Ergodic Theorems and Converses for PSPACE Functions

Satyadev Nandakumar
Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, India

Subin Pulari
Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, India

Abstract

We initiate the study of effective pointwise ergodic theorems in resource-bounded settings. Classically, the convergence of the ergodic averages for integrable functions can be arbitrarily slow [14]. In contrast, we show that for a class of PSPACE L^1 functions, and a class of PSPACE computable measure-preserving ergodic transformations, the ergodic average exists and is equal to the space average on every EXP random. We establish a partial converse that PSPACE non-randomness can be characterized as non-convergence of ergodic averages. Further, we prove that there is a class of resource-bounded randoms, *viz.* SUBEXP-space randoms, on which the corresponding ergodic theorem has an exact converse - a point x is SUBEXP-space random if and only if the corresponding effective ergodic theorem holds for x.

2012 ACM Subject Classification Theory of computation → Constructive mathematics; Theory of computation → Complexity theory and logic; Mathematics of computing → Probability and statistics

Keywords and phrases Ergodic Theorem, Resource-bounded randomness, Computable analysis, Complexity theory

Digital Object Identifier 10.4230/LIPIcs.MFCS.2021.80

Related Version *Full Version*: `https://arxiv.org/abs/2012.11266`

Acknowledgements The authors wish to thank anonymous reviewers for helpful suggestions.

1 Introduction

In Kolmogorov's program to found information theory on the theory of algorithms, we investigate whether individual "random" objects obey probabilistic laws, *i.e.*, properties which hold in sample spaces with probability 1. Indeed, a vast and growing literature establishes that *every* Martin-Löf random sequence (see for example, [4] or [19]) obeys the Strong Law of Large Numbers [24], the Law of Iterated Logarithm [25], and surprisingly, the Birkhoff Ergodic Theorem [26, 17, 10, 1] and the Shannon-McMillan-Breiman theorem [8, 9, 21]. In effective settings, the theorem for Martin-Löf random points implies the classical theorem since the set of Martin-Löf randoms has Lebesgue measure 1, and hence is stronger.

In this work, we initiate the study of ergodic theorems in resource-bounded settings. This is a difficult problem, since classically, the convergence speed in ergodic theorems is known to be arbitrarily slow (e.g. see Bishop [3], Krengel [14], and V'yugin [26]). However, we establish ergodic theorems in resource-bounded settings which hold on every resource-bounded random object of a particular class. The main technical hurdle we overcome is the lack of sharp tail bounds. The only general tail bound in ergodic settings is the maximal ergodic inequality. This yields only an inverse linear bound in the error bound, in contrast to the inverse exponential bounds in the Chernoff and the Azuma-Hoeffding inequalities.

46th International Symposium on Mathematical Foundations of Computer Science (MFCS 2021).
Editors: Filippo Bonchi and Simon J. Puglisi; Article No. 80; pp. 80:1–80:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We first establish an unconditional result. – For the entire class of PSPACE L^1 functions on Bernoulli systems, the ergodic average exists and is equal to the space average on all EXP randoms. We utilize a non-trivial connection with the theory of uniform distribution of sequences modulo 1 [15, 16, 20, 18] to prove this result.

In the general case, rapid L^1 convergence of subsequences of ergodic averages suffices to establish the same consequence that the ergodic average exists and is equal to the space average on all EXP randoms. In general, such assumptions are unavoidable since an adaptation of V'yugin's counterexample [26] shows that there are PSPACE computable ergodic Markov systems where the convergence rate to the ergodic average is not even computable.

Conversely, we ask whether we can characterize non-randomness using the failure of the PSPACE ergodic theorem. Franklin and Towsner [5] show that for every non-Martin-Löf random x, there is an effective ergodic system where the ergodic average at x does not converge to the space average. We first show that our PSPACE effective ergodic theorem admits a partial converse of this form. PSPACE non-randoms can be characterized as points where the PSPACE ergodic theorem fails.

We know that the set of EXP randoms is a subset of the set of PSPACE randoms. Since the forward direction holds on the smaller set of randoms, it is important to know whether there is a class of resource-bounded randoms on which an effective ergodic theorem holds with an exact converse. We show that the class of SUBEXP-space randoms is one such. We summarize our results in Table 1.

The proofs of these results are adapted from the techniques of Rute [21], Ko [13], Galatolo, Hoyrup & Rojas [7, 11], and Huang & Stull [12].[1] Our proofs involve several new quantitative estimates, which may of general interest.

Table 1 Summary of the results involving PSPACE/SUBEXP-space systems.

Class of functions	Convergence of ergodic averages (Theorems)	
	$\forall f(A_n^f \to \int f d\mu)$	$\exists f(A_n^f \not\to \int f d\mu)$
PSPACE L^1	EXP randoms (6.2)	PSPACE nonrandoms (7.1)
SUBEXP-space L^1	SUBEXP-space randoms (8.11)	SUBEXP-space nonrandoms (8.12)

2 Preliminaries

Let $\Sigma = \{0, 1\}$ be the binary alphabet. Denote the set of all finite binary strings by Σ^* and the set of infinite binary strings by Σ^∞. For $\sigma \in \Sigma^*$ and $y \in \Sigma^* \cup \Sigma^\infty$, we write $\sigma \sqsubseteq y$ if σ is a prefix of y. For any infinite string y and any finite string σ, $\sigma[n]$ and $y[n]$ denotes the character at the n^{th} position in y and σ respectively. For any infinite string y and any finite string σ, $\sigma[n, m]$ and $y[n, m]$ represents the strings $\sigma[n]\sigma[n+1]\ldots\sigma[m]$ and $y[n]y[n+1]\ldots y[m]$ respectively. We denote finite strings using small Greek letters like σ, α etc. The length of a finite binary string σ is denoted by $|\sigma|$.

[1] There are alternative approaches to the proof in Martin-Löf settings, like that of V'yugin [26]. However, the tool he uses for establishing the result is a lower semicomputable test defined on infinite sequences - this is difficult to adapt to resource-bounded settings requiring the output value within bounded time or space. Moreover, the functions in V'yugin's approach are continuous. We consider the larger class of L^1 functions, which can be discontinuous in general.

For $\sigma \in \Sigma^*$, the *cylinder* $[\sigma]$ is the set of all infinite sequences with σ as a prefix. χ_σ denotes the characteristic function of $[\sigma]$. For any set of strings $S \subseteq \Sigma^*$, $[S]$ is the union of $[\sigma]$ over all $\sigma \in S$. Extending the notation, χ_S denotes the characteristic function of $[S]$. The Borel σ-algebra generated by the set of all cylinders is denoted by $\mathcal{B}(\Sigma^\infty)$.

Unless specified otherwise, any $n \in \mathbb{N}$ is represented in the binary alphabet. As is typical in resource-bounded settings, some integer parameters are represented in unary. The set of unary strings is represented as 1^*, and the representation of $n \in \mathbb{N}$ in unary is 1^n, a string consisting of n ones. For any $n_1, n_2 \in \mathbb{N}$, $[n_1, n_2]$ represents the set $\{n \in \mathbb{N} : n_1 \leq n \leq n_2\}$.

Throughout the paper we take into account the number of cells used in the output tape and the working tape when calculating the space complexity of functions. We assume a finite representation for the set of rational numbers $\mathbb{Q}$ satisfying the following: there exists a $c \in \mathbb{N}$ such that if $r \in \mathbb{Q}$ has a representation of length l then $r \leq 2^{l^c}$. Following the works of Hoyrup, and Rojas [11], we introduce the notion of a PSPACE-probability Cantor space by endowing the Cantor space with a PSPACE-computable probability measure.

▶ **Definition 2.1.** *Consider the Cantor space* $(\Sigma^\infty, \mathcal{B}(\Sigma^\infty))$. *A Borel probability measure* $\mu : \mathcal{B}(\Sigma^\infty) \to [0,1]$, *is a* PSPACE-probability measure *if there is a* PSPACE *machine* $M : \Sigma^* \times 1^* \to \mathbb{Q}$ *such that for every* $\sigma \in \Sigma^*$, *and* $n \in \mathbb{N}$, *we have that* $|M(\sigma, 1^n) - \mu([\sigma])| \leq 2^{-n}$.

In order to define PSPACE (EXP) randomness using PSPACE (EXP) tests we require the following method for approximating sequences of open sets in Σ^∞ in polynomial space (exponential time).

▶ **Definition 2.2** (PSPACE/EXP sequence of open sets [12])**.** *A sequence of open sets* $\langle U_n \rangle_{n=1}^\infty$ *is a* PSPACE sequence of open sets *if there exists a sequence of sets* $\langle S_n^k \rangle_{k,n\in\mathbb{N}}$, *where* $S_n^k \subseteq \Sigma^*$ *such that*

1. $U_n = \cup_{k=1}^\infty [S_n^k]$, *where for any* $m > 0$, $\mu\left(U_n - \cup_{k=1}^m [S_n^k]\right) \leq 2^{-m}$.
2. *There exists a* controlling polynomial p *such that* $max\{|\sigma| : \sigma \in \cup_{k=1}^m S_n^k)\} \leq p(n+m)$.
3. *The function* $g : \Sigma^* \times 1^* \times 1^* \to \{0,1\}$ *such that* $g(\sigma, 1^n, 1^m) = 1$ *if* $\sigma \in S_n^m$, *and 0 otherwise, is decidable by a* PSPACE *machine.*

The definition of EXP sequence of open sets *is similar but the bound in condition 2 is replaced with* $2^{p(n+m)}$ *and the machine in condition 3 is an* EXP*-time machine.*

Henceforth, we study the notion of resource bounded randomness on (Σ^∞, μ).

▶ **Definition 2.3** (PSPACE/EXP randomness [23])**.** *A sequence of open sets* $\langle U_n \rangle_{n=1}^\infty$ *is a* PSPACE test *if it is a* PSPACE *sequence of open sets and for all* $n \in \mathbb{N}$, $\mu(U_n) \leq 2^{-n}$.

A set $A \subseteq \Sigma^\infty$ *is* PSPACE null *or* PSPACE non-random *if there is a* PSPACE *test* $\langle U_n \rangle_{n=1}^\infty$ *such that* $A \subseteq \cap_{n=1}^\infty U_n$, *and is* PSPACE random *otherwise. The* EXP *analogues of the above concepts are defined similarly except that* $\langle U_n \rangle_{n=1}^\infty$ *is an* EXP *sequence of open sets.*

By considering the sequence $\left\langle \cup_{i=1}^k S_n^i \right\rangle_{k,n\in\mathbb{N}}$ instead of $\left\langle S_n^k \right\rangle_{k,n\in\mathbb{N}}$, without loss of generality, we can assume that for each n, $\left\langle S_n^k \right\rangle_{k=1}^\infty$ is an increasing sequence of sets. Since every PSPACE test is an EXP test, every EXP random is PSPACE random.

In order to establish our ergodic theorem, it is convenient to define a PSPACE version of Solovay tests, where the relaxation is that the measures of the sets U_n can be any sufficiently fast convergent sequence. We later show that this captures the same set of randoms as PSPACE tests.

▶ **Definition 2.4** (PSPACE Solovay test). *A sequence of open sets* $\langle U_n \rangle_{n=1}^{\infty}$ *is a* PSPACE Solovay test *if it is a* PSPACE *sequence of open sets and there is a polynomial* p *such that* $\sum_{n=p(m)+1}^{\infty} \mu(U_n) \leq 2^{-m}$ *for all* $m \in \mathbb{N} \setminus \{0\}$ [2]. *A set* $A \subseteq \Sigma^\infty$ *is* PSPACE Solovay null *or* PSPACE Solovay non-random *if there exists a* PSPACE *Solovay test* $\langle U_n \rangle_{n=1}^{\infty}$ *such that* $A \subseteq \cap_{i=1}^{\infty} \cup_{n=i}^{\infty} U_n$*, and is* PSPACE Solovay random *otherwise.*

▶ **Theorem 2.5.** *A set* $A \subseteq \Sigma^\infty$ *is* PSPACE *null if and only if* A *is* PSPACE *Solovay null.*

The set of PSPACE Solovay randoms and PSPACE randoms are equal, hence to prove PSPACE randomness results, it suffices to form Solovay tests.

3 PSPACE L^1 computability

The resource-bounded ergodic theorems in our work hold for PSPACE-L^1 functions, the PSPACE analogue of integrable functions. In this section, we briefly recall standard definitions for PSPACE computable L^1 functions and measure-preserving transformations. The justifications and proofs of equivalences of various notions are present in Stull's thesis [22] and [23]. We initially define PSPACE sequence of simple functions, and define PSPACE integrable functions based on approximations using these functions.

▶ **Definition 3.1** (PSPACE sequence of simple functions [23]). *A sequence of simple functions* $\langle f_n \rangle_{n=1}^{\infty}$*, where each* $f_n : \Sigma^\infty \to \mathbb{Q}$*, is a* PSPACE sequence of simple functions *if*

1. *There is a* controlling polynomial p *such that for each* n*, there exists* $k(n) \in \mathbb{N}$*,* $\{d_1, d_2, \ldots, d_{k(n)}\} \subseteq \mathbb{Q}$ *and* $\{\sigma_1, \sigma_2, \ldots, \sigma_{k(n)}\} \subseteq \Sigma^{p(n)}$ *satisfying* $f_n = \sum_{i=1}^{k(n)} d_i \chi_{\sigma_i}$.
2. *There is a* PSPACE *machine* M *such that for each* $n \in \mathbb{N}$*, and* $\sigma \in \Sigma^*$*,* $M(1^n, \sigma)$ *outputs* $f_n(\sigma 0^\infty)$ *if* $|\sigma| \geq p(n)$ *and ? otherwise.*

Note that since M is a PSPACE machine, $\{d_1, d_2 \ldots d_{k(n)}\}$ is a set of PSPACE representable numbers. Now, we define PSPACE L^1-computable functions in terms of limits of convergent PSPACE sequence of simple functions.

▶ **Definition 3.2** (PSPACE L^1-computable functions [23]). *A function* $f \in L^1(\Sigma^\infty, \mu)$ *is* PSPACE L^1*-computable if there exists a* PSPACE *sequence of simple functions* $\langle f_n \rangle_{n=1}^{\infty}$ *such that for every* $n \in \mathbb{N}$*,* $\|f - f_n\| \leq 2^{-n}$*. The sequence* $\langle f_n \rangle_{n=1}^{\infty}$ *is called a* PSPACE L^1-approximation *of* f*.*

A sequence of L^1 functions $\langle f_n \rangle_{n=1}^{\infty}$ converging to f in the L^1-norm need not have pointwise limits. Hence the following concept ([21]) is important in studying the pointwise ergodic theorem in the setting of L^1-computability

▶ **Definition 3.3** ($\widetilde{f}$ for PSPACE L^1-computable f). *Let* $f \in L^1(\Sigma^\infty, \mu)$ *be* PSPACE L^1*-computable and with a* PSPACE L^1 *approximation* $\langle f_n \rangle_{n=1}^{\infty}$*. Define* $\widetilde{f} : \Sigma^\infty \to \mathbb{R} \cup \{undefined\}$ *by* $\widetilde{f}(x) = \lim_{n \to \infty} f_n(x)$ *if this limit exists, and is undefined otherwise.*[3]

To define ergodic averages, we restrict ourselves to the following class of transformations.

---

[2] This implies that $\sum_{n=1}^{\infty} \mu(U_n) < \infty$.

[3] The definition of $\widetilde{f}$ is dependent on the choice of the approximating sequence $\langle f_n \rangle_{n=1}^{\infty}$. However, due to Lemma 4.3, we use $\widetilde{f}$ in a sequence independent manner.

▶ **Definition 3.4** (PSPACE simple transformation). *A measurable function $T : (\Sigma^\infty, \mu) \to (\Sigma^\infty, \mu)$ is a* PSPACE *simple transformation if there is a controlling constant c and a* PSPACE *machine M such that such that for any $\sigma \in \Sigma^*$, $T^{-1}([\sigma]) = \cup_{i=1}^{k(\sigma)}[\sigma_i]$ where the following properties hold.*

1. *$\{\sigma_i\}_{i=1}^{k(\sigma)}$ is a prefix free set and for all $1 \leq i \leq k(\sigma)$, $|\sigma_i| \leq |\sigma| + c$*
2. *For each $\sigma, \alpha \in \Sigma^*$,*

$$M(\sigma, \alpha) = \begin{cases} 1 & \text{if } |\alpha| \geq |\sigma| + c \text{ and } \alpha 0^\infty \in T^{-1}([\sigma]) \\ 0 & \text{if } |\alpha| \geq |\sigma| + c \text{ and } \alpha 0^\infty \notin T^{-1}([\sigma]) \\ ? & \text{otherwise} \end{cases}$$

PSPACE computability as defined above, relates naturally to convergence of L^1 norms. But the pointwise ergodic theorem deals with almost everywhere convergence, and its resource-bounded versions deal with convergence on every random point. We introduce the modes of convergence we deal with in the present work.

▶ **Definition 3.5** (PSPACE-rapid limit point). *A real number a is a* PSPACE-rapid limit point *of the real number sequence $\langle a_n \rangle_{n=1}^\infty$ if there exists a polynomial p such that for all $m \in \mathbb{N}$, $\exists k \leq 2^{p(m)}$ such that $|a_k - a| \leq 2^{-m}$.*

Note that this requires rapid convergence only on a subsequence, which may not be a computable subsequence of the full sequence. The following definition is the L^1 version of the above.

▶ **Definition 3.6** (PSPACE-rapid L^1-limit point). *A function $f \in L^1(\Sigma^\infty, \mu)$ is a* PSPACE-rapid L^1-limit point *of a sequence $\langle f_n \rangle_{n=1}^\infty$ of functions in $L^1(\Sigma^\infty, \mu)$ if 0 is a* PSPACE*-rapid limit point of $\|f_n - f\|_1$.*

Now we define PSPACE analogue of almost everywhere convergence ([21]).

▶ **Definition 3.7** (PSPACE-rapid almost everywhere convergence). *A sequence of measurable functions $\langle f_n \rangle_{n=1}^\infty$ is* PSPACE-rapid almost everywhere convergent *to a measurable function f if there exists a polynomial p such that for all m_1 and m_2,*

$$\mu\left(\left\{x : \sup_{n \geq 2^{p(m_1+m_2)}} |f_n(x) - f(x)| \geq 2^{-m_1}\right\}\right) \leq 2^{-m_2}.$$

Notation. Let $A_n^{f,T} = \frac{f + f\circ T + f \circ T^2 + \ldots f \circ T^{n-1}}{n}$ denote the n^{th} Birkhoff average for any function f and transformation T. We prove the ergodic theorem in measure preserving systems where $\int f d\mu$ is a PSPACE-rapid L^1-limit point of $A_n^{f,T}$. In the rest of the paper we denote $A_n^{f,T}$ simply by A_n^f. The transformation T involved in the Birkhoff sum is implicit.

PSPACE rapidity of A_n^f is a stronger version of $\ln^2$-ergodicity introduced in [6].

▶ **Lemma 3.8.** *Let $T : \Sigma^\infty \to \Sigma^\infty$ be any measurable transformation and $f \in L^\infty(\Sigma^\infty, \mu)$. $\int f d\mu$ is a* PSPACE*-rapid L^1-limit point of A_n^f if and only if there exists $c > 0$ and $k \in \mathbb{N}$ such that for all $n > 0$,*

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} \int f \circ T^i . f - \left(\int f\right)^2 d\mu \right| \leq \frac{c}{2^{(\ln n)^{\frac{1}{k}}}}.$$

4 PSPACE-rapid almost everywhere convergence of ergodic averages

We present PSPACE versions of Theorem 2 and Proposition 5 from [7], relating the L^1 convergence of A_n^f to $\int f$ to its almost everywhere convergence. The main estimate which we require in this section is the maximal ergodic inequality, which we now recall.

▶ **Lemma 4.1** (Maximal ergodic inequality [2]). *If $f \in L^1(\Sigma^\infty, \mu)$ and $\delta > 0$ then* $\mu\left(\left\{x : \sup_{n\geq 1} |A_n^f(x)| > \delta\right\}\right) \leq (\|f\|_1)\delta^{-1}$.

Using this lemma, we now prove the almost everywhere convergence of ergodic averages.

▶ **Theorem 4.2.** *Let f be any function in $L^1(\Sigma^\infty, \mu)$ and let T be a measure preserving transformation. If $\int f d\mu$ is a* PSPACE*-rapid L^1-limit point of A_n^f then A_n^f is* PSPACE*-rapid almost everywhere convergent to $\int f d\mu$.*

If $f \in L^\infty$, the converse of Theorem 4.2 can be easily obtained by expanding $\|A_n^f - \int f d\mu\|_1$.

Now, we prove some auxiliary results that are useful in the proof of the PSPACE ergodic theorem. The following fact was shown in [12]. However, for our ergodic theorem we require an alternate proof of this fact using techniques from [21].

▶ **Lemma 4.3.** *Let $\langle f_n\rangle_{n=1}^\infty$, $\langle g_n\rangle_{n=1}^\infty$ be* PSPACE *sequence of simple functions which converges* PSPACE*-rapid almost everywhere to $f \in L^1(\Sigma^\infty, \mu)$. Then, for all* EXP *random x, $\lim_{n\to\infty} f_n(x)$ and $\lim_{n\to\infty} g_n(x)$ exist, and are equal.*

The following immediately follows from the above lemma.

▶ **Corollary 4.4.** *Let $f \in L^1(\Sigma^\infty, \mu)$ be a* PSPACE *L^1-computable function with L^1 approximating* PSPACE *sequences of simple functions $\langle f_n\rangle_{n=1}^\infty$ and $\langle g_n\rangle_{n=1}^\infty$. Then, for all* EXP *random x $\lim_{n\to\infty} f_n(x)$ and $\lim_{n\to\infty} g_n(x)$ exist, and are equal.*

The following properties satisfied by PSPACE simple transformations and PSPACE L^1-computable functions are useful in our proof of the PSPACE ergodic theorem.

▶ **Lemma 4.5.** *Let f be a* PSPACE *L^1-computable function over the Bernoulli space. Let $I_f : \Sigma^\infty \to \Sigma^\infty$ be the constant function taking the value $\int f d\mu$ over all $x \in \Sigma^\infty$. Then, I_f is* PSPACE *L^1-computable and $\widetilde{I_f}(x) = \int f d\mu$ for all* EXP *random x.*

▶ **Lemma 4.6.** *Let f be a* PSPACE *L^1-computable function with an L^1 approximating* PSPACE *sequence of simple functions $\langle f_n\rangle_{n=1}^\infty$. Let T be a* PSPACE *simple transformation and p be a polynomial. Then, $\left\langle A_n^{f_{p(n)}}\right\rangle_{n=1}^\infty$ is a* PSPACE *sequence of simple functions.*

5 Unconditional PSPACE ergodic theorem for the Bernoulli space

We now prove an unconditional version of our main result, namely, that for PSPACE L^1 computable functions, the ergodic average exists, and is equal to the space average, on every EXP random in the canonical setting of the Bernoulli space. We utilize the almost everywhere convergence results proved in the previous section, to prove the convergence on every PSPACE/EXP random. We first show that in the Bernoulli space, every PSPACE L^1 function exhibits PSPACE rapidity of A_n^f. The proof of this theorem is a non-trivial application of techniques from uniform distribution of sequences modulo 1 [15, 20, 16, 18].

▶ **Theorem 5.1.** *Let $f \in L^1(\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ where μ is the Bernoulli measure $\mu(\sigma) = \frac{1}{2^{|\sigma|}}$ and let T be the left shift transformation. If f is* PSPACE *L^1-computable, then there exists a polynomial q satisfying the following: given any $m \in \mathbb{N}$, for all $n \geq 2^{q(m)}$, $\|A_n^f - \int f d\mu\|_1 \leq 2^{-m}$.*

An equivalent statement is the following: The left-shift transformation on the Bernoulli probability measure is PSPACE ergodic[4]. Theorem 5.1 gives an explicit bound on the speed of convergence in the L^1 ergodic theorem for an interesting class of functions over the Bernoulli space. Such bounds do not exist in general for the L^1 ergodic theorem as demonstrated by Krengel in [14].

The above theorem can be obtained from the following assertion regarding PSPACE-rapid convergence of characteristic functions of long enough cylinders.

▶ **Lemma 5.2.** *Let T be the left shift transformation $T : (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu) \to (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ where μ is the Bernoulli measure $\mu(\sigma) = 2^{-|\sigma|}$. There exist polynomials q_1, q_2 such that for any $m \in \mathbb{N}$ and $\sigma \in \Sigma^*$ with $|\sigma| \geq q_1(m)$ we get $\|A_n^{\chi_\sigma} - \mu(\sigma)\|_1 \leq 2^{-m}$ for all $n \geq |\sigma|^3 2^{q_2(m)}$.*

Proof sketch. The major difficulty in directly approximating $\|A_n^{\chi_\sigma} - \mu(\sigma)\|_1$ is that for any $n, m \in \mathbb{N}$, $A_n^{\chi_\sigma}$ and $A_m^{\chi_\sigma}$ may not be *independent*. In order to overcome this, we use constructions similar to those used in proving Pillai's theorem (see [20], [16] for normal numbers, [18] for continued fractions) in order to approximate each $A_n^{\chi_\sigma}$ with sums of *disjoint* averages as follows.

$$A_n^{\chi_\sigma}(x) = \frac{\sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}(x)}{n} + \sum_{p=2}^{\lfloor \log_2(\frac{n}{k}) \rfloor} \sum_{j=1}^{k-1} \frac{\sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j}}{n} + \frac{(k-1).O(\log n)}{n}, \qquad \text{where}$$

$$X_i^{1,1}(x) = \begin{cases} 1 & \text{if } x[ik+1,(i+1)k] = \sigma \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and}$$

$$X_i^{p,j}(x) = \begin{cases} 1 & \text{if } x[2^{p-2}k - j + 1, 2^{p-2}k - j + k] = \sigma \\ 0 & \text{otherwise} \end{cases}$$

The first two terms on the right of the equation turns out to be averages of independent Bernoulli random variables. Hence, elementary results from probability theory regarding independent Bernoulli random variables can be used to show that $A_n^{\chi_\sigma}$ converges to $\int f d\mu$ sufficiently fast. ◀

We remark that since Lemma 5.2 is true with the L^1-norm replaced by the L^2-norm, Theorem 5.1 is also true in the L^2 setting. i.e, if a function f is PSPACE L^2-computable (replacing L^1 norms with L^2 norms in definition 3.2) then there exists a polynomial q satisfying the following: given any $m \in \mathbb{N}$, for all $n \geq 2^{q(m)}$, $\|A_n^f - \int f d\mu\|_2 \leq 2^{-m}$. Hence, for PSPACE L^2-computable functions and the left shift transformation T, we get bounds on the convergence speed in the von-Neumann's ergodic theorem.

It is easy to verify that if T is a PSPACE simple transformation then for any $n \geq 2$, T^n is also a PSPACE simple transformation. We need the following stronger assertion in the proof of the ergodic theorem.

[4] Equivalently, there exists a constant c such that for all $n > 0$, $\|A_n^f - \int f d\mu\|_1 \leq 2^{-\lfloor \log(n)^{\frac{1}{c}} \rfloor}$.

▶ **Lemma 5.3.** *Let $T : (\Sigma^\infty, \mu) \to (\Sigma^\infty, \mu)$ be a* PSPACE *simple transformation with controlling constant c. There exists a* PSPACE *machine N such that for each $n \in \mathbb{N}$ and $\sigma, \alpha \in \Sigma^*$,*

$$N(1^n, \sigma, \alpha) = \begin{cases} 1 & \text{if } |\alpha| \geq |\sigma| + cn \text{ and } \alpha 0^\infty \in T^{-n}([\sigma]) \\ 0 & \text{if } |\alpha| \geq |\sigma| + cn \text{ and } \alpha 0^\infty \notin T^{-n}([\sigma]) \\ ? & \text{otherwise} \end{cases}$$

Proof of Lemma 5.3. Let M be the machine witnessing the fact that T is a PSPACE simple transformation with the polynomial space complexity bound $p(n)$. Let the machine N do the following on input $(1^n, \sigma, \alpha)$:

1. If $\alpha < |\sigma| + cn$, then output ?.
2. If $n = 1$ then, run $M(\sigma, \alpha)$ and output the result of this simulation.
3. Else:
   a. For all strings α' of length $|\sigma| + c(n-1)$ do the following:
      i. If $N(1^{n-1}, \sigma, \alpha') = 1$ then, output 1 if $M(\alpha', \alpha) = 1$.
4. If no output is produced in the above steps, output 0.

When $n = 1$, N uses at most $p(|\sigma|+|\alpha|+cn)+O(1)$ space. Inductively, assume that for $n = k$, N uses at most $(2k-1)p(|\sigma|+|\alpha|+cn)+O(1)$ space. For $n = k+1$, the storage of α' and the two simulations inside step 3a can be done in $2p(|\sigma|+|\alpha|+cn)+(2k-1)p(|\sigma|+|\alpha|+cn)+O(1) = (2(k+1)-1)p(|\sigma|+|\alpha|+cn)+O(1)$ space. Hence, N is a PSPACE machine. ◀

Now, we prove the unconditional ergodic theorem for PSPACE L^1 functions over the Bernoulli space. The proof involves adaptations of techniques from Rute [21], together with new quantitative bounds which yield the result within prescribed resource bounds.

▶ **Theorem 5.4.** *Let T be the left shift transformation $T : (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu) \to (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ where μ is the Bernoulli measure $\mu(\sigma) = 2^{-|\sigma|}$. Then, for any* PSPACE *L^1-computable f, $\lim_{n\to\infty} \widetilde{A_n^f} = \int f d\mu$ on* EXP *randoms.*

Proof of Theorem 5.4. Let $\langle f_m \rangle_{m=1}^\infty$ be any PSPACE sequence of simple functions L^1 approximating f. We initially approximate A_n^f with a PSPACE sequence of simple functions $\langle g_n \rangle_{n=1}^\infty$ which converges to $\int f d\mu$ on EXP randoms. Then we show that $\widetilde{A_n^f}$ has the same limit as g_n on PSPACE randoms and hence on EXP randoms.

For each n, it is easy to verify that $\left\langle A_n^{f_m} \right\rangle_{m=1}^\infty$ is a PSPACE sequence of simple functions L^1 approximating A_n^f with the same rate of convergence. Using techniques similar to those in Lemma 4.3 and Corollary 4.4, we can obtain a polynomial p such that

$$\mu\left(\left\{x : \sup_{m \geq p(n+i)} |A_n^{f_m}(x) - A_n^{f_{p(n+i)}}(x)| \geq \frac{1}{2^{n+i+1}}\right\}\right) \leq \frac{1}{2^{n+i+1}}.$$

For every $n > 0$, let $g_n = A_n^{f_{p(n)}}$. We initially show that $\langle g_n \rangle_{n=1}^\infty$ converges to $\int f d\mu$ on EXP randoms. Let $m_1, m_2 \geq 0$. From Theorem 4.2, A_n^f is PSPACE-rapid almost everywhere convergent to $\int f d\mu$. Hence there is a polynomial q such that

$$\mu\left(\left\{x : \sup_{n \geq 2^{q(m_1+m_2)}} |A_n^f(x) - \int f d\mu| \geq \frac{1}{2^{m_1+1}}\right\}\right) \leq \frac{1}{2^{m_2+1}}.$$

Let $N(m_1,m_2) = \max\{2m_1, 2m_2, 2^{q(m_1+m_2)}\}$. Then,

$$\sum_{n \geq N(m_1,m_2)} \frac{1}{2^{k+1}} = \frac{1}{2^{N(m_1,m_2)}} \leq \min\left\{\frac{1}{2^{m_1+1}}, \frac{1}{2^{m_2+1}}\right\}.$$

Let

$$G_n = \left\{x : \sup_{n \geq N(m_1,m_2)} |g_n - \int f d\mu| > \frac{1}{2^{m_1}}\right\}.$$

Now, we have

$$\begin{aligned}
\mu\left(G_n\right) \leq & \sum_{n \geq N(m_1,m_2)} \mu\left(\left\{x : |g_n - A_n^f(x)| > \frac{1}{2^{m_1+1}}\right\}\right) \\
& + \mu\left(\left\{x : \sup_{n \geq 2^{q(m_1+m_2)}} |A_n^f(x) - \int f d\mu| \geq \frac{1}{2^{m_1+1}}\right\}\right) \\
\leq & \sum_{n \geq N(m_1,m_2)} \frac{1}{2^{n+1}} + \frac{1}{2^{m_2+1}} \\
\leq & \frac{1}{2^{m_2}}.
\end{aligned}$$

Note that $N(m_1,m_2)$ is bounded by $2^{(m_1+m_2)^c}$ for some $c \in \mathbb{N}$. Hence, g_n is PSPACE-rapid almost everywhere convergent to $\int f d\mu$. From Lemma 4.6 it follows that $\langle g_n \rangle_{n=1}^{\infty} = \left\langle A_n^{f_{p(n)}} \right\rangle_{n=1}^{\infty}$ is a PSPACE sequence of simple functions (in parameter n). Let $I_f : \Sigma^\infty \to \Sigma^\infty$ be the constant function taking the value $\int f d\mu$ over all $x \in \Sigma^\infty$. From the above observations and Lemma 4.3 we get that $\lim_{n\to\infty} g_n(x) = \widetilde{I}_f(x)$ for any x which is EXP random. From Lemma 4.5, we get that $\lim_{n\to\infty} g_n(x) = \int f d\mu$ for any x which is EXP random.

We now show that $\lim_{n\to\infty} \widetilde{A}_n^f = \lim_{n\to\infty} g_n$ on PSPACE randoms. Define

$$U_{n,i} = \left\{x : \max_{p(n+i) \leq m \leq p(n+i+1)} |A_n^{f_m}(x) - A_n^{f_{p(n+i)}}(x)| \geq \frac{1}{2^{n+i+1}}\right\}.$$

We already know $\mu(U_{n,i}) \leq \frac{1}{2^{n+i+1}}$. $U_{n,i}$ can be shown to be polynomial space approximable in parameters n and i in the following sense. There exists a sequence of sets of strings $\langle S_{n,i} \rangle_{i,n \in \mathbb{N}}$ and polynomial p satisfying the following conditions:

1. $U_{n,i} = [S_{n,i}]$.
2. There exists a *controlling polynomial* r such that $max\{|\sigma| : \sigma \in S_{n,i})\} \leq r(n+i)$.
3. The function $g : \Sigma^* \times 1^* \times 1^* \to \{0,1\}$ such that

$$g(\sigma, 1^n, 1^i) = \begin{cases} 1 & \text{if } \sigma \in S_{n,i} \\ 0 & \text{otherwise,} \end{cases}$$

   is decidable by a PSPACE machine.

The above claims can be established by using techniques similar to those in Lemma 4.6 and Lemma 4.3. We show the construction of a machine N computing the function g above. Let M_f be a computing machine and let q be a controlling polynomial for $\langle f_n \rangle_{n=1}^{\infty}$. Let c be a controlling constant for T. Let M' be the machine from Lemma 5.3. Machine N on input $(\sigma, 1^n, 1^i)$ does the following:

1. If $|\sigma| > q(p(n+i+1)) + cn$, then output 0.
2. Compute $A_n^{f_{p(n+i)}}(\sigma 0^\infty)$ as in Lemma 4.6 by using M_f and M' and store the result.
3. For each $m \in [p(n+i), p(n+i+1)]$ do the following:
   a. Compute $A_n^{f_m}(\sigma 0^\infty)$ as in Lemma 4.6 by using M_f and M' and store the result.
   b. Check if $|A_n^{f_m}(\sigma 0^\infty) - A_n^{f_{p(n+i)}}(\sigma 0^\infty)| \geq \frac{1}{2^{n+i+1}}$. If so, output 1.
4. Output 0.

It can be easily verified that N is a PSPACE machine. $r(n+i) = q(p(n+i+1)) + cn$ is a controlling polynomial for $\langle U_{n,i}\rangle_{n,i\in\mathbb{N}}$. Define

$$V_m = \bigcup_{\substack{n,i\geq 0\\ n+i=m}} U_{n,i}.$$

Note that

$$\mu(V_m) \leq \frac{m}{2^m}.$$

It can be shown that for any j,

$$\sum_{n>j} \frac{m}{2^m} = \frac{1}{2^{j-1}} + \frac{j}{2^j}.$$

Given any $k \geq 0$, let $p(k) = 3(k+1)$. Hence, we have

$$\sum_{n=p(k)+1}^{\infty} \frac{m}{2^m} = \frac{1}{2^{3(k+1)}} + \frac{3(k+1)}{2^{3(k+1)}} \;<\; \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}}\frac{3(k+1)}{2^{2(k+1)}} \;<\; \frac{2}{2^{k+1}} = \frac{1}{2^k}.$$

The last inequality holds since $3(k+1) < 2^{2(k+1)}$ for all $k \geq 0$. Since each V_m is a finite union of sets from $\langle U_{n,i}\rangle_{n,i\in\mathbb{N}}$, the machine computing $\langle U_{n,i}\rangle_{n,i\in\mathbb{N}}$ can be easily modified to construct a machine witnessing that $\langle V_m\rangle_{m=1}^{\infty}$ is a PSPACE approximable sequence of sets. From these observations, it follows that $\langle V_m\rangle_{m=1}^{\infty}$ is a PSPACE Solovay test. Let x be a PSPACE random. x is in at most finitely many V_m and hence in at most finitely many $U_{n,i}$. Hence for some large enough N for all $n \geq N$, $i \geq 0$ and for all m such that $p(n+i) \leq m \leq p(n+i+1)$, we have $|A_n^{f_m}(x) - A_n^{f_{p(n+i)}}(x)| < \frac{1}{2^{n+i+1}}$. It follows that for all $n \geq N$ and for all $m \geq p(n)$ that

$$|A_n^{f_m}(x) - g_n(x)| = |A_n^{f_m}(x) - A_n^{f_{p(n)}}(x)| \;\leq\; \sum_{i=0}^{\infty} \frac{1}{2^{n+i+1}} \;\leq\; 2^{-n}.$$

Therefore, $\lim_{n\to\infty} \widetilde{A}_n^f(x) = \lim_{n\to\infty} g_n(x)$ on all PSPACE random x and hence on all x which is EXP random.

Hence, we have shown that $\lim_{n\to\infty} \widetilde{A}_n^f = \int f d\mu$ on EXP randoms which completes the proof of the theorem. ◀

6 General PSPACE ergodic theorem

We now extend Theorem 5.4 into the setting of PSPACE-probability Cantor spaces. V'yugin [26] shows that the speed of a.e. convergence to ergodic averages in computable ergodic systems is not computable in general. This leads us to consider some assumption on the rapidity of convergence in resource-bounded settings. We show that the requirement on L^1 rapidity of convergence of A_n^f is sufficient to derive our result. Several probabilistic laws like the Law of Large Numbers, Law of Iterated Logarithm satisfy this criterion, hence the assumption is sufficiently general. Moreover, Theorem 5.1 shows that in the canonical example of Bernoulli systems with the left-shift, every PSPACE L^1 function exhibits PSPACE rapidity of A_n^f, showing that the latter property is not artificial. We prove the general PSPACE ergodic theorem for transformations which satisfy PSPACE ergodicity.

▶ **Definition 6.1** (PSPACE ergodic transformations). *A measurable function $T : (\Sigma^\infty, \mu) \to (\Sigma^\infty, \mu)$ is* PSPACE ergodic *if T is a* PSPACE *simple measure preserving transformation such that for any* PSPACE *L^1-computable $f \in L^1(\Sigma^\infty, \mu)$, $\int f d\mu$ is a* PSPACE*-rapid L^1 limit point of A_n^f.*

Now, we prove the main result of our work.

▶ **Theorem 6.2.** *Let $(\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ be a* PSPACE*-probability Cantor space. Let $T : (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu) \to (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ be a* PSPACE *ergodic measure preserving transformation. Then, for any* PSPACE *L^1-computable f, $\lim_{n\to\infty} \widetilde{A_n^f} = \int f d\mu$ on* EXP *randoms.*

Proof. Observe that Lemma 4.3, Corollary 4.4 and Lemma 4.6 are true in the setting of PSPACE-probability Cantor spaces. The proof of Lemma 4.5 can be extended to the setting of PSPACE-probability Cantor spaces in a straightforward manner. Since T is PSPACE ergodic, we get that we get that $\int f d\mu$ is a PSPACE-rapid L^1-limit point of A_n^f. Now, the theorem follows from these observations and the same techniques as in the proof of Theorem 5.4. ◀

The convergence notions involved in proving the PSPACE/SUBEXP-space ergodic theorems and their interrelationships are summarized in Figure 1.

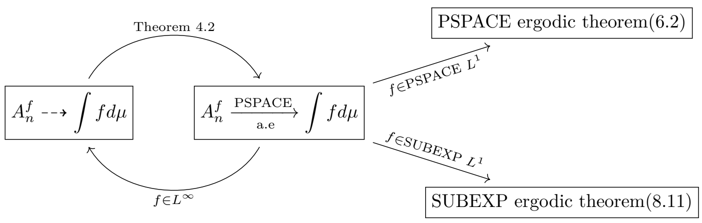

Figure 1 Relationships between the major convergence notions involving PSPACE simple measure preserving transformations. $A_n^f \dashrightarrow \int f d\mu$ denotes that $\int f d\mu$ is a PSPACE-rapid L^1-limit point of A_n^f. PSPACE/SUBEXP-space ergodicity is required only for obtaining the ergodic theorems from PSPACE a.e convergence.

7 A partial converse to the PSPACE Ergodic Theorem

In this section we give a partial converse to the PSPACE ergodic theorem (Theorem 6.2). We show that for any PSPACE null x, there exists a function f and transformation T satisfying all the conditions in Theorem 6.2 such that $\widetilde{A_n^f}(x)$ does not converge to $\int f d\mu$.

Let us first observe that due to Corollary 4.4, Theorem 6.2 is equivalent to the following:

▶ **Theorem.** *Let T be a* PSPACE *ergodic measure preserving transformation such that for any* PSPACE *L^1-computable f, $\int f d\mu$ is an* PSPACE*-rapid L^1-limit point of A_n^f. Let $\{g_{n,i}\}$ be any collection of simple functions such that for each n, $\langle g_{n,i}\rangle_{i=1}^\infty$ is a* PSPACE *L^1-approximation of $\widetilde{A_n^f}$. Then, $\lim_{n\to\infty} \lim_{i\to\infty} g_{n,i}(x) = \int f d\mu$ for any* EXP *random x.*

Hence, the ideal converse to Theorem 6.2 is the following:

▶ **Theorem.** *Given any* EXP *null x, there exists a* PSPACE *ergodic measure preserving transformation T and* PSPACE *L^1-computable $f \in L^1(\Sigma^\infty, \mu)$ such that the following conditions are true:*

1. *$\int f d\mu$ is an* PSPACE*-rapid limit point of A_n^f.*
2. *There exists a collection of simple functions $\{g_{n,i}\}$ such that for each n, $\langle g_{n,i}\rangle_{i=1}^\infty$ is a* PSPACE *L^1-approximation of A_n^f but $\lim_{n\to\infty} \lim_{i\to\infty} g_{n,i}(x) \neq \int f d\mu$.*

But, we show the following partial converse to Theorem 6.2.

▶ **Theorem 7.1.** *Given any* PSPACE *null x, there exists a* PSPACE *L^1-computable $f \in L^1(\Sigma^\infty, \mu)$ such that for any* PSPACE *simple measure preserving transformation, the following conditions are true:*

1. *For all $n \in \mathbb{N}$, $\|A_n^f - \int f d\mu\|_1 = 0$. Hence, $\int f d\mu$ is an* PSPACE*-rapid L^1-limit point of A_n^f.*
2. *There exists a collection of simple functions $\{g_{n,i}\}$ such that for each n, $\langle g_{n,i}\rangle_{i=1}^\infty$ is a* PSPACE *L^1-approximation of A_n^f but $\lim_{n\to\infty} \lim_{i\to\infty} g_{n,i}(x) \neq \int f d\mu$.*

A proof of the above theorem requires the construction in the following lemma.

▶ **Lemma 7.2.** *Let $\langle U_n\rangle_{n=1}^\infty$ be a* PSPACE *test. Then there exists a sequences of sets $\left\langle \widehat{S}_n \right\rangle_{n=1}^\infty$ such that for each $n \in \mathbb{N}$, $\widehat{S}_n \subseteq \Sigma^*$ satisfying the following conditions:*

1. *$\mu([\widehat{S}_n]) \leq 2^{-n}$.*
2. *$\cap_{m=1}^\infty \cup_{n=m}^\infty [\widehat{S}_n] \supseteq \cap_{n=1}^\infty U_n$.*
3. *There exists $c \in \mathbb{N}$ such that for all n, $\sigma \in \widehat{S}_n$ implies $|\sigma| \leq n^c$.*
4. *There exists a* PSPACE *machine N such that $N(\sigma, 1^n) = 1$ if $\sigma \in \widehat{S}_n$ and 0 otherwise.*

Proof of Theorem 7.1. Let $\langle V_n\rangle_{n=1}^\infty$ be any PSPACE test such that $x \in \cap_{n=1}^\infty V_n$. From Lemma 7.2, there exists a collection of sets $\left\langle \widehat{S}_n \right\rangle_{n=1}^\infty$ such that $\cap_{m=1}^\infty \cup_{n=m}^\infty [\widehat{S}_n] \supseteq \cap_{n=1}^\infty V_n$. Let,

$$U_n = \{\sigma : [\sigma] \in \widehat{S}_i \text{ for some } i \text{ such that } 2n+1 \leq i \leq 2(n+1)+1\}$$

Let $f_n = n\chi_{U_n}$. Since

$$\mu(U_n) \leq \sum_{i=2n+1}^{2(n+1)+1} \frac{1}{2^i} \leq \frac{1}{2^{2n}},$$

it follows that

$$\|f_n\|_1 \leq \frac{n}{2^{n+n}} \leq \frac{1}{2^n}.$$

Using the properties of $\left\langle \widehat{S}_n \right\rangle_{n=1}^\infty$, it can be shown that $\langle f_n\rangle_{n=1}^\infty$ is a PSPACE L^1-approximation of $f = 0$. We construct a machine M computing $\langle f_n\rangle_{n=1}^\infty$. The other conditions are easily verified. Let N be the machine from Lemma 7.2. On input $(1^n, \sigma)$, M does the following:

1. If $|\sigma| < (2(n+1)+1)^c$ then, output ?.
2. Else, for each $i \in [2n+1, 2(n+1)+1]$ do the following:
   a. For each $\alpha \subseteq \sigma$, do the following:
      i. If $N(1^i, \alpha) = 1$ then, output n.
3. Output 0.

M uses at most polynomial space and computes $\langle f_n \rangle_{n=1}^{\infty}$. Define

$$g_{n,i} = \frac{f_i + f_i \circ T + \cdots + f_i \circ T^{n-1}}{n}$$

For any fixed $n \in \mathbb{N}$, since T is a PSPACE simple transformation, as in Lemma 4.6 it can be shown that $\langle g_{n,i} \rangle_{i=1}^{\infty}$ is a PSPACE L^1-approximation of A_n^f. We know that there exist infinitely many m such that $x \in [\widehat{S}_m]$. For any such m, let i be the unique number such that $2i + 1 \leq m \leq 2(i+1) + 1$. For this i, $f_i(x) = i$. This shows that there exist infinitely many i such that $f_i(x) = i$. Since each f_i is a non-negative function, it follows that there are infinitely many i with $g_{n,i} \geq i/n$. Hence, if $\lim\limits_{i \to \infty} g_{n,i}(x)$ exists, then it is equal to ∞. It may be the case that $\lim\limits_{i \to \infty} g_{n,i}(x)$ does not exist. In either case, $\lim\limits_{n \to \infty} \lim\limits_{i \to \infty} g_{n,i}(x)$ cannot be equal to $\int f d\mu = 0$. Hence, our construction satisfies all the desired conditions. ◀

8 An ergodic theorem for SUBEXP-space randoms and its converse

In the previous sections, we demonstrated that for PSPACE L^1-computable functions and PSPACE simple transformations, the Birkhoff averages converge to the desired value over EXP randoms. However, the converse holds only over PSPACE non-randoms. The two major reasons for this *gap* are the following: PSPACE-rapid convergence necessitates exponential length cylinders while constructing the randomness tests, and PSPACE L^1-computable functions are not strong enough to *capture* all PSPACE randoms. In this section, we demonstrate that for a different notion of randomness - SUBEXP-space randoms and a larger class of L^1-computable functions (SUBEXP-space L^1-computable), we can prove the ergodic theorem on the randoms and obtain its converse on the non-randoms. Analogous to Towsner and Franklin [5], we demonstrate that the ergodic theorem for PSPACE simple transformations and SUBEXP-space L^1-computable functions satisfying PSPACE rapidity, fails for exactly this class of non-random points. We first introduce SUBEXP-space tests and SUBEXP-space randomness.

▶ **Definition 8.1** (SUBEXP-space sequence of open sets). *A sequence of open sets $\langle U_n \rangle_{n=1}^{\infty}$ is a* SUBEXP-space sequence of open sets *if there exists a sequence of sets $\langle S_n^k \rangle_{k,n \in \mathbb{N}}$, where $S_n^k \subseteq \Sigma^*$ such that*

1. *$U_n = \cup_{k=1}^{\infty} [S_n^k]$, where for any $m > 0$, $\mu\left(U_n - \cup_{k=1}^{m} [S_n^k]\right) \leq m^{-\log(m)}$.*
2. *There exists a* controlling polynomial *p such that max$\{|\sigma| : \sigma \in \cup_{k=1}^{m} S_n^k)\} \leq 2^{p(\log(n)+\log(m))}$.*
3. *The function $g : \Sigma^* \times 1^* \times 1^* \to \{0,1\}$ such that $g(\sigma, 1^n, 1^m) = 1$ if $\sigma \in S_n^m$, and 0 otherwise, is decidable by a* PSPACE *machine.*

▶ **Definition 8.2** (SUBEXP-space randomness). *A sequence of open sets $\langle U_n \rangle_{n=1}^{\infty}$ is a* SUBEXP-space test *if it is a* SUBEXP*-space sequence of open sets and for all $n \in \mathbb{N}$, $\mu(U_n) \leq n^{-\log(n)}$.*

A set $A \subseteq \Sigma^{\infty}$ is SUBEXP-space null *if there is a* SUBEXP*-space test $\langle U_n \rangle_{n=1}^{\infty}$ such that $A \subseteq \cap_{n=1}^{\infty} U_n$ and is* SUBEXP-space random *otherwise.*[5]

The slower decay rate of $n^{-\log(n)} = 2^{-\log(n)^2}$ enables us to obtain an ergodic theorem and an exact converse in the SUBEXP-space setting .The following result is useful in manipulating sums involving terms of the form $2^{-(\log(n))^k}$ for $k \geq 2$.

[5] It is easy to see that the set of SUBEXP-space randoms is smaller than the set of PSPACE randoms. But, we do not know if any inclusion holds between SUBEXP-space randoms and EXP-randoms.

▶ **Lemma 8.3.** *For any* $m \in \mathbb{N}$, $\sum_{n=2(2m^2+1)}^{\infty} \frac{n}{n^{\log(n)}} \leq \frac{1}{m^{\log(m)}}$.

A similar inequality holds on replacing $n/n^{\log(n)}$ with $1/n^{\log(n)}$. Now, we introduce the Solovay analogue of SUBEXP-space randomness and prove that these notions are analogous.

▶ **Definition 8.4** (SUBEXP-space Solovay test)**.** *A sequence of open sets* $\langle U_n \rangle_{n=1}^{\infty}$ *is a* SUBEXP-space Solovay test *if it is a* SUBEXP*-space sequence of open sets and there exists a polynomial* p *such that* $\forall m \geq 0$, $\sum_{n=p(m)+1}^{\infty} \mu(U_n) \leq m^{-\log(m)}$. *A set* $A \subseteq \Sigma^{\infty}$ *is* SUBEXP-space Solovay null *if there exists a* SUBEXP*-space Solovay test* $\langle U_n \rangle_{n=1}^{\infty}$ *such that* $A \subseteq \cap_{i=1}^{\infty} \cup_{n=i}^{\infty} U_n$, *and is* SUBEXP-space Solovay random *otherwise.*

▶ **Lemma 8.5.** *A set* $A \subseteq \Sigma^{\infty}$ *is* SUBEXP*-space null if and only if* A *is* SUBEXP*-space Solovay null.*

Now, we define SUBEXP-space analogues of concepts from Section 3.

▶ **Definition 8.6** (SUBEXP-space sequence of simple functions)**.** *A sequence of simple functions* $\langle f_n \rangle_{n=1}^{\infty}$ *where each* $f_n : \Sigma^{\infty} \to \mathbb{Q}$ *is a* SUBEXP-space sequence of simple functions *if*

1. *There is a* controlling polynomial p *such that for each* n, *there exists* $k(n) \in \mathbb{N}$, $\{d_1, d_2 \ldots, d_{k(n)}\} \subseteq \mathbb{Q}$ *and* $\{\sigma_1, \sigma_2 \ldots \sigma_{k(n)}\} \subseteq \Sigma^{2^{p(\log(n))}}$ *such that* $f_n = \sum_{i=1}^{k(n)} d_i \chi_{\sigma_i}$, *where* χ_{σ_i} *is the characteristic function of the cylinder* $[\sigma_i]$.
2. *There is a* PSPACE *machine* M *such that for each* $n \in \mathbb{N}, \sigma \in \Sigma^*$, $M(1^n, \sigma)$ *outputs* $f_n(\sigma 0^{\infty})$ *if* $|\sigma| \geq 2^{p(\log(n))}$ *and ? otherwise.*

▶ **Definition 8.7** (SUBEXP-space L^1-computable functions)**.** *A function* $f \in L^1(\Sigma^{\infty}, \mu)$ *is* SUBEXP*-space* L^1*-computable if there exists a* SUBEXP*-space sequence of simple functions* $\langle f_n \rangle_{n=1}^{\infty}$ *such that for every* $n \in \mathbb{N}$, $\|f - f_n\| \leq n^{-\log(n)}$. *The sequence* $\langle f_n \rangle_{n=1}^{\infty}$ *is called a* SUBEXP-space L^1-approximation *of* f.

We require the following equivalent definitions of PSPACE-rapid convergence notions for working in the setting of SUBEXP-space randomness.

▶ **Lemma 8.8.** *A real number* a *is a* PSPACE*-rapid limit point of the real number sequence* $\langle a_n \rangle_{n=1}^{\infty}$ *if and only if there exists a polynomial* p *such that for all* $m \in \mathbb{N}$, $\exists k \leq 2^{p(\log(m))}$ *such that* $|a_k - a| \leq m^{-\log(m)}$.

▶ **Lemma 8.9.** *A sequence of measurable functions* $\langle f_n \rangle_{n=1}^{\infty}$ *is* PSPACE*-rapid almost everywhere convergent to a measurable function* f *if and only if there exists a polynomial* p *such that for all* m_1 *and* m_2,

$$\mu\left(\left\{x : \sup_{n \geq 2^{p(\log(m_1)+\log(m_2))}} |f_n(x) - f(x)| \geq \frac{1}{m_1^{\log(m_1)}}\right\}\right) \leq \frac{1}{m_2^{\log(m_2)}}.$$

The same technique used in the proof of Lemma 8.8 can be used to prove this claim.

Before addressing the main result, let us define SUBEXP-space ergodicity.

▶ **Definition 8.10** (SUBEXP-space ergodic transformations)**.** *A measurable function* $T : (\Sigma^{\infty}, \mu) \to (\Sigma^{\infty}, \mu)$ *is* SUBEXP-space ergodic *if* T *is a* PSPACE *simple transformation such that for any* SUBEXP*-space* L^1*-computable* $f \in L^1(\Sigma^{\infty}, \mu)$, $\int f d\mu$ *is a* PSPACE*-rapid* L^1 *limit point of* $A_n^{f,T}$.

Lemma 4.3, Corollary 4.4 and Lemma 4.6 have analogous results in the SUBEXP-space setting. We prove the SUBEXP-space ergodic theorem below.

▶ **Theorem 8.11.** *Let* $T : (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu) \to (\Sigma^\infty, \mathcal{B}(\Sigma^\infty), \mu)$ *be a* SUBEXP*-space ergodic measure preserving transformation. Then, for any* SUBEXP*-space* L^1*-computable* f*,* $\lim_{n\to\infty} \widetilde{A_n^f} = \int f d\mu$ *on* SUBEXP*-space randoms.*

An important reason for investigating SUBEXP-space randomness is that the SUBEXP-space ergodic theorem has an exact converse unlike the PSPACE ergodic theorem which only seems to have a partial converse (Theorem 7.1).

▶ **Theorem 8.12.** *Given any* SUBEXP*-space null* x*, there exists a* SUBEXP*-space* L^1*-computable* $f \in L^1(\Sigma^\infty, \mu)$ *such that for any* PSPACE *simple measure preserving transformation, the following conditions are true:*

1. *For all* $n \in \mathbb{N}$*,* $\|A_n^f - \int f d\mu\|_1 = 0$*. Hence,* $\int f d\mu$ *is an* PSPACE*-rapid* L^1*-limit point of* A_n^f*.*
2. *There exists a collection of simple functions* $\{g_{n,i}\}$ *such that for each* n*,* $\langle g_{n,i}\rangle_{i=1}^\infty$ *is a* SUBEXP*-space* L^1*-approximation of* A_n^f *but* $\lim_{n\to\infty} \lim_{i\to\infty} g_{n,i}(x) \neq \int f d\mu$*.*

The proofs of both Theorem 8.11 and Theorem 8.12 are similar to those of Theorem 6.2 and Theorem 7.1, but requires Lemma 8.3 for minimizing summations of the form $\sum n^{-\log(n)}$ appearing in the error bounds.

References

1. L. Bienvenu, A. Day, M. Hoyrup, I. Mezhirov, and A. Shen. A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points. *Information and Computation*, 210:21–30, 2012.
2. P. Billingsley. *Probability and Measure,* second edition. John Wiley and Sons, New York, N.Y., 1986.
3. E. Bishop. *Foundations of constructive analysis.* McGraw-Hill, New York, 1967.
4. Rodney G Downey and Denis R Hirschfeldt. *Algorithmic randomness and complexity.* Springer Science & Business Media, 2010.
5. Johanna NY Franklin and Henry Towsner. Randomness and non-ergodic systems. *Moscow Mathematical Journal*, 14:711–744, 2014.
6. Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. A constructive Borel–Cantelli lemma. constructing orbits with required statistical properties. *Theoretical Computer Science*, 410(21-23):2207–2222, 2009.
7. S. Galotolo, M. Hoyrup, and C. Rojas. Computing the speed of convergence of ergodic averages and pseudorandom points in computable dynamical systems. In *7th International Conference on Computability and Complexity in Analysis*, pages 7–18, 2010.
8. Michael Hochman. Upcrossing ineualities for stationary sequences and applications. *Annals of Probability*, 37(6):2135–2149, 2009.
9. Mathieu Hoyrup. The dimension of ergodic random sequences. In *Proceedings of the Symposium on Theoretical Aspects of Computer Science*, pages 567–576, 2012.
10. Mathieu Hoyrup and Cristóbal Rojas. Applications of effective probability theory to Martin-Löf randomness. In *International Colloquium on Automata, Languages, and Programming*, pages 549–561. Springer, 2009.
11. Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and martin-löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.
12. Xiang Huang and D. M. Stull. Polynomial space randomness in analysis. In *MFCS*, 2016.
13. Ker-I Ko. *Complexity theory of real functions.* Springer Science & Business Media, 2012.
14. Ulrich Krengel. On the speed of convergence in the ergodic theorem. *Monatshefte für Mathematik*, 86:3–6, 1978.
15. L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. Pure and Applied Mathematics.
16. John E. Maxfield. A short proof of Pillai's theorem on normal numbers. *Pacific Journal of Mathematics*, 2(1):23–24, 1952.

17 S. Nandakumar. An effective ergodic theorem and some applications. In *Proceedings of the 40th Annual Symposium on the Theory of Computing*, pages 39–44, 2008.
18 Satyadev Nandakumar, Subin Pulari, Prateek Vishnoi, and Gopal Viswanathan. An analogue of pillai's theorem for continued fraction normality and an application to subsequences. *arXiv preprint*, 2019. `arXiv:1909.03431`.
19 André Nies. *Computability and randomness*, volume 51. OUP Oxford, 2009.
20 S. S. Pillai. On normal numbers. *Proceedings of the Indian Academy of Sciences Sect. A*, 11:73–80, 1940.
21 Jason Rute. *Topics in algorithmic randomness and computable analysis*. PhD thesis, Carnegie Mellon University, August 2013.
22 Donald M Stull. *Algorithmic randomness and analysis*. PhD thesis, Iowa State University, Ames, Iowa, USA, 2017.
23 Donald M. Stull. Resource bounded randomness and its applications. *Algorithmic Randomness: Progress and Prospects*, 50:301, 2020.
24 M. van Lambalgen. *Random Sequences*. PhD thesis, Department of Mathematics, University of Amsterdam, 1987.
25 V. G. Vovk. The law of the iterated logarithm for random Kolmogorov, or chaotic, sequences. *Theory of Probability and its Applications*, 32(3):413–425, 1988.
26 V. V. V'yugin. Effective convergence in probability and an ergodic theorem for individual random sequences. *Theory of Probability and Its Applications*, 42(1):39–50, 1997.

A Proof of Theorem 5.1

Proof of Theorem 5.1. Let $\langle f_n \rangle_{n=1}^{\infty}$ be a PSPACE sequence of simple functions witnessing the fact that f is PSPACE L^1-computable. Let p be a controlling polynomial and let t be a polynomial upper bound for the space complexity of the machine associated with $\langle f_n \rangle_{n=1}^{\infty}$. Let q_1,q_2 be the polynomials from Lemma 5.2. Let $c \in \mathbb{N}$ be any number such that if a $r \in \mathbb{Q}$ has a representation of length l then $r \leq 2^{l^c}$ (see Section 2). Observe that for any $m \in \mathbb{N}$,

$$\begin{aligned}\|A_n^f - \int f d\mu\|_1 &\leq \|A_n^f - A_n^{f_{q_1(m+3)}}\|_1 + \|A_n^{f_{q_1(m+3)}} - \int f_{q_1(m+3)} d\mu\|_1 + \|\int f_{q_1(m+3)} d\mu - \int f d\mu\|_1 \\ &\leq \frac{1}{2^{q_1(m+3)}} + \|A_n^{f_{q_1(m+3)}} - \int f_{q_1(m+3)} d\mu\|_1 + \frac{1}{2^{q_1(m+3)}}. \\ &\leq \frac{1}{2^{m+3}} + \|A_n^{f_{q_1(m+3)}} - \int f_{q_1(m+3)} d\mu\|_1 + \frac{1}{2^{m+3}}.\end{aligned}$$

We know that there exist $\{\sigma_1, \sigma_2 \ldots \sigma_k\} \subseteq \Sigma^{p(q_1(m+3))}$ such that $A_n^{f_{q_1(m+3)}} = \sum_{i=1}^{k(q_1(m+3))} d_i \chi_{\sigma_i}$ where each $d_i \leq 2^{t(q_1(m+3)+p(q_1(m+3)))^c}$. Hence,

$$\|A_n^{f_{q_1(m+3)}} - \int f_{q_1(m+3)} d\mu\|_1 \leq 2^{t(q_1(m+3)+p(q_1(m+3)))^c} \sum_{i=1}^{k(q_1(m+3))} \|A_n^{\chi_{\sigma_i}} - \mu(\sigma_i)\|_1$$

Since $|\sigma_i| \geq p(q_1(m+3)) \geq q_1(m+3)$, using Lemma 5.2, for

$$n \geq p(q_1(m+3))^3 2^{q_2(t(q_1(m+3)+p(q_1(m+3)))^c + p(q_1(m+3)) + m+3)}$$

we get that,

$$\begin{aligned}\|A_n^{f_{q_1(m+3)}} - \int f_{q_1(m+3)} d\mu\|_1 &\leq \frac{2^{t(q_1(m+3)+p(q_1(m+3)))^c + p(q_1(m+3))}}{2^{t(q_1(m+3)+p(q_1(m+3)))^c + p(q_1(m+3)) + m+3}} \\ &\leq \frac{1}{2^{m+3}}.\end{aligned}$$

Hence, for all $n \geq p(q_1(m+3))^3 2^{q_2(t(q_1(m+3)+p(q_1(m+3)))^c + p(q_1(m+3)) + m+3)}$ we have $\|A_n^f - \int f d\mu\|_1 \leq 3.2^{-(m+3)} < 2^{-m}$. ◀

Proof of Lemma 5.2. The major difficulty in directly approximating $\|A_n^{\chi_\sigma} - \mu(\sigma)\|_1$ is that for any $n, m \in \mathbb{N}$, $A_n^{\chi_\sigma}$ and $A_m^{\chi_\sigma}$ may not be *independent*. In order to overcome this, we use constructions similar to those used in proving Pillai's theorem (see [20], [16] for normal numbers, [18] for continued fractions) in order to approximate each $A_n^{\chi_\sigma}$ with sums of *disjoint* averages. These *disjoint* averages turns out to be averages of independent random variables. Hence, elementary results from probability theory regarding independent random variables can be used to show that $A_n^{\chi_\sigma}$ converges to $\int f d\mu$ sufficiently fast.

Observe that for any $x \in \Sigma^\infty$

$$A_n^{\chi_\sigma}(x) = \frac{|\{i \in [0, n-1] \mid T^i x \in [\sigma]\}|}{n}$$

Let $k = |\sigma|$. As in the proof of Theorem 3.1 from [18], the following is a decomposition of the above term as *disjoint* averages,

$$\frac{|\{i \in [0, n-1] \mid T^i x \in [\sigma]\}|}{n} = g_1(n) + g_2(n) + \cdots + g_{(1+\lfloor \log_2 \frac{n}{k} \rfloor)}(n) + \frac{(k-1).O(\log n)}{n}$$

where

$$g_p(n) = \begin{cases} \frac{|\{i \;\mid\; T^{ki}x \;\in\; [\sigma]\;,\; 0 \;\leq\; i \;\leq\; \lfloor n/k \rfloor\}|}{n} & \text{if } p = 1 \\ \frac{\sum_{j=1}^{k-1} |\{i \;\mid\; T^{(2^{p-1})ki}x \;\in\; [S_j], 0 \;\leq\; i \;\leq \lfloor n/2^{p-1}k \rfloor\}|}{n} & \text{if } 1 < p \leq (1 + \lfloor \log_2(n/k) \rfloor) \\ 0 & \text{otherwise,} \end{cases}$$

and S_j is the finite collection of $2^{(p-1)}k$ length blocks s.t σ occurs in it at starting position $(2^{(p-2)}k - j + 1)^{th}$ position i.e S_j is the set of strings of the form, $u\ a_1 a_2 \ldots a_k\ v$ where u is some string of length $2^{p-2}k - j$, and v is some string of length $2^{p-2}k - k + j$.

When $p = 1$,

$$g_1(n) = \frac{\sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}}{n}$$

where

$$X_i^{1,1}(x) = \begin{cases} 1 & \text{if } x[ik+1, (i+1)k] = \sigma \\ 0 & \text{otherwise} \end{cases}$$

When $1 < p \leq \lfloor \log_2(n/k) \rfloor$,

$$g_p(n) = \frac{\sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} \sum_{j=1}^{k-1} X_i^{p,j}}{n}$$

where,

$$X_i^{p,j}(x) = \begin{cases} 1 & \text{if } x[2^{p-2}k - j + 1, 2^{p-2}k - j + k] = \sigma \\ 0 & \text{otherwise} \end{cases}$$

Hence,

$$A_n^{\chi_\sigma}(x) = \frac{\sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}(x)}{n} + \sum_{p=2}^{\lfloor \log_2(\frac{n}{k}) \rfloor} \sum_{j=1}^{k-1} \frac{\sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j}}{n} + \frac{(k-1).O(\log n)}{n}$$

An important observation that we use later in the proof is that for any fixed p and j, $\{X_i^{p,j}\}_{i=1}^{\infty}$ is a collection of i.i.d Bernoulli random variables such that $\mu(\{x : X_i^{p,j}(x) = 1\}) = 2^{-|\sigma|}$. We will show that the conclusion of the lemma holds when $q_1(m) = 2(m+6)$ and $q_2(m) = 5(m+6)$. For any $m \in \mathbb{N}$,

$$\left\| \sum_{p=m+5+2}^{\infty} \sum_{j=1}^{k-1} \frac{1}{n} \sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j} \right\|_2 \leq \sum_{p=m+5+2}^{\infty} \frac{1}{2^{p-1}} \leq \frac{1}{2^{m+5}} \tag{1}$$

And for $n \geq |\sigma|^3 2^{q_2(m)} > |\sigma|^2 2^{2(m+5)}$,

$$\left\| \frac{(k-1)O(\log(n))}{n} \right\|_2 = \left\| \frac{(k-1)O(\log(n))}{\sqrt{n}\sqrt{n}} \right\|_2 \leq \left| \frac{k-1}{\sqrt{n}} \right| \leq \left| \frac{k-1}{k 2^{m+5}} \right| \leq \frac{1}{2^{m+5}} \tag{2}$$

Let,

$$D_{n,m}^{\sigma}(x) = \frac{\sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}(x)}{n} + \sum_{p=2}^{m+5+2} \sum_{j=1}^{k-1} \frac{\sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j}}{n}$$

From (1) and (2), we get that

$$\|A_n^{\chi_\sigma} - D_{n,m}^{\sigma}\|_2 \leq \frac{2}{2^{m+5}}.$$

Let,

$$E_{n,m}^{\sigma}(x) = \left(\frac{\sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}(x)}{\lfloor \frac{n}{k} \rfloor} - \frac{1}{2^k} \right) \frac{\lfloor \frac{n}{k} \rfloor}{n} + \sum_{p=2}^{m+5+2} \sum_{j=1}^{k-1} \left(\frac{\sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j}}{\lfloor \frac{n}{2^{p-1}k} \rfloor} - \frac{1}{2^k} \right) \frac{\lfloor \frac{n}{2^{p-1}k} \rfloor}{n}$$

Now,

$$D_{n,m}^{\sigma}(x) - E_{n,m}^{\sigma}(x) = \frac{1}{2^k k} + \sum_{p=2}^{m+5+2} \sum_{j=1}^{k-1} \frac{1}{2^k} \frac{\lfloor \frac{n}{2^{p-1}k} \rfloor}{n}$$

It follows that,

$$\begin{aligned}
\|D_{n,m}^{\sigma}(x) - E_{n,m}^{\sigma}\|_2 &\leq \frac{1}{2^k} + \sum_{p=2}^{m+5+2} \sum_{j=1}^{k-1} \frac{1}{2^k 2^{p-1} k} \\
&\leq \frac{1}{2^k} + \sum_{p=2}^{m+5+2} \frac{1}{2^k 2^{p-1}} \\
&\leq \frac{1}{2^k} + \sum_{p=2}^{m+5+2} \frac{1}{2^k} \\
&\leq \frac{m+5+2}{2^k}
\end{aligned}$$

Hence, if $|\sigma| = k \geq q_1(m) = m + 5 + 2 + m + 5$ then,

$$\|D^{\sigma}_{n,m}(x) - E^{\sigma}_{n,m}\|_2 \leq \frac{1}{2^{m+5}}$$

and,

$$\begin{aligned}\|A_n^{\chi_\sigma} - \mu(\sigma)\|_2 &\leq \|A_n^{\chi_\sigma} - D^{\sigma}_{n,m}\|_2 + \|D^{\sigma}_{n,m}(x) - E^{\sigma}_{n,m}\|_2 + \|E^{\sigma}_{n,m}\|_2 + \frac{1}{2^k} \\ &\leq \frac{3}{2^{m+5}} + \|E^{\sigma}_{n,m}\|_2 + \frac{1}{2^{2m+12}} \\ &\leq \frac{4}{2^{m+5}} + \|E^{\sigma}_{n,m}\|_2.\end{aligned}$$

Hence, in order to show that for all $n \geq |\sigma|^3 2^{q_2(m)}$, $\|A_n^{\chi_\sigma} - \mu(\sigma)\|_1 \leq \|A_n^{\chi_\sigma} - \mu(\sigma)\|_2 \leq 2^{-m}$, it is enough to show that for all $n \geq |\sigma|^3 2^{q_2(m)}$, $\|E^{\sigma}_{n,m}\|_2 \leq 2^{-(m+5)}$. Observe that

$$\|E^{\sigma}_{n,m}\|_2 \leq \left\| \frac{1}{\lfloor \frac{n}{k} \rfloor} \sum_{i=1}^{\lfloor \frac{n}{k} \rfloor} X_i^{1,1}(x) - \frac{1}{2^k} \right\|_2 + \sum_{p=2}^{m+5+2} \sum_{j=1}^{k-1} \left\| \frac{1}{\lfloor \frac{n}{2^{p-1}k} \rfloor} \sum_{i=1}^{\lfloor \frac{n}{2^{p-1}k} \rfloor} X_i^{p,j} - \frac{1}{2^k} \right\|_2 .$$

Let $Y_1, Y_2, \ldots Y_n$ be i.i.d Bernoulli random variables,

$$\begin{aligned}\left\| \frac{1}{n} \sum_{i=1}^{n} Y_i - \mathbf{E}(Y_1) \right\|_2 &= \sqrt{\mathbf{E}\left(\left(\frac{1}{n} \sum_{i=1}^{n} Y_i - \mathbf{E}(Y_1) \right)^2 \right)} \\ &= \sqrt{\mathrm{Var}\left(\frac{1}{n} \sum_{i=1}^{n} Y_i \right)} \\ &= \sqrt{\frac{1}{n^2} n \mathrm{Var}(Y_1)} \\ &\leq \frac{\sqrt{\mathrm{Var}(Y_1)}}{\sqrt{n}} \\ &\leq \frac{1}{2\sqrt{n}}\end{aligned}$$

The last inequality follows from the fact that the variance of Bernoulli random variables are always bounded by $\frac{1}{4}$. Hence, if $n \geq |\sigma|^3 2^{q_2(m)} = |\sigma|^3 2^{5(m+6)}$ then,

$$\left\lfloor \frac{n}{k} \right\rfloor > k^2 2^{4(m+6)}$$

and

$$\left\lfloor \frac{n}{2^{p-1}k} \right\rfloor \geq \frac{k^3 2^{5(m+6)}}{2^{m+5+1}k} > k^2 2^{4(m+6)}.$$

Hence for all $n \geq |\sigma|^3 2^{q_2(m)} = |\sigma|^3 2^{5(m+6)}$,

$$\begin{aligned}\|E^{\sigma}_{n,m}\|_2 &\leq \frac{1}{2k2^{2(m+6)}} + (m+6)k \frac{1}{2k2^{2(m+6)}} \\ &< \frac{1}{2^{m+6}} + \frac{1}{2^{m+6}} \\ &\leq \frac{1}{2^{m+5}}.\end{aligned}$$

Hence we obtain the desired conclusion. ◀