):31–52, 2005.
[28] S. Kakutani. On equivalence of infinite product measures. *Ann. of Math. (2)*, 49:214–224, 1948.
[29] S. Kalikow and R. McCutcheon. *An outline of ergodic theory*, volume 122 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
[30] M. Keane and M. Smorodinsky. A class of finitary codes. *Israel J. Math.*, 26:352–371, 1977.
[31] M. Keane and M. Smorodinsky. Bernoulli schemes of the same entropy are finitarily isomorphic. *Ann. of Math. (2)*, 109:397–406, 1979.
[32] M. Keane and M. Smorodinsky. Finitary isomorphisms of irreducible Markov shifts. *Israel J. Math.*, 34(4):281–286 (1980), 1979.
[33] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk SSSR (N.S.)*, 119:861–864, 1958.
[34] A. N. Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR*, 124:754–755, 1959.
[35] Z. Kosloff. On the K property for Maharam extensions of Bernoulli shifts and a question of Krengel. *Israel J. Math.*, 199(1):485–506, 2014.
[36] Z. Kosloff and T. Soo. The orbital equivalence of Bernoulli actions and their Sinai factors. *J. Mod. Dyn.*, 17:145–182, 2021.
[37] Z. Kosloff and T. Soo. Some factors of nonsingular Bernoulli shifts. *Studia Math.*, 262(1):23–43, 2022.
[38] U. Krengel. Darstellungssätze für Strömungen und Halbströmungen. II. *Math. Ann.*, 182:1–39, 1969.
[39] U. Krengel. Transformations without finite invariant measure have finite strong generators. In *Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus, Ohio, 1970)*, pages 133–157. Springer, Berlin, 1970.
[40] P. L'Ecuyer. History of uniform random number generation. In *2017 Winter Simulation Conference (WSC)*, pages 202–230, 2017.
[41] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.
[42] A. N. Livšic and J. G. Sinaĭ. Invariant measures that are compatible with smoothness for transitive C-systems. *Dokl. Akad. Nauk SSSR*, 207:1039–1041, 1972.
[43] L. D. Mešalkin. A case of isomorphism of Bernoulli schemes. *Dokl. Akad. Nauk SSSR*, 128:41–44, 1959.
[44] M. Morse and G. A. Hedlund. Symbolic Dynamics. *Amer. J. Math.*, 60(4):815–866, 1938.
[45] D. Ornstein. Newton's laws and coin tossing. *Notices Amer. Math. Soc.*, 60(4):450–459, 2013.
[46] D. S. Ornstein. The isomorphism theorem for Bernoulli flows. *Advances in Math.*, 10:124–142, 1973.
[47] D. S. Ornstein and P. C. Shields. Mixing Markov shifts of kernel type are Bernoulli. *Advances in Math.*, 10:143–146, 1973.
[48] D. S. Ornstein and B. Weiss. Geodesic flows are Bernoullian. *Israel J. Math.*, 14:184–198, 1973.
[49] D. S. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987.
[50] D. S. Ornstein and B. Weiss. Statistical properties of chaotic systems. *Bull. Amer. Math. Soc. (N.S.)*, 24(1):11–116, 1991. With an appendix by David Fried.

[51] Y. Peres. Iterating von Neumann's procedure for extracting random bits. *Ann. Statist.*, 20(1):590–597, 1992.

[52] K. Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. Corrected reprint of the 1983 original.

[53] I. Prigogine. Time, structure, and fluctuations. *Science*, 201(4358):777–785, 1978.

[54] A. Quas and T. Soo. A monotone Sinai theorem. *Ann. Probab.*, 44(1):107–130, 2016.

[55] M. Ratner. Anosov flows with Gibbs measures are also Bernoullian. *Israel J. Math.*, 17:380–391, 1974.

[56] V. A. Rohlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Translation*, 1952(71):55, 1952.

[57] J. Rosiński. Decomposition of stationary α-stable random fields. *Ann. Probab.*, 28(4):1797–1813, 2000.

[58] I. N. Sanov. On the probability of large deviations of random variables. In *Select. Transl. Math. Statist. and Probability, Vol. 1*, pages 213–244. Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., 1961.

[59] A. Sard. The measure of the critical values of differentiable maps. *Bull. Amer. Math. Soc.*, 48:883–890, 1942.

[60] J. Serafin. Finitary codes, a short survey. In *Dynamics & Stochastics*, volume 48 of *IMS Lecture Notes Monogr. Ser.*, pages 262–273. Inst. Math. Statist., Beachwood, OH, 2006.

[61] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[62] J. Sinaĭ. On the concept of entropy for a dynamic system. *Dokl. Akad. Nauk SSSR*, 124:768–771, 1959.

[63] J. G. Sinaĭ. On a weak isomorphism of transformations with invariant measure. *Mat. Sb. (N.S.)*, 63 (105):23–42, 1964.

[64] J. G. Sinaĭ. Markov partitions and u-diffeomorphisms. *Funkcional. Anal. i Priložen*, 2(1):64–89, 1968.

[65] J. G. Sinaĭ. Gibbs measures in ergodic theory. *Uspehi Mat. Nauk*, 27(4(166)):21–64, 1972.

[66] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.*, 73:747–817, 1967.

[67] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[68] H. Totoki. On a class of special flows. *Probab. Theory Related Fields*, 15:157–167, 1970.

[69] S. Vaes and J. Wahl. Bernoulli actions of type III_1 and L^2-cohomology. *Geom. Funct. Anal.*, 28(2):518–562, 2018.

[70] J. von Neumann. Various techniques used in connection with random digits. In A. S. Householder, G. E. Forsythe, and H. H. Germond, editors, *Monte Carlo Method*, volume 12 of *National Bureau of Standards Applied Mathematics Series*, chapter 13, pages 36–38. US Government Printing Office, Washington, DC, 1951.

[71] B. Weiss. The structure of Bernoulli systems. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 321–325, 1975.

[72] L.-S. Young. What are SRB measures, and which dynamical systems have them? *J. Statist. Phys.*, 108(5-6):733–754, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[73] H. Zhou and J. Bruck. Linear extractors for extracting randomness from noisy sources. In *2011 IEEE International Symposium on Information Theory Proceedings*, pages 1738–1742, 2011.

Einstein Institute of Mathematics, Hebrew University of Jerusalem, Edmund J. Safra Campus, Givat Ram. Jerusalem, 9190401, Israel.
Email address: zemer.kosloff@mail.huji.ac.il
URL: http://math.huji.ac.il/~zemkos/

Department of Statistical Science, University College London.
Email address: t.soo@ucl.ac.uk; math@terrysoo.com
URL: http://www.terrysoo.com