Ergod. Th. & Dynam. Sys., (2022), **42**, 917–973
doi:10.1017/etds.2021.73

Loosely Bernoulli odometer-based systems whose corresponding circular systems are not loosely Bernoulli

MARLIES GERBER† and PHILIPP KUNDE‡

†*Department of Mathematics, Indiana University, Bloomington, IN 47405, USA*
(e-mail: gerber@indiana.edu)
‡*Department of Mathematics, University of Hamburg, Bundesstraße 55, 20146 Hamburg, Germany*
(e-mail: philipp.kunde@math.uni-hamburg.de)

(*Received* 20 *October* 2019 *and accepted in revised form* 24 *July* 2020)

Dedication: We dedicate this paper to the memory of Anatole Katok. It utilizes two major contributions of Katok, namely the introduction of standard automorphisms (also called zero-entropy loosely Bernoulli) and his work together with D.V. Anosov on the construction of smooth ergodic diffeomorphisms on the disk.

Abstract. Foreman and Weiss [Measure preserving diffeomorphisms of the torus are unclassifiable. *Preprint*, 2020, arXiv:1705.04414] obtained an anti-classification result for smooth ergodic diffeomorphisms, up to measure isomorphism, by using a functor $\mathcal{F}$ (see [Foreman and Weiss, From odometers to circular systems: a global structure theorem. *J. Mod. Dyn.* **15** (2019), 345–423]) mapping odometer-based systems, $\mathcal{OB}$, to circular systems, $\mathcal{CB}$. This functor transfers the classification problem from $\mathcal{OB}$ to $\mathcal{CB}$, and it preserves weakly mixing extensions, compact extensions, factor maps, the rank-one property, and certain types of isomorphisms. Thus it is natural to ask whether $\mathcal{F}$ preserves other dynamical properties. We show that $\mathcal{F}$ does *not* preserve the loosely Bernoulli property by providing positive and zero-entropy examples of loosely Bernoulli odometer-based systems whose corresponding circular systems are not loosely Bernoulli. We also construct a loosely Bernoulli circular system whose corresponding odometer-based system has zero entropy and is not loosely Bernoulli.

Key words: loosely Bernoulli, Kakutani equivalence, odometer
2020 Mathematics Subject Classification: 37A05 (Primary); 37A35, 37A20, 37B10 (Secondary)

1. *Introduction*

An important development in ergodic theory that began in the late 1990s is the emergence of *anti-classification* results for measure-preserving transformations (MPTs) up to isomorphism. Here, an MPT is a measure-preserving automorphism of a standard non-atomic probability space, and two such MPTs, T and S, are said to be isomorphic if there is a measure-preserving isomorphism between the underlying probability spaces that intertwines the actions of T and S. We denote by X the set of MPTs on a fixed standard non-atomic probability space $(\Omega, \mathcal{M}, \mu)$, and let the equivalence relation $\mathcal{R} \subset X \times X$ be defined by $\mathcal{R} := \{(T, S) : T \text{ and } S \text{ are isomorphic}\}$. We endow X with the weak topology. (Recall that $T_n \to T$ in the weak topology if and only if $\mu(T_n(A) \triangle T(A)) \to 0$ for every $A \in \mathcal{M}$.) The first anti-classification theorem in ergodic theory is due to Beleznay and Foreman [BF96], who showed that a certain natural class of measure-distal transformations is not a Borel set in X. In the present context of isomorphism of MPTs, the first result is due to Hjorth [Hj01], who proved that $\mathcal{R}$ is not a Borel subset of $X \times X$. However, this left open the question of what happens if we replace X by the subset $\widetilde{X}$ consisting of ergodic MPTs, with the relative topology. Foreman, Rudolph, and Weiss [FRW11] proved that the equivalence relation $\widetilde{R} := \mathcal{R} \cap (\widetilde{X} \times \widetilde{X})$ is also not a Borel set. These results show that the problem of classifying MPTs (or ergodic MPTs) up to isomorphism, which goes back to von Neumann's 1932 paper [Ne32], is inaccessible to countable methods that use countable amounts of information. (See [FW1, FW2, FW3] for further discussion of this interpretation of these anti-classification results.)

The most important positive results consist of Halmos and von Neumann's classification of ergodic MPTs with pure point spectrum [HN42] and the classification of Bernoulli shifts by their (metric) entropy due to Kolmogorov [KH95, §4.3], Sinai [Si62], and Ornstein [Or70]. Yet many open questions remain. For example, the rank-one transformations, which have been studied extensively, form a dense G_δ subset of X, and the restriction of the equivalence relation $\mathcal{R}$ to rank-one transformations is Borel [FRW11]. However, there is still no known classification of rank-one transformations up to isomorphism.

In view of the anti-classification results mentioned above and, in general, the difficulty of classifying ergodic MPTs, other versions of the classification problem have been considered. One possibility is to restrict the attempted classification to smooth ergodic diffeomorphisms of a compact manifold M with respect to a smooth measure μ. Except in dimensions one and two, there are no known obstructions to realizing an arbitrary ergodic MPT as a diffeomorphism of a compact manifold except the requirement, proved by Kushnirenko [Ku65], that the ergodic MPT have finite entropy. Thus, it is not clear that this restricted classification problem is any easier. Indeed, in this context there is also an anti-classification result due to Foreman and Weiss [FW3]. They showed that if X is replaced by the collection $\widehat{X}$ of smooth Lebesgue-measure-preserving diffeomorphisms of the two-dimensional torus and $\widehat{X}$ is given the C^∞ topology, then the equivalence relation $\widehat{R}$ consisting of pairs of isomorphic elements of $\widehat{X}$ still fails to be a Borel set.

Another modification of the classification problem is to consider Kakutani equivalence instead of isomorphism. Two ergodic MPTs are said to be Kakutani equivalent if they are isomorphic to measurable cross-sections of the same ergodic flow. It follows

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

from Abramov's formula that two Kakutani-equivalent MPTs have the same entropy type: zero entropy, finite entropy, or infinite entropy. Until the work of Katok [**Ka75**, **Ka77**] in the case of zero entropy, and Feldman [**Fe76**] in the general case, no other restrictions were known for achieving Kakutani equivalence. Ornstein, Rudolph, and Weiss [**ORW82**] showed, by building on the work of Feldman, that there are uncountably many non-Kakutani equivalent ergodic MPTs of each entropy type. Thus there is a rich variety of Kakutani equivalence classes, and classification of ergodic MPTs up to Kakutani equivalence also remains an open problem. It is not known whether anti-classification results analogous to those in [**FRW11**, **FW3**] can be obtained for Kakutani equivalence, either in the original setting of ergodic MPTs or in the setting of smooth diffeomorphisms preserving a smooth measure.

In transferring the results of [**FRW11**] to the smooth setting, Foreman and Weiss [**FW2**] introduced a continuous functor $\mathcal{F}$ that maps odometer-based systems to circular systems. (See §3 for definitions of these terms.) According to an announcement in [**FW2**], any finite-entropy system that has an odometer factor can be represented as an odometer-based system. It is a difficult open question whether any transformation with a non-trivial odometer factor can be realized as a smooth diffeomorphism on a compact manifold. Foreman and Weiss [**FW2**] were able to circumvent this difficulty by using the functor $\mathcal{F}$ to transfer the classification problem for odometer-based systems to circular systems. Under mild growth conditions on the parameters, circular systems can be realized as smooth diffeomorphisms of the two-dimensional torus using the Anosov–Katok method [**AK70**]. The functor $\mathcal{F}$ preserves weakly mixing extensions, compact extensions, factor maps, the rank-one property, and certain types of isomorphisms, as well as numerous other properties (see [**FW2**]). While all circular systems have zero entropy, there exist positive-entropy odometer-based systems, and thus $\mathcal{F}$ does not preserve the entropy type. In connection with a possible Kakutani-equivalence version of the results in [**FW3**], a natural question is whether $\mathcal{F}$ preserves Kakutani equivalence (at least for zero-entropy odometer-based systems). In particular, J.-P. Thouvenot asked whether $\mathcal{F}$ maps loosely Bernoulli automorphisms (those Kakutani-equivalent to an irrational rotation of the circle in case of zero entropy, or those Kakutani-equivalent to a Bernoulli shift in case of positive entropy) to loosely Bernoulli automorphisms. We provide examples to show that the answer to both of these questions is 'no'. We also obtain an example which shows that $\mathcal{F}^{-1}$ fails to preserve the loosely Bernoulli property. Our examples suggest that a different approach may be needed for Kakutani-equivalence versions of anti-classification results in the diffeomorphism setting.

In §§4.1 and 4.2 we give an example of a positive-entropy odometer-based system $\mathbb{E}$ that is loosely Bernoulli, but the circular system $\mathcal{F}(\mathbb{E})$ is not loosely Bernoulli. In this example, $(n+1)$-blocks of the odometer-based system are constructed mostly by independent concatenation of n-blocks. Because of this, $\mathbb{E}$ satisfies the positive-entropy version of the loosely Bernoulli property. However, using techniques of Rothstein [**Ro80**], we show that this independent concatenation, when transferred to $\mathcal{F}(\mathbb{E})$, causes the zero-entropy version of the loosely Bernoulli property to fail.

The zero-entropy odometer-based system $\mathbb{K}$ constructed in §§5.1–5.4, is of greater interest in connection with [**FW2**, **FW3**], because the anti-classification results of Foreman

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

and Weiss can be obtained by considering only zero-entropy odometer-based systems. Our zero-entropy example is more difficult to construct than our positive-entropy example, and it uses some delicate refinements of the methods in [ORW82]. However, the heuristics of the construction can be described fairly easily, as illustrated in Figure 2. This example is loosely Bernoulli, but its image under $\mathcal{F}$ is not loosely Bernoulli. There is also a simple example (see Example 3.12) of a zero-entropy odometer-based system that is loosely Bernoulli and whose image under $\mathcal{F}$ is again loosely Bernoulli. This example, together with our example $\mathbb{K}$, shows that $\mathcal{F}$ does not preserve Kakutani equivalence.

Finally, in §§6.1–6.4, we give an example $\mathbb{M}$ of a zero-entropy non-loosely Bernoulli odometer-based system whose corresponding circular system is loosely Bernoulli. Figure 3 shows the idea for this construction. Our §§5.1–6.4 with the zero-entropy odometer-based systems can be read independently of §§4.1 and 4.2.

Our results may also be of interest as another way that non-loosely Bernoulli transformations arise naturally from loosely Bernoulli transformations. Previous examples in this spirit include non-loosely Bernoulli Cartesian products in which the factors are loosely Bernoulli [KR, KW19, ORW82, Ra78, Ra79]. The functor $\mathcal{F}$ in [FW2] changes the way $(n+1)$-blocks are built out of n-blocks according to a scheme that seems, upon first consideration, likely to preserve the loosely Bernoulli property. In this sense, our zero-entropy examples $\mathbb{K}$ and $\mathbb{M}$ were unexpected.

2. *The $\overline{f}$ metric and the loosely Bernoulli property*

Feldman [Fe76] introduced a notion of distance, now called $\overline{f}$, between strings of symbols. He replaced the Hamming metric in Ornstein's very weak Bernoulli property [Or] to define *loosely Bernoulli* transformations (see Definitions 2.2 and 2.3 below). A zero-entropy version of this property was introduced independently by Katok [Ka77].

Definition 2.1. A *match* between two strings of symbols $a_1a_2\ldots a_n$ and $b_1b_2\ldots b_m$, from a given alphabet Σ, is a collection I of pairs of indices (i_s, j_s), $s = 1, \ldots, r$, such that $1 \le i_1 < i_2 < \cdots < i_r \le n$, $1 \le j_1 < j_2 < \cdots < j_r \le m$ and $a_{i_s} = b_{j_s}$ for $s = 1, 2, \ldots, r$. Then

$$\overline{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m) = 1 - \frac{2\sup\{|I| : I \text{ is a match between } a_1a_2\ldots a_n \text{ and } b_1b_2\ldots b_m\}}{n+m}. \tag{2.1}$$

We will refer to $\overline{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m)$ as the '$\overline{f}$-distance' between $a_1a_2\ldots a_n$ and $b_1b_2\ldots b_m$, even though $\overline{f}$ does not satisfy the triangle inequality unless the strings are all of the same length. A match I is called a *best possible match* if it realizes the supremum in the definition of $\overline{f}$.

Suppose $(T, \mathcal{P})$ is a process, that is, T is a measurable automorphism of a measurable space $(\Omega, \mathcal{M})$ and $\mathcal{P} = \{P_\sigma : \sigma \in \Sigma\}$ is a finite measurable partition of Ω. For $x, y \in \Omega$ with $T^i(x) \in P_{a_i}$ and $T^i(y) \in P_{b_i}$ for $i = 1, \ldots, K$, we define $\overline{f}_K(x, y) := \overline{f}(a_1a_2\ldots a_K, b_1b_2\ldots b_K)$. If ν and ω are probability measures on $(\Omega, \mathcal{M})$ then we say $\overline{f}_K(\nu, \omega) < \epsilon$ if there is a measure-preserving invertible map $\phi : (\Omega, \mathcal{M}, \nu) \to$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$(\Omega, \mathcal{M}, \omega)$ such that there exists a set $G \subset \Omega$ with $\nu(G) > 1 - \varepsilon$ and $\overline{f}_K(x, \phi(x)) < \varepsilon$ for all $x \in G$.

We now define loosely Bernoulli in the general case (no assumptions on the entropy of the process).

Definition 2.2. (Loosely Bernoulli in the general case) A measure-preserving process $(T, \mathcal{P}, \nu)$ is *loosely Bernoulli* if for every $\varepsilon > 0$, there exists a positive integer $K = K(\varepsilon)$ such that for every positive integer M the following holds: there exists a collection $\mathcal{G}$ of 'good' atoms of $\vee_{-M}^{0} T^{-i}\mathcal{P}$ whose union has measure greater than $1 - \varepsilon$, so that for each pair A, B of atoms in $\mathcal{G}$ of positive ν-measure, the measures ν_A and ν_B satisfy $\overline{f}_K(\nu_A, \nu_B) < \varepsilon$. Here ν_A and ν_B denote the conditional measures on Ω defined by $\nu_A(C) = \nu(C|A) = \nu(C \cap A)/\nu(A)$, and similarly for ν_B.

An ergodic measure-preserving transformation (T, ν) is *loosely Bernoulli* if $(T, \mathcal{P}, \nu)$ is a loosely Bernoulli process for every partition $\mathcal{P}$. In fact, it suffices for $(T, \mathcal{P}, \nu)$ to be loosely Bernoulli for a generating partition $\mathcal{P}$. (By Krieger's generator theorem [**Kr70**] every ergodic measure-preserving transformation of finite entropy has a finite generator.)

As was pointed out in [**Fe76**], Definition 2.2 is equivalent to the definition obtained by replacing 'there exists a positive integer $K = K(\varepsilon)$' by 'for any sufficiently large positive integer K (how large depends on ε)'. Moreover, according to [**Fe76**, Corollary 2], 'every positive integer M' can be replaced by 'for every sufficiently large positive integer M'. In fact, our Lemma 4.5 implies [**Fe76**, Corollary 2], and the proof is similar.

In the case of zero entropy, no conditioning on the past is needed, and there is a simpler definition of loosely Bernoulli. That is, the definition reduces to the following version.

Definition 2.3. (Loosely Bernoulli in the case of zero entropy) A measure-preserving process $(T, \mathcal{P}, \nu)$ is *zero-entropy loosely Bernoulli* if for every $\varepsilon > 0$, there exist a positive integer $K = K(\varepsilon)$ and a collection $\mathcal{G}$ of 'good' atoms of $\vee_1^K T^{-i}\mathcal{P}$ with total measure greater than $1 - \varepsilon$ such that for each pair A, B of atoms in $\mathcal{G}$, $\overline{f}_K(x, y) < \varepsilon$ for $x \in A$, $y \in B$.

If this condition is satisfied, then routine estimates show that the $(T, \mathcal{P}, \nu)$ process indeed has zero entropy.

Remark 2.4. For infinite strings of symbols $a_0a_1 \ldots, b_0b_1 \ldots$, we can define

$$\overline{f}(a_0a_1 \ldots, b_0b_1 \ldots) = \limsup_{n\to\infty} \overline{f}(a_0a_1 \ldots a_n, b_0b_1 \ldots b_n).$$

There is an alternate definition of zero-entropy loosely Bernoulli (which we do not make use of) that can be formulated as follows. The process $(T, \mathcal{P}, \nu)$ is *zero-entropy loosely Bernoulli* if there is a $G \in \vee_0^\infty T^{-i}\mathcal{P}$ with $\nu(G) = 1$ such that for $x, y \in G$, and for $T^i x \in P_{a_i}, T^i y \in P_{b_i}$ for $i = 0, 1, 2, \ldots$, we have $\overline{f}(a_0a_1 \ldots, b_0b_1 \ldots) = 0$.

The following simple properties of $\overline{f}$, which were already used in [**Fe76**, **ORW82**], will appear frequently in our arguments. These properties can be proved easily by considering the *fit*, $1 - \overline{f}(a, b)$, between two strings a and b.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

Property 2.5. Suppose a and b are strings of symbols of length n and m, respectively, from an alphabet Σ. If $\tilde{a}$ and $\tilde{b}$ are strings of symbols obtained by deleting at most $\lfloor \gamma(n+m) \rfloor$ terms from a and b altogether, where $0 < \gamma < 1$, then

$$\overline{f}(a, b) \geq \overline{f}(\tilde{a}, \tilde{b}) - 2\gamma. \tag{2.2}$$

Moreover, if there exists a best possible match between a and b such that no term that is deleted from a and b to form $\tilde{a}$ and $\tilde{b}$ is matched with a non-deleted term, then

$$\overline{f}(a, b) \geq \overline{f}(\tilde{a}, \tilde{b}) - \gamma. \tag{2.3}$$

Likewise, if $\tilde{a}$ and $\tilde{b}$ are obtained by adding at most $\lfloor \gamma(n+m) \rfloor$ symbols to a and b, then (2.3) holds.

Property 2.6. Suppose $x = x_1 x_2 \ldots x_n$ and $y = y_1 y_2 \ldots y_n$ are decompositions of the strings of symbols x and y into substrings such that there exists a best possible match between x and y where terms in x_i are only matched with terms in y_i (if they are matched with any term in y). Then

$$\overline{f}(x, y) = \sum_{i=1}^{n} \overline{f}(x_i, y_i) v_i,$$

where

$$v_i = \frac{|x_i| + |y_i|}{|x| + |y|}. \tag{2.4}$$

Property 2.7. If x and y are strings of symbols such that $\overline{f}(x, y) \leq \gamma$, for some $0 \leq \gamma < 1$, then

$$\left(\frac{1-\gamma}{1+\gamma}\right)|x| \leq |y| \leq \left(\frac{1+\gamma}{1-\gamma}\right)|x|. \tag{2.5}$$

We often use this property with $\gamma = 1/7$, in which case the conclusion can be formulated as

$$\frac{3|x|}{4} \leq |y| \leq \frac{4|x|}{3}. \tag{2.6}$$

3. *Odometer-based and circular symbolic systems*

In this section we review the notation and definitions for odometer-based and circular symbolic systems. We also present the functor $\mathcal{F}$ of [**FW2**] between these two systems.

3.1. *Symbolic systems.*

An *alphabet* is a countable or finite collection of symbols. In the following, let Σ be a finite alphabet endowed with the discrete topology. Then $\Sigma^{\mathbb{Z}}$ with the product topology is a separable, totally disconnected and compact space. The shift

$$sh : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}, \ sh(f)(n) = f(n+1)$$

is a homeomorphism. If μ is a shift-invariant Borel measure, then the measure-preserving dynamical system $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, sh)$ is called a *symbolic system*. The closed support of μ is a shift-invariant subset of $\Sigma^{\mathbb{Z}}$ called a *symbolic shift* or *subshift*.

Symbolic shifts are often described by giving a collection of words that constitute a basis for the support of an invariant measure. A word w in Σ is a finite sequence of elements of Σ, and we denote its length by $|w|$. A *language* (over Σ) is a subset of the set of all words.

Definition 3.1. A sequence of collection of words $(W_n)_{n\in\mathbb{N}}$, where $\mathbb{N} = \{0, 1, 2, \ldots\}$, satisfying the following properties is called a *construction sequence*.

(1) For every $n \in \mathbb{N}$ all words in W_n have the same length h_n.

(2) Each $w \in W_n$ occurs at least once as a subword of each $w' \in W_{n+1}$.

(3) There is a summable sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of positive numbers such that for every $n \in \mathbb{N}$, every word $w \in W_{n+1}$ can be uniquely parsed into segments $u_0 w_1 u_1 w_1 \ldots w_l u_{l+1}$ such that each $w_i \in W_n$, each u_i (called spacer or boundary) is a word in Σ of finite length, and for this parsing

$$\frac{\sum_{i=0}^{l+1}|u_i|}{h_{n+1}} < \varepsilon_{n+1}.$$

We will often call words in W_n *n-words* or *n-blocks*, while a general concatenation of symbols from Σ is called a *string*. We also associate a symbolic shift with a construction sequence. Let $\mathbb{K}$ be the collection of $x \in \Sigma^{\mathbb{Z}}$ such that every finite contiguous substring of x occurs inside some $w \in W_n$. Then $\mathbb{K}$ is a closed shift-invariant subset of $\Sigma^{\mathbb{Z}}$ that is compact if Σ is finite. In order to be able to unambiguously parse elements of $\mathbb{K}$ we will use construction sequences consisting of uniquely readable words.

Definition 3.2. Let Σ be a language and W be a collection of finite words in Σ. Then W is *uniquely readable* if and only if whenever $u, v, w \in W$ and $uv = pws$ with p and s strings of symbols in Σ, then either p or s is the empty word.

Moreover, our $(n + 1)$-words will be uniform in the n-words as defined below.

Definition 3.3. We call a construction sequence $(W_n)_{n\in\mathbb{N}}$ *uniform* if for each $n \in \mathbb{N}$ there is a constant $c > 0$ such that for all words $w' \in W_{n+1}$ and $w \in W_n$ the number of occurrences of w in w' is equal to c.

Remark 3.4. In [**FW2**] such construction sequences are called 'strongly uniform'. Since we will only deal with this strong notion of uniformity in this paper, we abbreviate that terminology.

To check the zero-entropy loosely Bernoulli property for our symbolic systems we will use the following criterion, which follows from Rothstein's Lemma 2.6 in [**Ro80**] and the fact that the given condition in terms of n-blocks implies entropy zero. (See p. 18 of [**ORW82**] for the type of estimate that is needed.)

LEMMA 3.5. *Suppose $\mathbb{K}$ is a symbolic system with uniform and uniquely readable construction sequence. Then $\mathbb{K}$ is zero-entropy loosely Bernoulli if and only if for every $\varepsilon > 0$ there exists N such that for $n \geq N$, there is a set of n-blocks $\mathcal{G}_n$ with cardinality $|\mathcal{G}_n| > (1 - \varepsilon)|W_n|$ such that for $A, B \in \mathcal{G}_n$, $\overline{f}(A, B) < \varepsilon$.*

3.2. *Odometer-based systems.* Let $(k_n)_{n\in\mathbb{N}}$ be a sequence of natural numbers $k_n \geq 2$ and

$$O = \prod_{n\in\mathbb{N}} (\mathbb{Z}/k_n\mathbb{Z})$$

be the $(k_n)_{n\in\mathbb{N}}$-adic integers. Then O has a compact abelian group structure and hence carries a Haar measure λ. We define a transformation $T : O \to O$ to be addition by 1 in the $(k_n)_{n\in\mathbb{N}}$-adic integers (that is, the map that adds one in $\mathbb{Z}/k_0\mathbb{Z}$ and carries right). Then T is a λ-preserving invertible transformation called *odometer transformation* which is ergodic and has discrete spectrum.

We now define the collection of symbolic systems that have odometer systems as their timing mechanism to parse typical elements of the system.

Definition 3.6. Let $(\mathtt{W}_n)_{n\in\mathbb{N}}$ be a uniquely readable construction sequence with $\mathtt{W}_0 = \Sigma$ and $\mathtt{W}_{n+1} \subseteq (\mathtt{W}_n)^{k_n}$ for every $n \in \mathbb{N}$. The associated symbolic shift will be called an *odometer-based system*.

Thus, odometer-based systems are those built from construction sequences $(\mathtt{W}_n)_{n\in\mathbb{N}}$ such that the words in $\mathtt{W}_{n+1}$ are concatenations of a fixed number k_n of words in $\mathtt{W}_n$. Hence, the words in $\mathtt{W}_n$ have length $\mathtt{h}_n$, where

$$\mathtt{h}_n = \prod_{i=0}^{n-1} k_i$$

if $n > 0$, and $\mathtt{h}_0 = 1$. Moreover, the spacers in part (3) of Definition 3.1 are all the empty words (that is, an odometer-based transformation can be built by a cut-and-stack construction using no spacers).

3.3. *Circular systems.* A *circular coefficient sequence*is a sequence of pairs of integers $(k_n, l_n)_{n\in\mathbb{N}}$ such that $k_n \geq 2$ and $\sum_{n\in\mathbb{N}}(1/l_n) < \infty$. From these numbers we inductively define numbers

$$q_{n+1} = k_n l_n q_n^2$$

and

$$p_{n+1} = p_n k_n l_n q_n + 1,$$

where we set $p_0 = 0$ and $q_0 = 1$. Obviously, p_{n+1} and q_{n+1} are relatively prime. Moreover, let Σ be a non-empty finite alphabet and b, e be two additional symbols (called *spacers*). Then, given a circular coefficient sequence $(k_n, l_n)_{n\in\mathbb{N}}$, we build collections of words $\mathcal{W}_n$ in the alphabet $\Sigma \cup \{b, e\}$ by induction as follows.

- Set $\mathcal{W}_0 = \Sigma$.
- Having built $\mathcal{W}_n$, choose a set $P_{n+1} \subseteq (\mathcal{W}_n)^{k_n}$ of so-called *prewords* and form $\mathcal{W}_{n+1}$ by taking all words of the form

$$\mathcal{C}_n(w_0, w_1, \ldots, w_{k_n-1}) = \prod_{i=0}^{q_n-1} \prod_{j=0}^{k_n-1} (b^{q_n-j_i} w_j^{l_n-1} e^{j_i})$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

with $w_0 \dots w_{k_n-1} \in P_{n+1}$. If $n = 0$ take $j_0 = 0$, and for $n > 0$ let $j_i \in \{0, \dots, q_n - 1\}$ be such that

$$j_i \equiv (p_n)^{-1} i \mod q_n.$$

We note that each word in $\mathcal{W}_{n+1}$ has length $k_n l_n q_n^2 = q_{n+1}$.

Definition 3.7. A construction sequence $(\mathcal{W}_n)_{n\in\mathbb{N}}$ will be called *circular* if it is built in this manner using the $\mathcal{C}$-operators and a circular coefficient sequence, and each P_{n+1} is uniquely readable in the alphabet with the words from $\mathcal{W}_n$ as letters (this last property is called the *strong readability assumption*).

Remark 3.8. By [**FW2**, Lemma 45] each $\mathcal{W}_n$ in a circular construction sequence is uniquely readable even if the prewords are not uniquely readable. However, the definition of a circular construction sequence requires this stronger readability assumption.

The definition of a circular construction sequence stems from a symbolic representation for untwisted Anosov–Katok diffeomorphisms on $\mathbb{T}^2$ (and similarly on $\mathbb{D}^2$ and $\mathbb{S}^1 \times [0, 1]$) described in [**FW1**]. Such a diffeomorphism is constructed inductively as the limit of a sequence $T_{n+1} = H_{n+1} \circ R_{\alpha_{n+1}} \circ H_{n+1}^{-1}$, where $R_{\alpha_{n+1}}(x, y) = (x + \alpha_{n+1}, y)$ is the rotation by $\alpha_{n+1} = (p_{n+1}/q_{n+1})$ with the numbers p_{n+1}, q_{n+1} from above, and $H_{n+1} = H_n \circ h_{n+1}$ with area-preserving diffeomorphisms h_{n+1} satisfying $h_{n+1} \circ R_{1/q_n} = R_{1/q_n} \circ h_{n+1}$. In the untwisted version of the Anosov–Katok method the fundamental domain $[0, 1/q_n] \times [0, 1]$ is required to map to itself under h_{n+1}. The conjugation map h_{n+1} approximately permutes sets of the form $[i/(k_n q_n), (i + 1)/(k_n q_n)) \times [s/s_{n+1}, (s + 1)/s_{n+1})$, where $s_{n+1} \geq 2$ is the number of $(n + 1)$-words. The combinatorics of h_{n+1} on the fundamental domain is determined by s_{n+1} different concatenations $w_0 \dots w_{k_n-1}$ of n-words, and each n-word codes a T_n trajectory of length q_n. Since $k_n l_n q_n \alpha_{n+1} = 1/q_n \mod 1$ and h_{n+1} commutes with R_{1/q_n}, the code repeats but starting at a different position within the n-words. Thus one uses the spacer symbols b and e to label the incomplete beginning and end segments of words. We refer to [**FW1**, §7.5] for a detailed exposition. The newly introduced spacers b, e will not matter for our arguments since they occupy a small proportion of the $(n + 1)$-words and we will often neglect them in our estimates with the aid of Property 2.5.

Definition 3.9. A symbolic shift $\mathbb{K}$ built from a circular construction sequence is called a *circular system*. For emphasis we will often denote it by $\mathbb{K}^c$.

For a word $w \in \mathcal{W}_{n+1}$ we introduce the following subscales as in [**FW2**, §3.3].

- Subscale 0 is the scale of the individual powers of $w_j \in \mathcal{W}_n$ of the form w_j^{l-1}, and each such occurrence of a w_j^{l-1} is called a 0 *-subsubsection*.
- Subscale 1 is the scale of each term in the product $\prod_{j=0}^{k_n-1} (b^{q_n-j_i} w_j^{l_n-1} e^{j_i})$ that has the form $(b^{q_n-j_i} w_j^{l_n-1} e^{j_i})$ and these terms are called 1-*subsections*.
- Subscale 2 is the scale of each term of $\prod_{i=0}^{q_n-1} \prod_{j=0}^{k_n-1} (b^{q_n-j_i} w_j^{l_n-1} e^{j_i})$ that has the form $\prod_{j=0}^{k_n-1} (b^{q_n-j_i} w_j^{l_n-1} e^{j_i})$, and these terms are called 2-*subsections*.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

3.4. *The functor $\mathcal{F}$.* For a fixed circular coefficient sequence $(k_n, l_n)_{n\in\mathbb{N}}$ we consider two categories $\mathcal{OB}$ and $\mathcal{CB}$ whose objects are odometer-based and circular systems, respectively. The morphisms in these categories are (synchronous and anti-synchronous) graph joinings. In [**FW2**] Foreman and Weiss define a functor taking odometer-based systems to circular systems that preserve the factor and conjugacy structure. In this subsubsection we review the definition of the functor from the odometer-based symbolic systems to the circular symbolic systems.

For this purpose, we fix a circular coefficient sequence $(k_n, l_n)_{n\in\mathbb{N}}$. Let Σ be an alphabet and $(\mathtt{W}_n)_{n\in\mathbb{N}}$ be a construction sequence for an odometer-based system with coefficients $(k_n)_{n\in\mathbb{N}}$. Then we define a circular construction sequence $(\mathcal{W}_n)_{n\in\mathbb{N}}$ and bijections $c_n : \mathtt{W}_n \to \mathcal{W}_n$ by induction.

- Let $\mathcal{W}_0 = \Sigma$ and c_0 be the identity map.
- Suppose that $\mathtt{W}_n$, $\mathcal{W}_n$ and c_n have already been defined. Then we define

$$\mathcal{W}_{n+1} = \{\mathcal{C}_n(c_n(\mathtt{w}_0), c_n(\mathtt{w}_1), \ldots, c_n(\mathtt{w}_{k_n-1})) \;:\; \mathtt{w}_0\mathtt{w}_1 \ldots \mathtt{w}_{k_n-1} \in \mathtt{W}_{n+1}\}$$

and the map c_{n+1} by setting

$$c_{n+1}(\mathtt{w}_0\mathtt{w}_1 \ldots \mathtt{w}_{k_n-1}) = \mathcal{C}_n(c_n(\mathtt{w}_0), c_n(\mathtt{w}_1), \ldots, c_n(\mathtt{w}_{k_n-1})).$$

In particular, the prewords are

$$P_{n+1} = \{c_n(\mathtt{w}_0)c_n(\mathtt{w}_1) \ldots c_n(\mathtt{w}_{k_n-1}) \;:\; \mathtt{w}_0\mathtt{w}_1 \ldots \mathtt{w}_{k_n-1} \in \mathtt{W}_{n+1}\}.$$

Definition 3.10. Suppose that $\mathbb{K}$ is built from a construction sequence $(\mathtt{W}_n)_{n\in\mathbb{N}}$ and $\mathbb{K}^c$ has the circular construction sequence $(\mathcal{W}_n)_{n\in\mathbb{N}}$ as constructed above. Then we define a map $\mathcal{F}$ from the set of odometer-based systems (viewed as subshifts) to circular systems (viewed as subshifts) by

$$\mathcal{F}(\mathbb{K}) = \mathbb{K}^c.$$

Remark 3.11. The map $\mathcal{F}$ is a bijection between odometer-based symbolic systems with coefficients $(k_n)_{n\in\mathbb{N}}$ and circular symbolic systems with coefficients $(k_n, l_n)_{n\in\mathbb{N}}$ that preserves uniformity. Since the construction sequences for our odometer-based systems will be uniquely readable, the corresponding circular construction sequences will automatically satisfy the strong readability assumption.

In the following we will denote blocks in the odometer-based system by letters in typewriter font (for example, $\mathtt{A}$). For the corresponding block in the circular system we will use calligraphic letters (for example, $\mathcal{A}$). As already noted, the length of a n-block $\mathtt{w}$ in the odometer-based system is $\mathtt{h}_n = \prod_{i=0}^{n-1} k_i$ if $n > 0$, and $\mathtt{h}_0 = 1$, while the length of a n-block in the circular system is q_n, that is, $|c_n(\mathtt{w})| = q_n$. Moreover, we will use the following map from substrings of the underlying odometer-based system to the circular system:

$$\mathcal{C}_{n,i}(\mathtt{w}_s\mathtt{w}_{s+1} \ldots \mathtt{w}_t) = \prod_{j=s}^{t} (b^{q_n - j_i}(c_n(\mathtt{w}_j))^{l_n-1} e^{j_i})$$

for any $0 \le i \le q_n - 1$ and $0 \le s \le t \le k_n - 1$.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

Example 3.12. We give an example of a loosely Bernoulli odometer-based system $\mathbb{K}$ of zero measure-theoretic entropy with uniform and uniquely readable construction sequence such that $\mathcal{F}(\mathbb{K})$ is also loosely Bernoulli. For this purpose, let Σ be an alphabet with two symbols and $\varepsilon_n \searrow 0$. Assume that we have two n-blocks $\mathtt{w}_0$ and $\mathtt{w}_1$ in the odometer-based system. Then we define two $(n+1)$-blocks by the following rule:

$$\mathtt{B}_0^{(n+1)} = \mathtt{w}_1\mathtt{w}_1 \underbrace{\mathtt{w}_0\mathtt{w}_1\mathtt{w}_0\mathtt{w}_1 \ldots \mathtt{w}_0\mathtt{w}_1}_{2s_n \text{ blocks}} \mathtt{w}_0\mathtt{w}_0,$$

$$\mathtt{B}_1^{(n+1)} = \mathtt{w}_1\mathtt{w}_1\mathtt{w}_1 \underbrace{\mathtt{w}_0\mathtt{w}_1 \ldots \mathtt{w}_0\mathtt{w}_1}_{2s_n-2 \text{ blocks}} \mathtt{w}_0\mathtt{w}_0\mathtt{w}_0,$$

where we choose the integer s_n sufficiently large to guarantee that $\mathtt{B}_0^{(n+1)}$ and $\mathtt{B}_1^{(n+1)}$ are ε_n-close to each other in $\overline{f}$. Clearly, the construction sequence defined like this is uniform and uniquely readable. Moreover, the corresponding $(n+1)$-blocks $\mathcal{B}_0^{(n+1)}$ and $\mathcal{B}_1^{(n+1)}$ in the circular system are also ε_n-close to each other in $\overline{f}$. Hence, $\mathbb{K}$ and $\mathcal{F}(\mathbb{K})$ are loosely Bernoulli by Lemma 3.5.

4. *Positive-entropy example*

4.1. *Positive-entropy loosely Bernoulli odometer-based system.*

In this section we construct a uniquely readable uniform odometer-based system $\mathbb{E}$ of positive entropy that is loosely Bernoulli. In the next section we will prove that $\mathcal{F}(\mathbb{E})$ is not loosely Bernoulli. The main idea in the construction of $\mathbb{E}$ is to concatenate n-blocks independently in long initial segments of $(n+1)$-blocks, and use a relatively small final segment of the $(n+1)$-block to achieve uniformity and unique readability. If we used only the independent concatenation, then $\mathbb{E}$ would be Bernoulli (and hence loosely Bernoulli), and $\mathcal{F}(\mathbb{E})$ would still be non-loosely Bernoulli. In this case the proof given in the next section that $\mathcal{F}(\mathbb{E})$ is not loosely Bernoulli could be simplified, but we want to achieve uniformity and unique readability to make our example fit the framework of the odometer-based constructions in [**FW2**, **FW3**].

Our approach takes advantage of the fact that in the case of positive entropy, in particular for the system $\mathbb{E}$, there will be many n-blocks that are bounded apart in $\overline{f}$-distance and the loosely Bernoulli property can still be satisfied. However, all circular systems, as described in §2, and in particular $\mathcal{F}(\mathbb{E})$, have entropy zero. In this case the loosely Bernoulli property fails to hold if most of the n-blocks are bounded apart in the $\overline{f}$ metric. In the next section we will use the approach of Rothstein [**Ro80**] to prove that the independent concatenation of n-blocks that are mostly bounded apart in $\overline{f}$ distance leads to $(n+1)$-blocks that are also mostly bounded apart in $\overline{f}$ distance, and the lower bound on the $\overline{f}$ distance decreases only slightly in going from n-blocks to $(n+1)$-blocks.

We begin by describing some of the conditions on the parameters involved in the construction of $\mathbb{E}$. Further requirements on the lower bound on the k_n will be imposed in the next section and after the first lemma in the present section. First we choose positive rational numbers ε_n such that $\varepsilon_n < 2^{-(n+12)}$. Then we choose k_n (depending on $\ell_n, N(n), \varepsilon_n$) so that $\sum_{n=1}^{\infty} N(n)^2/(\varepsilon_n^2 k_n) < 1/8$. Furthermore, we require that $\varepsilon_n N(n) > 2$, $\varepsilon_n k_n$ is an integer, and k_n is a multiple of $N(n)$. Let $k_n' = (1-\varepsilon_n)k_n$.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

We now describe the construction sequence $(\mathtt{W}_n)_{n\in\mathbb{N}}$ for $\mathbb{E}$. Recall that $\mathtt{W}_0 = \Sigma$. Suppose there are $N(n)$ distinct n-blocks of length h_n in $\mathtt{W}_n$, say $\mathtt{W}_n = \{\mathtt{y}_1, \ldots, \mathtt{y}_{N(n)}\}$. Then $\mathtt{W}_{n+1}$ consists of all words of the form $\mathtt{w}_1\mathtt{w}_2 \ldots \mathtt{w}_{k_n}$, where each $\mathtt{w}_j = \mathtt{y}_{i(j)}$ for some $i(j) \in \{1, \ldots, N_n(n)\}$, subject to the following conditions.

(1) For each $i \in \{1, 2, \ldots, N(n) - 1\}$, $\mathrm{card}\{j \in \{1, 2, \ldots, k'_n\} : \mathtt{w}_j = \mathtt{y}_i\} \leq k_n/N(n)$, and $\mathtt{w}_j \neq \mathtt{y}_{N(n)}$ for $j \in \{1, 2, \ldots, k'_n\}$.

(2) For $j \in \{k_n - (k_n/N(n)), k_n - (k_n/N(n)) + 1, \ldots, k_n - 1\}$, $\mathtt{w}_j = \mathtt{y}_{N(n)}$.

(3) The substring $\mathtt{w}_{(1-\varepsilon_n)k_n} \ldots \mathtt{w}_{k_n-(k_n/N(n))-1}$ consists of a finite string of $\mathtt{y}_1$'s, followed by a finite string of $\mathtt{y}_2$'s, etc. ending with a finite string of $\mathtt{y}_{N(n)-1}$'s such that $\mathrm{card}\{j \in \{1, 2, \ldots k_n\} : \mathtt{w}_j = \mathtt{y}_i\} = k_n/N(n)$ for every $i \in \{1, 2, \ldots, N(n)\}$.

Whenever condition (1) on the first k'_n n-blocks is satisfied, there is a unique way of completing the $(n + 1)$-block so that conditions (2) and (3) are satisfied. Condition (2) implies unique readability, and condition (3) implies uniformity. Our block construction and Lemma 4.1 below are essentially a special case of the techniques in the substitution lemma in [**FRW11**].

LEMMA 4.1. (Chebyshev application) *Suppose $k'_n = (1 - \varepsilon_n)k_n$ symbols are chosen independently from $\{1, 2, \ldots, N(n) - 1\}$, where each symbol is equally likely to be chosen. Then the probability τ_n that there exists a symbol that is chosen more than $k_n/N(n)$ times satisfies $\tau_n < 4N(n)^2/(\varepsilon_n^2 k_n)$.*

Proof. For a fixed $i_0 \in \{1, \ldots, N(n) - 1\}$, let $S = S_{k'_n}$ be the number of times i_0 is chosen in k'_n Bernoulli trials, where the probability of i_0 being chosen in any one trial is $1/(N(n) - 1)$. Then the expected value of S is $E(S) = k'_n/(N(n) - 1)$ and the standard deviation of S is $\sigma = \sqrt{k'_n(N(n) - 2)}/(N(n) - 1)$. Note that the condition $\varepsilon_n N(n) > 2$ implies that $1/N(n) > (1 - (\varepsilon_n/2))/(N(n) - 1)$. We have the following estimates on the probabilities, where we apply Chebyshev's inequality in the last step:

$$\begin{aligned}\Pr(S > k_n/N(n)) &\leq \Pr\bigg(|S - E(S)| > \frac{k_n}{N(n)} - \frac{(1-\varepsilon_n)k_n}{N(n)-1}\bigg)\\ &\leq \Pr\bigg(|S - E(S)| > \frac{(1-(\varepsilon_n/2))k_n}{N(n)-1} - \frac{(1-\varepsilon_n)k_n}{N(n)-1}\bigg)\\ &= \Pr\bigg(|S - E(S)| > \frac{(\varepsilon_n/2)k_n}{N(n)-1}\bigg)\\ &= \Pr(|S - E(S)| > \alpha\sigma)\\ &< 1/\alpha^2,\end{aligned}$$

where $\alpha = \varepsilon_n k_n/(2\sqrt{k'_n(N(n) - 2)})$. Thus $\Pr(S > k_n/N(n)) < 4N(n)/(\varepsilon_n^2 k_n)$. Since i_0 was only one of $N(n) - 1$ possible symbols, the upper bound in the statement of the lemma is obtained by multiplying the upper bound on $\Pr(S > k_n/N(n))$ by $N(n)$. □

We require the k_n's to be sufficiently large so that $\sum_{n=1}^{\infty} \tau_n^{(2^{-n})} < \infty$. From Lemma 4.1, this is possible because k_n is chosen after ε_n and $N(n)$ are determined.

We apply Lemma 4.1 to the independent choice of k'_n n-blocks from the first $N(n) - 1$ n-blocks. Suppose E is a string of k'_n n-blocks chosen from the first $N(n) - 1$ n-blocks.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

If this string of k'_n n-blocks satisfies condition (1) above, then E is a possible initial string of k'_n n-blocks in an $(n+1)$-block. If $\Pr_n(E)$ is the probability of E in the process $\mathbb{E}$, and $\widetilde{\Pr}_n(E)$ is the probability of E in the process that consists of concatenating k'_n n-blocks chosen independently from the first $N(n)-1$ n-blocks, then

$$\Pr_n(E) = \bigg(\frac{1}{1-\tau_n}\bigg)\widetilde{\Pr}_n(E) \tag{4.1}$$

if E is a possible initial string, and $\Pr_n(E)=0$ if E is not a possible initial string. Note that if we compare the probability distributions $\Pr_n$ and $\widetilde{\Pr}_n$ on the collection $\mathcal{A}$ of all strings of k'_n n-blocks, and the collection $\mathcal{A}'$ of possible initial strings of k'_n n-blocks chosen from the first $N(n)-1$ n-blocks, we obtain

$$\begin{aligned}\sum_{E\in\mathcal{A}} |\Pr_n(E)-\widetilde{\Pr}_n(E)| &= \tau_n + \sum_{E\in\mathcal{A}'} |\Pr_n(E)-\widetilde{\Pr}_n(E)| \\ &= \tau_n + \bigg(\frac{1}{1-\tau_n}-1\bigg)(1-\tau_n) = 2\tau_n. \end{aligned} \tag{4.2}$$

For $i=1,2,\ldots,|\Sigma|$, let P_i be the set of points in $\Sigma^{\mathbb{Z}}$ with the symbol i in position 0. Then $\mathcal{P} := \{P_1, P_2, \ldots, P_{|\Sigma|}\}$ is a generating partition for the odometer system. For integers a and b with $a \le b$, let $\mathcal{P}_a^b$ denote the partition of $\Sigma^{\mathbb{Z}}$ into sets with the same $\mathcal{P}$-name from time a to time b. As before, we let μ denote the invariant measure on $\Sigma^{\mathbb{Z}}$ corresponding to the process $\mathbb{E}$. We let $H = H(\mathrm{sh},\mathcal{P})$ denote the measure entropy of the left shift on $\Sigma^{\mathbb{Z}}$ with respect to $\mathcal{P}$, and we let $H_{\mathrm{top}}(\mathrm{sh},\mathcal{P})$ denote the topological entropy. Note that uniformity and unique readability imply that $\mathrm{sh}:\Sigma^{\mathbb{Z}}\to\Sigma^{\mathbb{Z}}$ is uniquely ergodic. Therefore, by the variational principle [**KH95**, Theorem 4.5.3], $H = H_{\mathrm{top}}(\mathrm{sh},\mathcal{P})$.

LEMMA 4.2. *If $\mathbb{E}$ is the odometer-based system constructed above, then the entropy of $\mathbb{E}$ is positive.*

Proof. The number of elements in $\mathcal{P}_1^{\mathtt{h}_n}$ is at least $N(n)$, while the number of elements in $\mathcal{P}_1^{k\mathtt{h}_n}$ is at most $\mathtt{h}_n N(n)^{k+1}$. These estimates show that

$$H = \lim_{n\to\infty} \frac{\log N(n)}{\mathtt{h}_n}.$$

Here $\mathtt{h}_n = \prod_{i=0}^{n-1} k_i$ and $N(n+1) = N(n)^{k'_n}(1-\tau_n) > N(n)^{k'_n-1}$. Thus

$$\frac{\log N(n+1)}{\mathtt{h}_{n+1}} \ge \frac{(k'_n-1)\log N(n)}{k_n\mathtt{h}_n} \ge \frac{(1-2\varepsilon_n)\log N(n)}{\mathtt{h}_n}.$$

Therefore

$$\frac{\log N(n)}{\mathtt{h}_n} \ge (\log|\Sigma|)\prod_{i=0}^{n-1}(1-2\varepsilon_i).$$

Since $\prod_{i=0}^{\infty}(1-2\varepsilon_i)$ converges to a positive value, it follows that $H>0$. □

LEMMA 4.3. (Conditioning lemma) *Suppose* Pr *and* Pr′ *are two probability distributions defined on the join $\mathcal{Q}\vee\mathcal{R}$ of two partitions $\mathcal{Q}$ and $\mathcal{R}$ of the same space. Suppose that for*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

some $0 < \varepsilon < 1$,

$$\sum_{Q \in \mathcal{Q}, R \in \mathcal{R}} |\Pr(Q \cap R) - \Pr'(Q \cap R)| < \varepsilon. \tag{4.3}$$

Also assume that $\Pr'(Q) > 0$ *whenever* $\Pr(Q) > 0$. *Then for all but* Pr *at most* $\sqrt{\varepsilon}$ *of the* Q*s in* $\mathcal{Q}$, *the conditional probabilities corresponding to* Pr *and* Pr′ *satisfy*

$$\sum_{R \in \mathcal{R}} |\Pr(R|Q) - \Pr'(R|Q)| < 2\sqrt{\varepsilon}.$$

Proof. It follows from (4.3) that for all but Pr at most $\sqrt{\varepsilon}$ of the Qs in $\mathcal{Q}$,

$$\sum_{R \in \mathcal{R}} |\Pr(Q \cap R) - \Pr'(Q \cap R)| < \sqrt{\varepsilon}\Pr(Q). \tag{4.4}$$

Suppose $Q \in \mathcal{Q}$ is chosen so that (4.4) holds. Then

$$|\Pr(Q) - \Pr'(Q)| = \left| \sum_{R \in \mathcal{R}} [\Pr(Q \cap R) - \Pr'(Q \cap R)] \right| < \sqrt{\varepsilon}\Pr(Q).$$

If we divide (4.4) by $\Pr(Q)$, we obtain

$$\sum_{R \in \mathcal{R}} \left| \frac{\Pr(Q \cap R)}{\Pr(Q)} - \frac{\Pr'(Q \cap R)}{\Pr(Q)} \right| < \sqrt{\varepsilon}. \tag{4.5}$$

We also have

$$\begin{aligned} \sum_{R \in \mathcal{R}} \left| \frac{\Pr'(Q \cap R)}{\Pr(Q)} - \frac{\Pr'(Q \cap R)}{\Pr'(Q)} \right| &= \sum_{R \in \mathcal{R}} \Pr'(Q \cap R) \left| \frac{1}{\Pr(Q)} - \frac{1}{\Pr'(Q)} \right| \\ &= \Pr'(Q) \frac{|\Pr(Q) - \Pr'(Q)|}{\Pr(Q)\Pr'(Q)} < \sqrt{\varepsilon}. \end{aligned} \tag{4.6}$$

The lemma now follows from (4.5) and (4.6). □

Remark 4.4. The above proof also holds in case Pr′ is a probability distribution on a larger space that contains $\cup_{Q \in \mathcal{Q}, R \in \mathcal{R}}$. That is, $\sum_{Q \in \mathcal{Q}, R \in \mathcal{R}} \Pr'(Q \cap R)$ can be less than 1.

LEMMA 4.5. (Finer partitioning lemma) *Let* $0 < \varepsilon < 1$. *Suppose* $\mathcal{Q}$ *is a refinement of* $\mathcal{P}^0_{-M}$ *and there is a probability measure* ω *on* $\Sigma^{\mathbb{Z}}$ *such that there is a collection* $\tilde{\mathcal{G}}$ *of 'good atoms' in* $\mathcal{Q}$ *with total* μ*-measure greater than* $1 - \varepsilon^2/16$ *such that for* $\tilde{Q} \in \tilde{\mathcal{G}}$ *we have*

$$\overline{f}_K(\mu(\cdot|\tilde{Q}), \omega) < \varepsilon/4.$$

Then there is a collection $\mathcal{G}$ *of 'good atoms' in* $\mathcal{P}^0_{-M}$ *with total* μ*-measure greater than* $1 - \varepsilon/4$ *such that for* $Q \in \mathcal{G}$ *we have*

$$\overline{f}_K(\mu(\cdot|Q), \omega) < \varepsilon/2.$$

Consequently,

$$\overline{f}_K(\mu(\cdot|Q), \mu(\cdot|R)) < \varepsilon,$$

for $Q, R \in \mathcal{G}$.

Proof. We let $\mathcal{G}$ consist of those atoms Q in $\mathcal{P}^0_{-M}$ such that a subset of Q of measure greater than $(1-\varepsilon/4)\mu(Q)$ is a union of atoms in $\tilde{\mathcal{G}}$. □

The following definition is due to Ornstein [Or70] .

Definition 4.6. A partition $\mathcal{R}$ is said to be ε-independent of a partition $\mathcal{Q}$ (with respect to a given measure ν) if for a collection of atoms $Q \in \mathcal{Q}$ of total ν-measure at least $1-\varepsilon$,

$$\sum_{R\in\mathcal{R}} |\nu(R|Q) - \nu(R)| \leq \varepsilon.$$

In this case we write $\mathcal{R} \perp^{\varepsilon}_{\nu} \mathcal{Q}$. If the measure ν is understood, we may omit the subscript ν.

Remark 4.7. If $\mathcal{Q}$, $\mathcal{R}$, and $\mathcal{S}$ are partitions such that $\mathcal{R}$ refines $\mathcal{S}$ and $\mathcal{R} \perp^{\varepsilon}_{\nu} \mathcal{Q}$, then by the triangle inequality, $\mathcal{S} \perp^{\varepsilon}_{\nu} \mathcal{Q}$.

Remark 4.8. The definition of ε-independence is not symmetric in $\mathcal{Q}$ and $\mathcal{R}$, but $\mathcal{R} \perp^{\varepsilon}_{\nu} \mathcal{Q}$ implies $\mathcal{Q} \perp^{\sqrt{3\varepsilon}}_{\nu} \mathcal{R}$. (See [Sm71].)

LEMMA 4.9. (Epsilon independence lemma) *Let τ_n be as in the Chebyshev application, and let* Pr_n *be the probability distribution on possible initial strings of k'_n n-blocks within $(n+1)$-blocks. Let $\mathcal{Q}$ and $\mathcal{R}$ be partitions of the union of all $(n+1)$-blocks such that $\mathcal{Q}$ is the partition into sets that have the same a_n initial n-blocks and $\mathcal{R}$ is the partition into sets that have the same b_n n-blocks appearing as the (a_n+1)th through (a_n+b_n)th n-blocks of an $(n+1)$-block. Assume that $a_n + b_n \leq k'_n$. Then $\mathcal{R} \perp^{3\sqrt{\tau_n}}_{\mathrm{Pr}_n} \mathcal{Q}$.*

Proof. Let $\widetilde{\mathrm{Pr}}_n$ be the probability distribution for k'_n n-blocks chosen independently from the first $N(n)-1$ n-blocks, with each of these n-blocks equally likely. Then by equation (4.2),

$$\sum_{Q\cap R\in\mathcal{Q}\vee\mathcal{R}} |\mathrm{Pr}_n(Q\cap R) - \widetilde{\mathrm{Pr}}_n(Q\cap R)| \leq \tau_n.$$

Note that we are only summing over those $Q \cap R$ that actually occur as initial strings of some $(n+1)$-block(s). Therefore by Lemma 4.3 and Remark 4.4, for a collection $\mathcal{G}$ of $Q \in \mathcal{Q}$ of total Pr_n measure at least $1-\sqrt{\tau_n}$,

$$\sum_{R\in\mathcal{R}} |\mathrm{Pr}_n(R|Q) - \widetilde{\mathrm{Pr}}_n(R|Q)| \leq 2\sqrt{\tau_n}.$$

But $\widetilde{\mathrm{Pr}}_n(R|Q) = \widetilde{\mathrm{Pr}}_n(R)$, and $\sum_{R\in\mathcal{R}} |\mathrm{Pr}_n(R) - \widetilde{\mathrm{Pr}}_n(R)| \leq \tau_n$. Therefore for $Q \in \mathcal{G}$,

$$\sum_{R\in\mathcal{R}} |\mathrm{Pr}_n(R|Q) - \mathrm{Pr}_n(R)| \leq 2\sqrt{\tau_n} + \tau_n < 3\sqrt{\tau_n}.$$ □

Remark 4.10. If $b_n = 1$, then $\mathrm{Pr}_n(R) = \widetilde{\mathrm{Pr}}_n(R)$ and $\mathcal{R} \perp^{2\sqrt{\tau_n}}_{\mathrm{Pr}_n} \mathcal{Q}$.

The following lemma will be used in the inductive step of the proof that the odometer-based system $\mathbb{E}$ is loosely Bernoulli.

LEMMA 4.11. (Inductive step) *Suppose ν is a probability measure, and $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{Q}_1$, $\mathcal{Q}_2$ are measurable partitions. If $\mathcal{R}_1 \perp_\nu^{\varepsilon_1} \mathcal{Q}_1$ and $\mathcal{R}_1 = \mathcal{Q}_2 \vee \mathcal{R}_2$, where $\mathcal{R}_2 \perp_\nu^{\varepsilon_2} \mathcal{Q}_2$ and $0 < \varepsilon_1, \varepsilon_2 < 1$, then $\mathcal{R}_2 \perp_\nu^{2\sqrt{\varepsilon_1}+2\sqrt{\varepsilon_2}} \mathcal{Q}_1 \vee \mathcal{Q}_2$.*

Proof. Since $\mathcal{R}_2 \perp_\nu^{\varepsilon_2} \mathcal{Q}_2$, for a collection $\mathcal{G}_2$ of atoms Q_2 of $\mathcal{Q}_2$ with $\nu(\cup_{Q_2 \in \mathcal{G}_2} Q_2) \geq 1 - \varepsilon_2$,

$$\sum_{R_2 \in \mathcal{R}_2} |\nu(R_2|Q_2) - \nu(R_2)| \leq \varepsilon_2. \tag{4.7}$$

For at least total ν-measure $1 - \sqrt{\varepsilon_2}$ of the Q_1 in $\mathcal{Q}_1$,

$$\nu(Q_1 \cap (\cup_{Q_2 \in \mathcal{G}_2} Q_2)) \geq (1 - \sqrt{\varepsilon_2})\nu(Q_1). \tag{4.8}$$

That is, $\mathcal{G}_2$ has total $\nu(\cdot|Q_1)$ measure at least $1 - \sqrt{\varepsilon_2}$. Since $\mathcal{R}_1 = \mathcal{Q}_2 \vee \mathcal{R}_2$ and $\mathcal{R}_1 \perp_\nu^{\varepsilon_1} \mathcal{Q}_1$, for a collection of atoms $Q_1 \in \mathcal{Q}_1$ of total ν-measure at least $1 - \varepsilon_1$,

$$\sum_{Q_2 \cap R_2 \in \mathcal{Q}_2 \vee \mathcal{R}_2} |\nu(Q_2 \cap R_2|Q_1) - \nu(Q_2 \cap R_2)| \leq \varepsilon_1. \tag{4.9}$$

Let $\mathcal{G}_1$ be the collection of $Q_1 \in \mathcal{Q}_1$ such that (4.8) and (4.9) hold. Then we have $\nu(\cup_{Q_1 \in \mathcal{G}_1} Q_1) \geq 1 - \varepsilon_1 - \sqrt{\varepsilon_2}$. Fix a choice of $Q_1 \in \mathcal{G}_1$. By Lemma 4.3 applied to $\Pr = \nu(\cdot|Q_1)$ and $\Pr' = \nu$, it follows from (4.9) that for a collection $\mathcal{G}_3 = \mathcal{G}_3(Q_1)$ of atoms $Q_2 \in \mathcal{Q}_2$ with total $\nu(\cdot|Q_1)$ measure at least $1 - \sqrt{\varepsilon_1}$,

$$\sum_{R_2 \in \mathcal{R}_2} |\nu(R_2|Q_1 \cap Q_2) - \nu(R_2|Q_2)| \leq 2\sqrt{\varepsilon_1}. \tag{4.10}$$

Then $\nu(\cup_{Q_2 \in \mathcal{G}_2 \cap \mathcal{G}_3} Q_2|Q_1) \geq 1 - \sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}$, and for $Q_2 \in \mathcal{G}_2 \cap \mathcal{G}_3$, it follows from (4.7) and (4.10) that

$$\sum_{R_2 \in \mathcal{R}_2} |\nu(R_2|Q_1 \cap Q_2) - \nu(R_2)| \leq 2\sqrt{\varepsilon_1} + \varepsilon_2. \tag{4.11}$$

Since $\nu(\cup_{Q_1 \in \mathcal{G}_1} Q_1) \geq 1 - \varepsilon_1 - \sqrt{\varepsilon_2}$ and for any $Q_1 \in \mathcal{G}_1$, $\nu(\cup_{Q_2 \in \mathcal{G}_2 \cap \mathcal{G}_3} Q_2|Q_1) \geq 1 - \sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}$, the collection $\mathcal{G}$ of $Q_1 \cap Q_2 \in \mathcal{Q}_1 \vee \mathcal{Q}_2$ such that (4.11) holds has total ν-measure at least $(1 - \varepsilon_1 - \sqrt{\varepsilon_2})(1 - \sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}) > 1 - 2\sqrt{\varepsilon_1} - 2\sqrt{\varepsilon_2}$. Therefore $\mathcal{R}_2 \perp_\nu^{2\sqrt{\varepsilon_1}+2\sqrt{\varepsilon_2}} \mathcal{Q}_1 \vee \mathcal{Q}_2$. □

THEOREM 4.12. *The odometer-based system $\mathbb{E}$ constructed in this section is loosely Bernoulli.*

Proof. Let $0 < \varepsilon < 1$. Fix a choice of $n \geq 2$ sufficiently large so that $\sum_{j \geq n} \varepsilon_j < \varepsilon^2/100$, $\sum_{j \geq n} \tau_j^{(2^{-j})} < \varepsilon^2/100$, and $(\varepsilon_n k_n)^{-1} < \varepsilon^2/16$. Let $K = (\varepsilon_n k_n + 1)h_n$ and $M \geq h_{n+1}$. We will show that the conclusion of Definition 2.2 holds for these choices of K and M. Let ω be any probability measure on $\Sigma^{\mathbb{Z}}$ such that

$$\omega\{(x_k) : x_1 x_2 \cdots x_{\varepsilon_n k_n h_n} = c_1 c_2 \cdots c_{\varepsilon_n k_n h_n}\} = \Pr_n(c_1 c_2 \cdots c_{\varepsilon_n k_n h_n}),$$

which is the probability that the string $c_1 c_2 \cdots c_{\varepsilon_n k_n h_n}$ comprises the $\varepsilon_n k_n$ initial n-blocks in an $(n+1)$-block. By Lemma 4.5, it suffices to show that for some refinement $\mathcal{Q}$ of

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

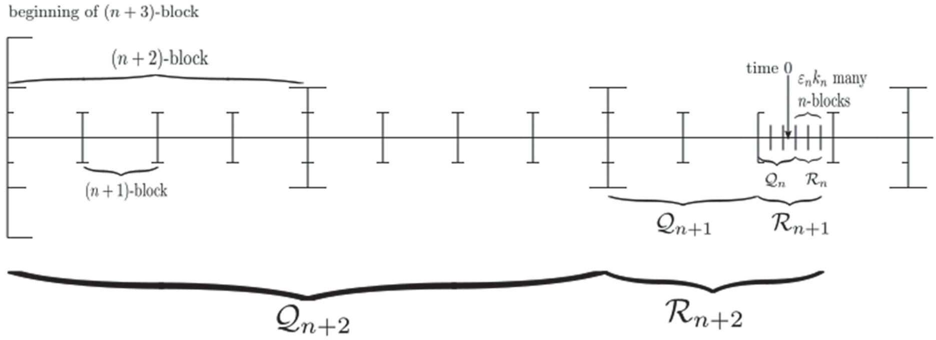

FIGURE 1. For $j = n, n+1, n+2$, the $\mathcal{Q}_j$ and $\mathcal{R}_j$ are partitions into sets according to the $\mathcal{P}$-names that appear in the indicated j-blocks. The actual numbers of j-blocks are much larger than can be depicted in the figure.

$\mathcal{P}^0_{-M}$, there is a collection $\mathcal{G}$ of good atoms of $\mathcal{Q}$ of total measure at least $1 - \varepsilon^2/16$ such that for $Q \in \mathcal{G}$,

$$\overline{f}_K(\mu(\cdot|Q), \omega) < \varepsilon/4. \tag{4.12}$$

Fix a choice of m such that $h_{n+m-1} \geq M$. Let $\mathcal{S}$ be the partition of the space $\Sigma^{\mathbb{Z}}$ into sets that have the same $(n+m)$-block structure, that is, time 0 is in the same position within the $(n+m)$-block.

We now describe the collection $\tilde{\mathcal{G}}$ of good atoms in $\mathcal{S}$. First we eliminate those atoms in $\mathcal{S}$ such that for any $j = 2, \ldots, m$, the deterministic part of the $(n+j)$-block containing time 0 overlaps with the time interval $[1, K]$. (The deterministic part of an $(n+j)$-block consists of the last $\varepsilon_{n+j-1}k_{n+j-1}$ $(n+j-1)$-blocks within the $(n+j)$-block.) We also eliminate those atoms in $\mathcal{S}$ such that time 0 lies in the first $(n+m-1)$-block within the $(n+m)$-block containing time 0. Moreover, we eliminate those atoms in $\mathcal{S}$ such that time 0 occurs in any of the last $2\varepsilon_n k_n$ n-blocks within an $(n+1)$-block. The total measure of the sets eliminated is less than $2(\varepsilon_n + \cdots + \varepsilon_{n+m-1}) < \varepsilon^2/50$. Let $\tilde{\mathcal{G}}$ be the collection of atoms in $\mathcal{S}$ that remain, and fix a choice $S \in \tilde{\mathcal{G}}$. For this S and $j = 0, \ldots, m-1$, let a_{n+j} be the number of $(n+j)$-blocks preceding the $(n+j)$-block containing time 0 within the $(n+j+1)$-block containing time 0. Since $a_{n+m-1} \geq 1$ and $h_{n+m-1} \geq M$, the beginning of the $(n+m)$-block containing time 0 occurs at or before time $-M$. Let $\nu = \nu_S$ be the normalized restriction of μ to S.

Let $\mathcal{Q}_n$ be the partition of S into sets with the same collection of n-blocks appearing in positions 1 to $a_n + 1$ at the beginning of the $(n+1)$-block containing time 0, and let $\mathcal{R}_n$ be the partition of S into sets with the same collection of n-blocks appearing in the $\varepsilon_n k_n$ n-blocks that follow the n-block containing time 0. For $k = 1, 2, \ldots, m-1$, let $\mathcal{Q}_{n+m-k}$ be the partition of S into sets with the same collection of $(n+m-k)$-blocks comprising the a_{n+m-k} $(n+m-k)$-blocks that precede the $(n+m-k)$-block that contains time 0, and let $\mathcal{R}_{n+m-k}$ be the partition of S into sets with $(n+m-k)$-blocks containing time 0 agreeing up to the position of the last symbol in the $\varepsilon_n k_n$ n-blocks that follow the n-block containing time 0 (see Figure 1).

CLAIM. *Let* $\eta_{n+m-k} = 4(\tau_{n+m-1}^{2^{-(k+1)}} + \tau_{n+m-2}^{2^{-k}} + \cdots + \tau_{n+m-(k-1)}^{2^{-3}} + \tau_{n+m-k}^{2^{-2}})$ *for* $k = 1, 2, \ldots, m$. *Then* $\mathcal{R}_{n+m-k} \perp_\nu^{\eta_{n+m-k}} (\mathcal{Q}_{n+m-1} \vee \mathcal{Q}_{n+m-2} \vee \cdots \vee \mathcal{Q}_{n+m-k})$, *for* $k = 1, 2, \ldots, m$.

Proof. According to Lemma 4.9 and Remark 4.7, $\mathcal{R}_{n+m-1} \perp_\nu^{3\sqrt{\tau_{n+m-1}}} \mathcal{Q}_{n+m-1}$. Since $3\sqrt{\tau_{n+m-1}} < 4\tau_{n+m-1}^{2^{-2}} = \eta_{n+m-1}$, the claim holds for $k = 1$. Now suppose the claim holds for some $k = 1, 2, \ldots, m-1$, that is, $\mathcal{R}_{n+m-k} \perp_\nu^{\eta_{n+m-k}} (\mathcal{Q}_{n+m-1} \vee \mathcal{Q}_{n+m-2} \vee \cdots \vee \mathcal{Q}_{n+m-k})$. We have $\mathcal{R}_{n+m-k} = \mathcal{Q}_{n+m-(k+1)} \vee \mathcal{R}_{n+m-(k+1)}$, and by Lemma 4.9 and Remark 4.7, $\mathcal{R}_{n+m-(k+1)} \perp_\nu^{3\sqrt{\tau_{n+m-(k+1)}}} \mathcal{Q}_{n+m-(k+1)}$. Therefore, by Lemma 4.11, $\mathcal{R}_{n+m-(k+1)} \perp_\nu^{\eta} (\mathcal{Q}_{n+m-1} \vee \mathcal{Q}_{n+m-2} \vee \cdots \vee \mathcal{Q}_{n+m-(k+1)})$, where

$$\begin{aligned}\eta &= 2\sqrt{\eta_{n+m-k}} + 2\sqrt{3}\tau_{n+m-(k+1)}^{2^{-2}} \\ &= 4[\tau_{n+m-1}^{2^{-(k+1)}} + \tau_{n+m-2}^{2^{-k}} + \cdots + \tau_{n+m-(k+1)}^{2^{-3}} + \tau_{n+m-k}^{2^{-2}}]^{1/2} + 2\sqrt{3}\tau_{n+m-(k+1)}^{2^{-2}} \\ &\leq 4[\tau_{n+m-1}^{2^{-(k+2)}} + \tau_{n+m-2}^{2^{-(k+1)}} + \cdots + \tau_{n+m-(k+1)}^{2^{-4}} + \tau_{n+m-k}^{2^{-3}} + \tau_{n+m-(k+1)}^{2^{-2}}] \\ &= \eta_{n+m-(k+1)}.\end{aligned}$$

This completes the inductive step. Therefore the claim holds. □

Applying the claim with $k = m$, we obtain $\mathcal{R}_n \perp_\nu^{\eta_n} (\mathcal{Q}_{n+m-1} \vee \cdots \vee \mathcal{Q}_n)$. Note that $\mathcal{Q}_{n+m-1} \vee \cdots \vee \mathcal{Q}_n$ is a refinement of $S \cap \mathcal{P}_{-M}^0$ and $\mathcal{R}_n$ is the partition into the possible $\varepsilon_n k_n$ n-blocks comprising a fraction $\varepsilon_n k_n h_n / K = \varepsilon_n k_n / (\varepsilon_n k_n + 1) > 1 - \varepsilon^2/16$ of the $\mathcal{P}_1^K$ names. Since $\mathcal{R}_n \perp_\nu^{\eta_n} (\mathcal{Q}_{n+m-1} \vee \cdots \vee \mathcal{Q}_n)$, for a set $\mathcal{G}_S$ of atoms $Q \in \mathcal{Q}_{n+m-1} \vee \cdots \vee \mathcal{Q}_n$ of ν-measure at least $1 - \eta_n$, there is a measure-preserving map $\phi_S : (S, \nu) \to (S \cap Q, \nu(\cdot|Q))$ such that on a set of ν-measure at least $1 - \eta_n$, $\phi_S(x)$ is contained in the intersection with Q of that atom of $\mathcal{R}_n$ that contains x, and the $\mathcal{R}_n$ part of the $\mathcal{P}_1^K$ name of x is the same as that of $\phi_S(x)$. Thus, $\overline{f}_K(\nu(\cdot|Q), \omega) < \varepsilon/4$ for $Q \in \mathcal{G}_S$. Finally, we let $\mathcal{G} = \cup_{S \in \tilde{\mathcal{G}}} \mathcal{G}_S$. Then the total μ-measure of atoms in $\mathcal{G}$ is greater than $1 - \varepsilon^2/50 - \eta_n > 1 - \varepsilon^2/16$. Note that for $Q \in \mathcal{G}_S$, $\nu_S(\cdot|Q)$ is the same as $\mu(\cdot|Q)$. Therefore, for $Q \in \mathcal{G}_S$, $\overline{f}_K(\mu(\cdot|Q), \omega) < \varepsilon/4$. Then Lemma 4.5 implies that Definition 2.2 is satisfied. □

4.2. *Non-loosely Bernoulli circular system arising from positive-entropy loosely Bernoulli odometer-based system.*

THEOREM 4.13. *If* $\mathbb{E}$ *is the positive-entropy loosely Bernoulli odometer-based system constructed in the previous section, and* $\mathcal{F}$ *is the map from odometer-based systems to circular systems defined in §3, then* $\mathcal{F}(\mathbb{E})$ *is non-loosely Bernoulli.*

We will prove $\mathcal{F}(\mathbb{E})$ is not loosely Bernoulli by proving that the condition in Lemma 3.5 does not hold.

In the construction of $\mathcal{W}_{n+1}$ words in the circular system, we have many repetitions of $\mathcal{W}_n$ words. To get lower bounds on the $\overline{f}$ distance between $\mathcal{W}_{n+1}$ words, we will make use of Definition 4.14 and Lemma 4.15 below.

Definition 4.14. If $b_1 b_2 \ldots b_s$ is any string of s symbols, let $\mathcal{T}(b_1 b_2 \ldots b_s)$ denote the collection of all finite consecutive substrings of $(b_1 b_2 \ldots b_s)^t$ for any $t \geq 1$.

LEMMA 4.15. (Repeated substring matching lemma) *Suppose $a_1 \ldots a_r$ and $b_1 \ldots b_s$ are strings of symbols. Then for any $\ell, \tilde{\ell} \geq 1$,*

$$\overline{f}((a_1 \ldots a_r)^{\ell}, (b_1 \ldots b_s)^{\tilde{\ell}}) \geq \overline{f}(a_1 \ldots a_r, \mathcal{T}(b_1 \ldots b_s)),$$

where the right-hand side denotes the infimum of the $\overline{f}$ distance from $a_1 \ldots a_r$ to any element of $\mathcal{T}(b_1 \ldots b_s)$.

Proof. Apply Property 2.6 with $x_1 = x_2 = \cdots = x_\ell = a_1 a_2 \ldots a_r$. □

The estimate below is proved in [**Ro80**] by a simple argument using just the binomial theorem. A similar estimate can be obtained from Stirling's formula.

LEMMA 4.16. *If m is a positive integer and $0 < \sigma < 1$, then we have the following inequality for the binomial coefficient:*

$$\binom{m}{\lfloor \sigma m \rfloor} < 2^{3m\sqrt{\sigma}}.$$

We make the following choices of parameters (in addition to those already described in the previous section). Let $b_n = 2^{-(n+10)}$, $\ell_n > 2^{n+10}$. Recall that $\varepsilon_n < 2^{-(n+12)}$.

LEMMA 4.17. (Inductive step in Rothstein's argument to obtain a lower bound on the $\overline{f}$ distance) *For $n \in \mathbb{N}$ and $0 < \delta_n < 1$ define*

$$\xi_n = \xi_n(\delta_n) := 2^8 (3\delta_n)^{((b_n/2)-\varepsilon_n)(1-\varepsilon_n)} 2^{3(1-\varepsilon_n)\sqrt{(b_n/2)}}. \tag{4.13}$$

Fix a particular $n \in \mathbb{N}$ and suppose δ_n is sufficiently small that $\xi_n < 1$. Assume that for at least $(1-\delta_n)$ of the n-blocks in $\mathcal{W}_n$ the $\overline{f}$ distance from the n-block to any specific n-block or substrings of its extensions (as in Definition 4.14) is greater than a_n, where $0 < a_n < 1/3$. Then for k_n sufficiently large, depending only on parameters with subscript n, there exists δ_{n+1} such that $\xi_{n+1} = \xi_{n+1}(\delta_{n+1}) < 1$, and for at least $(1-\delta_{n+1})$ of the $(n+1)$-blocks in $\mathcal{W}_{n+1}$ the $\overline{f}$ distance from the $(n+1)$-block to any specific $(n+1)$-block or substrings of its extensions is greater than $a_{n+1} := a_n(1-b_n) - 15\ell_n^{-1}$.

Proof. Fix a particular choice

$$\prod_{i=0}^{q_n-1} \prod_{j=0}^{k_n-1} (b^{q_n-j_i} \mathcal{B}_j^{\ell_n-1} e^{j_i})$$

of $(n+1)$-block in $\mathcal{W}_{n+1}$, where $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_{k_{n-1}}$ are n-blocks in $\mathcal{W}_n$. Let $\mathcal{T}_0 := \mathcal{T}(\prod_{i=0}^{q_n-1} \prod_{j=0}^{k_n-1} (b^{q_n-j_i} \mathcal{B}_j^{\ell_n-1} e^{j_i}))$. We will prove that the inequality

$$\overline{f}\left(\prod_{i=0}^{q_n-1} \prod_{j=0}^{k_n-1} (b^{q_n-j_i} \mathcal{A}_j^{\ell_n-1} e^{j_i}), \mathcal{T}_0 \right) > a_n(1-b_n) - 15\ell_n^{-1} \tag{4.14}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

holds for at least $(1-\delta_{n+1})$ of the $(n+1)$-blocks $\prod_{i=0}^{q_n-1}\prod_{j=0}^{k_n-1}(b^{q_n-j_i}\mathcal{A}_j^{\ell_n-1}e^{j_i})$ in $\mathcal{W}_{n+1}$, where δ_{n+1} will be specified later in the proof. The b's and e's that are newly added in constructing $(n+1)$-blocks from n-blocks in the circular system make up a fraction ℓ_n^{-1} of the symbols in any $(n+1)$-block. We may assume that the smallest $\overline{f}$ distance between $\prod_{i=0}^{q_n-1}\prod_{j=0}^{k_n-1}(b^{q_n-j_i}\mathcal{A}_j^{\ell_n-1}e^{j_i})$ and any element of $\mathcal{T}_0$ occurs for an element of $\mathcal{T}_0$ of length at least $q_n^2 k_n\ell_n/2$, because otherwise it follows from (2.5) that the $\overline{f}$ distance in (4.14) is greater than 1/3. For such an element of $\mathcal{T}_0$, the number of newly added b's and e's is a fraction less than $2\ell_n^{-1}$ of the length of that element. Therefore by Property 2.5 and Lemma 4.15, to obtain (4.14) it suffices to show that

$$\overline{f}(\mathcal{A}_1^{\ell_n-1}\mathcal{A}_2^{\ell_n-1}\ldots\mathcal{A}_{k_n}^{\ell_n-1},\mathcal{T}(\mathcal{B}_1^{\ell_n-1}\mathcal{B}_2^{\ell_n-1}\ldots\mathcal{B}_{k_n}^{\ell_n-1}))>a_n(1-b_n)-9\ell_n^{-1}, \quad (4.15)$$

without repeating the strings q_n times. Suppose to the contrary of (4.15) that

$$\begin{aligned}&\overline{f}(\mathcal{A}_1^{\ell_n-1}\mathcal{A}_2^{\ell_n-1}\ldots\mathcal{A}_{k_n}^{\ell_n-1},\mathcal{T}(\mathcal{B}_1^{\ell_n-1}\mathcal{B}_2^{\ell_n-1}\ldots\mathcal{B}_{k_n}^{\ell_n-1}))\\ &\quad\le a_n(1-b_n)-9\ell_n^{-1}<(a_n-9\ell_n^{-1})(1-b_n)<1/3,\end{aligned} \quad (4.16)$$

for some k_n n-blocks $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_{k_n}$. We will show that this happens for at most δ_{n+1} of the $(n+1)$-blocks in $\mathcal{W}_{n+1}$. Choose a match between the $\mathcal{A}$-string and a $\mathcal{B}$-string that realizes the $\overline{f}$ distance in (4.16). For each substring $\mathcal{A}_i^{\ell_n-1}$ of the $\mathcal{A}$-string, let f_i be the $\overline{f}$ distance between $\mathcal{A}_i^{\ell_n-1}$ and the corresponding part of the $\mathcal{B}$-string. Let v_i be the ratio of the number of symbols in $\mathcal{A}_i^{\ell_n-1}$ plus the number of symbols in the corresponding part of the $\mathcal{B}$-string to the total length of the $\mathcal{A}$- and $\mathcal{B}$-strings. Then by Property 2.6, the $\overline{f}$ distance for the entire strings is $\sum_{i=1}^{k_n} f_i v_i$, that is, a weighted average of the f_i with weights v_i. Since this weighted average is less than $(a_n-9\ell_n^{-1})(1-b_n)$, the weights v_i for those f_i with $f_i<a_n-9\ell_n^{-1}<1/3$ must have sum at least b_n. For these weights v_i, Property 2.7 and the assumptions that the $\overline{f}$ distance in (4.16) and f_i are both less than 1/3 imply that $v_i<2k_n^{-1}$. Thus there must be at least $(b_n/2)k_n$ indices i such that $f_i<a_n-9\ell_n^{-1}$. Then for at least a fraction $(b_n/2)-\varepsilon_n$ of the indices $i\in\{1,2,\ldots,k_n'\}$ the $\overline{f}$ distance between $\mathcal{A}_i^{\ell_n-1}$ and the corresponding part of the $\mathcal{B}$-block is less than $a_n-9\ell_n^{-1}$. For each such i, let $\sigma(i)$ be the first index such that the part of the $\mathcal{B}$-block corresponding to $\mathcal{A}_i^{\ell_n-1}$ starts with a substring of a $\mathcal{B}_{\sigma(i)}^{\ell_n-1}$. Then $\mathcal{A}_i^{\ell_n-1}$ may correspond just to a substring of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}$ or to a substring of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}\mathcal{B}_{\sigma(i)+1}^{\ell_n-1}$ or to a substring of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}\mathcal{B}_{\sigma(i)+1}^{\ell_n-1}\mathcal{B}_{\sigma(i)+2}^{\ell_n-1}$. Here the addition in the subscripts is modulo k_n. Any correspondence between $\mathcal{A}_i^{\ell_n-1}$ and strings of four or more $\mathcal{B}_j^{\ell_n-1}$ would lead to the $\overline{f}$ distance between $\mathcal{A}_i^{\ell_n-1}$ and the corresponding part of the $\mathcal{B}$-string being greater than 1/3, and therefore we may disregard this possibility. The number of ways of choosing the $\sigma(i)$'s and deciding whether to use just $\sigma(i)$ or to continue with just $\sigma(i)+1$ or to continue with both $\sigma(i)+1$ and $\sigma(i)+2$ is at most $k_n\binom{7k_n}{k_n}$. Here we estimate $7k_n>3(2k_n+1)$, which is an upper bound on the number of possible $\sigma(i)$ combinations corresponding to $\mathcal{A}_i^{\ell_n-1}$ strings, allowing for the $\mathcal{B}$-string to be up to twice the length of the $\mathcal{A}$-string (and thus contained in at most $2k_n+1$ consecutive $\mathcal{B}_j^{l_n-1}$ strings) and allowing for the three choices: just $\sigma(i)$, just $\sigma(i)$ and $\sigma(i)+1$, and all of $\sigma(i)$, $\sigma(i)$+1, and $\sigma(i)+2$. The additional factor of k_n in front is due to being able

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

to start the $\mathcal{B}$-string with $\sigma(1)$ being any of $1, 2, \ldots, k_n$. According to the estimate on binomial coefficients given in Lemma 4.16, $\binom{7k_n}{k_n} \leq 2^{21k_n\sqrt{1/7}} \leq 2^{8k_n}$. When we match $\mathcal{A}_i^{l_n-1}$ with a substring of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}\mathcal{B}_{\sigma(i)+1}^{\ell_n-1}\mathcal{B}_{\sigma(i)+2}^{\ell_n-1}$, we divide $\mathcal{A}_i^{\ell_n-1}$ into as many as three substrings according to which part corresponds to each of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}$, $\mathcal{B}_{\sigma(i)+1}^{\ell_n-1}$, or $\mathcal{B}_{\sigma(i)+2}^{\ell_n-1}$. By removing at most a fraction $4(\ell_n-1)^{-1}$ of the symbols in each $\mathcal{A}_i^{\ell_n-1}$ string, we may assume that full $\mathcal{A}_i$ strings correspond to each corresponding substring of $\mathcal{B}_{\sigma(i)}^{\ell_n-1}$, $\mathcal{B}_{\sigma(i)+1}^{\ell_n-1}$, and $\mathcal{B}_{\sigma(i)+2}^{\ell_n-1}$. By Property 2.5 this removal will increase the $\overline{f}$ distance from $\mathcal{A}_i^{\ell_n-1}$ to the corresponding part of the $\mathcal{B}$-string by at most $9\ell_n^{-1}$. Thus by Lemma 4.15, for a fraction of at least $(b_n/2)-\varepsilon_n$ of the indices in $\{1, 2, \ldots, k_n'\}$, at least one of the three $\overline{f}$ distances from $\mathcal{A}_i$ to a string in $\mathcal{T}(\mathcal{B}_{\sigma(i)})$ or from $\mathcal{A}_i$ to a string in $\mathcal{T}(\mathcal{B}_{\sigma(i)+1})$ or from $\mathcal{A}_i$ to a string in $\mathcal{T}(\mathcal{B}_{\sigma(i)+2})$ must be less than a_n. By assumption, the probability that the n-block $\mathcal{A}_i$ satisfies at least one of these three conditions is less than $3\delta_n$. Thus if the first k_n' n-blocks in the odometer $(n+1)$-block corresponding to the $\mathcal{A}$-string were selected independently, then the probability δ_{n+1} that

$$\overline{f}(\mathcal{A}_1^{\ell_n-1}\mathcal{A}_2^{\ell_n-1}\ldots\mathcal{A}_{k_n}^{\ell_n-1}, \mathcal{T}(\mathcal{B}_1^{\ell_n-1}\mathcal{B}_2^{\ell_n-1}\ldots\mathcal{B}_{k_n}^{\ell_n-1})) < a_n - 9\ell_n^{-1}$$

would be less than $k_n 2^{8k_n}(3\delta_n)^{((b_n/2)-\varepsilon_n)k_n'}\binom{k_n'}{(b_n/2)k_n'}$. Since the selection of the k_n' n-blocks is not quite independent, we apply (4.1), and multiply our bound on the probability by $(1-\tau_n)^{-1}$. Therefore from Lemma 4.16, we obtain $\delta_{n+1} < k_n(1-\tau_n)^{-1}\xi_n^{k_n}$. By assumption $\xi_n < 1$. Thus we can choose k_n sufficiently large so that δ_{n+1} is sufficiently small to imply $\xi_{n+1} < 1$. For at least $1-\delta_{n+1}$ of the $(n+1)$-blocks in $\mathcal{W}_{n+1}$ the $\overline{f}$ distance from the $(n+1)$-block to any specific $(n+1)$-block or substrings of its extensions is greater than $a_{n+1} := a_n(1-b_n) - 15\ell_n^{-1}$. □

Proof of Theorem 4.13. The inductive step is contained in Lemma 4.17. For the base case, we recall that 0-blocks are single symbols $1, 2, \ldots, |\Sigma|$. Choose $0 < \delta_0 < 1$ so that $\xi_0(\delta_0) < 1$. Then require $|\Sigma|$ to be sufficiently large that $\delta_0 > |\Sigma|^{-1}$. We let $a_1 = 1/4$. According to the recursive formula for a_n, we have $a_n > 1/8$ for all n. Thus, if $\varepsilon = 1/8$, the condition in Lemma 3.5 is not satisfied. Therefore $\mathcal{F}(\mathbb{E})$ is not loosely Bernoulli. □

5. *Zero-entropy example*

In this section we prove the following theorem, which gives a zero-entropy version of the example constructed in §§4.1 and 4.2.

THEOREM 5.1. *There exist circular coefficients (l_n) and a loosely Bernoulli odometer-based system $\mathbb{K}$ of zero measure-theoretic entropy with uniform and uniquely readable construction sequence such that $\mathcal{F}(\mathbb{K})$ is not loosely Bernoulli.*

Outline of the proof of Theorem 5.1. In §5.4 we give a precise description of the inductive building process of the uniform and uniquely readable construction sequence for the odometer-based system $\mathbb{K}$ such that $\mathbb{K}$ will be loosely Bernoulli but $\mathcal{F}(\mathbb{K})$ will not be loosely Bernoulli. The creation of this sequence relies on two mechanisms. On the one hand, we will use what we will call the *Feldman mechanism* presented in §5.2. This will

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

allow us to produce an arbitrarily large number of blocks that in the circular system remain almost as far apart in $\overline{f}$ as the building blocks. In particular, we can produce sufficiently many blocks to apply the second mechanism, the so-called *shifting mechanism* introduced in §5.3 (see Figure 2 for a sketch of its idea). This mechanism requires not only sufficiently many n-words to start with but also a sufficiently large number of stages p. Then we can produce $(n+p)$-words in our construction sequence in such a way that these are close to each other in the $\overline{f}$ metric and that the corresponding blocks in the circular construction sequence stay apart from each other in the $\overline{f}$ metric. Both mechanisms will make use of *Feldman patterns* for which we prove a general statement in §5.1. □

5.1. *Feldman patterns for blocks.* In [**Fe76**] Feldman constructed the first example of an ergodic zero-entropy automorphism that is not loosely Bernoulli. The construction is based on the observation that no pair of the strings

$$abababab$$
$$aabbaabb$$
$$aaaabbbb$$

can be matched very well. We use his construction of blocks (which we call *Feldman patterns*) frequently in our two mechanisms in §§5.2 and 5.3. The basic Feldman patterns are displayed in Lemma 5.8. In applying these patterns, we substitute blocks of symbols for the individual symbols to produce a large number of strings that are almost as far apart in $\overline{f}$ as their building blocks.

Our presentation of the Feldman patterns is similar to the one in [**ORW82**], but we apply the patterns in the odometer system and then examine the $\overline{f}$ distance between strings in the corresponding circular system. We also allow the consideration of different families of strings and a preliminary concatenation of blocks (which will prove useful when dealing with grouped blocks in §5.3). While we make a statement about substrings of different Feldman patterns from either the same or different families in Proposition 5.10, we focus on the situation of the same Feldman pattern but different families in Lemma 5.11.

In order to obtain lower bounds on the $\overline{f}$ distance between strings that are built from blocks of symbols (as in Proposition 5.4 and Corollary 5.5), it is convenient to introduce a notion of approximate $\overline{f}$ distance that we call $\tilde{f}$.

Definition 5.2. If (i, j) and $(i', j') \in \mathbb{N} \times \mathbb{N}$, then we define $(i, j) \preceq (i', j')$ if $i \leq i'$ and $j \leq j'$. If $(i, j) \preceq (i', j')$ and $(i, j) \neq (i', j')$, then we say $(i, j) \prec (i', j')$. An *approximate match* between two strings of symbols $a_1 a_2 \ldots a_n$ and $b_1 b_2 \ldots b_m$ from a given alphabet Σ is a collection $\tilde{I}$ of pairs of indices (i_s, j_s), $s = 1, \ldots, r$, such that the following conditions hold.

- $(1, 1) \preceq (i_1, j_1) \prec (i_2, j_2) \prec \cdots \prec (i_r, j_r) \preceq (n, m)$.
- $a_{i_s} = b_{j_s}$ for $s = 1, 2, \ldots, r$.
- If $(i, j) \in \tilde{I}$, then there exist $s, t \in \{1, \ldots, r-2\}$ such that $\{s' : (i, s') \in \tilde{I}\} \subset \{s, s+1, s+2\}$ and $\{t' : (t', j) \in \tilde{I}\} \subset \{t, t+1, t+2\}$.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

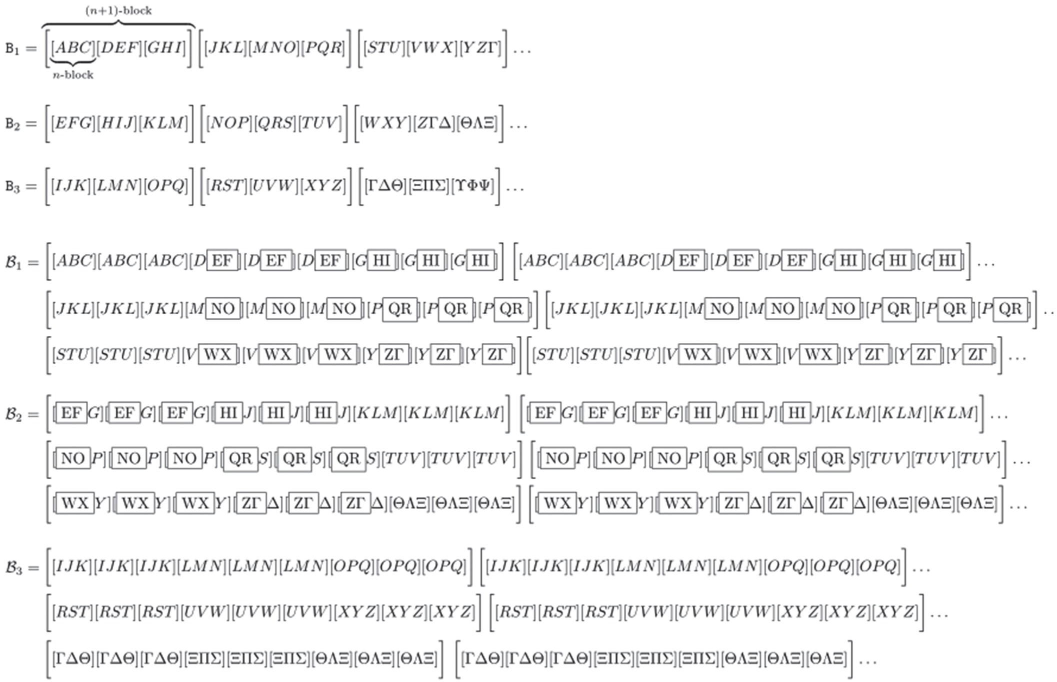

FIGURE 2. Heuristic representation of two stages of the shifting mechanism. Parts of three $(n+2)$-blocks $\mathtt{B}_1$, $\mathtt{B}_2$, $\mathtt{B}_3$ in the odometer-based system and parts of their images $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$ under the circular operator are represented. The marked letters indicate a best possible $\overline{f}$ match between $\mathcal{B}_1$ and $\mathcal{B}_2$ with a fit of approximately $\left(1-\frac{1}{3}\right)^2$ (ignoring spacers and boundary effects) while the blocks $\mathtt{B}_1$ and $\mathtt{B}_2$ have a very good fit in the odometer-based system.

Then

$$\tilde{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m)$$
$$= \max\left(0, 1 - \frac{2\sup\{|\tilde{I}| : \tilde{I} \text{ is an approx. match between } a_1a_2\ldots a_n \text{ and } b_1b_2\ldots b_m\}}{n+m}\right).$$

Clearly $\overline{f}(a_1a_2\ldots a_n, b_1, b_2\ldots b_m) \geq \tilde{f}(a_1a_2\ldots a_n, b_1, b_2\ldots b_m)$. Moreover, if every three consecutive symbols a_s, a_{s+1}, a_{s+2} are distinct and every three consecutive symbols b_s, b_{s+1}, b_{s+2} are distinct, then $\overline{f}(a_1a_2\ldots a_n, b_1, b_2\ldots b_m) = \tilde{f}(a_1a_2\ldots a_n, b_1, b_2\ldots b_m)$. Note that it is possible for $\tilde{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m)$ to be zero, even for $a_1a_2\ldots a_n \neq b_1b_2\ldots b_m$; for example, $\tilde{f}(11000, 11100) = 0$.

LEMMA 5.3. *Suppose* $\overline{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m) = 1-\varepsilon$, *where* $0 \leq \varepsilon < \frac{1}{3}$. *Then* $\tilde{f}(a_1a_2\ldots a_n, b_1b_2, \ldots b_m) \geq 1-3\varepsilon$.

Proof. Suppose $\tilde{I} = \{(i_1, j_1), \ldots, (i_r, j_r)\}$ is an approximate match between $a_1a_2\ldots a_n$ and $b_1b_2\ldots b_m$, as in Definition 5.2. We construct a match I with $|I| \geq |\tilde{I}|/3$. Select the first element (i_1, j_1) in $\tilde{I}$ as an element of I and discard those at most two other elements of $\tilde{I}$ that have the same first coordinate or the same second coordinate as (i_1, j_1). Then select the next element (i_s, j_s) in $\tilde{I}$ that has not already been selected for I or discarded. We again retain this (i_s, j_s) for I and discard those at most two other elements of $\tilde{I}$ that have not been discarded previously and that have the same first coordinate or the same second coordinate as (i_s, j_s). Continue in this way until all elements of $\tilde{I}$ have either been discarded or retained for I. Then $|\tilde{I}| \leq 3|I|$. □

PROPOSITION 5.4. (Symbol by block replacement) *Suppose* $A_{a_1}, A_{a_2}, \ldots, A_{a_n}$ *and* $B_{b_1}, B_{b_2}, \ldots, B_{b_m}$ *are blocks of symbols, with each block of length* L. *Assume that* $\alpha \in (0, \frac{1}{7})$, $\beta \in [0, \frac{1}{7})$, $\alpha \geq \beta$, $R \geq 1$, *and for all substrings* C *and* D *consisting of consecutive symbols from* A_{a_i} *and* B_{b_j}, *respectively, with* $|C|, |D| \geq L/R$ *we have*

$$\overline{f}(C, D) \geq \alpha \quad \text{if } a_i \neq b_j$$

and

$$\overline{f}(C, D) \geq \beta \quad \text{if } a_i = b_j.$$

Let $\tilde{f} = \tilde{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m)$. *Then*

$$\overline{f}(A_{a_1}A_{a_2}\ldots A_{a_n}, B_{b_1}B_{b_2}\ldots B_{b_m}) \geq \alpha\left(\tilde{f} - \frac{3}{R}\right) + \beta(1-\tilde{f}) > \alpha\tilde{f} + \beta(1-\tilde{f}) - \frac{1}{R}$$
$$> \alpha - (1-\tilde{f}) + \beta(1-\tilde{f}) - \frac{1}{R}.$$

Proof. We may assume $R \geq 3$; otherwise the conclusion holds trivially. We may also assume that $n \leq m$. We decompose $A_{a_1}A_{a_2}\ldots A_{a_n}$ into the substrings $A_{a_1}, A_{a_2}, \ldots, A_{a_n}$ and decompose $B_{b_1}B_{b_2}\ldots B_{b_m}$ into corresponding substrings $\tilde{B}_1, \tilde{B}_2, \ldots, \tilde{B}_n$ according to a best possible match between the two strings. Then we decompose each A_{a_i} into at most three further substrings $A_{a_i,0}, A_{a_i,1}, A_{a_i,2}$ corresponding to substrings $B_{i,b_{j_i}}, B_{i,b_{j_i+1}}, B_{i,b_{j_i+2}}$ of $\tilde{B}_i$ that lie entirely in $B_{b_{j_i}}, B_{b_{j_i+1}}, B_{b_{j_i+2}}$ respectively, to obtain a best possible match between A_{a_i} and $\tilde{B}_i$. We will apply Property 2.6 to this decomposition. We may ignore any A_{a_i} and corresponding $\tilde{B}_i$ for which $\tilde{B}_i$ fails to lie in three consecutive blocks $B_{b_{j_i}}, B_{b_{j_i+1}}, B_{b_{j_i+2}}$, because in this case it follows from equation (2.6) in Property 2.7, that $\overline{f}(A_{a_i}, \tilde{B}_i) > \frac{1}{7} > \alpha$. For the same reason we may also ignore any B_{b_j} whose corresponding substring in $A_{a_1}A_{a_2}\ldots A_{a_n}$ fails to lie in three consecutive blocks $A_{a_i}, A_{a_{i+1}}, A_{a_{i+2}}$. We let $\tilde{I}$ consist of those remaining pairs $(i', j') \in \cup_{i=1}^{n}\{(i, j_i), (i, j_i+1), (i, j_i+2)\}$ such that $a_{i'} = b_{j'}$. Then $\tilde{I}$ gives an approximate match between $a_1a_2\ldots a_n$ and $b_1b_2\ldots b_m$. The symbols in all $A_{a_{i'}} \cup B_{b_{j'}}$ for all $(i', j') \in \tilde{I}$ form a fraction of at most $\min(1, (2|\tilde{I}|/(n+m))) \leq 1-\tilde{f}$ of the total number of symbols in $A_{a_1}A_{a_2}\ldots A_{a_n}$ and $B_{b_1}B_{b_2}\ldots B_{b_m}$.

For any substring $A_{a_i,s}$ corresponding to a substring $B_{i,b_{j_i+s}}$ such that the length of at least one of these two substrings is less than L/R, the other substring has length less than $4L/3R$ (unless $\overline{f}(A_{a_i,s}, B_{i,b_{j_i+s}}) > \alpha$, a case that we again ignore). Moreover, if $A_{a_i,0}$ and $A_{a_i,2}$ are both non-empty, then $B_{i,b_{j_i+1}} = B_{j_i}$ and $|A_{a_i,1}| \geq (3/4)L \geq L/R$ (unless $\overline{f}(A_{a_i,1}, B_{j_i}) > \alpha$). Hence if we eliminate substrings $A_{a_i,s}$ and $B_{i,b_{j_i+s}}$ such that the length of at least one of the two substrings is less than L/R, we eliminate at most $14Ln/3R$ symbols from the two strings $A_{a_1}A_{a_2}\ldots A_{a_n}$ and $B_{b_1}B_{b_2}\ldots B_{b_m}$ whose total combined length is $L(n+m) \geq 2Ln$. Thus the fraction of symbols that are eliminated due to such short substrings is at most $7/3R < 3/R$. For those pairs (i, j_i+s), $s \in \{0, 1, 2\}$, such that neither of the strings $A_{a_i,s}$ and $B_{b_{j_i+s}}$ has length less than L/R, it follows from the hypothesis that

$$\overline{f}(A_{a_i,s}, B_{b_{j_i+s}}) \geq \alpha \quad \text{if } (i, j_i+s) \notin \tilde{I},$$

and

$$\overline{f}(A_{a_i,s}, B_{b_{j_i+s}}) \geq \beta \quad \text{if } (i, j_i+s) \in \tilde{I}.$$

Therefore by Property 2.6, we obtain

$$\begin{aligned}\overline{f}(A_{a_1}A_{a_2}\ldots A_{a_n}, B_{b_1}B_{b_2}\ldots B_{b_m}) &\geq \alpha\left(1-(1-\tilde{f})-\frac{3}{R}\right)+\beta(1-\tilde{f})\\ &> \alpha\tilde{f}+\beta(1-\tilde{f})-\frac{1}{R}\\ &\geq \alpha-(1-\tilde{f})+\beta(1-\tilde{f})-\frac{1}{R}.\end{aligned}$$

□

COROLLARY 5.5 *Suppose* $A_{a_1}, A_{a_2}, \ldots, A_{a_n}$ *and* $A_{b_1}, A_{b_2}, \ldots, A_{b_m}$ *are blocks of symbols with each block of length* L. *Assume that* $\alpha \in \left(0, \frac{1}{7}\right)$, $R \geq 1$, *and*

$$\overline{f}(C, D) \geq \alpha$$

for all substrings C *and* D *consisting of consecutive symbols from* A_{a_i} *and* A_{b_j} *respectively, where* $a_i \neq b_j$, *and* $|C|, |D| \geq L/R$. *Then for* $\tilde{f} = \tilde{f}(a_1a_2\ldots a_n, b_1b_2\ldots b_m)$, *we have*

$$\overline{f}(A_{a_1}A_{a_2}\ldots A_{a_n}, A_{b_1}A_{b_2}\ldots A_{b_m}) > \alpha\tilde{f}-\frac{1}{R} \geq \alpha-(1-\tilde{f})-\frac{1}{R}.$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

Proof. Let $B_{b_j} = A_{b_j}$ for $j = 1, 2, \ldots, m$, and $\beta = 0$ in Proposition 5.4. □

Remark 5.6. In fact, Corollary 5.5 and the case $\beta = 0$ of Proposition 5.4 hold with $\tilde{f}$ replaced by $\overline{f}(a_1 a_2 \ldots a_n, b_1 b_2 \ldots b_m)$. We omit the proof because these versions are not needed for our results.

Remark 5.7. The following lemma is essentially the same as [**Fe76**, Theorem 4] and [**ORW82**, Proposition 1.1 in Ch. 10]. The proof is also essentially the same, but the estimates are more suited to our applications.

LEMMA 5.8. *Suppose $a_1, a_2, \ldots, a_N$ are distinct symbols in Σ. Let*

$$\begin{aligned} B_1 &= (a_1^{N^2} a_2^{N^2} \ldots a_N^{N^2})^{N^{2M}}, \\ B_2 &= (a_1^{N^4} a_2^{N^4} \ldots a_N^{N^4})^{N^{2M-2}}, \\ \vdots\ &\ \ \vdots \\ B_M &= (a_1^{N^{2M}} a_2^{N^{2M}} \ldots a_N^{N^{2M}})^{N^2}. \end{aligned}$$

Suppose B and $\overline{B}$ are strings of consecutive symbols in B_j and B_k respectively, where $|B| \geq N^{2M+2}$, $|\overline{B}| \geq N^{2M+2}$, and $j \neq k$. Assume that $N \geq 20$ and $M \geq 2$. Then

$$\overline{f}(B, \overline{B}) > 1 - \frac{4}{\sqrt{N}} \quad \text{and} \quad \tilde{f}(B, \overline{B}) > 1 - \frac{12}{\sqrt{N}}.$$

Proof. We may assume that $j > k$. By removing fewer than $2N^{2j}$ symbols from the beginning and end of B, we can decompose the remaining part of B into strings $C_1, C_2, \ldots, C_r$ each of the form $a_i^{N^{2j}}$. Since $2N^{2j} \leq (|B| + |\overline{B}|)/N^2$, it follows from Property 2.5 that removing these symbols increases the $\overline{f}$ distance between B and $\overline{B}$ by less than $2/N^2$. Let $\overline{C}_1, \overline{C}_2, \ldots, \overline{C}_r$ be the decomposition of $\overline{B}$ into substrings corresponding to $C_1, C_2, \ldots, C_r$ under a best possible match between $C_1 C_2 \ldots C_r$ and $\overline{B}$.

Let $i \in \{1, 2, \ldots, r\}$.

Case 1. $|\overline{C}_i| < (3/2\sqrt{N})|C_i|$. Then by Property 2.7, $\overline{f}(C_i, \overline{C}_i) > 1 - (3/\sqrt{N})$.

Case 2. $|\overline{C}_i| \geq (3/2\sqrt{N})|C_i| = (3/2)N^{2j-(1/2)}$. A cycle $a_1^{N^{2k}} a_2^{N^{2k}} \ldots a_N N^{2k}$ in B_k has length at most N^{2j-1}. Therefore $\overline{C}_i$ contains at least $\lfloor 3\sqrt{N}/2 \rfloor - 1 > \sqrt{N}$ complete cycles. Thus deleting any partial cycles at the beginning and end of $\overline{C}_i$ would increase the $\overline{f}$ distance between C_i and $\overline{C}_i$ by less than $2/\sqrt{N}$. On the rest of $\overline{C}_i$, only $1/N$ of the symbols in $\overline{C}_i$ can match the symbol in C_i. Thus $\overline{f}(C_i, \overline{C}_i) > 1 - (2/\sqrt{N}) - (4/N) > 1 - (3/\sqrt{N})$.

Therefore, by Property 2.6, $\overline{f}(C_1 C_2 \ldots C_r, \overline{C}_1, \overline{C}_2 \ldots, \overline{C}_r) > 1 - (3/\sqrt{N})$. By Lemma 5.3 the claimed $\tilde{f}$ inequality holds as well. □

Remark 5.9. If we replace each symbol a_i by a constant number of repetitions a_i^l, then the same conclusion still holds for substrings of length at least lN^{2M+2}.

PROPOSITION 5.10. *Let $\alpha \in \left(0, \frac{1}{7}\right)$, $n \in \mathbb{N}$, and $K, R, S, N, M \in \mathbb{N} \setminus \{0\}$ with $N \geq 20$ and $M \geq 2$. For $1 \leq s \leq S$, let $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ be a family of strings, where each $\mathtt{A}_j^{(s)}$*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

is a concatenation of K n-blocks. Assume that for all $0 \le i_1, i_2 < q_n$, all $1 \le s_1, s_2 \le S$ and all $j_1, j_2 \in \{1, \ldots, N\}$, $j_1 \ne j_2$, we have $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ for all sequences $\mathcal{A}$, $\overline{\mathcal{A}}$ each consisting of at least $K l_n q_n / R$ consecutive symbols from $\mathcal{C}_{n,i_1}(\mathtt{A}_{j_1}^{(s_1)})$ and $\mathcal{C}_{n,i_2}(\mathtt{A}_{j_2}^{(s_2)})$, respectively.

Then for $1 \le s \le S$, we can construct a family of strings $\mathtt{B}_1^{(s)}, \ldots, \mathtt{B}_M^{(s)}$ (of equal length $N^{2M+3} K \mathtt{h}_n$ and containing each block $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ exactly N^{2M+2} times) such that for all $0 \le i_1, i_2 < q_n$, all $1 \le s_1, s_2 \le S$, all $j, k \in \{1, \ldots, M\}$, $j \ne k$, and all sequences $\mathcal{B}$, $\overline{\mathcal{B}}$ of at least $N^{2M+2} l_n K q_n$ consecutive symbols from $\mathcal{C}_{n,i_1}(\mathtt{B}_j^{(s_1)})$ and $\mathcal{C}_{n,i_2}(\mathtt{B}_k^{(s_2)})$ we have

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \alpha - \frac{13}{\sqrt{N}} - \frac{1}{R}.$$

Proof. For every $1 \le s \le S$ we define

$$\begin{aligned}
\mathtt{B}_1^{(s)} &= \Big((\mathtt{A}_1^{(s)})^{N^2} (\mathtt{A}_2^{(s)})^{N^2} \ldots (\mathtt{A}_N^{(s)})^{N^2}\Big)^{N^{2M}}, \\
\mathtt{B}_2^{(s)} &= \Big((\mathtt{A}_1^{(s)})^{N^4} (\mathtt{A}_2^{(s)})^{N^4} \ldots (\mathtt{A}_N^{(s)})^{N^4}\Big)^{N^{2M-2}}, \\
&\vdots \quad \vdots \\
\mathtt{B}_M^{(s)} &= \Big((\mathtt{A}_1^{(s)})^{N^{2M}} (\mathtt{A}_2^{(s)})^{N^{2M}} \ldots (\mathtt{A}_N^{(s)})^{N^{2M}}\Big)^{N^2}.
\end{aligned}$$

Let $\mathcal{B}_{j,i_1}^{(s_1)} = \mathcal{C}_{n,i_1}(\mathtt{B}_j^{(s_1)})$, $\mathcal{B}_{j,i_2}^{(s_2)} = \mathcal{C}_{n,i_2}(\mathtt{B}_j^{(s_1)})$, $\mathcal{A}_{j,i_1}^{(s_1)} = \mathcal{C}_{n,i_1}(\mathtt{A}_j^{(s_1)})$, and $\mathcal{A}_{j,i_2}^{(s_2)} = \mathcal{C}_{n,i_2}(\mathtt{A}_j^{(s_2)})$. Then the formulas for the $\mathcal{B}_{j,i_1}^{(s_1)}$, $j = 1, \ldots, M$, in terms of the $\mathcal{A}_{1,i_1}^{(s_1)}, \ldots, \mathcal{A}_{N,i_1}^{(s_1)}$ can be obtained from the formulas for the $\mathtt{B}_j^{(s_1)}$ in terms of the $\mathtt{A}_1^{(s_1)}, \ldots, \mathtt{A}_N^{(s_1)}$ by replacing each typewriter font $\mathtt{A}$, $\mathtt{B}$ by the calligraphic $\mathcal{A}$, $\mathcal{B}$ with the corresponding sub- and superscripts, and the analogous statement is true for the $\mathcal{B}_{j,i_2}^{(s_2)}$, $j = 1, \ldots, M$.

By adding fewer than $2 l_n K q_n$ symbols to each of $\mathcal{B}$ and $\overline{\mathcal{B}}$ we can complete any partial $\mathcal{A}_{j,i_1}^{(s_1)}$ at the beginning and end of $\mathcal{B}$ and any partial $\mathcal{A}_{j,i_2}^{(s_2)}$ at the beginning and end of $\overline{\mathcal{B}}$. Let $\mathcal{B}_{\text{aug}}$ and $\overline{\mathcal{B}}_{\text{aug}}$ be the augmented $\mathcal{B}$ and $\overline{\mathcal{B}}$ strings obtained in this way. By Property 2.5, $\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \overline{f}(\mathcal{B}_{\text{aug}}, \overline{\mathcal{B}}_{\text{aug}}) - (4 l_n K q_n)/(2N^{2M+2} l_n K q_n) = \overline{f}(\mathcal{B}_{\text{aug}}, \overline{\mathcal{B}}_{\text{aug}}) - 2/N^{2M+2}$. Then we are comparing two different Feldman patterns of blocks, and by Lemma 5.8 and Corollary 5.5 with $\tilde{f} > 1 - (12/\sqrt{N})$, we have

$$\overline{f}(\mathcal{B}_{\text{aug}}, \overline{\mathcal{B}}_{\text{aug}}) > \alpha - \frac{12}{\sqrt{N}} - \frac{1}{R}.$$

Therefore

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \alpha - \frac{12}{\sqrt{N}} - \frac{1}{R} - \frac{2}{N^{2M+2}} > \alpha - \frac{13}{\sqrt{N}} - \frac{1}{R}. \qquad \square$$

For an application in §5.3 we will need the following result on the $\overline{f}$ distance between strings that can be built as the same or different Feldman patterns but with building blocks from different families.

LEMMA 5.11. *Let $\alpha \in \left(0, \frac{1}{7}\right)$, $n \in \mathbb{N}$, and $K, R, S, N, M \in \mathbb{N} \setminus \{0\}$ with $N \ge 20$, and M, S at least 2. For $1 \le s \le S$, let $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ be a family of strings, where each $\mathtt{A}_j^{(s)}$*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

is a concatenation of K n-blocks. Assume that for all $0 \le i_1, i_2 < q_n$, *all* $s_1 \ne s_2$, *and all* $j_1, j_2 \in \{1, \ldots, N\}$ *we have* $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ *for all strings* $\mathcal{A}, \overline{\mathcal{A}}$ *of at least* Kl_nq_n/R *consecutive symbols from* $\mathcal{C}_{n,i_1}(\mathtt{A}_{j_1}^{(s_1)})$ *and* $\mathcal{C}_{n,i_2}(\mathtt{A}_{j_2}^{(s_2)})$ *respectively.*

Then for $1 \le s \le S$ *we can construct a family of strings* $\mathtt{B}_1^{(s)}, \ldots \mathtt{B}_M^{(s)}$, *as in Proposition 5.10, and obtain that for all* $0 \le i_1, i_2 < q_n$, *all* $j_1, j_2 \in \{1, \ldots, M\}$, *and all sequences* $\mathcal{B}$, $\overline{\mathcal{B}}$ *of at least* $N^{2M+2}l_nKq_n$ *consecutive symbols from* $\mathcal{C}_{n,i_1}(\mathtt{B}_{j_1}^{(s_1)})$ *and* $\mathcal{C}_{n,i_2}(\mathtt{B}_{j_2}^{(s_2)})$, $s_1 \ne s_2$, *we have*

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \alpha - \frac{2}{N^{2M+2}} - \frac{1}{R}. \tag{5.1}$$

Proof. The same argument as in the proof of Proposition 5.10 applies, except we use Corollary 5.5 with $\tilde{f} = 1$. □

5.2. *Feldman mechanism to produce sufficiently many blocks.* Recall that the images of odometer n-blocks under $\mathcal{C}_{n,i}$ are part of the circular $(n+1)$-block. The following statement allows us to obtain lower bounds on the $\overline{f}$ distance between substrings of images of odometer n-blocks under $\mathcal{C}_{n,i}$, given a lower bound on the $\overline{f}$ distance between substrings of the circular n-blocks. Since we will apply this result several times in the proofs of Propositions 5.13 and 5.15, we state it as a separate lemma.

LEMMA 5.12. *Let* $\alpha \in \left(0, \frac{1}{7}\right)$ *and* $n, N, R \in \mathbb{N} \setminus \{0\}$. *Moreover, let N odometer n-blocks* $\mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ *be given such that for all* $j, k \in \{1, \ldots, N\}$, $j \ne k$, *we have* $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ *for any sequences* $\mathcal{A}, \overline{\mathcal{A}}$ *of at least* q_n/R *consecutive symbols from the circular n-blocks* $\mathcal{B}_j^{(n)}$ *and* $\mathcal{B}_k^{(n)}$ *respectively. Then for all* $j, k \in \{1, \ldots, N\}$, $j \ne k$, *any* $0 \le i_1, i_2 < q_n$, *and any sequences* $\mathcal{D}, \overline{\mathcal{D}}$ *of at least* q_nl_n/R *consecutive symbols from* $\mathcal{C}_{n,i_1}(\mathtt{B}_j^{(n)})$ *and* $\mathcal{C}_{n,i_2}(\mathtt{B}_k^{(n)})$ *respectively, we have*

$$\overline{f}(\mathcal{D}, \overline{\mathcal{D}}) > \alpha - \frac{1}{R} - \frac{4R}{l_n}.$$

Proof. Let $\mathcal{D}, \overline{\mathcal{D}}$ be arbitrary sequences of at least q_nl_n/R consecutive symbols from $\mathcal{C}_{n,i_1}(\mathtt{B}_j^{(n)}) = b^{q_n-j_{i_1}}(\mathcal{B}_j^{(n)})^{l_n-1}e^{j_{i_1}}$ and from $\mathcal{C}_{n,i_2}(\mathtt{B}_k^{(n)}) = b^{q_n-j_{i_2}}(\mathcal{B}_j^{(n)})^{l_n-1}e^{j_{i_2}}$, respectively, for any $j \ne k$ and any $0 \le i_1, i_2 < q_n$. We modify $\mathcal{D}$ and $\overline{\mathcal{D}}$ by first completing any partial blocks $\mathcal{B}_j^{(n)}$ and $\mathcal{B}_k^{(n)}$, which can be accomplished by adding fewer than $2q_n$ symbols to each of $\mathcal{D}$ and $\overline{\mathcal{D}}$. Then we remove any of the b's preceding the $\mathcal{B}_j^{(n)}$s and any of the e's following the $\mathcal{B}_j^{(n)}$s that are included in $\mathcal{D}$s, and similarly for such b's and e's in $\overline{\mathcal{D}}$. At most q_n symbols are removed from each of $\mathcal{D}$ and $\overline{\mathcal{D}}$. Let $\mathcal{D}_{\mathrm{mod}}$ and $\overline{\mathcal{D}}_{\mathrm{mod}}$ be these modified versions of $\mathcal{D}$ and $\overline{\mathcal{D}}$. Then by Property 2.5,

$$\overline{f}(\mathcal{D}, \overline{\mathcal{D}}) > \overline{f}(\mathcal{D}_{\mathrm{mod}}, \overline{\mathcal{D}}_{\mathrm{mod}}) - \frac{2R}{l_n} - \frac{2R}{l_n}.$$

We have $\mathcal{D}_{\mathrm{mod}} = (\mathcal{B}_j^{(n)})^l$ and $\overline{\mathcal{D}}_{\mathrm{mod}} = (\mathcal{B}_k^{(n)})^{\bar{l}}$, for some positive integers l and $\bar{l}$. We are given that $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ for any strings of at least q_n/R consecutive symbols from $\mathcal{B}_j^{(n)}$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

and $\mathcal{B}_k^{(n)}$, respectively. Thus it follows from Corollary 5.5 with $\tilde{f} = 1$ that

$$\overline{f}(\mathcal{D}_{\mathrm{mod}}, \overline{\mathcal{D}}_{\mathrm{mod}}) > \alpha - \frac{1}{R}. \qquad \square$$

In the proofs of both Propositions 5.13 and 5.15 we will also use the sequence $(R_n)_{n=1}^{\infty}$, where $R_1 = N(0)$ (with $N(0) + 1$ the number of symbols in our alphabet) and $R_n = k_{n-2} \cdot q_{n-2}^2$ for $n \geq 2$. We note that for $n \geq 2$,

$$\frac{q_n}{R_n} = \frac{k_{n-1} \cdot l_{n-1} \cdot (k_{n-2} \cdot l_{n-2} \cdot q_{n-2}^2) \cdot q_{n-1}}{k_{n-2} \cdot q_{n-2}^2} = l_{n-2} \cdot k_{n-1} \cdot l_{n-1} \cdot q_{n-1}. \tag{5.2}$$

Hence, for $n \geq 2$ a substring of at least q_n/R_n consecutive symbols in a circular n-block contains at least $l_{n-2} - 1$ complete 2-subsections which have length $k_{n-1} l_{n-1} q_{n-1}$ (recall the notion of a 2-subsubsection from the end of §3.3). This will allow us to ignore incomplete 2-subsections at the ends of the substring.

PROPOSITION 5.13. *Let $\alpha \in \left(0, \frac{1}{7}\right)$ and $n, N, M \in \mathbb{N}$ with $N \geq 100$ and $M \geq 2$. Suppose $\mathtt{A}_0, \ldots, \mathtt{A}_N$ is the collection of n-blocks, which have equal length $\mathtt{h}_\mathtt{n}$ and satisfy the unique readability property. Furthermore, if $n > 0$ assume that for all $j_1, j_2 \in \{0, \ldots, N\}$, $j_1 \neq j_2$, we have $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ for any sequences $\mathcal{A}, \overline{\mathcal{A}}$ of at least q_n/R_n consecutive symbols from $\mathcal{A}_{j_1}$ and $\mathcal{A}_{j_2}$, respectively. Then we can construct M $(n+1)$-blocks $\mathtt{B}_1, \ldots, \mathtt{B}_M$ of equal length $\mathtt{h}_{n+1}$ (which are uniform in the n-blocks and satisfy the unique readability property) such that for all $j, k \in \{1, \ldots, M\}$, $j \neq k$, and any sequences $\mathcal{B}, \overline{\mathcal{B}}$ of at least q_{n+1}/R_{n+1} consecutive symbols from $\mathcal{B}_j$ and $\mathcal{B}_k$ we have*

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \begin{cases} \alpha - \left(\dfrac{4}{R_n} + \dfrac{4R_n}{l_n} + \dfrac{14}{\sqrt{N}} + \dfrac{2}{l_{n-1}}\right) & \text{if } n > 0, \\ 1 - \dfrac{5}{\sqrt{N}} - \dfrac{2}{l_0} & \text{if } n = 0. \end{cases}$$

Proof. We choose the block $\mathtt{A}_0$ as a 'marker', that is, an n-block whose appearances can be used to identify the end of an $(n+1)$-block. We distribute the marker blocks over the new words and modify the classical Feldman patterns on the building blocks $\mathtt{A}_1, \ldots, \mathtt{A}_N$ in the following way to define the $(n+1)$-blocks:

$$\begin{aligned}
\mathtt{B}_1 &= \left(\left((\mathtt{A}_1)^{N^2} (\mathtt{A}_2)^{N^2} \ldots (\mathtt{A}_N)^{N^2}\right)^{N^{2M-1}} (\mathtt{A}_0)^{N^{2M+1}-1}\right)^N (\mathtt{A}_0)^N, \\
\mathtt{B}_2 &= \left(\left((\mathtt{A}_1)^{N^4} (\mathtt{A}_2)^{N^4} \ldots (\mathtt{A}_N)^{N^4}\right)^{N^{2M-3}} (\mathtt{A}_0)^{N^{2M+1}-1}\right)^N (\mathtt{A}_0)^N, \\
&\vdots \quad \vdots \\
\mathtt{B}_M &= \left(\left((\mathtt{A}_1)^{N^{2M}} (\mathtt{A}_2)^{N^{2M}} \ldots (\mathtt{A}_N)^{N^{2M}}\right)^{N} (\mathtt{A}_0)^{N^{2M+1}-1}\right)^N (\mathtt{A}_0)^N.
\end{aligned}$$

We note that every $(n+1)$-block $\mathtt{B}_k$ contains each n-block $\mathtt{A}_l$ exactly N^{2M+2} times and has length $(N+1) \cdot N^{2M+2} \cdot \mathtt{h}_n$. Moreover, the new blocks are uniquely readable because the string $(\mathtt{A}_0)^{N^{2M+1}+N-1}$ only occurs at the end of an $(n+1)$-block. We also observe that

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$\mathtt{B}_k$ is built with $N^{2\cdot(M-k+1)}$ *cycles*

$$\mathtt{F}_k := (\mathtt{A}_1)^{N^{2k}} (\mathtt{A}_2)^{N^{2k}} \ldots (\mathtt{A}_N)^{N^{2k}}.$$

Let $\mathcal{B}$ and $\overline{\mathcal{B}}$ be sequences of at least q_{n+1}/R_{n+1} consecutive symbols from $\mathcal{B}_j$ and $\mathcal{B}_k$, $j \neq k$. In the case $n > 0$ we note that $\mathcal{B}$ and $\overline{\mathcal{B}}$ have at least the length of l_{n-1} complete 2-subsections by equation (5.2). By adding fewer than $2l_n k_n q_n$ symbols to each of $\mathcal{B}$ and $\overline{\mathcal{B}}$, we can complete any partial 2-subsections at the beginning and end of $\mathcal{B}$ and $\overline{\mathcal{B}}$. This change can increase the $\overline{f}$ distance between $\mathcal{B}$ and $\overline{\mathcal{B}}$, but by less than $2/l_{n-1}$. In addition, we remove the marker blocks, possibly increasing $\overline{f}$ by at most $2/(N+1)$. The modified strings $\mathcal{B}_{\text{mod}}$ and $\overline{\mathcal{B}}_{\text{mod}}$ obtained satisfy $\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \overline{f}(\mathcal{B}_{\text{mod}}, \overline{\mathcal{B}}_{\text{mod}}) - (2/l_{n-1}) - (2/(N+1))$.

By Lemma 5.12, $\overline{f}(\mathcal{D}, \overline{\mathcal{D}}) > \alpha - (2/R_n) - (4R_n/l_n)$ for any substrings $\mathcal{D}, \overline{\mathcal{D}}$ of at least $q_n l_n/R_n$ consecutive symbols from $\mathcal{C}_{n,i_1}(\mathtt{A}_{j_1})$ and $\mathcal{C}_{n,i_2}(\mathtt{A}_{j_2})$ with $j_1 \neq j_2$. If we let Φ_{j,i_1} and Φ_{k,i_2} be the jth and kth Feldman patterns built from $\mathcal{C}_{n,i_1}(\mathtt{A}_1), \mathcal{C}_{n,i_1}(\mathtt{A}_2), \ldots, \mathcal{C}_{n,i_1}(\mathtt{A}_N)$ and $\mathcal{C}_{n,i_2}(\mathtt{A}_1), \mathcal{C}_{n,i_2}(\mathtt{A}_2), \ldots, \mathcal{C}_{n,i_2}(\mathtt{A}_N)$ respectively, then the same argument as in Proposition 5.10 shows that for any substrings $\mathcal{E}, \overline{\mathcal{E}}$ consisting of at least $|\Phi_{j,i_1}|/N = |\Phi_{k,i_1}|/N$ consecutive symbols from Φ_{j,i_2} and Φ_{k,i_2}, we have $\overline{f}(\mathcal{E}, \overline{\mathcal{E}}) > \alpha - (4/R_n) - (4R_n/l_n) - (13/\sqrt{N})$.

Note that $\mathcal{B}_{\text{mod}}$ and $\overline{\mathcal{B}}_{\text{mod}}$ consist respectively of a string of $\Phi_{j,i}$s and a string of $\Phi_{k,i}$s (j and k fixed, i varying). Therefore, by Corollary 5.5 with $\tilde{f} = 1$, $\overline{f}(\mathcal{B}_{\text{mod}}, \overline{\mathcal{B}}_{\text{mod}}) > \alpha - (4/R_n) - (4R_n/l_n) - (13/\sqrt{N}) - (2/N)$. Thus in the case $n > 0$ we obtain

$$\begin{aligned}\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) &> \alpha - (4/R_n) - (4R_n/l_n) - (13/\sqrt{N}) - (2/N) - (2/l_{n-1}) - (2/N) \\ &> \alpha - (4/R_n) - (4R_n/l_n) - (14/\sqrt{N}) - (2/l_{n-1}).\end{aligned}$$

In the case $n = 0$ we complete strings $\mathcal{C}_{0,0}(\mathtt{F}_j)$ and $\mathcal{C}_{0,0}(\mathtt{F}_k)$ at the beginning and end of $\mathcal{B}$ and $\overline{\mathcal{B}}$ respectively by adding fewer than $2l_0 N^{2M+1}$ symbols to each of $\mathcal{B}$ and $\overline{\mathcal{B}}$. This corresponds to a fraction of at most $2/(N+1)$ of the total length. In the next step we remove the marker blocks and spacers b, possibly increasing $\overline{f}$ by at most $(6/N) + (2/l_0)$. On the remaining strings we apply Remark 5.9 (note that we have enough symbols by our completion above) to obtain the claim for $n = 0$. □

Remark 5.14. In the proof above we cannot put all markers at the end of the new $(n+1)$-block since these markers would cover a fraction $\mathtt{h}_{n+1}/(N+1)$ of that block due to uniformity. Thus, the conclusion would not hold true for $n = 0$. We will also need the chosen form of the $(n+1)$-blocks in an analogous statement for the odometer-based system in Proposition 6.2.

5.3. *Mechanism to produce closeness in odometer-based system and separation in corresponding circular system.* We impose the following conditions on the circular coefficients $(l_n)_{n\in\mathbb{N}}$:

$$l_{n+1} \geq l_n^2 \text{ and } l_n \geq 4R_{n+1} \quad \text{for every } n \in \mathbb{N}. \tag{5.3}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

PROPOSITION 5.15. *Let $K \geq 2$, $0 < \varepsilon < \alpha$, $\delta > 0$, and $\alpha \in \left(0, \frac{1}{7}\right)$. Then there are numbers $N, p \in \mathbb{N}$ such that for $N+1$ uniquely readable n-blocks $\mathtt{B}_0^{(n)}, \mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ with $\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) > \alpha$ for all sequences $\mathcal{A}, \overline{\mathcal{A}}$ of at least q_n/R_n consecutive symbols from $\mathcal{B}_i^{(n)}$ and $\mathcal{B}_j^{(n)}$, $i > j$, we can build K $(n+p)$-blocks $\mathtt{B}_1^{(n+p)}, \ldots, \mathtt{B}_K^{(n+p)}$ of equal length $\mathtt{h}_{n+p}$ (satisfying the unique readability property and uniformity in all blocks from stage n through $n+p$) with the following properties:*

(1) $\overline{f}(\mathtt{B}_i^{(n+p)}, \mathtt{B}_j^{(n+p)}) < \delta$ *for all* i, j;

(2) $\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) > \alpha - \varepsilon - \sum_{s=0}^{p-1}(6/R_{n+s})$ *for all sequences $\mathcal{B}, \overline{\mathcal{B}}$ of at least q_{n+p}/R_{n+p} consecutive symbols from $\mathcal{B}_i^{(n+p)}$ and $\mathcal{B}_j^{(n+p)}$, $i > j$.*

The number of stages p will be the least integer such that

$$\left(1 - \frac{1}{K}\right)^p < \frac{\varepsilon}{2}.$$

The proof is based upon an inductive construction (called a *shifting mechanism*; see Figure 2 for a sketch of its idea) and a final step to align blocks in the odometer-based system. For this construction process and the given $n \in \mathbb{N}$, let $(u_{n+m})_{m\in\mathbb{N}}$ and $(e_{n+m})_{m\in\mathbb{N}}$ be increasing sequences of natural numbers such that

$$\sum_{m\in\mathbb{N}} \frac{1}{u_{n+m}^2} < \frac{\delta}{4} \tag{5.4}$$

and

$$\sum_{m\in\mathbb{N}} \left(\frac{8}{u_{n+m}} + \frac{17}{\sqrt{e_{n+m}}}\right) < \frac{\varepsilon}{8}. \tag{5.5}$$

Moreover, we will use the sequence $(d_{n+m})_{m\in\mathbb{N}}$, where

$$d_{n+m} = u_{n+m}^2. \tag{5.6}$$

In the following, we will also use the notation

$$\lambda_{n+m} = d_{n+m} \cdot e_{n+m}$$

and $N(n+m) + 1 = K\lambda_{n+m} + 1$ will be the number of $(n+m)$-blocks. In particular, we start with $N(n) + 1 = N + 1$ n-blocks and we require N to satisfy

$$N > \max(2/\delta, (100/\varepsilon)^2). \tag{5.7}$$

5.3.1. *Initial stage of the shifting mechanism: construction of $(n+1)$-blocks.* First of all, we choose one n-block $\mathtt{B}_0^{(n)}$ as a marker. Then we apply Proposition 5.10 on the remaining n-blocks $\mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ to build $\tilde{N}(n+1) := 2K\lambda_{n+1}$ *pre-$(n+1)$-blocks* denoted by $\mathtt{A}_{i,j}$, $i = 1, \ldots, K$, $j = 1, \ldots, 2\lambda_{n+1}$. In particular, these have length $\tilde{\mathtt{h}}_{n+1} = N^{2\cdot\tilde{N}(n+1)+3} \cdot \mathtt{h}_n$ and are uniform in the n-blocks $\mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ by construction. More precisely, every pre-$(n+1)$-block contains each n-block $\mathtt{B}_j^{(n)}$, $1 \leq j \leq N$, exactly $N^{2\cdot\tilde{N}(n+1)+2}$ times, and pre-$(n+1)$-blocks $\mathcal{A}_{i,j}$ in the circular system (that is, images of $\mathtt{A}_{i,j}$ under the operator $\mathcal{C}_{n,k}$ for some $k \in \{1, \ldots, q_n\}$, where the value of k does not matter

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

for the following investigation) have length $\tilde{q}_{n+1} = N^{2\cdot\tilde{N}(n+1)+3} \cdot l_n \cdot q_n$. Moreover, with the aid of Lemma 5.12, Proposition 5.10 also implies that different pre-$(n+1)$-blocks in the circular system are at least

$$\alpha - \beta_{n+1} \ \text{ where } \beta_{n+1} := \frac{4}{R_n} + \frac{13}{\sqrt{N}} + \frac{4R_n}{l_n}, \tag{5.8}$$

$\overline{f}$ apart on substantial substrings of length at least $N^{2\cdot\tilde{N}(n+1)+2} \cdot l_n \cdot q_n = \tilde{q}_{n+1}/N$. Finally, we introduce the abbreviation

$$\mathtt{a}_n = (\mathtt{B}_0^{(n)})^{K\cdot N^{2\tilde{N}(n+1)+2}}.$$

We use these pre-$(n+1)$-blocks to construct $(n+1)$-*blocks* $\mathtt{B}_{i,j}^{(n+1)}$ of K different types (the index i indicates the type, and $j = 1, \ldots, \lambda_{n+1} = d_{n+1}e_{n+1}$ numbers the $(n+1)$-blocks of that type consecutively):

$$\begin{aligned}
&(n+1)\text{-blocks of type 1:} \quad && \mathtt{B}_{1,1}^{(n+1)} = \mathtt{A}_{1,1}\mathtt{A}_{2,1} \ldots \mathtt{A}_{K,1}\mathtt{a}_n,\\
& && \mathtt{B}_{1,2}^{(n+1)} = \mathtt{A}_{1,2}\mathtt{A}_{2,2} \ldots \mathtt{A}_{K,2}\mathtt{a}_n,\\
& && \mathtt{B}_{1,3}^{(n+1)} = \mathtt{A}_{1,3}\mathtt{A}_{2,3} \ldots \mathtt{A}_{K,3}\mathtt{a}_n,\\
& && \vdots\\
& && \mathtt{B}_{1,\lambda_{n+1}}^{(n+1)} = \mathtt{A}_{1,\lambda_{n+1}}\mathtt{A}_{2,\lambda_{n+1}} \ldots \mathtt{A}_{K,\lambda_{n+1}}\mathtt{a}_n,\\
&(n+1)\text{-blocks of type 2:} && \mathtt{B}_{2,1}^{(n+1)} = \mathtt{A}_{2,1}\mathtt{A}_{3,1} \ldots \mathtt{A}_{K,1}\mathtt{A}_{1,2}\mathtt{a}_n,\\
& && \mathtt{B}_{2,2}^{(n+1)} = \mathtt{A}_{2,2}\mathtt{A}_{3,2} \ldots \mathtt{A}_{K,2}\mathtt{A}_{1,3}\mathtt{a}_n,\\
& && \mathtt{B}_{2,3}^{(n+1)} = \mathtt{A}_{2,3}\mathtt{A}_{3,3} \ldots \mathtt{A}_{K,3}\mathtt{A}_{1,4}\mathtt{a}_n,\\
& && \vdots\\
& && \mathtt{B}_{2,\lambda_{n+1}}^{(n+1)} = \mathtt{A}_{2,\lambda_{n+1}}\mathtt{A}_{3,\lambda_{n+1}} \ldots \mathtt{A}_{1,\lambda_{n+1}+1}\mathtt{a}_n,\\
&\vdots && \vdots\\
&(n+1)\text{-blocks of type } K: && \mathtt{B}_{K,1}^{(n+1)} = \mathtt{A}_{K,1}\mathtt{A}_{1,2} \ldots \mathtt{A}_{K-1,2}\mathtt{a}_n,\\
& && \mathtt{B}_{K,2}^{(n+1)} = \mathtt{A}_{K,2}\mathtt{A}_{1,3} \ldots \mathtt{A}_{K-1,3}\mathtt{a}_n,\\
& && \mathtt{B}_{K,3}^{(n+1)} = \mathtt{A}_{K,3}\mathtt{A}_{1,4} \ldots \mathtt{A}_{K-1,4}\mathtt{a}_n,\\
& && \vdots\\
& && \mathtt{B}_{K,\lambda_{n+1}}^{(n+1)} = \mathtt{A}_{K,\lambda_{n+1}}\mathtt{A}_{1,\lambda_{n+1}+1} \ldots \mathtt{A}_{K-1,\lambda_{n+1}+1}\mathtt{a}_n.
\end{aligned}$$

Moreover, we define an additional $(n+1)$-block $\mathtt{B}_0^{(n+1)}$ which will play the role of a marker:

$$\mathtt{B}_0^{(n+1)} = \mathtt{A}_{1,\lambda_{n+1}+2} \ldots \mathtt{A}_{K,\lambda_{n+1}+2}\mathtt{a}_n,$$

where the pre-$(n+1)$-blocks $\mathtt{A}_{i,\lambda_{n+1}+2}$, $i = 1, \ldots, K$, are not used in any other $(n+1)$-block. In total, there are $Kd_{n+1}e_{n+1} + 1$ $(n+1)$-blocks. Since each of the pre-$(n+1)$-blocks contains each n-block $\mathtt{B}_i^{(n)}$, $1 \leq i \leq N$, exactly $N^{2\tilde{N}(n+1)+2}$ times, and each $(n+1)$-block contains $\mathtt{B}_0^{(n)}$ exactly $KN^{2\tilde{N}(n+1)+2}$ times, every $(n+1)$-block is uniform in the n-blocks.

LEMMA 5.16. (Distance between $(n+1)$-blocks in the odometer-based system) *For every $i_1, i_2 \in \{1, \ldots, K\}$ and every $j \in \{1, \ldots, d_{n+1}e_{n+1}\}$ we have*

$$\overline{f}(\mathtt{B}_{i_1,j}^{(n+1)}, \mathtt{B}_{i_2,j}^{(n+1)}) \leq \left(\frac{N}{N+1}\right) \cdot \frac{|i_2 - i_1|}{K}. \tag{5.9}$$

Proof. Observe that

$$|\mathtt{a}_n| = \frac{\mathtt{h}_{n+1}}{N+1},$$

due to the uniformity of n-blocks within the $(n+1)$-blocks. Thus the pre-$(n+1)$-block part of each $\mathtt{B}_{i,j}^{(n+1)}$, that is, the part before $\mathtt{a}_n$, forms a fraction $N/(N+1)$ of $\mathtt{B}_{i,j}^{(n+1)}$. Without loss of generality, let $i_2 > i_1$. We note that $\mathtt{B}_{i_1,j}^{(n+1)}$ and $\mathtt{B}_{i_2,j}^{(n+1)}$ have

$$\mathtt{A}_{i_2,j}\mathtt{A}_{i_2+1,j} \ldots \mathtt{A}_{K,j}\mathtt{A}_{1,j+1} \ldots \mathtt{A}_{i_1-1,j+1}$$

as a common substring of their pre-$(n+1)$-block parts. This substring forms a fraction $(K-(i_2-i_1))/K$ of the pre-$(n+1)$-block part of each of $\mathtt{B}_{i_1,j}^{(n+1)}$ and $\mathtt{B}_{i_2,j}^{(n+1)}$. Therefore (5.9) holds. □

LEMMA 5.17. (Distance between $(n+1)$-blocks in the circular system) *Let $\mathtt{B}_1^{(n+1)}$ and $\mathtt{B}_2^{(n+1)}$ be $(n+1)$-blocks and J be the number of pre-$(n+1)$-blocks both have in common. Then for any sequences $\mathcal{B}$ and $\overline{\mathcal{B}}$ of at least q_{n+1}/R_{n+1} consecutive symbols in $\mathcal{B}_1^{(n+1)}$ and $\mathcal{B}_2^{(n+1)}$ we have*

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \left(1 - \frac{J}{K}\right)\alpha - E_{n+1}, \quad \textit{where } E_{n+1} := \beta_{n+1} + \frac{4}{N} + \frac{4}{l_{n-1}}$$

with β_{n+1} as in equation (5.8).

Proof. In order to get the estimate in the circular system we recall that under the circular operator $\mathcal{C}_n$ the whole $(n+1)$-word consists of q_n many 2-subsections which differ from each other just in the exponents of the newly introduced spacers b and e. Then we consider $\mathcal{B}$ and $\overline{\mathcal{B}}$ to be a concatenation of complete 2-subsections ignoring incomplete ones at the ends which constitute a fraction of at most $2/l_{n-1}$ of the total length of $\mathcal{B}$ and $\overline{\mathcal{B}}$ by equation (5.2). In the following consideration we ignore the marker segments which amount to a fraction $1/(N+1)$ of the total length. Accordingly, we consider $\mathcal{B}$ and $\overline{\mathcal{B}}$ to be a concatenation of complete pre-$(n+1)$-blocks $\mathcal{A}_h$. Using the estimate from equation (5.8) we apply Corollary 5.5 with $\tilde{f} \geq 1 - (J/K)$ and obtain

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \left(1 - \frac{J}{K}\right) \cdot (\alpha - \beta_{n+1}) - \frac{2}{N} - \frac{4}{l_{n-1}} - \frac{2}{N+1},$$

which yields the claim. □

5.3.2. *Induction step: construction of $(n+m)$-blocks.* We will follow the inductive scheme for the construction of $(n+m)$-blocks described here for $2 \leq m < p$, where

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$p > 2$ is the smallest number such that

$$\left(1 - \frac{1}{K}\right)^p < \frac{\varepsilon}{2}. \tag{5.10}$$

Assume that in our inductive construction we have constructed $K\lambda_{n+m-1}$ $(n+m-1)$ *-blocks* $\mathtt{B}_{i,j}^{(n+m-1)}$ of K different types (once again, the index $1 \leq i \leq K$ indicates the type, and $1 \leq j \leq \lambda_{n+m-1}$ numbers the $(n+m-1)$-blocks of that type consecutively), where for $m = 2$ the $(n+1)$-blocks are the ones constructed in §5.3.1 and for $m > 2$ the $(n+m-1)$-blocks are constructed according to the following formula (with $\lambda = \lambda_{n+m-1}$)

$(n+m-1)$-blocks of type 1:

$$\begin{aligned}
\mathtt{B}_{1,1}^{(n+m-1)} &= \mathtt{A}_{1,1}^{(n+m-1)}\mathtt{A}_{2,2}^{(n+m-1)} \ldots \mathtt{A}_{K,K}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
\mathtt{B}_{1,2}^{(n+m-1)} &= \mathtt{A}_{1,K+1}^{(n+m-1)}\mathtt{A}_{2,K+2}^{(n+m-1)} \ldots \mathtt{A}_{K,2K}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
&\vdots\\
\mathtt{B}_{1,\lambda}^{(n+m-1)} &= \mathtt{A}_{1,(\lambda-1)K+1}^{(n+m-1)}\mathtt{A}_{2,(\lambda-1)K+2}^{(n+m-1)} \ldots \mathtt{A}_{K,\lambda K}^{(n+m-1)}\mathtt{a}_{n+m-2},
\end{aligned}$$

$(n+m-1)$-blocks of type 2:

$$\begin{aligned}
\mathtt{B}_{2,1}^{(n+m-1)} &= \mathtt{A}_{1,2}^{(n+m-1)}\mathtt{A}_{2,3}^{(n+m-1)} \ldots \mathtt{A}_{K,K+1}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
\mathtt{B}_{2,2}^{(n+m-1)} &= \mathtt{A}_{1,K+2}^{(n+m-1)}\mathtt{A}_{2,K+3}^{(n+m-1)} \ldots \mathtt{A}_{K,2K+1}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
&\vdots\\
\mathtt{B}_{2,\lambda}^{(n+m-1)} &= \mathtt{A}_{1,(\lambda-1)K+2}^{(n+m-1)}\mathtt{A}_{2,(\lambda-1)K+3}^{(n+m-1)} \ldots \mathtt{A}_{K,\lambda K+1}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
&\vdots
\end{aligned}$$

$(n+m-1)$-blocks of type K:

$$\begin{aligned}
\mathtt{B}_{K,1}^{(n+m-1)} &= \mathtt{A}_{1,K}^{(n+m-1)}\mathtt{A}_{2,K+1}^{(n+m-1)} \ldots \mathtt{A}_{K,2K-1}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
\mathtt{B}_{K,2}^{(n+m-1)} &= \mathtt{A}_{1,2K}^{(n+m-1)}\mathtt{A}_{2,2K+1}^{(n+m-1)} \ldots \mathtt{A}_{K,3K-1}^{(n+m-1)}\mathtt{a}_{n+m-2},\\
&\vdots\\
\mathtt{B}_{K,\lambda}^{(n+m-1)} &= \mathtt{A}_{1,\lambda K}^{(n+m-1)}\mathtt{A}_{2,\lambda K+1}^{(n+m-1)} \ldots \mathtt{A}_{K,(\lambda+1)K-1}^{(n+m-1)}\mathtt{a}_{n+m-2},
\end{aligned}$$

with marker segment $\mathtt{a}_{n+m-2} = (\mathtt{B}_0^{(n+m-2)})^{\bar{N}(n+m-1)}$, where $\bar{N}(n+m-1)$ is chosen according to (5.16) to guarantee uniformity of $(n+m-2)$-blocks in the $(n+m-1)$-blocks. Moreover, we have an additional marker block,

$$\mathtt{B}_0^{(n+m-1)} = \mathtt{A}_{1,(\lambda_{n+m-1}+1)K}^{(n+m-1)}\mathtt{A}_{2,(\lambda_{n+m-1}+1)K}^{(n+m-1)} \ldots \mathtt{A}_{K,(\lambda_{n+m-1}+1)K}^{(n+m-1)}\mathtt{a}_{n+m-2}.$$

These blocks are defined using pre-$(n+m-1)$-blocks $\mathtt{A}_{i,j}^{(n+m-1)}$ with

$$\overline{f}(\mathtt{A}_{i_1,j}^{(n+m-1)}, \mathtt{A}_{i_2,j}^{(n+m-1)}) \leq \sum_{u=1}^{m-2}\left(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}}\right) \tag{5.11}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

if $m > 2$ (note that the assumption is vacuous if $m = 2$). Moreover, for any sequences $\mathcal{B}$ and $\overline{\mathcal{B}}$ of at least q_{n+m-1}/R_{n+m-1} consecutive symbols in $\mathcal{B}^{(n+m-1)}_{i_1,j_1}$ and $\mathcal{B}^{(n+m-1)}_{i_2,j_2}$ for some $i_1, i_2 \in \{0, 1, \ldots, K\}$, $i_1 \leq i_2$, and $j_1, j_2 \in \{1, \ldots, \lambda_{n+m-1}\}$, we assume

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \alpha - E_{n+m-1} \quad \text{for } i_1 = i_2 \text{ and } j_1 \neq j_2, \tag{5.12}$$

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \left(1 - \left(1 - \frac{1}{K}\right)^{m-1}\right)\alpha - E_{n+m-1} \quad \text{for } i_1 < i_2 \text{ and } j_1 = j_2 \text{ or } j_1 = j_2 + 1, \tag{5.13}$$

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \alpha - E_{n+m-1} \quad \text{for } i_1 < i_2 \text{ and all other cases of } j_1 \neq j_2, \tag{5.14}$$

where

$$E_{n+m-1} = E_{n+1} + \sum_{i=1}^{m-2} \left(\frac{6}{R_{n+i}} + \frac{8}{u_{n+i}} + \frac{17}{\sqrt{e_{n+i}}}\right). \tag{5.15}$$

We point out that assumptions (5.12)-(5.15) hold for $m = 2$ by Lemma 5.17.

In order to continue the inductive construction we define *grouped* $(n + m - 1)$ *-blocks* of type i by concatenating d_{n+m-1} $(n + m - 1)$-blocks of type i:

$$\mathtt{G}^{(n+m-1)}_{i,s} = \mathtt{B}^{(n+m-1)}_{i,s\cdot d_{n+m-1}+1}\mathtt{B}^{(n+m-1)}_{i,s\cdot d_{n+m-1}+2} \ldots \mathtt{B}^{(n+m-1)}_{i,(s+1)\cdot d_{n+m-1}}, \quad \text{for } s = 0, \ldots, e_{n+m-1} - 1.$$

LEMMA 5.18. (Distance between grouped $(n + m - 1)$-blocks in the circular system) *Let $\mathcal{G}$ and $\overline{\mathcal{G}}$ be sequences of at least $u_{n+m-1}l_{n+m-1}q_{n+m-1}$ consecutive symbols in $\mathcal{G}^{(n+m-1)}_{i_1,s_1}$ and $\mathcal{G}^{(n+m-1)}_{i_2,s_2}$ for some $i_1, i_2 \in \{1, \ldots, K\}$ and $s_1, s_2 \in \{0, \ldots, e_{n+m-1} - 1\}$.*

(1) *For $i_1 = i_2$ and $s_1 \neq s_2$ we have*

$$\overline{f}(\mathcal{G}, \overline{\mathcal{G}}) \geq \alpha - E_{n+m-1} - \frac{4}{R_{n+m-1}} - \frac{4R_{n+m-1}}{l_{n+m-1}} - \frac{4}{u_{n+m-1}}.$$

(2) *For $i_1 \neq i_2$ and $s_1 = s_2$ we have*

$$\overline{f}(\mathcal{G}, \overline{\mathcal{G}}) \geq \left(1 - \left(1 - \frac{1}{K}\right)^{m-1}\right) \cdot \alpha - E_{n+m-1} - \frac{4}{R_{n+m-1}} - \frac{4R_{n+m-1}}{l_{n+m-1}} - \frac{4}{u_{n+m-1}}.$$

(3) *For $i_1 \neq i_2$ and $s_1 \neq s_2$ we have*

$$\overline{f}(\mathcal{G}, \overline{\mathcal{G}}) \geq \alpha - E_{n+m-1} - \frac{4}{R_{n+m-1}} - \frac{4R_{n+m-1}}{l_{n+m-1}} - \frac{6}{u_{n+m-1}}.$$

Proof. We factor $\mathcal{G}$ and $\overline{\mathcal{G}}$ into 2-subsections $\mathcal{C}_{n,j_1}(\mathtt{B}^{(n+m-1)}_{i_1,t})$ and $\mathcal{C}_{n,j_2}(\mathtt{B}^{(n+m-1)}_{i_2,u})$ for some $j_1, j_2 \in \{0, 1, \ldots, q_n - 1\}$, omitting partial blocks at the ends which constitute a portion of at most $2/u_{n+m-1}$ of the total length of $\mathcal{G}$ and $\overline{\mathcal{G}}$. With the aid of Lemma 5.12 we can transfer the estimates in equations (5.12)–(5.14) to estimates on the $\overline{f}$ distance of substrings of at least $q_{n+m-1}l_{n+m-1}/R_{n+m-1}$ consecutive symbols in these strings of the form $\mathcal{C}_{n,j_1}(\mathtt{B}^{(n+m-1)}_{i_1,t})$ and $\mathcal{C}_{n,j_2}(\mathtt{B}^{(n+m-1)}_{i_2,u})$, respectively.

We now examine the particular situation of each part of the lemma.

(1) Since the $(n + m - 1)$-blocks in $\mathcal{G}^{(n+m-1)}_{i_1,s_1}$ and $\mathcal{G}^{(n+m-1)}_{i_1,s_2}$ are of same type but different pattern, Corollary 5.5 (with $\tilde{f} = 1$) and the modified version of equation (5.12)

yield

$$\overline{f}(\mathcal{G}, \overline{\mathcal{G}}) \geq \alpha - E_{n+m-1} - \frac{2}{R_{n+m-1}} - \frac{4R_{n+m-1}}{l_{n+m-1}} - \frac{2}{R_{n+m-1}} - \frac{4}{u_{n+m-1}}.$$

(2) For $m = 2$ we note that $\mathtt{B}_{i_1,t}^{(n+1)}$ and $\mathtt{B}_{i_2,u}^{(n+1)}$ in $\mathtt{G}_{i_1,s_1}^{(n+1)}$ and $\mathtt{G}_{i_2,s_1}^{(n+1)}$ respectively have at most $J = \max(|i_2 - i_1|,\ K - |i_2 - i_1|)$ pre-$(n+1)$-blocks in common. Then we apply Lemma 5.17, Lemma 5.12, and Corollary 5.5 (with $\tilde{f} = 1$) to conclude the inequality.

In order to obtain a lower bound on the $\overline{f}$ distance for $m > 2$ we note that $\mathtt{B}_{i_1,t}^{(n+m-1)}$ and $\mathtt{B}_{i_2,u}^{(n+m-1)}$ in $\mathtt{G}_{i_1,s_1}^{(n+m-1)}$ and $\mathtt{G}_{i_2,s_1}^{(n+m-1)}$ could fall under the situation of the worst possible estimate in equation (5.13). By another application of Lemma 5.12 and Corollary 5.5 (with $\tilde{f} = 1$) we get the claim.

(3) Without loss of generality let $i_2 > i_1$. Once again, we start with the proof for $m = 2$. There is at most one pair of $(n+1)$-blocks in $\mathtt{G}_{i_1,s_1}^{(n+1)}$ and $\mathtt{G}_{i_2,s_2}^{(n+1)}$ respectively that have $i_2 - i_1$ pre-$(n+1)$-blocks in common while all other pairs have no pair in common. This corresponds to a proportion of at most $2/u_{n+1}$. Then we use Lemma 5.17, Lemma 5.12, and Proposition 5.4 (with $\tilde{f} \geq 1 - (2/u_{n+1})$) to obtain

$$\begin{aligned}\overline{f}(\mathcal{G}, \overline{\mathcal{G}}) &\geq \bigg(1 - \frac{2}{u_{n+1}} + \frac{2}{u_{n+1}}\bigg(1 - \frac{i_2 - i_1}{K}\bigg)\bigg)\alpha - E_{n+1} - \frac{4}{R_{n+1}} - \frac{4R_{n+1}}{l_{n+1}} - \frac{4}{u_{n+1}} \\ &\geq \alpha - E_{n+1} - \frac{4}{R_{n+1}} - \frac{4R_{n+1}}{l_{n+1}} - \frac{6}{u_{n+1}}.\end{aligned}$$

We proceed with the case of $m > 2$. There is at most one pair of n-blocks $\mathtt{B}_{i_1,j_1}^{(n+m-1)}$ and $\mathtt{B}_{i_2,j_2}^{(n+m-1)}$ in $\mathtt{G}_{i_1,s_1}^{(n+m-1)}$ and $\mathtt{G}_{i_2,s_2}^{(n+m-1)}$ respectively with $j_1 = j_2 + 1$, while for all other pairs $j_1 \neq j_2$ and $j_1 \neq j_2 + 1$. Since this corresponds to a proportion of at most $2/u_{n+m-1}$, we use the modified versions of equations (5.13) and (5.14) and Proposition 5.4 (with $\tilde{f} \geq 1 - (2/u_{n+m-1})$) to obtain the estimate on $\overline{f}(\mathcal{G}, \overline{\mathcal{G}})$ by the same calculation as above, replacing the subscripts $n+1$ by $n+m-1$ and the term $1 - (i_2 - i_1/K)$ by $1 - (1 - (1/K))^{m-1}$.

In all other cases $s_1 \neq s_2$ all pairs of n-blocks $\mathtt{B}_{i_1,j_1}^{(n+m-1)}$ and $\mathtt{B}_{i_2,j_2}^{(n+m-1)}$ in $\mathtt{G}_{i_1,s_1}^{(n+m-1)}$ and $\mathtt{G}_{i_2,s_2}^{(n+m-1)}$ respectively have $j_1 \neq j_2$ and $j_1 \neq j_2 + 1$. Then we use the same estimate as in part (1). □

Let us introduce the notation

$$\gamma_{n+m-1} := E_{n+m-1} + \frac{4}{R_{n+m-1}} + \frac{4R_{n+m-1}}{l_{n+m-1}} + \frac{6}{u_{n+m-1}}.$$

We continue the inductive construction by building $2K^2\lambda_{n+m} = 2K^2 d_{n+m}e_{n+m}$ *pre-*$(n+m)$*-blocks* $\mathtt{A}_{i,j}^{(n+m)}$, where $j \in \{1, \ldots, 2\lambda_{n+m}K\}$ stands for the jth Feldman pattern and $i \in \{1, \ldots, K\}$ indicates the type of $(n+m-1)$-blocks used. We let $\mathtt{A}_{i,j}^{(n+m)}$ be the jth Feldman pattern built of the grouped $(n+m-1)$-blocks of type i. We point out that $\mathtt{A}_{i,j}^{(n+m)}$ contains each $(n+m-1)$-block of type i

$$\bar{N}(n+m) = (e_{n+m-1})^{4K\lambda_{n+m}+2} \tag{5.16}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

times because it is uniform in the building blocks $\mathtt{G}_{i,s}^{(n+m-1)}$ by construction of the Feldman patterns in Proposition 5.10 and each $(n+m-1)$-block of type i is contained in exactly one grouped $(n+m-1)$-block. Moreover, this number of occurrences is the same for every chosen Feldman pattern j. We will denote the length of the circular image $\mathcal{A}_{i,j}^{(n+m)} = \mathcal{C}_{n+m-1,k}(\mathtt{A}_{i,j}^{(n+m)})$, $k \in \{0, \ldots, q_{n+m-1}-1\}$, by $\tilde{q}_{n+m}$.

LEMMA 5.19. (Closeness and separation of pre-$(n+m)$-blocks of same Feldman pattern) *For every $j \in \{1, \ldots, 2Kd_{n+m}e_{n+m}\}$ and $i_1, i_2 \in \{1, \ldots, K\}$ we have*

$$\overline{f}(\mathtt{A}_{i_1,j}^{(n+m)}, \mathtt{A}_{i_2,j}^{(n+m)}) \leq \sum_{u=1}^{m-1} \left(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}} \right)$$

in the odometer-based system. In the circular system we have for $i_1 \neq i_2$ and for any sequences $\mathcal{A}$ and $\overline{\mathcal{A}}$ of at least $\tilde{q}_{n+m}/e_{n+m-1}$ consecutive symbols in $\mathcal{A}_{i_1,j}^{(n+m)}$ and $\mathcal{A}_{i_2,j}^{(n+m)}$ respectively that

$$\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) \geq \left(1 - \left(1 - \frac{1}{K}\right)^{m-1}\right)\alpha - \gamma_{n+m-1} - \frac{2}{e_{n+m-1}} - \frac{2}{u_{n+m-1}}.$$

Proof. Without loss of generality, let $i_2 > i_1$.

We start with the proof of the first statement for $m = 2$. Ignoring the marker segment $\mathtt{a}_n$, we note that for every $s = 0, \ldots, e_{n+1}-1$ the grouped n-block $\mathtt{G}_{i_2,s}^{(n+1)}$ is obtained from $\mathtt{G}_{i_1,s}^{(n+1)}$ by a shift of $i_2 - i_1$ pre-$(n+1)$-blocks. Since a grouped $(n+1)$-block consists of Kd_{n+1} pre-$(n+1)$-blocks, the grouped $(n+1)$-blocks $\mathtt{G}_{i_1,s}^{(n+1)}$, $\mathtt{G}_{i_2,s}^{(n+1)}$ can be matched on a portion of at least $1 - ((i_2 - i_1)/Kd_{n+1})$ of the part of pre-$(n+1)$-blocks. Thus, we have

$$\overline{f}(\mathtt{G}_{i_1,s}^{(n+1)}, \mathtt{G}_{i_2,s}^{(n+1)}) \leq \frac{1}{N+1} + \frac{N}{N+1} \cdot \frac{i_2 - i_1}{Kd_{n+1}},$$

which yields the claim because $\mathtt{A}_{i_1,j}^{(n+2)}$ and $\mathtt{A}_{i_2,j}^{(n+2)}$ are constructed as the same Feldman pattern with these grouped $(n+1)$-blocks of different types as building blocks.

Turning to the proof of the first statement for $m > 2$, we ignore the marker blocks which occupy a fraction

$$\frac{1}{K\lambda_{n+m-2}+1} = \frac{1}{N(n+m-2)+1}$$

of each $(n+m-1)$-block (since there are $1 + K\lambda_{n+m-2}$ $(n+m-2)$-blocks, it follows from the uniformity of $(n+m-2)$-blocks within each $(n+m-1)$-block that the marker segment occupies a fraction $1/(1+K\lambda_{n+m-2})$ of the $(n+m-1)$-block). Let M denote the right-hand side from inequality (5.11), that is, an upper bound for the $\overline{f}$ distance between pre-$(n+m-1)$-blocks of type i_1 and i_2 and the same Feldman pattern. Since in the definition of $(n+m-1)$-blocks of types i_1 and i_2 the Feldman patterns used are shifted by $i_2 - i_1$, we obtain, for every $s \in \{0, \ldots, e_{n+m-1}-1\}\mathtt{G}_{i_1,s}^{(n+m-1)}$ and $\mathtt{G}_{i_2,s}^{(n+m-1)}$,

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$$\overline{f}(\mathtt{G}_{i_1,s}^{(n+m-1)}, \mathtt{G}_{i_2,s}^{(n+m-1)}) \le \frac{1}{N(n+m-2)+1} + \frac{(Kd_{n+m-1} - |i_1 - i_2|)}{Kd_{n+m-1}} \cdot M + \frac{|i_1 - i_2|}{Kd_{n+m-1}}$$
$$\le \sum_{u=1}^{m-1} \left(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}} \right)$$

using equation (5.11). This yields the claim because $\mathtt{A}_{i_1,j}^{(n+m)}$ and $\mathtt{A}_{i_2,j}^{(n+m)}$ are constructed as the same Feldman pattern with these grouped $(n+m-1)$-blocks of different type as building blocks.

The second statement follows from Lemmas 5.11 and 5.18. □

We will need a statement on the $\overline{f}$ distance of different Feldman patterns in the circular system only, but not in the odometer-based system.

LEMMA 5.20. (Separation of pre-$(n+m)$-blocks of different Feldman patterns) *For any sequences $\mathcal{A}$ and $\overline{\mathcal{A}}$ of at least $\tilde{q}_{n+m}/e_{n+m-1}$ consecutive symbols in $\mathcal{A}_{i_1,j_1}^{(n+m)}$ and $\mathcal{A}_{i_2,j_2}^{(n+m)}$ for some $i_1, i_2 \in \{1, \ldots, K\}$ and $j_1, j_2 \in \{1, \ldots, 2Kd_{n+m}e_{n+m}\}$, $j_1 \neq j_2$, we have*

$$\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) \ge \alpha - \gamma_{n+m-1} - \frac{13}{\sqrt{e_{n+m-1}}} - \frac{2}{u_{n+m-1}}.$$

Proof. The result follows from Proposition 5.10 and Lemma 5.18. □

Then we set

$$\beta_{n+m} := \gamma_{n+m-1} + \frac{13}{\sqrt{e_{n+m-1}}} + \frac{2}{u_{n+m-1}}.$$

We continue our construction process by building $(n+m)$-*blocks* $\mathtt{B}_{i,j}^{(n+m)}$ of K different types (once again, the index $1 \le i \le K$ indicates the type, and $1 \le j \le \lambda = \lambda_{n+m}$ numbers the $(n+m)$-blocks of that type consecutively) using the formula from the beginning of §5.3.2 (with $n+m-1$ replaced by $n+m$). Moreover, we define an additional $(n+m)$-block $\mathtt{B}_0^{(n+m)}$ which will play the role of a marker again:

$$\mathtt{B}_0^{(n+m)} = \mathtt{A}_{1,(\lambda_{n+m}+1)K}^{(n+m)} \mathtt{A}_{2,(\lambda_{n+m}+1)K}^{(n+m)} \ldots \mathtt{A}_{K,(\lambda_{n+m}+1)K}^{(n+m)} \mathtt{a}_{n+m-1},$$

where the pre-$(n+m)$-blocks $\mathtt{A}_{i,(\lambda_{n+m}+1)K}^{(n+m)}$ are not used in any other $(n+m)$-block. Hence, there are $K\lambda_{n+m} + 1$ $(n+m)$-blocks in total. We also note that each $(n+m)$-block $\mathtt{B}_{i,s}^{(n+m)}$ contains exactly one pre-$(n+m)$-block of each type. Thus the $(n+m)$-blocks are uniform in the $(n+m-1)$-blocks by our observation above.

LEMMA 5.21. (Distance between $(n+m)$-blocks in the circular system) *Let $\mathcal{B}$ and $\overline{\mathcal{B}}$ be sequences of at least q_{n+m}/R_{n+m} consecutive symbols in $\mathcal{B}_{i_1,j_1}^{(n+m)}$ and $\mathcal{B}_{i_2,j_2}^{(n+m)}$ for some $i_1, i_2 \in \{0, 1, \ldots, K\}$ and $j_1, j_2 \in \{1, \ldots, \lambda_{n+m}\}$.*

(1) *For blocks of same type, for $i_1 = i_2$ and $j_1 \neq j_2$ we have*

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \ge \alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}.$$

(2) *For blocks of different type, for $i_1 < i_2$ we have*

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \left(1 - \left(1 - \frac{1}{K}\right)^m\right) \cdot \alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}$$

if $j_1 = j_2$ or $j_1 = j_2 + 1$. For all other cases of $j_1 \neq j_2$ we have

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}.$$

Proof. As in the proof of Lemma 5.17 we consider $\mathcal{B}$ and $\overline{\mathcal{B}}$ to be a concatenation of complete 2-subsections. ignoring incomplete ones at the beginnings and ends which constitute a fraction of at most $2/l_{n+m-2}$ of the total length of $\mathcal{B}$ and $\overline{\mathcal{B}}$ by equation (5.2). In the following consideration we ignore the marker segments $\mathtt{a}_{n+m-1}$ which amount to a fraction $1/(1 + K\lambda_{n+m-1}) < (1/e_{n+m-1})$ of the total length due to uniformity. Accordingly, we consider $\mathcal{B}$ and $\overline{\mathcal{B}}$ to be a concatenation of complete pre-$(n + m)$-blocks $\mathcal{A}^{(n+m)}_{i_1,h_1}$ and $\mathcal{A}^{(n+m)}_{i_2,h_2}$, respectively. Finally, we examine the particular situation of each case of this lemma:

(1) We note that all Feldman patterns for pre-$(n + m)$-blocks used in $\mathtt{B}^{(n+m)}_{i_1,j_1}$ and $\mathtt{B}^{(n+m)}_{i_1,j_2}$ are different from each other. Accordingly, we apply Lemma 5.20 and Corollary 5.5 (with $\tilde{f} = 1$).

(2) For $i_1 < i_2$ the blocks $\mathtt{B}^{(n+m)}_{i_1,j_2}$ and $\mathtt{B}^{(n+m)}_{i_2,j_2}$ have $K - (i_2 - i_1)$ Feldman patterns in common. Then we use the second statement of Lemma 5.19 to estimate their $\overline{f}$ distance, while we use Lemma 5.20 for the $i_2 - i_1$ differing Feldman patterns. Altogether we conclude with the aid of Proposition 5.4 that $\overline{f}(\mathcal{B}, \overline{\mathcal{B}})$ is at least

$$\begin{aligned}&\left(\frac{i_2 - i_1}{K} + \left(1 - \left(1 - \frac{1}{K}\right)^{m-1}\right)\frac{K - (i_2 - i_1)}{K}\right)\alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}\\ &\quad\geq \left(1 - \left(1 - \frac{1}{K}\right)^m\right)\alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}.\end{aligned}$$

In our case of $i_1 < i_2$ the blocks $\mathtt{B}^{(n+m)}_{i_1,j_2+1}$ and $\mathtt{B}^{(n+m)}_{i_2,j_2}$ have $i_2 - i_1$ Feldman patterns in common. With the aid of the second part of Lemma 5.19, Lemma 5.20, and Proposition 5.4 again we obtain that $\overline{f}(\mathcal{B}, \overline{\mathcal{B}})$ is at least

$$\begin{aligned}&\left(\frac{i_2 - i_1}{K}\left(1 - \left(1 - \frac{1}{K}\right)^{m-1}\right) + 1 - \frac{i_2 - i_1}{K}\right)\alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}\\ &\quad\geq \left(1 - \left(1 - \frac{1}{K}\right)^m\right)\alpha - \beta_{n+m} - \frac{4}{e_{n+m-1}} - \frac{4}{l_{n+m-2}}.\end{aligned}$$

In all other cases of $j_1 \neq j_2$, $\mathtt{B}^{(n+m)}_{i_1,j_1}$ and $\mathtt{B}^{(n+m)}_{i_2,j_2}$ do not have any Feldman pattern in common. Hence, we use the same estimate as for the statement in part (1). □

By the conditions on the sequence $(l_n)_{n\in\mathbb{N}}$ from equation (5.3) we have

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$$\begin{aligned}\beta_{n+m} &+ \frac{4}{e_{n+m-1}} + \frac{4}{l_{n+m-2}} \\ &= E_{n+m-1} + \frac{4}{R_{n+m-1}} + \frac{8}{u_{n+m-1}} + \frac{13}{\sqrt{e_{n+m-1}}} + \frac{4}{e_{n+m-1}} + \frac{4}{l_{n+m-2}} + \frac{4R_{n+m-1}}{l_{n+m-1}} \\ &\leq E_{n+m-1} + \frac{6}{R_{n+m-1}} + \frac{8}{u_{n+m-1}} + \frac{17}{\sqrt{e_{n+m-1}}}.\end{aligned}$$

Accordingly, we set

$$\begin{aligned}E_{n+m} &:= E_{n+m-1} + \frac{6}{R_{n+m-1}} + \frac{8}{u_{n+m-1}} + \frac{17}{\sqrt{e_{n+m-1}}} \\ &= E_{n+1} + \sum_{i=1}^{m-1}\bigg(\frac{6}{R_{n+i}} + \frac{8}{u_{n+i}} + \frac{17}{\sqrt{e_{n+i}}}\bigg).\end{aligned}$$

Remark 5.22. We note that equations (5.11)–(5.15) are satisfied at stage $n+m$. Hence, the induction step has been accomplished successfully.

5.3.3. *Final step: construction of $(n+p)$-blocks.* Recall that we follow the inductive scheme described in the previous subsubsection until

$$\bigg(1-\frac{1}{K}\bigg)^p < \frac{\varepsilon}{2},$$

and we have constructed pre-$(n+p)$-blocks $\mathtt{A}_{i,j}^{(n+p)}$ (of type i and Feldman pattern j) with the following properties:

$$\overline{f}(\mathtt{A}_{i_1,j}^{(n+p)}, \mathtt{A}_{i_2,j}^{(n+p)}) \leq \sum_{u=1}^{p-1}\bigg(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}}\bigg); \tag{5.17}$$

for any sequences $\mathcal{A}$ and $\overline{\mathcal{A}}$ of at least $\tilde{q}_{n+p}/e_{n+p-1}$ consecutive symbols in $\mathcal{A}_{i_1,j_1}^{(n+p)}$ and $\mathcal{A}_{i_2,j_2}^{(n+p)}$ respectively, for some $i_1, i_2 \in \{1, \ldots, K\}$ and $j_1, j_2 \in \{1, \ldots, 2K\lambda_{n+p}\}$, we have

$$\overline{f}(\mathcal{A}, \overline{\mathcal{A}}) \geq \begin{cases} \Big(1-\frac{\varepsilon}{2}\Big)\cdot\alpha - \beta_{n+p} & \text{for all } i_1 \neq i_2 \text{ and } j_1 = j_2, \\ \alpha - \beta_{n+p} & \text{for all } i_1, i_2 \text{ and } j_1 \neq j_2, \end{cases} \tag{5.18}$$

where

$$\beta_{n+p} = E_{n+p-1} + \frac{4}{R_{n+p-1}} + \frac{8}{u_{n+p-1}} + \frac{13}{\sqrt{e_{n+p-1}}} + \frac{4R_{n+p-1}}{l_{n+p-1}} \tag{5.19}$$

and

$$E_{n+p-1} = E_{n+1} + \sum_{i=1}^{p-2}\bigg(\frac{6}{R_{n+i}} + \frac{8}{u_{n+i}} + \frac{17}{\sqrt{e_{n+i}}}\bigg). \tag{5.20}$$

By construction every pre-$(n+p)$-block $\mathtt{A}_{i,j}^{(n+p)}$ contains each $(n+p-1)$-block of type i exactly $\bar{N}(n+p)$ times and this number of occurrences is the same for every chosen Feldman pattern j.

Then we construct $K(n+p)$-blocks as follows:

$$
\begin{aligned}
(n+p)\text{-block of type } 1\colon\ \mathtt{B}_1^{(n+p)} &= \mathtt{A}_{1,1}^{(n+p)}\mathtt{A}_{2,2}^{(n+p)}\ldots\mathtt{A}_{K,K}^{(n+p)}(\mathtt{B}_0^{(n+p-1)})^{\bar{N}(n+p)},\\
(n+p)\text{-block of type } 2\colon\ \mathtt{B}_2^{(n+p)} &= \mathtt{A}_{2,1}^{(n+p)}\mathtt{A}_{3,2}^{(n+p)}\ldots\mathtt{A}_{1,K}^{(n+p)}(\mathtt{B}_0^{(n+p-1)})^{\bar{N}(n+p)},\\
(n+p)\text{-block of type } 3\colon\ \mathtt{B}_3^{(n+p)} &= \mathtt{A}_{3,1}^{(n+p)}\mathtt{A}_{4,2}^{(n+p)}\ldots\mathtt{A}_{2,K}^{(n+p)}(\mathtt{B}_0^{(n+p-1)})^{\bar{N}(n+p)},\\
&\vdots\\
(n+p)\text{-block of type } K\colon\ \mathtt{B}_K^{(n+p)} &= \mathtt{A}_{K,1}^{(n+p)}\mathtt{A}_{1,2}^{(n+p)}\ldots\mathtt{A}_{K-1,K}^{(n+p)}(\mathtt{B}_0^{(n+p-1)})^{\bar{N}(n+p)}.
\end{aligned}
$$

Remark 5.23. We note that every $(n+p)$-block contains exactly one pre-$(n+p)$-block of each type. Hence, it is uniform in the $(n+p-1)$-blocks by our observation above.

LEMMA 5.24. (Closeness of $(n+p)$-blocks in the odometer-based system) *For every $i_1, i_2 \in \{1,\ldots,K\}$ we have*

$$
\overline{f}(\mathtt{B}_{i_1}^{(n+p)}, \mathtt{B}_{i_2}^{(n+p)}) \le \sum_{u=1}^{p-1}\left(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}}\right). \tag{5.21}
$$

Proof. Let M denote the right-hand side of inequality (5.17). We observe that the Feldman patterns of pre-$(n+p)$-blocks are aligned in all the $(n+p)$-blocks by construction. Moreover, the marker segments are aligned as well and these occupy a fraction $1/(N(n+p-1)+1)$ of each $(n+p)$-block by uniformity. Hence, by equation (5.17) we have

$$
\overline{f}(\mathtt{B}_{i_1}^{(n+p)}, \mathtt{B}_{i_2}^{(n+p)}) \le \left(1 - \frac{1}{N(n+p-1)+1}\right)\cdot M \le M.
$$

□

LEMMA 5.25. (Distance between $(n+p)$-blocks in the circular system) *For every $i_1, i_2 \in \{1,\ldots,K\}$, $i_1 \neq i_2$, and any sequences $\mathcal{B}$ and $\overline{\mathcal{B}}$ of at least q_{n+p}/R_{n+p} consecutive symbols in $\mathcal{B}_{i_1}^{(n+p)}$ and $\mathcal{B}_{i_2}^{(n+p)}$ we have*

$$
\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \ge \left(1 - \frac{\varepsilon}{2}\right)\cdot\alpha - E_{n+p}. \tag{5.22}
$$

Proof. By the same proof as in Lemma 5.17 (with $J = 0$), as well as case (1) of Lemma 5.21, and the observation that all pre-$(n+p)$-blocks used in the construction of the $(n+p)$-blocks are distinct, (5.22) follows from equations (5.18)–(5.20). □

Proof of Proposition 5.15. By Lemma 5.24 and equations (5.4), (5.6), and (5.7), we have

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$$\overline{f}(\mathtt{B}_i^{(n+p)}, \mathtt{B}_j^{(n+p)}) \leq \sum_{u=1}^{p-1} \left(\frac{1}{N(n+u-1)+1} + \frac{1}{d_{n+u}} \right)$$
$$< \frac{1}{N+1} + \sum_{u=1}^{p-1} \left(\frac{1}{K d_{n+u} e_{n+u} + 1} + \frac{1}{d_{n+u}} \right) < \delta,$$

which is the first statement of Proposition 5.15. In order to prove the second statement we note that

$$E_{n+p} = E_{n+1} + \sum_{i=1}^{p-1} \left(\frac{6}{R_{n+i}} + \frac{8}{u_{n+i}} + \frac{17}{\sqrt{e_{n+i}}} \right)$$
$$\leq \frac{14}{\sqrt{N}} + \sum_{i=1}^{p-1} \left(\frac{8}{u_{n+i}} + \frac{17}{\sqrt{e_{n+i}}} \right) + \sum_{i=0}^{p-1} \frac{6}{R_{n+i}}$$
$$< \frac{\varepsilon}{2} + \sum_{i=0}^{p-1} \frac{6}{R_{n+i}}$$

by equations (5.5), (5.7), and our assumption on the circular coefficients $(l_n)_{n\in\mathbb{N}}$ in (5.3). We conclude for any sequences $\mathcal{B}$ and $\overline{\mathcal{B}}$ of at least q_{n+p}/R_{n+p} consecutive symbols in $\mathcal{B}_i^{(n+p)}$ and $\mathcal{B}_j^{(n+p)}$, respectively, that

$$\overline{f}(\mathcal{B}, \overline{\mathcal{B}}) \geq \left(1 - \frac{\varepsilon}{2}\right) \cdot \alpha - E_{n+p} \geq \alpha - \frac{\varepsilon}{2} - E_{n+p} > \alpha - \varepsilon - \sum_{i=0}^{p-1} \frac{6}{R_{n+i}}$$

with the aid of Lemma 5.25. □

5.4. *Proof of Theorem 5.1.* We define the construction sequence for the odometer-based system inductively. We begin by choosing an integer $R_1 \geq 400$, an increasing sequence $(K_s)_{s\in\mathbb{N}}$ of positive integers such that

$$\sum_{s\in\mathbb{N}} \frac{14}{\sqrt{K_s}} < \frac{1}{32}, \tag{5.23}$$

and two decreasing sequences $(\varepsilon_s)_{s\in\mathbb{N}}$ and $(\delta_s)_{s\in\mathbb{N}}$ of positive real numbers such that $\delta_s \searrow 0$ and

$$\sum_{s\in\mathbb{N}} \varepsilon_s < \frac{1}{64}. \tag{5.24}$$

In addition to (5.3) we impose the condition

$$\frac{6}{R_1} + \sum_{n\in\mathbb{N}} \frac{6}{k_n} < \frac{1}{32}$$

on the circular coefficients $(k_n, l_n)_{n\in\mathbb{N}}$. In particular, this yields

$$\sum_{n=1}^{\infty} \frac{6}{R_n} < \frac{1}{32} \tag{5.25}$$

by $R_n = k_{n-2} q_{n-2}^2$ for $n \geq 2$. We start with $R_1 + 1$ symbols and let $\alpha_0 = \frac{1}{8}$.

The first application of Proposition 5.15 will be on 1-blocks and we will apply it for $\varepsilon = \varepsilon_1$, $\delta = \delta_1$, and $K = K_1 + 1$. In order to apply the proposition we need sufficiently many 1-blocks. Moreover, we want the 1-blocks in the circular system to be at least $\alpha_0 - \varepsilon_1$ apart in the $\overline{f}$ metric on substantial substrings of at least q_1/R_1 consecutive symbols. To produce such a family of 1-blocks we apply Proposition 5.13.

After the application of Proposition 5.15 on the 1-blocks we have $K_1 + 1$ n_1-blocks (where $n_1 = 1 + p_1$ with p_1 from Proposition 5.15, that is, the least integer such that $(1 - (1/K_1 + 1))^{p_1} < \varepsilon_1/2)$ which are δ_1-close in the odometer-based system and at least $\alpha_1 := \alpha_0 - 2\varepsilon_1 - \sum_{s=1}^{n_1-1}(6/R_s)$ apart in the $\overline{f}$ metric on substantial substrings of at least q_{n_1}/R_{n_1} consecutive symbols. The next application of Proposition 5.15 will be on $(n_1 + 1)$-blocks and we will apply it for $\varepsilon = \varepsilon_2$, $\delta = \delta_2$, and $K = K_2 + 1$. Once again this will require sufficiently many $(n_1 + 1)$-blocks, and we apply Proposition 5.13 to produce such a family of $(n_1 + 1)$-blocks that are at least $\alpha_1 - (14/\sqrt{K_1}) - (4/R_{n_1}) - \varepsilon_2$ apart on substantial substrings. This imposes the condition $(2/l_{n_1-1}) + (4R_{n_1}/l_{n_1}) < \varepsilon_2$ which by condition (5.3) is fulfilled if $(3/l_{n_1-1}) < \varepsilon_2$, and we choose l_{n_1-1} sufficiently large to satisfy this requirement.

Continuing like this, we use Propositions 5.13 and 5.15 alternately to produce $K_s + 1$ n_s-blocks which are at least

$$\alpha_s = \alpha_0 - \sum_{i=1}^{s-1} \frac{14}{\sqrt{K_i}} - \sum_{i=1}^{s} 2\varepsilon_i - \sum_{i=1}^{n_s-1} \frac{6}{R_i} \tag{5.26}$$

apart on substantial substrings of length at least $h_{n_s}/(K_s + 1)$ in the circular system and δ_s-close in the odometer-based system. In the next step, we want to apply Proposition 5.15 on $(n_s + 1)$-blocks with $\varepsilon = \varepsilon_{s+1}$, $\delta = \delta_{s+1}$, and $K = K_{s+1} + 1$. In order to have sufficiently many $(n_s + 1)$-blocks for this application we make use of Proposition 5.13 (imposing the condition on l_{n_s-1} as described above) and produce the required number of $(n_s + 1)$-blocks which are at least $\alpha_s - (14/\sqrt{K_s}) - (4/R_{n_s}) - \varepsilon_{s+1}$ apart on substantial substrings in the circular system. After the intended application of Proposition 5.15 we have $K_{s+1} + 1$ n_{s+1}-blocks (where $n_{s+1} = n_s + 1 + p_{s+1}$ with p_{s+1} the least integer such that $(1 - (1/(K_{s+1} + 1)))^{p_{s+1}} < (\varepsilon_{s+1}/2))$ which are at least

$$\alpha_{s+1} = \alpha_s - \frac{14}{\sqrt{K_s}} - 2\varepsilon_{s+1} - \sum_{i=n_s}^{n_{s+1}-1} \frac{6}{R_i} \tag{5.27}$$

apart on substantial substrings in the circular system and δ_{s+1}-close in the odometer-based system. This completes the inductive step.

By the requirements (5.23)–(5.25) this shows that any two distinct n_s-blocks in the circular system are at least

$$\alpha_0 - \sum_{s=1}^{\infty} \left(\frac{14}{\sqrt{K_s}} + 2\varepsilon_s + \frac{6}{R_s} \right) > \frac{1}{8} - \frac{1}{32} - \frac{1}{32} - \frac{1}{32} = \frac{1}{32} \tag{5.28}$$

apart on substantial substrings. Hence, the circular system cannot be loosely Bernoulli by Lemma 3.5.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

On the other hand, the $\overline{f}$ distance between n_s-blocks in the odometer-based system goes to zero because $\delta_s \searrow 0$. Thus, the odometer-based system is zero-entropy loosely Bernoulli by Lemma 3.5. Since the blocks constructed by Propositions 5.13 and 5.15 satisfy the properties of uniformity and unique readability, our construction sequence satisfies those as well.

6. *Non-loosely Bernoulli odometer-based system whose corresponding circular system is loosely Bernoulli*

In the previous two sections we showed that $\mathcal{F}$ does not preserve the loosely Bernoulli property. In this section we will show that $\mathcal{F}^{-1}$ also does not preserve the loosely Bernoulli property.

THEOREM 6.1. *There exist circular coefficients $(l_n)_{n\in\mathbb{N}}$ and a non-loosely Bernoulli odometer-based system $\mathbb{M}$ of zero measure-theoretic entropy with uniform and uniquely readable construction sequence such that $\mathcal{F}(\mathbb{M})$ is loosely Bernoulli.*

The proof of Theorem 6.1 is based upon two mechanisms. On the one hand, we again use Feldman patterns to produce an arbitrarily large number of new blocks whose substantial substrings are almost as far apart in $\overline{f}$ as the building blocks. This time we obtain lower bounds on the $\overline{f}$-distance between blocks in the odometer-based system. On the other hand, we develop the so-called *cycling mechanism* in §6.3 to produce an arbitrarily large number of blocks that are still $\overline{f}$-apart from each other in the odometer-based system but arbitrarily close to each other in the circular system. See Figure 3 for a sketch of this idea.

6.1. *Feldman patterns revisited.*

PROPOSITION 6.2. *Let $\alpha \in \left(0, \frac{1}{7}\right)$ and $n \in \mathbb{N}$, $K, R, S, N, M \in \mathbb{N}\setminus\{0\}$ with $N \geq 20$, and $M \geq 2$. For $1 \leq s \leq S$, let $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ be a family of strings, where each $\mathtt{A}_j^{(s)}$ is a concatenation of K n-blocks. Assume that for all $1 \leq s_1, s_2 \leq S$ and all $j_1, j_2 \in \{1, \ldots, N\}$ with $j_1 \neq j_2$, we have $\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) > \alpha$ for all strings $\mathtt{A}, \overline{\mathtt{A}}$ of at least $K\mathtt{h}_n/R$ consecutive symbols from $\mathtt{A}_{j_1}^{(s_1)}$ and $\mathtt{A}_{j_2}^{(s_2)}$, respectively.*

Then for $1 \leq s \leq S$, we can construct a family of strings $\mathtt{B}_1^{(s)}, \ldots, \mathtt{B}_M^{(s)}$ (of equal length $N^{2M+3}\cdot K\cdot \mathtt{h}_n$ and containing each block $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ exactly N^{2M+2} times) such that for all $1 \leq s_1, s_2 \leq S$, all $j, k \in \{1, \ldots, M\}$ with $j \neq k$, and all strings $\mathtt{B}$, $\overline{\mathtt{B}}$ of at least $N^{2M+2}\cdot K\cdot \mathtt{h}_n$ consecutive symbols from $\mathtt{B}_j^{(s_1)}$ and $\mathtt{B}_k^{(s_2)}$ respectively, we have

$$\overline{f}(\mathtt{B}, \overline{\mathtt{B}}) > \alpha - \frac{13}{\sqrt{N}} - \frac{1}{R}.$$

Proof. The proof follows along the lines of the statement on Feldman patterns for the circular system in Proposition 5.10. (Here $\mathcal{C}_{n,i_1}$ and $\mathcal{C}_{n,i_2}$ are not applied since we remain in the odometer-based system.) □

We will also need a statement on the $\overline{f}$ distance between the identical Feldman pattern but with building blocks from different families (compare with Lemma 5.11).

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

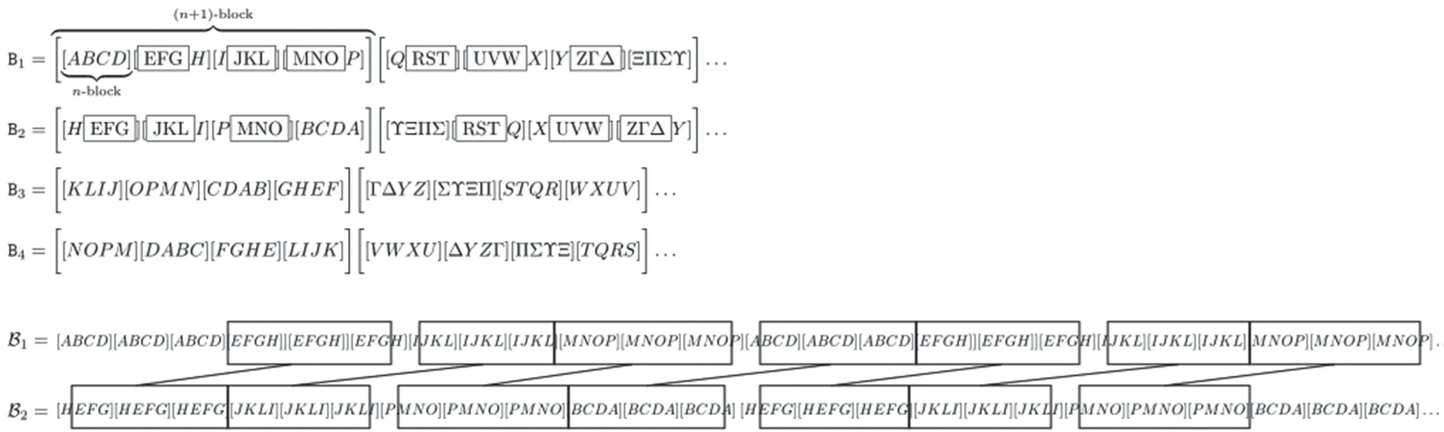

FIGURE 3. Heuristic representation of two stages of the cycling mechanism. Parts of four $(n+2)$-blocks $\mathtt{B}_i$, $i = 1, \ldots, 4$, in the odometer-based system and parts of the images $\mathcal{B}_1$ and $\mathcal{B}_2$ under the circular operator (omitting the spacers) are represented. The marked letters indicate a best possible $\overline{f}$ match between $\mathtt{B}_1$ and $\mathtt{B}_2$ with a fit of approximately $\left(1 - \frac{1}{4}\right)^2$ (ignoring boundary effects), while the blocks $\mathcal{B}_1$ and $\mathcal{B}_2$ have a very good fit in the circular system (the lines indicating a best possible $\overline{f}$ match).

LEMMA 6.3. *Let* $\alpha \in \left(0, \frac{1}{7}\right)$ *and* $n \in \mathbb{N}$, $K, R, S, N, M \in \mathbb{N} \setminus \{0\}$ *with* $N \geq 20$, *and* M, S *at least* 2. *For* $1 \leq s \leq S$, *let* $\mathtt{A}_1^{(s)}, \ldots, \mathtt{A}_N^{(s)}$ *be a family of strings, where each* $\mathtt{A}_j^{(s)}$ *is a concatenation of* K n*-blocks. Assume that for all* $j_1, j_2 \in \{1, \ldots, N\}$ *and all* $s_1, s_2 \in \{1, \ldots, S\}$ *with* $s_1 \neq s_2$, *we have* $\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) > \alpha$ *for all sequences* $\mathtt{A}, \overline{\mathtt{A}}$ *of at least* $K\mathtt{h}_n/R$ *consecutive symbols from* $\mathtt{A}_{j_1}^{(s_1)}$ *and* $\mathtt{A}_{j_2}^{(s_2)}$, *respectively. Then for* $1 \leq s \leq S$, *we can construct a family of strings* $\mathtt{B}_1^{(s)}, \ldots, \mathtt{B}_M^{(s)}$ *as in Proposition 6.2 such that for all* $j, k \in \{1, \ldots, M\}$, *all* $s_1, s_2 \in \{1, \ldots, S\}$ *with* $s_1 \neq s_2$, *and all strings* $\mathtt{B}, \overline{\mathtt{B}}$ *of at least* $N^{2M+2}K\mathtt{h}_n$ *consecutive symbols from* $\mathtt{B}_j^{(s_1)}$ *and* $\mathtt{B}_k^{(s_2)}$ *respectively, we have*

$$\overline{f}(\mathtt{B}, \overline{\mathtt{B}}) > \alpha - \frac{2}{N^{2M+2}} - \frac{1}{R}.$$

6.2. *Feldman mechanism in the odometer-based system.* Again we use the *Feldman mechanism* to produce an arbitrarily large number of new blocks whose substantial substrings are almost as far apart in $\overline{f}$ as the building blocks. In contrast to Proposition 6.2, the constructed words also satisfy the unique readability property.

PROPOSITION 6.4. *Let* $\alpha \in \left(0, \frac{1}{7}\right)$ *and* $n, N, M, R \in \mathbb{N}$ *with* $R \geq 1$, $N \geq 100$, *and* $M \geq 2$. *Suppose there are* $N + 1$ n*-blocks* $\mathtt{A}_0, \ldots, \mathtt{A}_N$, *which have equal length* $\mathtt{h}_n$ *and satisfy the unique readability property. Furthermore, if* $n > 0$ *assume that for all* $j \neq k$, *we have* $\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) > \alpha$ *for all strings* $\mathtt{A}, \overline{\mathtt{A}}$ *of at least* $\mathtt{h}_n/R$ *consecutive symbols from* $\mathtt{A}_j$ *and* $\mathtt{A}_k$, *respectively. Then we can construct* M $(n + 1)$*-blocks* $\mathtt{B}_1, \ldots, \mathtt{B}_M$ *of equal length* $\mathtt{h}_{n+1}$ *(which are uniform in the n-blocks and satisfy the unique readability property) such that for all* $j \neq k$ *and all strings* $\mathtt{B}, \overline{\mathtt{B}}$ *of at least* $\mathtt{h}_{n+1}/N$ *consecutive symbols from* $\mathtt{B}_j$ *and* $\mathtt{B}_k$ *respectively, we have*

$$\overline{f}(\mathtt{B}, \overline{\mathtt{B}}) > \alpha - \frac{12}{\sqrt{N}} - \frac{1}{R} - \frac{8}{N} \geq \alpha - \frac{13}{\sqrt{N}} - \frac{1}{R}.$$

Proof. We define the $(n + 1)$-blocks as in Proposition 5.13. As in its proof for $n = 0$ we complete cycles at the beginning and end of each $\mathtt{B}$ and $\overline{\mathtt{B}}$. Once again, we remove the marker blocks and apply Corollary 5.5 and $\tilde{f} > 1 - (12/\sqrt{N})$ from Lemma 5.8. □

6.3. *Cycling mechanism.*

PROPOSITION 6.5. *If* $n, K, T \in \mathbb{N}$, $K \geq 2$, $T > 0$, $\alpha \in (4/T, 1/7)$, $\delta > 0$, *and* $\varepsilon \in (0, \alpha/(4K))$, *then there exist* $N, m_n \in \mathbb{N}$ *and circular coefficients* $(k_{n+m}, \ell_{n+m})_{m=0}^{m_n-1}$ *satisfying the following condition. If we are given* $N + 1$ *uniquely readable n-blocks* $\mathtt{B}_0^{(n)}, \mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ *in the odometer-based system such that for all* $i \neq j$ *and all sequences* $\mathtt{A}$ *and* $\overline{\mathtt{A}}$ *of at least* $\mathtt{h}_n/T$ *consecutive symbols from* $\mathtt{B}_i^{(n)}$ *and* $\mathtt{B}_j^{(n)}$ *respectively we have* $\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) \geq \alpha$, *then we can build* K $(n + m_n)$*-blocks* $\mathtt{B}_1^{(n+m_n)}, \ldots, \mathtt{B}_K^{(n+m_n)}$ *of equal length* $\mathtt{h}_{n+m_n}$ *(with the unique readability property and uniformity in all blocks from stage n through* $n + m_n$*) satisfying the following properties.*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

(1) *For all $i \neq j$ and all sequences $\mathtt{B}$ and $\overline{\mathtt{B}}$ of at least $\mathtt{h}_{n+m_n}/K$ consecutive symbols from $\mathtt{B}_i^{(n+m_n)}$ and $\mathtt{B}_j^{(n+m_n)}$ respectively, we have*

$$\overline{f}(\mathtt{B}, \overline{\mathtt{B}}) \geq \alpha - \frac{2}{T} - \varepsilon.$$

(2) *If $\mathcal{B}_1^{(n+m_n)}, \ldots, \mathcal{B}_K^{(n+m_n)}$ are the corresponding circular $(n+m_n)$-blocks, then for all $1 \leq i, j \leq K$ we have*

$$\overline{f}(\mathcal{B}_i^{(n+m_n)}, \mathcal{B}_j^{(n+m_n)}) \leq \delta.$$

Remark 6.6. Thus we obtain a mechanism to produce an arbitrarily large number of $(n+m_n)$-blocks that are still apart from each other in the odometer-based system but arbitrarily close to each other in the circular system. The proof is based on an inductive construction that we call the *cycling mechanism* (indicated in Figure 3) and a final step to guarantee closeness of blocks in the circular system. In each step of the cycling mechanism the blocks of different types are constructed by cycling the K pre-blocks used. On the one hand, this will yield closeness in the circular system under the repetitions in the circular operator. On the other hand, in a matching in the odometer-based system at least one pair of pre-blocks will have a $\overline{f}$ distance close to α and at most $K-1$ pairs will have a smaller $\overline{f}$ distance. Over the course of the construction this distance will increase towards α (see equation (6.9)).

For this construction, let $(u_{n+m})_{m\in\mathbb{N}}$ and $(e_{n+m})_{m\in\mathbb{N}}$ be increasing sequences of positive integers such that

$$\sum_{m\in\mathbb{N}} \frac{2}{K u_{n+m}^2 e_{n+m}} < \frac{\delta}{2} \tag{6.1}$$

and

$$\sum_{m\in\mathbb{N}} \left(\frac{4}{u_{n+m}} + \frac{14}{\sqrt{e_{n+m}}} \right) < \frac{\varepsilon}{8}. \tag{6.2}$$

Additionally, we define the sequences $(\lambda_{n+m})_{m\in\mathbb{N}}$ and $(d_{n+m})_{m\in\mathbb{N}}$ by

$$d_{n+m} = u_{n+m}^2 \tag{6.3}$$

and

$$\lambda_{n+m} = 2 d_{n+m} e_{n+m}.$$

Moreover, we choose N sufficiently large such that

$$\frac{14}{\sqrt{N}} < \frac{\varepsilon}{8}, \tag{6.4}$$

and the positive integers $(l_{n+m})_{m\in\mathbb{N}}$ sufficiently large such that

$$\sum_{m\in\mathbb{N}} \frac{4}{l_{n+m}} < \frac{\delta}{2}. \tag{6.5}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

The terms $l_n, l_{n+1}, \ldots, l_{n+m_n-1}$ are the coefficients in the circular $(n+1)$-, $(n+2)$-, $\ldots$, $(n+m_n)$-blocks, where m_n is defined below. The proof of Proposition 6.5 utilizes the parameters $(\alpha_{n+m})_{m=2}^{m=\infty}$ and $(\beta_{n+m})_{m=2}^{m=\infty}$ defined inductively via

$$\alpha_{n+2} = \alpha - \frac{14}{\sqrt{N}} - \frac{2}{T} - \frac{4}{u_{n+1}} - \frac{13}{\sqrt{e_{n+1}}}, \tag{6.6}$$

$$\beta_{n+2} = \frac{1}{K}\alpha - \frac{14}{\sqrt{N}} - \frac{2}{KT} - \frac{4}{u_{n+1}} - \frac{2}{e_{n+1}}, \tag{6.7}$$

$$\alpha_{n+m+1} = \alpha_{n+2} - \sum_{i=2}^{m}\left(\frac{2}{\lambda_{n+i-1}K+1} + \frac{2}{e_{n+i-1}} + \frac{4}{u_{n+i}} + \frac{13}{\sqrt{e_{n+i}}}\right), \tag{6.8}$$

and

$$\beta_{n+m+1} = \beta_{n+m} + \frac{1}{K}(\alpha_{n+m} - \beta_{n+m}) - \frac{2}{e_{n+m-1}} - \frac{4}{u_{n+m}} - \frac{2}{\lambda_{n+m-1}K+1} - \frac{2}{e_{n+m}}. \tag{6.9}$$

Since $1/K(\alpha - (2/T)) > \alpha/2K > 2\varepsilon$, assumptions (6.2) and (6.4) imply that $\beta_{n+2} > \varepsilon$. We also note that for every $m \geq 2$,

$$\alpha_{n+m} \geq \alpha - \frac{14}{\sqrt{N}} - \frac{2}{T} - \sum_{i=1}^{\infty}\left(\frac{14}{\sqrt{e_{n+i}}} + \frac{4}{u_{n+i}}\right) > \alpha - \frac{2}{T} - \frac{\varepsilon}{8} - \frac{\varepsilon}{8},$$

by equation (6.8) and our assumptions (6.2) and (6.4). Similarly, assumption (6.2) implies that in our β_{n+m} equation (6.9), the terms we subtract from β_2 over the whole course of the construction are bounded by $\varepsilon/8$. Due to these bounds, we can choose m_n as the least integer such that

$$\beta_{n+m_n} > \alpha - \frac{2}{T} - \frac{\varepsilon}{2}, \tag{6.10}$$

and we will apply our inductive construction until stage $n + m_n$.

6.3.1. *Initial step: construction of* $(n+1)$*-blocks.* First of all, we choose one n-block $\mathtt{B}_0^{(n)}$ as a marker. Then we apply Proposition 6.2 on the remaining n-blocks $\mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ to build $\tilde{N}(n+1) := K \cdot (\lambda_{n+1} + 1)$ Feldman patterns denoted by $\mathtt{A}_{i,j}$, $i = 1, \ldots, K$, $j = 1, \ldots, \lambda_{n+1} + 1$. We will call them *pre-*$(n+1)$*-blocks*. In particular, these have length $\tilde{\mathtt{h}}_{n+1} = N^{2\cdot\tilde{N}(n+1)+3} \cdot \mathtt{h}_n$ and are uniform in the n-blocks $\mathtt{B}_1^{(n)}, \ldots, \mathtt{B}_N^{(n)}$ by construction. More precisely, every pre-$(n+1)$-block contains each n-block $\mathtt{B}_i^{(n)}$, $1 \leq i \leq N$, exactly

$$\bar{N}(n) := N^{2\cdot\tilde{N}(n+1)+2} \tag{6.11}$$

times and pre-$(n+1)$-blocks in the circular system have length $\tilde{q}_{n+1} = N^{2\cdot\tilde{N}(n+1)+3} \cdot l_n \cdot q_n$. In order to obtain uniformity, we will define the marker segment by

$$\mathtt{a}_n = (\mathtt{B}_0^{(n)})^{K\bar{N}(n)}.$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

Moreover, we have

$$\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) \geq \alpha - \frac{13}{\sqrt{N}} - \frac{2}{T} \tag{6.12}$$

for any strings $\mathtt{A}$ and $\overline{\mathtt{A}}$ of at least $\tilde{\mathtt{h}}_{n+1}/N = N^{2 \cdot \tilde{N}(n+1)+2} \cdot \mathtt{h}_n$ consecutive symbols in different pre-$(n+1)$-blocks by Proposition 6.2.

Finally, we define the $(n+1)$-blocks:

$$\begin{aligned}
(n+1)\text{-blocks of type 1:}\quad \mathtt{B}_{1,1}^{(n+1)} &= \mathtt{A}_{1,1}\mathtt{A}_{2,1}\ldots\mathtt{A}_{K-1,1}\mathtt{A}_{K,1}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{1,2}^{(n+1)} &= \mathtt{A}_{1,2}\mathtt{A}_{2,2}\ldots\mathtt{A}_{K-1,2}\mathtt{A}_{K,2}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{1,3}^{(n+1)} &= \mathtt{A}_{1,3}\mathtt{A}_{2,3}\ldots\mathtt{A}_{K-1,3}\mathtt{A}_{K,3}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{1,4}^{(n+1)} &= \mathtt{A}_{1,4}\mathtt{A}_{2,4}\ldots\mathtt{A}_{K-1,4}\mathtt{A}_{K,4}\mathtt{a}_{\mathtt{n}},\\
&\vdots\\
\mathtt{B}_{1,\lambda_{n+1}}^{(n+1)} &= \mathtt{A}_{1,\lambda_{n+1}}\mathtt{A}_{2,\lambda_{n+1}}\ldots\mathtt{A}_{K-1,\lambda_{n+1}}\mathtt{A}_{K,\lambda_{n+1}}\mathtt{a}_{\mathtt{n}},\\
(n+1)\text{-blocks of type 2:}\quad \mathtt{B}_{2,1}^{(n+1)} &= \mathtt{A}_{2,1}\mathtt{A}_{3,1}\ldots\mathtt{A}_{K,1}\mathtt{A}_{1,1}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{2,2}^{(n+1)} &= \mathtt{A}_{K,2}\mathtt{A}_{1,2}\ldots\mathtt{A}_{K-1,2}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{2,3}^{(n+1)} &= \mathtt{A}_{2,3}\mathtt{A}_{3,3}\ldots\mathtt{A}_{K,3}\mathtt{A}_{1,3}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{2,4}^{(n+1)} &= \mathtt{A}_{K,4}\mathtt{A}_{1,4}\ldots\mathtt{A}_{K-1,4}\mathtt{a}_{\mathtt{n}},\\
&\vdots\\
\mathtt{B}_{2,\lambda_{n+1}}^{(n+1)} &= \mathtt{A}_{K,\lambda_{n+1}}\mathtt{A}_{1,\lambda_{n+1}}\ldots\mathtt{A}_{K-1,\lambda_{n+1}}\mathtt{a}_{\mathtt{n}},\\
&\vdots\;\vdots\\
(n+1)\text{-blocks of type K:}\quad \mathtt{B}_{K,1}^{(n+1)} &= \mathtt{A}_{K,1}\mathtt{A}_{1,1}\ldots\mathtt{A}_{K-1,1}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{K,2}^{(n+1)} &= \mathtt{A}_{2,2}\mathtt{A}_{3,2}\ldots\mathtt{A}_{K,2}\mathtt{A}_{1,2}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{K,3}^{(n+1)} &= \mathtt{A}_{K,3}\mathtt{A}_{1,3}\ldots\mathtt{A}_{K-1,3}\mathtt{a}_{\mathtt{n}},\\
\mathtt{B}_{K,4}^{(n+1)} &= \mathtt{A}_{2,4}\mathtt{A}_{3,4}\ldots\mathtt{A}_{K,4}\mathtt{A}_{1,4}\mathtt{a}_{\mathtt{n}},\\
&\vdots\\
\mathtt{B}_{K,\lambda_{n+1}}^{(n+1)} &= \mathtt{A}_{2,\lambda_{n+1}}\mathtt{A}_{3,\lambda_{n+1}}\ldots\mathtt{A}_{K,\lambda_{n+1}}\mathtt{A}_{1,\lambda_{n+1}}\mathtt{a}_{\mathtt{n}}.
\end{aligned}$$

Here, the index $i \in \{1, \ldots, K\}$ in $\mathtt{B}_{i,j}^{(n+1)}$ indicates the type and $j = 1, \ldots, \lambda_{n+1}$ numbers the $(n+1)$-blocks of that type consecutively. We note that for j odd the block $\mathtt{B}_{i+1,j}^{(n+1)}$ is obtained from $\mathtt{B}_{i,j}^{(n+1)}$ by cycling the pre-$(n+1)$-blocks to the left. On the other hand, for j even the block $\mathtt{B}_{i+1,j}^{(n+1)}$ is obtained from $\mathtt{B}_{i,j}^{(n+1)}$ by cycling the pre-$(n+1)$-blocks to the right. Additionally, we define the next marker block.

$$\mathtt{B}_0^{(n+1)} = \mathtt{A}_{1,\lambda_{n+1}+1}^{(n+1)}\mathtt{A}_{2,\lambda_{n+1}+1}^{(n+1)}\ldots\mathtt{A}_{K,\lambda_{n+1}+1}^{(n+1)}\mathtt{a}_n,$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

where $\mathtt{A}^{(n+1)}_{i,\lambda_{n+1}+1}$, $i = 1, \ldots, K$, have not been used in any of the other $(n+1)$-blocks. Hence, there are are $N(n+1) + 1 = \lambda_{n+1}K + 1(n+1)$-blocks in total. We also note that every $(n+1)$-block is uniform in the n-blocks by equation (6.11).

6.3.2. *Inductive step: construction of $(n+m)$-blocks.* In an inductive process we construct $(n+m)$-blocks for $m \geq 2$. Assume that in our inductive construction we have constructed $K\lambda_{n+m-1}$ $(n+m-1)$-blocks $\mathtt{B}^{(n+m-1)}_{i,j}$ of K different types, where for $m = 2$ the $(n+1)$-blocks are the ones constructed in §6.3.1 and for $m \geq 3$ the $(n+m-1)$ -blocks are constructed according to the following formula (with $\lambda = \lambda_{n+m-1}$):

$(n+m-1)$-blocks of type 1:

$$\mathtt{B}^{(n+m-1)}_{1,1} = \mathtt{A}^{(n+m-1)}_{1,1}\mathtt{A}^{(n+m-1)}_{2,2} \ldots \mathtt{A}^{(n+m-1)}_{K,K}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{1,2} = \mathtt{A}^{(n+m-1)}_{1,K+1}\mathtt{A}^{(n+m-1)}_{2,K+2} \ldots \mathtt{A}^{(n+m-1)}_{K,2K}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{1,3} = \mathtt{A}^{(n+m-1)}_{1,2K+1}\mathtt{A}^{(n+m-1)}_{2,2K+2} \ldots \mathtt{A}^{(n+m-1)}_{K,3K}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{1,4} = \mathtt{A}^{(n+m-1)}_{1,3K+1}\mathtt{A}^{(n+m-1)}_{2,3K+2} \ldots \mathtt{A}^{(n+m-1)}_{K,4K}\mathtt{a}_{n+m-2},$$
$$\vdots$$
$$\mathtt{B}^{(n+m-1)}_{1,\lambda} = \mathtt{A}^{(n+m-1)}_{1,(\lambda-1)K+1}\mathtt{A}^{(n+m-1)}_{2,(\lambda-1)K+2} \ldots \mathtt{A}^{(n+m-1)}_{K,\lambda K}\mathtt{a}_{n+m-2},$$

$(n+m-1)$-blocks of type 2:

$$\mathtt{B}^{(n+m-1)}_{2,1} = \mathtt{A}^{(n+m-1)}_{1,2}\mathtt{A}^{(n+m-1)}_{2,3} \ldots \mathtt{A}^{(n+m-1)}_{K-1,K}\mathtt{A}^{(n+m-1)}_{K,1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{2,2} = \mathtt{A}^{(n+m-1)}_{1,2K}\mathtt{A}^{(n+m-1)}_{2,K+1} \ldots \mathtt{A}^{(n+m-1)}_{K,K-1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{2,3} = \mathtt{A}^{(n+m-1)}_{1,2K+2}\mathtt{A}^{(n+m-1)}_{2,2K+3} \ldots \mathtt{A}^{(n+m-1)}_{K-1,3K}\mathtt{A}^{(n+m-1)}_{K,2K+1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{2,4} = \mathtt{A}^{(n+m-1)}_{1,4K}\mathtt{A}^{(n+m-1)}_{2,3K+1} \ldots \mathtt{A}^{(n+m-1)}_{K,4K-1}\mathtt{a}_{n+m-2},$$
$$\vdots$$
$$\mathtt{B}^{(n+m-1)}_{2,\lambda} = \mathtt{A}^{(n+m-1)}_{1,\lambda K}\mathtt{A}^{(n+m-1)}_{2,(\lambda-1)K+1} \ldots \mathtt{A}^{(n+m-1)}_{K,\lambda K-1}\mathtt{a}_{n+m-2},$$
$$\vdots$$

$(n+m-1)$-blocks of type K:

$$\mathtt{B}^{(n+m-1)}_{K,1} = \mathtt{A}^{(n+m-1)}_{1,K}\mathtt{A}^{(n+m-1)}_{2,1} \ldots \mathtt{A}^{(n+m-1)}_{K,K-1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{K,2} = \mathtt{A}^{(n+m-1)}_{1,K+2}\mathtt{A}^{(n+m-1)}_{2,K+3} \ldots \mathtt{A}^{(n+m-1)}_{K-1,2K}\mathtt{A}^{(n+m-1)}_{K,K+1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{K,3} = \mathtt{A}^{(n+m-1)}_{1,3K}\mathtt{A}^{(n+m-1)}_{2,2K+1} \ldots \mathtt{A}^{(n+m-1)}_{K,3K-1}\mathtt{a}_{n+m-2},$$
$$\mathtt{B}^{(n+m-1)}_{K,4} = \mathtt{A}^{(n+m-1)}_{1,3K+2}\mathtt{A}^{(n+m-1)}_{2,3K+3} \ldots \mathtt{A}^{(n+m-1)}_{K-1,4K}\mathtt{A}^{(n+m-1)}_{K,3K+1}\mathtt{a}_{n+m-2},$$
$$\vdots$$
$$\mathtt{B}^{(n+m-1)}_{K,\lambda} = \mathtt{A}^{(n+m-1)}_{1,(\lambda-1)K+2} \ldots \mathtt{A}^{(n+m-1)}_{K-1,\lambda K}\mathtt{A}^{(n+m-1)}_{K,(\lambda-1)K+1}\mathtt{a}_{n+m-2},$$

using pre-$(n+m-1)$-blocks $\mathtt{A}_{i,j}^{(n+m-1)}$ of length $\tilde{h}_{n+m-1}$ and a marker segment $\mathtt{a}_{n+m-2} = (\mathtt{B}_0^{(n+m-2)})^{\bar{N}(n+m-2)}$ with $\bar{N}(n+m-2)$ chosen according to equation (6.17) such that the $(n+m-1)$-blocks are uniform in $(n+m-2)$-blocks. We note that for j odd the block $\mathtt{B}_{i+1,j}^{(n+m-1)}$ is obtained from $\mathtt{B}_{i,j}^{(n+m-1)}$ by cycling the second index to the left. On the other hand, for j even the block $\mathtt{B}_{i+1,j}^{(n+m-1)}$ is obtained from $\mathtt{B}_{i,j}^{(n+m-1)}$ by cycling the second index to the right. Additionally, we have a marker block

$$\mathtt{B}_0^{(n+m-1)} = \mathtt{A}_{1,\lambda_{n+m-1}K+1}^{(n+m-1)} \mathtt{A}_{2,\lambda_{n+m-1}K+1}^{(n+m-1)} \cdots \mathtt{A}_{K,\lambda_{n+m-1}K+1}^{(n+m-1)} \mathtt{a}_{n+m-2},$$

where the pre-$(n+m-1)$-blocks $\mathtt{A}_{i,\lambda_{n+m-1}K+1}^{(n+m-1)}$ have not been used in any other $(n+m-1)$-block. Hence, there are $N(n+m-1)+1 = \lambda_{n+m-1}K+1$ $(n+m-1)$-blocks in total.

In our inductive construction process for $m \geq 3$, for any strings $\mathtt{A}, \overline{\mathtt{A}}$ of at least $\tilde{h}_{n+m-1}/e_{n+m-2}$ consecutive symbols in $\mathtt{A}_{i_1,j_1}^{(n+m-1)}$ and $\mathtt{A}_{i_2,j_2}^{(n+m-1)}$, we assume

$$\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) \geq \beta_{n+m-1} \quad \text{in case of } i_1 \neq i_2,\ \ j_1 = j_2, \tag{6.13}$$

and

$$\overline{f}(\mathtt{A}, \overline{\mathtt{A}}) \geq \alpha_{n+m-1} \quad \text{in case of } j_1 \neq j_2 \text{ for all } i_1, i_2, \tag{6.14}$$

with the numbers α_{n+m-1} and β_{n+m-1} from equations (6.6)–(6.9). In the corresponding circular system we have

$$\overline{f}(\mathcal{A}_{i_1,j}^{(n+m-1)}, \mathcal{A}_{i_2,j}^{(n+m-1)}) \leq \sum_{i=1}^{m-2} \left(\frac{4}{l_{n+i}} + \frac{2}{N(n+i-1)+1} \right). \tag{6.15}$$

Note that this assumption is vacuous if $m = 2$. In the odometer-based system we will use equation (6.12) for the first inductive step.

In the inductive step starting with $m \geq 2$, we use $(n+m-1)$-blocks to define grouped $(n+m-1)$-blocks $\mathtt{G}_{i,j}^{(n+m-1)}$ (where $i = 1, \ldots, K$ indicates the type of $(n+m-1)$-blocks used and $j = 0, \ldots, e_{n+m-1}-1$ enumerates the grouped $(n+m-1)$-blocks of that type) as follows:

$$\mathtt{G}_{i,j}^{(n+m-1)} = \mathtt{B}_{i,j\cdot 2d_{n+m-1}+1}^{(n+m-1)} \mathtt{B}_{i,j\cdot 2d_{n+m-1}+2}^{(n+m-1)} \cdots \mathtt{B}_{i,(j+1)\cdot 2d_{n+m-1}}^{(n+m-1)},$$

that is, it is a concatenation of $2d_{n+m-1}$ $(n+m-1)$-blocks of the same type. In the following lemmas we see that different grouped blocks are still apart from each other in the odometer-based system but grouped blocks with coinciding index j can be made arbitrarily close to each other in the circular system.

LEMMA 6.7. (Distance between grouped $(n+m-1)$-blocks in the odometer-based system) *Let $i_1, i_2 \in \{1, \ldots, K\}$, $j_1, j_2 \in \{0, \ldots, e_{n+m-1}-1\}$ and $\mathtt{G}, \overline{\mathtt{G}}$ be strings of at least $u_{n+m-1}\mathtt{h}_{n+m-1}$ consecutive symbols from grouped $(n+m-1)$-blocks $\mathtt{G}_{i_1,j_1}^{(n+m-1)}$ and $\mathtt{G}_{i_2,j_2}^{(n+m-1)}$ respectively.*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

(1) *For $i_1 \neq i_2$ and $j_1 = j_2$ we have*

$$\overline{f}(\mathtt{G}, \overline{\mathtt{G}}) \geq \begin{cases} \frac{1}{K}\alpha - \frac{14}{\sqrt{N}} - \frac{2}{KT} - \frac{2}{u_{n+1}} & \text{for } m = 2, \\ \beta_{n+m-1} + \frac{1}{K}(\alpha_{n+m-1} - \beta_{n+m-1}) - \frac{2}{e_{n+m-2}} - \frac{2}{u_{n+m-1}} - \frac{2}{N(n+m-2)+1} & \text{for } m \geq 3. \end{cases}$$

(2) *For $j_1 \neq j_2$ and all i_1, i_2 we have*

$$\overline{f}(\mathtt{G}, \overline{\mathtt{G}}) \geq \begin{cases} \alpha - \frac{14}{\sqrt{N}} - \frac{2}{T} - \frac{2}{u_{n+1}} & \text{for } m = 2, \\ \alpha_{n+m-1} - \frac{2}{e_{n+m-2}} - \frac{2}{u_{n+m-1}} - \frac{2}{N(n+m-2)+1} & \text{for } m \geq 3. \end{cases}$$

Proof. In the first case we have $j_1 = j_2$. We treat $\mathtt{G}$ and $\overline{\mathtt{G}}$ as strings of complete $(n+m-1)$-blocks by adding fewer than $2h_{n+m-1}$ symbols to complete partial blocks at the beginning and end of $\mathtt{G}$ and $\overline{\mathtt{G}}$. These constitute a fraction of at most $2/u_{n+m-1}$ of the total length. Additionally, we ignore the marker segments $\mathtt{a}_{n+m-2}$ which form a fraction $1/(N(n+m-2)+1)$ of the length of each $(n+m-1)$-block due to uniformity and so of $\mathtt{G}$ as well as $\overline{\mathtt{G}}$. On the remaining strings for $m = 2$ we apply Corollary 5.5 with $\tilde{f} \geq 1/K$ and equation (6.12) which yields

$$\overline{f}(\mathtt{G}, \overline{\mathtt{G}}) \geq \frac{1}{K}\left(\alpha - \frac{13}{\sqrt{N}} - \frac{2}{T}\right) - \frac{2}{N} - \frac{2}{N+1} - \frac{2}{u_{n+1}} \geq \frac{1}{K}\alpha - \frac{14}{\sqrt{N}} - \frac{2}{KT} - \frac{2}{u_{n+1}}.$$

On the remaining strings $\mathtt{G}_{\mathrm{mod}}$ and $\overline{\mathtt{G}}_{\mathrm{mod}}$ for $m \geq 3$ we use Proposition 5.4 with $\tilde{f} \geq 1/K$ and equations (6.13) and (6.14) to obtain

$$\overline{f}(\mathtt{G}_{\mathrm{mod}}, \overline{\mathtt{G}}_{\mathrm{mod}}) \geq \frac{1}{K}\alpha_{n+m-1} + \left(1 - \frac{1}{K}\right)\beta_{n+m-1} - \frac{2}{e_{n+m-2}},$$

which implies the claim.

In the second case we observe that $\mathtt{G}$ and $\overline{\mathtt{G}}$ do not have any Feldman pattern of pre-$(n+m-1)$-blocks in common due to $j_1 \neq j_2$. As before we complete partial blocks at the beginning and end of $\mathtt{G}$ and $\overline{\mathtt{G}}$ and remove the marker segments $\mathtt{a}_{n+m-2}$. On the remaining strings for $m = 2$ we apply Corollary 5.5 with $\tilde{f} = 1$ and equation (6.12), while for $m \geq 3$ we use Corollary 5.5 with $\tilde{f} = 1$ and equation (6.14). □

LEMMA 6.8. (Distance between grouped $(n+m-1)$-blocks in the circular system) *For all $i_1, i_2 \in \{1, \ldots, K\}$ and $j \in \{0, \ldots, e_{n+m-1} - 1\}$, we have*

$$\overline{f}(\mathcal{G}_{i_1,j}^{(n+m-1)}, \mathcal{G}_{i_2,j}^{(n+m-1)}) \leq \sum_{i=1}^{m-1} \left(\frac{4}{l_{n+i}} + \frac{2}{N(n+i-1)+1}\right).$$

Proof. We recall that the marker segment $\mathtt{a}_{n+m-2}$ occupies a fraction $(N(n+m-2)+1)^{-1}$ of the total length of each $(n+m-1)$-block. Moreover, for every $j = 1, \ldots, \lambda_{n+m-1}$ the $(n+m-1)$-block $\mathtt{B}_{i_2,j}^{(n+m-1)}$ is obtained from $\mathtt{B}_{i_1,j}^{(n+m-1)}$ by a cycling permutation of the Feldman patterns used for the pre-$(n+m-1)$-blocks. Under the cyclic operator $\mathcal{C}_{n+m-1}$ each $(n+m-1)$-block is repeated $l_{n+m-1}-1$ times. Hence, the $\overline{f}$ distance between $\mathcal{G}^{(n+m-1)}_{i_1,j}$ and $\mathcal{G}^{(n+m-1)}_{i_2,j}$ is at most

$$M+\frac{4}{l_{n+m-1}}+\frac{2}{N(n+m-2)+1},$$

where M is the $\overline{f}$ distance of pre-$(n+m-1)$-blocks of the same pattern in the circular system. For $m=2$ this distance $M=0$, while for $m\geq 3$ we obtain the claim with the aid of equation (6.15). □

For each type $i\in\{1,\ldots,K\}$ we use the e_{n+m-1} grouped $(n+m-1)$-blocks of type i as building blocks for the Feldman patterns $\mathtt{A}^{(n+m)}_{i,j}$, $j=1,\ldots,(\lambda_{n+m}+1)K$, which are the pre-$(n+m)$-blocks with length $\tilde{h}_{n+m}$. Thus there are $\tilde{N}(n+m):=(\lambda_{n+m}+1)K^2$ pre-$(n+m)$-blocks. For every $i\in\{1,\ldots,K\}$ each pattern $\mathtt{A}^{(n+m)}_{i,j}$, $j=1,\ldots,(\lambda_{n+m}+1)K$, contains each $(n+m-1)$-block of type i exactly

$$\bar{N}(n+m-1)=(e_{n+m-1})^{2(\lambda_{n+m}+1)K+2} \tag{6.16}$$

times by the construction in Proposition 6.2.

Lemma 6.9. (Separation and closeness of pre-$(n+m)$-blocks of the same Feldman pattern) *Let $j\in\{1,\ldots,(\lambda_{n+m}+1)K\}$ and $i_1,i_2\in\{1,\ldots,K\}$, $i_1\neq i_2$. For all strings $\mathtt{A}$ and $\overline{\mathtt{A}}$ of at least $\tilde{h}_{n+m}/e_{n+m-1}$ consecutive symbols in $\mathtt{A}^{(n+m)}_{i_1,j}$ and $\mathtt{A}^{(n+m)}_{i_2,j}$ respectively, we have for $m=2$,*

$$\overline{f}(\mathtt{A},\overline{\mathtt{A}})\geq\frac{1}{K}\alpha-\frac{14}{\sqrt{N}}-\frac{2}{TK}-\frac{4}{u_{n+1}}-\frac{2}{e_{n+1}};$$

while for $m\geq 3$, we have $\overline{f}(\mathtt{A},\overline{\mathtt{A}})$ greater than or equal to

$$\beta_{n+m-1}+\frac{1}{K}(\alpha_{n+m-1}-\beta_{n+m-1})-\frac{2}{e_{n+m-2}}-\frac{4}{u_{n+m-1}}-\frac{2}{N(n+m-2)+1}-\frac{2}{e_{n+m-1}}.$$

For the corresponding strings in the circular system we have

$$\overline{f}(\mathcal{A}^{(n+m)}_{i_1,j},\mathcal{A}^{(n+m)}_{i_2,j})\leq\sum_{i=1}^{m-1}\left(\frac{4}{l_{n+i}}+\frac{2}{N(n+i-1)+1}\right).$$

Proof. The statement in the odometer-based system follows from the first part of Lemma 6.7 and Lemma 6.3 (with $R=u_{n+m-1}$) because $\mathtt{A}^{(n+m)}_{i_1,j}$ and $\mathtt{A}^{(n+m)}_{i_2,j}$ are constructed as the same Feldman pattern with the grouped $(n+m-1)$-blocks of different type but the same pattern as building blocks. This also yields the statement in the circular system as a direct consequence of Lemma 6.8. □

We will also need a statement on the $\overline{f}$ distance between different Feldman patterns in the odometer-based system.

Lemma 6.10. (Separation of pre-$(n+m)$-blocks of different Feldman patterns) *Let $j_1,j_2\in\{1,\ldots,(\lambda_{n+m}+1)K\}$, $j_1\neq j_2$, and $i_1,i_2\in\{1,\ldots,K\}$. For all strings $\mathtt{A}$ and*

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

$\overline{\mathtt{A}}$ *of at least* $\tilde{h}_{n+m}/e_{n+m-1}$ *consecutive symbols in* $\mathtt{A}^{(n+m)}_{i_1,j_1}$ *and* $\mathtt{A}^{(n+m)}_{i_2,j_2}$ *respectively, we have*

$$\overline{f}(\mathtt{A},\overline{\mathtt{A}}) \geq \begin{cases} \alpha - \dfrac{14}{\sqrt{N}} - \dfrac{2}{T} - \dfrac{4}{u_{n+1}} - \dfrac{13}{\sqrt{e_{n+1}}} & \textit{for } m=2, \\ \alpha_{n+m-1} - \dfrac{2}{e_{n+m-2}} - \dfrac{4}{u_{n+m-1}} - \dfrac{2}{N(n+m-2)+1} - \dfrac{13}{\sqrt{e_{n+m-1}}} & \textit{for } m \geq 3. \end{cases}$$

Proof. Since we consider different Feldman pattern with the grouped $(n+m-1)$-blocks as building blocks, we obtain the result from the second part of Lemma 6.7 and Proposition 6.2 (with $R = u_{n+m-1}$). □

In the next step, we use these Feldman patterns $\mathtt{A}^{(n+m)}_{i,j}$ to define $(n+m)$-blocks for $i=1,\ldots,K$ and $j=1,\ldots,\lambda_{n+m}$ as in the formula at the beginning of §6.3.2 with $n+m-1$ replaced by $n+m$, and an additional marker block $\mathtt{B}^{(n+m)}_0 = \mathtt{A}^{(n+m)}_{1,\lambda_{n+m}K+1}\mathtt{A}^{(n+m)}_{2,\lambda_{n+m}K+1}\ldots\mathtt{A}^{(n+m)}_{K,\lambda_{n+m}K+1}\mathtt{a}_{n+m-1}$ with the marker segment $\mathtt{a}_{n+m-1} = (\mathtt{B}^{(n+m-1)}_0)^{\bar{N}(n+m-1)}$. We also note that every $(n+m)$-block contains exactly one pattern $\mathtt{A}^{(n+m)}_{i,j}$ of each type $i \in \{1,\ldots,K\}$ and thus it is uniform in the $(n+m-1)$-blocks. Thus, the inductive step has been accomplished.

6.3.3. *Final step: construction of* $(n+m_n)$*-blocks.* As foreshadowed in equation (6.10), we follow the inductive construction scheme until $\beta_{n+m_n} > \alpha - (2/T) - (\varepsilon/2)$ and we have constructed Feldman patterns $\mathtt{A}^{(n+m_n)}_{i,j}$, $j=1,\ldots,(\lambda_{n+m_n}+1)K$, $i=1,\ldots,K$, of length $\tilde{h}_{n+m_n}$. In particular, we have

$$\overline{f}(\mathtt{A},\overline{\mathtt{A}}) \geq \beta_{n+m_n} > \alpha - \frac{2}{T} - \frac{\varepsilon}{2} \tag{6.17}$$

for all strings $\mathtt{A}$, $\overline{\mathtt{A}}$ of at least $\tilde{h}_{n+m_n}/e_{n+m_n-1}$ consecutive symbols in $\mathtt{A}^{(n+m_n)}_{i_1,j_1}$ and $\mathtt{A}^{(n+m_n)}_{i_2,j_2}$ respectively, for $i_1 \neq i_2$ or $j_1 \neq j_2$. On the other hand, we have

$$\overline{f}(\mathcal{A}^{(n+m_n)}_{i_1,j}, \mathcal{A}^{(n+m_n)}_{i_2,j}) \leq \sum_{i=1}^{m_n-1}\left(\frac{4}{l_{n+i}} + \frac{2}{N(n+i-1)+1}\right) \tag{6.18}$$

in the circular system. Moreover, we recall that for every $i \in \{1,\ldots,K\}$ each pattern $\mathtt{A}^{(n+m_n)}_{i,j}$, $j=1,\ldots,(\lambda_{n+m_n}+1)K$, contains each $(n+m_n-1)$-block of type i exactly

$$\bar{N}(n+m_n-1) = (e_{n+m_n-1})^{2\cdot(\lambda_{n+m_n}+1)\cdot K+2}$$

times. In the final step, we define K $(n+m_n)$-blocks as follows (with $\lambda = \lambda_{n+m_n}$):

$$\begin{aligned} \mathtt{B}^{(n+m_n)}_1 &= \mathtt{A}^{(n+m_n)}_{1,1}\mathtt{A}^{(n+m_n)}_{2,2}\ldots\mathtt{A}^{(n+m_n)}_{K,K}\mathtt{A}^{(n+m_n)}_{1,K+1}\ldots\mathtt{A}^{(n+m_n)}_{K,2K}\ldots\mathtt{A}^{(n+m_n)}_{K,\lambda K}\mathtt{a}_{n+m_n-1}, \\ \mathtt{B}^{(n+m_n)}_2 &= \mathtt{A}^{(n+m_n)}_{2,1}\mathtt{A}^{(n+m_n)}_{3,2}\ldots\mathtt{A}^{(n+m_n)}_{1,K}\mathtt{A}^{(n+m_n)}_{2,K+1}\ldots\mathtt{A}^{(n+m_n)}_{1,2K}\ldots\mathtt{A}^{(n+m_n)}_{1,\lambda K}\mathtt{a}_{n+m_n-1}, \\ &\vdots \\ \mathtt{B}^{(n+m_n)}_K &= \mathtt{A}^{(n+m_n)}_{K,1}\mathtt{A}^{(n+m_n)}_{1,2}\ldots\mathtt{A}^{(n+m_n)}_{K-1,K}\mathtt{A}^{(n+m_n)}_{K,K+1}\ldots\mathtt{A}^{(n+m_n)}_{K-1,2K}\ldots\mathtt{A}^{(n+m_n)}_{K-1,\lambda K}\mathtt{a}_{n+m_n-1}, \end{aligned}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

with

$$\mathtt{a}_{n+m_n-1} = (\mathtt{B}_0^{(n+m_n-1)})^{\lambda_{n+m}\cdot \bar{N}(n+m_n-1)}.$$

We note that each $(n+m_n)$-block contains exactly λ_{n+m_n} patterns of each type. Hence, it is uniform in the $(n+m_n-1)$-blocks. We prove the statement in Proposition 6.5 on the $\overline{f}$ distance in the odometer-based system.

Proof of part (1) of Proposition 6.5. By adding fewer than $2\tilde{h}_{n+m_n}$ symbols to each $\mathtt{B}$ and $\overline{\mathtt{B}}$ we can complete any partial pre-$(n+m_n)$-blocks at the beginning and end of $\mathtt{B}$ and $\overline{\mathtt{B}}$. This change increases the $\overline{f}$ distance between $\mathtt{B}$ and $\overline{\mathtt{B}}$ by at most

$$\frac{4\tilde{h}_{n+m_n}}{2\mathrm{h}_{n+m_n}/K} < \frac{2\tilde{h}_{n+m_n}}{\lambda_{n+m_n}\tilde{h}_{n+m_n}} = \frac{2}{\lambda_{n+m_n}}.$$

In the next step, we ignore the marker segment $\mathtt{a}_{n+m_n-1}$ which occupies a fraction of $1/(N(n+m_n-1)+1)$ in each $(n+m_n)$-block due to uniformity and so a fraction of at most $K(N(n+m_n-1)+1)^{-1} < e_{n+m_n-1}^{-1}$ of the total length of $\mathtt{B}$ and $\overline{\mathtt{B}}$. On the remaining strings all pre-$(n+m_n)$-blocks are different from each other. Hence, they are at least $\alpha - (2/T) - (\varepsilon/2)$ apart in $\overline{f}$ on substantial substrings of at least $\tilde{h}_{n+m_n}/e_{n+m_n-1}$ consecutive symbols by equation (6.18). We apply Corollary 5.5 with $\tilde{f} = 1$ to obtain

$$\overline{f}(\mathtt{B}, \overline{\mathtt{B}}) \geq \alpha - \frac{2}{T} - \frac{\varepsilon}{2} - \frac{4}{e_{n+m_n-1}} - \frac{2}{\lambda_{n+m_n}},$$

which yields the claim. □

By choosing the circular coefficients $(l_{n+m})_{m\in\mathbb{N}}$ to grow sufficiently fast as in equation (6.5) we also obtain the second statement in Proposition 6.5.

Proof of part (2) of Proposition 6.5. Since the marker segments and the used Feldman patterns of the pre-$(n+m_n)$-blocks are aligned and only the used type of blocks differs, we use equation (6.19) to obtain

$$\overline{f}(\mathcal{B}_i^{(n+m_n)}, \mathcal{B}_j^{(n+m_n)}) \leq \sum_{i=1}^{m_n-1} \frac{4}{l_{n+i}} + \sum_{i=1}^{m_n-1} \frac{2}{Ku_{n+i-1}^2 e_{n+i-1}+1} < \delta$$

with the aid of assumptions (6.1) and (6.5) in the last step. □

Hence the proof of Proposition 6.5 has been accomplished.

6.4. *Proof of Theorem 6.1.* To define the construction sequence for the odometer-based system inductively we choose an increasing sequence $(K_s)_{s\in\mathbb{N}}$ of positive integers with

$$\sum_{k\in\mathbb{N}} \frac{15}{\sqrt{K_s}} < \frac{1}{32} \tag{6.19}$$

and two decreasing sequences $(\varepsilon_s)_{s\in\mathbb{N}}$ and $(\delta_s)_{s\in\mathbb{N}}$ of positive real numbers such that $\delta_s \searrow 0$, $\varepsilon_s < 1/(64K_s)$, and

$$\sum_{s\in\mathbb{N}} \varepsilon_s < \frac{1}{32}. \tag{6.20}$$

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press

We start by applying Proposition 6.4 on K_0+1 symbols to obtain $N(1)+1$ uniform and uniquely readable 1-blocks that are $\alpha_1 := \frac{1}{8} - 13K_0^{-1/2}$ apart in $\overline{f}$ on substantial substrings of length at least h_1/K_0. Here, the number $N(1)+1$ is chosen such that it allows the application of the cycling mechanism from Proposition 6.5 to obtain K_1+1 n_1-blocks (where $n_1 := 1+m_1$ with m_1 from Proposition 6.5) that are $\alpha_2 = \alpha_1 - 2K_0^{-1} - \varepsilon_0$ apart on substantial subshifts of length at least $\mathrm{h}_{n_1}/(K_1+1)$ in the odometer-based system and are δ_0-close in the corresponding circular system.

We continue by applying Proposition 6.4 on those blocks to obtain sufficiently many (n_1+1)-blocks (which are $\alpha_3 = \alpha_2 - 13K_1^{-1/2}$ apart on substantial substrings of length at least h_{n_1+1}/K_1 in the odometer-based system) such that we can apply Proposition 6.5 again to get K_2+1 n_2-blocks (with $n_2 := n_1 + 1 + m_{n_1+1}$) that are $\alpha_4 = \alpha_3 - 2K_1^{-1} - \varepsilon_1$ apart on substantial subshifts of length at least $\mathrm{h}_{n_2}/(K_2+1)$ in the odometer-based system and δ_1-close in the circular system.

Continuing like this we produce n-blocks that are at least

$$\frac{1}{8} - \sum_{s\in\mathbb{N}} \left(\frac{15}{\sqrt{K_s}} + \varepsilon_s \right) > \frac{1}{16}$$

apart from each other by the requirements (6.20) and (6.21). Hence, the odometer-based system cannot be loosely Bernoulli by Lemma 3.5, provided that the metric entropy H of the odometer-based system is zero. As in Lemma 4.2, we have the formula

$$H = \lim_{n\to\infty} \frac{N(n)}{\mathrm{h}_n}.$$

In our construction of $(n+1)$-blocks, we start by producing $\tilde{N}(n+1)$ pre-$(n+1)$-blocks of length $\tilde{\mathrm{h}}_{n+1} \geq 2^{2\cdot\tilde{N}(n+1)+3}\mathrm{h}_n$. There are fewer $(n+1)$-blocks than pre-$(n+1)$-blocks while the $\mathrm{h}_{n+1} > \tilde{\mathrm{h}}_{n+1.}$ Therefore

$$\frac{\log N(n+1)}{\mathrm{h}_{n+1}} \leq \frac{\log \tilde{N}(n+1)}{2^{2\cdot\tilde{N}(n+1)+3}},$$

which converges to zero as n goes to infinity, because $\lim_{n\to\infty}\tilde{N}(n+1) = \infty$. Therefore $H = 0$.

On the other hand, the $\overline{f}$ distance between n-blocks in the circular system goes to zero because $\delta_s \searrow 0$. Thus, the odometer-based system is loosely Bernoulli by Lemma 3.5. Since the blocks constructed by Propositions 6.4 and 6.5 satisfy the properties of uniformity and unique readability, our construction sequence satisfies those as well.

Acknowledgements. We thank M. Foreman and J.-P. Thouvenot for helpful conversations. In particular, we thank Foreman for his patient explanations of his work with B. Weiss during his visit to Indiana University. We are grateful for Thouvenot's suggestion to use Rothstein's argument in our positive-entropy example. We thank the referee for helpful comments. P.K. acknowledges financial support from a DFG Forschungsstipendium under grant no. 405305501 and partial support from the visiting research professorship at IU.

Note added in proof: The open questions regarding anti-classification results for Kakutani equivalence mentioned in the introduction were answered recently [**GK**].

REFERENCES

[AK70] D. Anosov and A. Katok. New examples in smooth ergodic theory. Ergodic diffeomorphisms. *Trudy Moskov. Mat. Obsc.* **23** (1970), 3–36.

[BF96] F. Beleznay and M. Foreman. The complexity of the collection of measure-distal transformations. *Ergod. Th. & Dynam. Sys.* **16**(5) (1996), 929–962.

[Fe76] J. Feldman. New K-automorphisms and a problem of Kakutani. *Israel J. Math.* **24**(1) (1976), 16–38.

[FRW11] M. Foreman, D. Rudolph and B. Weiss. The conjugacy problem in ergodic theory. *Ann. of Math. (2)* **173**(3) (2011), 1529–1586.

[FW1] M. Foreman and B. Weiss. A symbolic representation of Anosov–Katok systems. *J. Anal. Math.* **137**(2) (2019), 603–661.

[FW2] M. Foreman and B. Weiss. From odometers to circular systems: a global structure theorem. *J. Mod. Dyn.* **15** (2019), 345–423.

[FW3] M. Foreman and B. Weiss. Measure preserving diffeomorphisms of the torus are unclassifiable. *J. Eur. Math. Soc*. to appear. *Preprint*, 2020, arXiv:1705.04414.

[GK] M. Gerber and P. Kunde. Anti-classification results for the Kakutani equivalence relation. Preprint, 2021, arXiv:2109.06086.

[Hj01] G. Hjorth. On invariants for measure preserving transformations. *Fund. Math.* **169**(1) (2001), 51–84.

[HN42] P. Halmos and J. von Neumann. Operator methods in classical mechanics. II. *Ann. of Math. (2)* **43** (1942), 332–350.

[Ka75] A. Katok. Time change, monotone equivalence, and standard dynamical systems. *Dokl. Akad. Nauk SSSR* **223** (1975), 784–792.

[Ka77] A. Katok. Monotone equivalence in ergodic theory. *Math. USSR Izv.* **11**(1) (1977), 99–146.

[KH95] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems (Encyclopedia of Mathematics and Its Applications, 54)*. Cambridge University Press, Cambridge, 1995.

[KR] A. Kanigowski and T. De La Rue. Product of two staircase rank one transformations that is not loosely Bernoulli. *J. Anal. Math.* **143** (2021), 535–553.

[Kr70] W. Krieger. On entropy and generators of measure-preserving transformations. *Trans. Amer. Math. Soc.* **199** (1970), 453–464; Erratum. *Trans. Amer. Math. Soc.* **168** (1972), 549.

[Ku65] A. Kushnirenko. An upper bound for the entropy of a classical dynamical system (in Russian). *Dokl. Akad. Nauk SSSR* **161** (1965), 37–38.

[KW19] A. Kanigowski and D. Wei. Product of two Kochergin flows with different exponents is not standard. *Studia Math.* **244** (2019), 265–283.

[Ne32] J. von Neumann. Zur Operatorenmethode in der klassischen Mechanik. *Ann. of Math. (2)* **33**(3) (1932), 587–642.

[Or] D. Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems*. Yale University Press, New Haven, CT, 1970.

[Or70] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.* **4** (1970), 337–352.

[ORW82] D. Ornstein, D. Rudolph and B. Weiss. Equivalence of measure preserving transformations. *Mem. Amer. Math. Soc.* **37**(262) (1982), 120pp.

[Ra78] M. Ratner. Horocycle flows are loosely Bernoulli. *Israel J. Math.* **31**(2) (1978), 122–132.

[Ra79] M. Ratner. The Cartesian square of the horocycle flow is not loosely Bernoulli. *Israel J. Math.* **34**(1) (1979), 72–96.

[Ro80] A. Rothstein. Vershik processes: first steps. *Israel J. Math.* **36**(3–4) (1980), 205–224.

[Si62] Ya. Sinai. A weak isomorphism of transformations with invariant measure. *Dokl. Akad. Nauk SSSR* **147** (1962), 797–800.

[Sm71] M. Smorodinsky. *Ergodic Theory, Entropy (Lecture Notes in Mathematics, 214)*. Springer, Berlin, 1971.

https://doi.org/10.1017/etds.2021.73 Published online by Cambridge University Press