Ergod. Th. & Dynam. Sys., (2026), **46**, 3123–3140 © The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

doi:10.1017/etds.2026.10322

The pointwise ergodic theorem on finitely additive spaces

MORENIKEJI NERI

Department of Mathematics, Technische Universität Darmstadt, Germany
(e-mail: neri@mathematik.tu-darmstadt.de, keji21@live.co.uk)

(*Received* 3 *November* 2025 *UTC and accepted in revised form* 20 *May* 2026 *UTC*)

Abstract. The almost sure convergence of ergodic averages in Birkhoff's pointwise ergodic theorem is known to fail in the finitely additive setting. We introduce a natural reformulation of almost sure convergence suitable for finitely additive measures, which we call *finite almost sure convergence*. Unlike the classical formulation, finite almost sure convergence only involves measures of finite unions and intersections, making it well adapted to finitely additive spaces. Using this notion, we extend the pointwise ergodic theorem to finitely additive probability spaces. Our proof relies on demonstrating that several quantitative generalizations of the pointwise ergodic theorem remain valid in the finitely additive setting via an extension of the Calderón transference principle. The result then follows by exploiting the relationships between quantitative notions of almost sure convergence developed by the author and Powell [On quantitative convergence for stochastic processes: crossings, fluctuations and martingales. *Trans. Amer. Math. Soc. Series B* **12** (2025), 974–1019].

Key words: Birkhoff's ergodic theorem, finitely additive probability theory, Calderón transference, metastability, proof mining
2020 Mathematics Subject Classification: 37A30 (Primary); 03F10 (Secondary)

1. *Introduction*

Let $(\Omega, \mathcal{F}, \mu, \tau)$ be a dynamical system, where $(\Omega, \mathcal{F}, \mu)$ is a probability space and $\tau : \Omega \to \Omega$ is a measure-preserving automorphism. Birkhoff's pointwise ergodic theorem asserts that for any $f \in L^1(\Omega, \mathcal{F}, \mu)$, the ergodic averages

$$A_n f(\omega) := \frac{1}{n} \sum_{i=1}^{n} f(\tau^i(\omega))$$

converge almost surely to the integral $\int f \, d\mu$.

It is well known that this convergence may fail when μ is only finitely additive (see, for example, [35]). However, in [35], Ramakrishnan shows that the pointwise ergodic theorem

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

does hold in the finitely additive setting, provided that the sequence of functions $\{A_n f\}$ satisfies a certain technical regularity condition with respect to μ (which is automatically satisfied in the σ-additive case), and that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = 0 \quad \text{whenever } \mu(A_i) = 0 \text{ for all } i,$$

where $\{A_i\}$ is a sequence of sets in the sub-σ-algebra of τ-invariant sets, that is, $\tau^{-1}(A_i) = A_i$ for all i.

Many other limit theorems in probability theory also fail in the finitely additive setting, such as the strong law of large numbers and the martingale convergence theorem. A substantial body of work, primarily from the 1970s and 1980s, is devoted to recovering classical results of probability theory in the finitely additive context by imposing additional structural or regularity assumptions on the space or the functions involved. In particular, motivated by the insights of Dubins and Savage in their book *How to Gamble If You Must* [10] and by the technical developments of Purves and Sudderth [33], there has been considerable interest [8, 9, 15, 34] in establishing various limit theorems over products of finitely additive probability spaces that possess suitable continuity properties.

In this paper, we argue that the almost sure convergence in the pointwise ergodic theorem actually *does* hold for finitely additive probability spaces, provided one adopts the *correct* notion of almost sure convergence in this setting.

1.1. *Motivation of main results.* Recall that a sequence of random variables $\{X_n\}$ on a probability space $(\Omega, \mathcal{F}, \mu)$ converges almost surely if:

(A) $\mu(\{\omega \in \Omega : \forall \varepsilon > 0, \exists N \in \mathbb{N}\, \forall i, j \geq N\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) = 1.$

By invoking the continuity of measure (see [29, Lemma 3.1]), this definition is equivalent to each of the following formulations:

(B) $(\forall \varepsilon > 0,\ (\mu(\{\omega \in \Omega : \exists N \in \mathbb{N}\, \forall i, j \geq N\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) = 1));$

(C) $(\forall \varepsilon, \lambda > 0, \exists N \in \mathbb{N}\ (\mu(\{\omega \in \Omega : \forall i, j \geq N\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) > 1 - \lambda));$

(D) $(\forall \varepsilon, \lambda > 0, \exists N \in \mathbb{N}\, \forall k \in \mathbb{N}, (\mu(\{\omega \in \Omega : \forall i, j \in [N; N + k]\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) > 1 - \lambda)).$

Here, and throughout the rest of this article, we write $[a; b] := [a, b] \cap \mathbb{Z}$.

However, these equivalences break down if we assume that μ is only finitely additive. One can readily verify that in this case,

$$\text{(A)} \Rightarrow \text{(B)}, \quad \text{(C)} \Rightarrow \text{(B)}, \quad \text{and} \quad \text{(C)} \Rightarrow \text{(D)},$$

but no other implication holds in general, as explicit counterexamples can be constructed for each of the remaining directions.

We have already noted that the pointwise ergodic theorem does not hold when almost sure convergence is interpreted as property (A). We can strengthen this result by showing that the theorem also fails under property (B) (and, hence, under property (C)).

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Example 1.1. Recall that a set $A \subseteq \mathbb{Z} =: \Omega$ is said to have *asymptotic density* if the limit

$$d(A) := \lim_{N\to\infty} \underbrace{\frac{|A \cap [-N; N]|}{2N+1}}_{=:d_N(A)}$$

exists. It is straightforward to verify that d defines a finitely additive probability measure on the algebra of all subsets of $\mathbb{Z}$ for which the above limit exists. Using a Banach limit, this measure can be extended to a finitely additive probability measure μ on the entire power set $\mathcal{P}(\mathbb{Z})$ (cf. [1]).

Fix such a measure μ. Let $f := I_A$ denote the indicator function of a set $A \subseteq \mathbb{Z}^+$ that does not possess an asymptotic density and let $\tau(\omega) := \omega + 1$ denote the shift map, which can easily be shown to be a measurable, invertible, measure-preserving transformation of $\mathbb{Z}$. Since A does not have an asymptotic density, we can take $\varepsilon > 0$ such that for all N, there exists $i, j \geq N$ satisfying

$$|d_i(A) - d_j(A)| > \frac{3\varepsilon}{2}.$$

Now, for $n \in \mathbb{N}$ and $\omega \in \mathbb{Z}$ with $n \geq -\omega$, we have

$$\begin{aligned} A_n f(\omega) &= \frac{1}{n}\sum_{i=1}^{n} f(\tau^i(\omega)) = \frac{|A \cap [\omega+1; \omega+n]|}{n} \\ &= \frac{2(n+\omega)+1}{n}\frac{|A \cap [-(\omega+n); \omega+n]|}{2(n+\omega)+1} - \frac{|A \cap [1; \omega]|}{n}. \end{aligned}$$

Thus, if $\omega \leq 0$, we have

$$A_n f(\omega) = 2d_{n+\omega}(A) + \frac{2\omega+1}{n} d_{n+\omega}(A)$$

and if $\omega > 0$, we have

$$A_n f(\omega) = 2d_{n+\omega}(A) + \frac{2\omega+1}{n}(d_{n+\omega}(A) - d_\omega(A)).$$

In both cases, since $d_k(A) \in [0, 1]$ for all $k \in \mathbb{N}$, it follows that

$$|A_n f(\omega) - 2d_{n+\omega}(A)| \leq \frac{2\omega+1}{n}.$$

Fix $\omega \in \mathbb{N}$. Suppose there exists $N \in \mathbb{N}$ such that for all $i, j \geq N$,

$$|A_i f(\omega) - A_j f(\omega)| \leq \varepsilon.$$

If $i, j \geq \max\{N, (2\omega+1)/\varepsilon\}$, the reverse triangle inequality gives

$$\begin{aligned} &|2d_{j+\omega}(A) - 2d_{i+\omega}(A)| - |A_i f(\omega) - 2d_{i+\omega}(A) + 2d_{j+\omega}(A) - A_j f(\omega)| \\ &\quad \leq |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon. \end{aligned}$$

Hence,

$$2|d_{j+\omega}(A) - d_{i+\omega}(A)| \leq \varepsilon + \frac{2\omega+1}{i} + \frac{2\omega+1}{j} \leq 3\varepsilon,$$

which contradicts the way we picked ε.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Therefore,

$$\{\omega \in \Omega : \exists N \in \mathbb{N}\, \forall i, j \geq N,\ |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon\} = \emptyset,$$

and so property (B) is not satisfied for the sequence $\{A_n f\}$.

Thus, the only formulation of almost sure convergence we are left with is property (D), which motivates the following definition.

Definition 1.1. Let Ω be a set, $\mathcal{F}$ be an algebra of subsets of Ω, and μ a finitely additive probability measure on $\mathcal{F}$. A sequence $\{X_n\}$ of real valued functions on Ω is said to be *finitely almost surely convergent* if

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})) \tag{1.1}$$

and

$$(\forall \varepsilon, \lambda > 0,\ \exists N \in \mathbb{N}, \forall k \in \mathbb{N},\ (\mu(\{\omega \in \Omega : \forall i, j \in [N; N+k]\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) > 1 - \lambda)).$$

Remark 1.2. If $\mathcal{F}$ is a σ-algebra and $\{X_n\}$ are random variables (which is the setting considered in [35] and also the context of Example 1.1), the measurability condition in (1.1) can be omitted. In general, this condition corresponds to the notion of *weak Borel measurability* introduced in [25], which represents the minimal assumption ensuring that the definition remains meaningful when $\mathcal{F}$ is merely an algebra.

1.2. *Main results.* We show that finite almost sure convergence furnishes the *appropriate* analog of almost sure convergence for extending the pointwise ergodic theorem to finitely additive probability spaces. Our main result is the following theorem.

THEOREM 1.3. *Let $(\Omega, \mathcal{F}, \mu)$ be a finitely additive probability space, $\tau : \Omega \to \Omega$ a measure-preserving automorphism, and $f \in L^1 := L^1(\Omega, \mathcal{F}, \mu)$. If $\{A_n f\}$ satisfies*

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon\} \in \mathcal{F})),$$

then $\{A_n f\}$ is finitely almost surely convergent.

The proof of Theorem 1.3 is quite subtle. First, we observe that the quantitative results on the fluctuations of ergodic averages established in [12] remain valid in the finitely additive setting. This is achieved by verifying that a suitable version of the Calderón transference principle [7], which we present in Theorem 3.1, holds for finitely additive spaces. We then combine these results with the insights developed in [29] on the relationships between various quantitative formulations of almost sure convergence to obtain the desired conclusion.

Furthermore, we obtain several quantitative strengthenings of Theorem 1.3, including a bound on the *uniform metastability* of the ergodic averages (cf. [3, 29, 38]).

THEOREM 1.4. *Let $(\Omega, \mathcal{F}, \mu)$ be a finitely additive probability space, $\tau : \Omega \to \Omega$ a measure-preserving automorphism, and $f \in L^1(\Omega, \mathcal{F}, \mu)$. Suppose that the sequence $\{A_n f\}$ satisfies*

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Then,

$$(\forall \varepsilon, \lambda \in (0,1), \forall g : \mathbb{N} \to \mathbb{N},\ \exists n \leq \tilde{g}^{(\lceil \delta(\lambda,\varepsilon) \rceil)}(0)$$
$$(\mu(\forall i, j \in [n, n+g(n)]\ (|A_i f - A_j f| \leq \varepsilon)) > 1 - \lambda)),$$

for some numerical constant c, where

$$\delta(\lambda, \varepsilon) := \frac{c}{\lambda}\left(\frac{\|f\|_1}{\varepsilon\lambda}\right)^2 \quad \textit{and} \quad \tilde{g}(n) := n + g(n),$$

with $\tilde{g}^{(i)}$ denoting the ith iterate of $\tilde{g}$.

1.3. *Connections with mathematical logic and the proof mining program.* This work can be seen as a contribution to Kohlenbach's *proof mining* program [16]. Proof mining employs tools from mathematical logic to analyze proofs in mainstream mathematics, yielding significant quantitative and qualitative improvements to known results.

There has been increasing interest in extending the proof mining program to probability theory. Recent developments include results on the Laws of Large Numbers [22, 23], studies on the asymptotic behavior of stochastic processes [26, 29], and work on stochastic optimization [27, 30, 31].

A key feature of analyzing proofs within the framework of proof mining is that one is often forced to strip a proof to its essential structure, which opens the door to generalizations. The results in this paper also benefit from this aspect of proof mining. Studying convergence results in probability theory within this framework, particularly in light of the logical system introduced in [25] (see also [24]), reveals that σ-additivity is typically used only sparingly in many quantitative analyses. Consequently, many arguments can be lifted to the finitely additive setting. This was already observed in the investigation of quantitative aspects of the ergodic theorem in [29] and motivated the belief that an appropriate generalization of the pointwise ergodic theorem must exist for finitely additive probability spaces, which we present in this article.

1.4. *Organization of the rest of the paper.* The structure of the paper is as follows. In §2.1, we recall the theory of integration for finitely additive measures. Section 2.2 reviews the relevant notions and results concerning quantitative formulations of almost sure convergence and their connection to the pointwise ergodic theorem, as presented in [12, 29]. The proof of our main results are presented in §3.

2. *Preliminary definitions and lemmas*

Throughout this section, fix a set Ω, an algebra of subsets $\mathcal{F}$, and a finitely additive probability measure μ.

2.1. *Integration on finitely additive spaces.* We review the theory of integration on finitely additive probability spaces and recall the properties we need in this paper. Our exposition closely follows the seminal monograph by K. Bhaskara Rao and M. Bhaskara Rao [5].

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

The first step is to define the integral of *simple functions*.

Definition 2.1. A real-valued function f on Ω is said to be *simple* if there exist sets $\{A_i\}_{i=1}^n \subseteq \mathcal{F}$ forming a partition of Ω and real numbers $a_i \in \mathbb{R}$ such that for all $\omega \in \Omega$,

$$f(\omega) = \sum_{i=1}^{n} a_i I_{A_i}(\omega).$$

For such a function f, the integral with respect to μ is defined by

$$\mu(f) := \sum_{i=1}^{n} a_i \mu(A_i).$$

It is straightforward to verify that $\mu(f)$ is independent of the particular representation of f. Moreover, linear combinations of simple functions are simple, and the absolute value of a simple function is simple (see [5, Propositions 4.4.2 and 4.4.4]).

As in the σ-additive case, a general *integrable* function $f : \Omega \to \mathbb{R}$ will be defined as the limit of simple functions. In the σ-additive case, taking this limit to be almost surely or in probability leads to equivalent definitions. In the finitely additive case, we adopt convergence in probability.

Definition 2.2. (Cf. [5, Definition 4.3.1]) A sequence $\{f_n\}$ of real-valued functions on Ω is said to *converge hazily* to a function f if

$$\Big(\forall \varepsilon > 0,\ \Big(\lim_{n\to\infty} \mu^*(\{\omega \in \Omega : |f_n(\omega) - f(\omega)| > \varepsilon\}) = 0\Big)\Big),$$

where $\mu^* : \mathcal{P}(\Omega) \to [0, 1]$ denotes the outer measure of μ, defined by

$$\mu^*(A) := \inf\{\mu(B) : A \subseteq B,\ B \in \mathcal{F}\}.$$

We can now define the integral for general functions.

Definition 2.3. (Cf. [5, Definition 4.4.11]) A real-valued function f on Ω is said to be *integrable* if there exists a sequence $\{f_n\}$ of simple functions such that $\{f_n\}$ converges hazily to f and

$$\lim_{n,m\to\infty} \mu(|f_n - f_m|) = 0.$$

Such a sequence $\{f_n\}$ is called a *determining sequence* for f. The integral of f is then defined by

$$\mu(f) := \lim_{n\to\infty} \mu(f_n).$$

Furthermore, we define

$$L_1(\Omega, \mathcal{F}, \mu) := \{f : \Omega \to \mathbb{R} \mid |f| \text{ is integrable}\}$$

and write $\|f\|_1$ for $\mu(|f|)$.

The fact that the above definition of the integral is well defined (that is, independent of the choice of determining sequence and finite) follows from [5, Proposition 4.4.10].

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Furthermore, the finitely additive integral satisfies the expected fundamental structural properties (cf. [5, Proposition 4.4.13]). In particular, linear combinations of integrable functions are themselves integrable, the integral defines a linear operator on the space of integrable functions, and it is monotone: if f, g are integrable and $f(\omega) \leq g(\omega)$ for all $\omega \in \Omega$, then $\mu(f) \leq \mu(g)$.

The final result we require concerning integration on finitely additive probability spaces involves the composition of integrable functions with measure-preserving automorphisms. Specifically, we consider an invertible map $\tau\colon \Omega \to \Omega$ satisfying $\tau^{-1}(A) \in \mathcal{F}$ and $\mu(\tau^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{F}$.

PROPOSITION 2.1. *If $f : \Omega \to \mathbb{R}$ is integrable and $\tau : \Omega \to \Omega$ a measure-preserving automorphism, then the function $\tilde{f} : \Omega \to \mathbb{R}$ defined by*

$$\tilde{f}(\omega) := f(\tau(\omega))$$

is integrable and $\mu(\tilde{f}) = \mu(f)$.

Proof. The result is immediate if f is simple. Let $\{f_n\}$ be a determining sequence for f and define $\{\tilde{f}_n\}$ by

$$\tilde{f}_n(\omega) := f_n(\tau(\omega)).$$

Since $\{f_n\}$ and $\{\tilde{f}_n\}$ are sequences of simple functions, and τ is measure preserving, we have

$$\mu(f_n) = \mu(\tilde{f}_n) \quad \text{for all } n,$$

and, hence, if $\{\tilde{f}_n\}$ is a determining sequence for $\tilde{f}$, we have

$$\mu(\tilde{f}) = \lim_{n\to\infty} \mu(\tilde{f}_n) = \lim_{n\to\infty} \mu(f_n) = \mu(f).$$

Thus, it remains to show that $\{\tilde{f}_n\}$ is indeed a determining sequence for $\tilde{f}$.

We first note the following.

CLAIM. *If τ is a measure-preserving automorphism, then for every $E \subseteq \Omega$,*

$$\mu^*(\tau^{-1}(E)) = \mu^*(E).$$

Proof of claim. For any $B \in \mathcal{F}$ with $E \subseteq B$, we have $\tau^{-1}(E) \subseteq \tau^{-1}(B)$ and $\mu(\tau^{-1}(B)) = \mu(B)$. Taking the infimum over all such B yields

$$\mu^*(\tau^{-1}(E)) \leq \inf\{\mu(\tau^{-1}(B)) : E \subseteq B,\ B \in \mathcal{F}\} = \mu^*(E).$$

The reverse inequality follows by applying the same argument to τ^{-1}. □

Now, for all $\varepsilon > 0$ and $n \in \mathbb{N}$,

$$\begin{aligned}\mu^*(\{\omega : |\tilde{f}_n(\omega) - \tilde{f}(\omega)| > \varepsilon\}) &= \mu^*(\tau^{-1}(\{\omega : |f_n(\omega) - f(\omega)| > \varepsilon\})) \\ &= \mu^*(\{\omega : |f_n(\omega) - f(\omega)| > \varepsilon\}).\end{aligned}$$

Thus, since $\{f_n\}$ converges hazily to f, it follows that $\{\tilde{f}_n\}$ converges hazily to $\tilde{f}$.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Finally, for all $n, m \in \mathbb{N}$,

$$\mu(|\tilde{f}_n - \tilde{f}_m|) = \mu(|f_n - f_m|),$$

and, hence,

$$\lim_{n,m\to\infty} \mu(|\tilde{f}_n - \tilde{f}_m|) = \lim_{n,m\to\infty} \mu(|f_n - f_m|) = 0.$$

This shows that $\{\tilde{f}_n\}$ is indeed a determining sequence for $\tilde{f}$, completing the proof. □

2.2. *Quantitative notions of stochastic convergence.* Quantitative aspects of the pointwise ergodic theorem have been studied extensively in the σ-additive setting. As we shall see in the following section, many of these quantitative results carry over to the finitely additive setting since one can establish a version of the Calderón transference principle [7] for finitely additive spaces. In this section, we review the relevant quantitative notions of almost sure convergence that are pertinent to our work.

The most natural quantitative interpretation of the convergence of a sequence of real numbers $\{x_n\}$ is known as a *rate of convergence*, which is a function $\phi : (0, 1) \to \mathbb{R}^+$ satisfying

$$(\forall \varepsilon \in (0, 1), \forall i, j \geq \phi(\varepsilon),\ (|x_i - x_j| \leq \varepsilon)).$$

There are many convergence results in analysis where one cannot obtain a uniform rate of convergence for the sequences satisfying the premises of these results; a notable example is the monotone convergence theorem, where no uniform rate of convergence exists for all monotone sequences in [0, 1], since such sequences can converge arbitrarily slowly. This remains true even for the class of computable sequences of rational numbers [36].

In situations where a uniform rate of convergence cannot be obtained, one can often obtain uniform bounds on an equivalent reformulation, known as *metastable convergence*:

$$(\forall \varepsilon > 0, \forall g : \mathbb{N} \to \mathbb{N},\ \exists n\, \forall i, j \in [n; n + g(n)]\ (|x_i - x_j| \leq \varepsilon)). \tag{2.1}$$

It is easy to see that (2.1) is equivalent to the convergence of $\{x_n\}$. Indeed, if $\{x_n\}$ does not converge, then there exists some $\varepsilon > 0$ such that

$$(\forall n,\ \exists i, j \geq n\ (|x_i - x_j| > \varepsilon)),$$

so we can take a function $g : \mathbb{N} \to \mathbb{N}$ bounding i, j depending on n, that is,

$$(\forall n,\ \exists i, j \in [n; n + g(n)]\ (|x_i - x| > \varepsilon)),$$

which contradicts (2.1). The converse implication is immediate.

A *rate of metastability* is a functional Φ bounding n in (2.1), that is,

$$(\forall \varepsilon \in (0, 1), \forall g : \mathbb{N} \to \mathbb{N},\ \exists n \leq \Phi(\varepsilon, g), \forall i, j \in [n; n + g(n)],\ (|x_i - x_j| \leq \varepsilon)).$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Remark 2.2. One can easily verify that if $\{x_n\} \subseteq [0, 1]$ is monotone, then

$$\Phi(\varepsilon, g) := \tilde{g}^{(\lceil 1/\varepsilon \rceil)}(0)$$

is a rate of metastability for $\{x_n\}$, where $\tilde{g}(n) := n + g(n)$ and $\tilde{g}^{(i)}$ denotes the ith iteration of $\tilde{g}$.

Remark 2.3. The term *metastability* originates from Tao [38], where he discussed metastable versions of the monotone convergence principle and other 'infinitary' statements. From the perspective of logic, metastability (2.1) corresponds to the Herbrand normal form of convergence, and a rate of metastability provides a solution to the so-called 'no-counterexample interpretation' of the convergence statement [20, 21]. Furthermore, there are theorems from mathematical logic guaranteeing the extractibility of uniform rates of metastability for large classes of proofs. Such uniformity was required in the proof of the convergence of the ergodic averages given in [39], where a uniform quantitative version of the dominated convergence theorem was required. Indeed, extractions of rates of metastability by proof-theoretic methods are common results in proof mining (e.g. [17, 18, 28, 32]). Furthermore, as we shall see later, metastability is intimately related to the notion of fluctuations of sequences, a concept that has attracted considerable attention in ergodic theory and probability theory (see, for example, [4, 13, 14]).

In the stochastic setting, the analog of a rate of convergence is a function $\phi : (0, 1) \times (0, 1) \to \mathbb{R}^+$ satisfying

$$\begin{aligned}&(\forall \varepsilon,\ \lambda \in (0, 1)\\ &\quad \forall k \in \mathbb{N},\ (\mu(\forall i, j \in [\phi(\lambda, \varepsilon); \phi(\lambda, \varepsilon) + k](|X_i - X_j| \leq \varepsilon)) > 1 - \lambda)),\end{aligned}$$

where $\{X_n\}$ is a sequence of real-valued functions on Ω satisfying

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon \in (0, 1),\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

Such a function is sometimes called a *rate of almost sure convergence*. In [3, Theorem 5.1], it is demonstrated that no computable such rate can exist for the convergence of the (σ-additive) pointwise ergodic theorem. Motivated by the deterministic case, the notion of *uniform metastable convergence* (there is also a notion of pointwise metastable convergence, cf. Remark 3.5) was introduced in [3].

Definition 2.4. Let $\{X_n\}$ be a sequence of real-valued functions on Ω satisfying

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

We say that $\{X_n\}$ is *uniformly metastable* if

$$\begin{aligned}&(\forall \varepsilon, \lambda > 0, \forall g : \mathbb{N} \to \mathbb{N},\ \exists n \in \mathbb{N}\\ &\quad (\mu(\forall i, j \in [n, n + g(n)]\ (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda)).\end{aligned}$$

Remark 2.4. In [2], uniform metastability was introduced under the name of λ-*uniform* ε-*metastable convergence* and given in a slightly different form. The formulation we present here is from [29] (see [29, discussion after Definition 4.10]).

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

As in the deterministic case, uniform metastability is equivalent to finite almost sure convergence.

PROPOSITION 2.5. *Let $\{X_n\}$ be a sequence of real-valued functions on Ω. Then, $\{X_n\}$ is uniformly metastable if and only if it is finitely almost surely convergent.*

Proof. Suppose first that $\{X_n\}$ is finitely almost surely convergent. Then, for every $\varepsilon, \lambda > 0$, there exists $N \in \mathbb{N}$ such that

$$(\forall k \in \mathbb{N},\ (\mu(\forall i, j \in [N; N+k]\ (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda)).$$

Let $g : \mathbb{N} \to \mathbb{N}$ be arbitrary. Taking $k = g(N)$ in the above yields

$$\mu(\forall i, j \in [N; N + g(N)]\ |X_i - X_j| \leq \varepsilon) > 1 - \lambda,$$

which shows that $\{X_n\}$ is uniformly metastable.

Conversely, suppose $\{X_n\}$ is not finitely almost surely convergent. Then, there exist $\varepsilon, \lambda > 0$ such that for all $N \in \mathbb{N}$, there exists $k \in \mathbb{N}$ with

$$\mu(\forall i, j \in [N; N+k]\ |X_i - X_j| \leq \varepsilon) \leq 1 - \lambda.$$

Define $g(N)$ to be the least such k. Then, for this choice of g, we have, for all $N \in \mathbb{N}$,

$$\mu(\forall i, j \in [N; N + g(N)]\ |X_i - X_j| \leq \varepsilon) \leq 1 - \lambda,$$

which shows that $\{X_n\}$ fails to be uniformly metastable. □

To prove Theorem 1.3, we shall demonstrate that the ergodic averages are uniformly metastable. Furthermore, we shall obtain a quantitative strengthening of this result through the construction of explicit so-called *rates of uniform metastability*.

Definition 2.5. Let $\{X_n\}$ be a sequence of real-valued functions on Ω satisfying

$$(\forall i, j \in \mathbb{N},\ \text{for all } \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

A functional $\Phi : (0,1) \times (0,1) \times \mathbb{N}^{\mathbb{N}} \to \mathbb{R}^+$ is called a *rate of uniform metastability* for $\{X_n\}$ if

$$(\forall \varepsilon, \lambda \in (0,1), \forall g : \mathbb{N} \to \mathbb{N},\ \exists n \leq \Phi(\lambda, \varepsilon, g)\ (\mu(\forall i, j \in [n, n + g(n)]\ (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda)).$$

In [3], rates of uniform metastability were computed for the pointwise ergodic theorem for L_2 functions, with the treatment of the full theorem later given by the author and Powell in [29]. Moreover, in [29], a stronger quantitative result is established by introducing the concept of *uniform learnability*.

Definition 2.6. Let $\{X_n\}$ be a sequence of real-valued functions on Ω satisfying

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

A function $\phi : (0,1) \times (0,1) \to \mathbb{R}^+$ is called a *learnable rate of uniform convergence* if, for any $\varepsilon, \lambda \in (0,1)$, and sequences $a_0 < b_0 \leq a_1 < b_1 \leq \cdots$, there exists $n \leq \phi(\lambda, \varepsilon)$ such that

$$\mu(\exists i, j \in [a_n, b_n]\ |X_i - X_j| \geq \varepsilon) \leq \lambda.$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

It turns out that a learnable rate of uniform convergence corresponds to a rate of metastability of a particularly nice form.

PROPOSITION 2.6. (Cf. [29]) *Let* $\{X_n\}$ *be a sequence of real-valued functions on* Ω *satisfying*

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

Then,

$$\Phi(\lambda, \varepsilon, g) = \tilde{g}^{(\lceil \phi(\lambda,\varepsilon)\rceil)}(0) \quad \textit{with } \tilde{g}(n) := n + g(n)$$

is a rate of uniform metastability for $\{X_n\}$ *if and only if* $\lceil \phi(\lambda, \varepsilon)\rceil$ *is a learnable rate of uniform convergence.*

This result is stated in the context of σ-additive spaces, immediately following [29, Definition 4.15], as an easy consequence of an abstract result given in [29, Lemma 3.5]. The finitely additive case follows similarly as a direct consequence of the same lemma.

We prove a quantitatively stronger result than that stated in Theorem 1.3 by computing a learnable rate of uniform convergence for the ergodic averages. Theorem 1.3 will then follow from Propositions 2.5 and 2.6.

To compute a learnable rate of uniform convergence for the pointwise ergodic theorem, we require a result from [12] concerning the fluctuations of ergodic averages on σ-additive probability spaces. First, we introduce the notion of the number of ε-fluctuations of a sequence of real numbers.

Definition 2.7. Let $\{x_n\}$ be a sequence of real numbers. We define $J_{N,\varepsilon}\{x_n\}$ to be the number of ε-fluctuations that occur in the initial segment $\{x_0, \ldots, x_{N-1}\}$, that is, the maximal $k \in \mathbb{N}$ such that there exist indices

$$i_1 < j_1 \leq i_2 < j_2 \leq \cdots \leq i_k < j_k < N$$

with

$$|x_{i_l} - x_{j_l}| \geq \varepsilon \quad \text{for all } l = 1, \ldots, k.$$

We write

$$J_\varepsilon\{x_n\} = \lim_{N\to\infty} J_{N,\varepsilon}\{x_n\}$$

for the total number of ε-fluctuations that occur in $\{x_n\}$. Note that $J_\varepsilon\{x_n\}$ could be infinite.

Remark 2.7. It is clear that $\{x_n\}$ converges if and only if

$$(\forall \varepsilon > 0, (J_\varepsilon\{x_n\} < \infty)).$$

In [12], Jones *et al* obtained the following quantitative strengthening of the pointwise ergodic theorem.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

THEOREM 2.8. *Let $(\Omega, \mathcal{F}, \mu)$ be a σ-additive probability space and let $\tau : \Omega \to \Omega$ be a measure-preserving automorphism. For any $f \in L_1$ and $k, \varepsilon > 0$, we have*

$$\mu(J_\varepsilon\{A_n f\} \geq k) \leq \frac{C\|f\|_1}{\varepsilon\sqrt{k}}$$

for some numerical constant C.

This result was originally conjectured by Ivanov [13], who, using entirely different methods, was only able to show

$$\mu(J_\varepsilon\{A_n f\} \geq k) \leq C\sqrt{\frac{\log(k)}{k}},$$

where C is a constant depending only on $\|f\|_1/\varepsilon$ (in [29], the authors were unaware of the bound of Jones, Kaufman, Rosenblatt, and Wierdl, and by improving the methods of Ivanov, they were able to give an improvement to his bound, which was still weaker than that in 12]).

By establishing a version of the Calderón transference principle [7] for finitely additive spaces, we shall show that a version of Theorem 2.8 also holds in the finitely additive setting.

It is clear from Theorem 2.8 that the ergodic averages satisfy

$$(\forall \varepsilon > 0,\ (J_\varepsilon\{A_n f\} < \infty \text{ almost surely})),$$

and thus converge almost surely. However, as in the case of convergence, bounded fluctuations can have different, non-equivalent formulations in the finitely additive setting. If $\mathcal{F}$ is a σ-algebra and $\{X_n\}$ is a sequence of random variables on Ω (so that for all k and ε, $J_{k,\varepsilon}\{X_n\}$ is itself a random variable), we have the following formulations of having finitely many fluctuations almost surely:

(J1) $\mu(\{\omega \in \Omega : \forall \varepsilon > 0,\ \exists N \in \mathbb{N}, \forall k \in \mathbb{N}\ J_{k,\varepsilon}\{X_n\}(\omega) < N\}) = 1$;

(J2) $(\forall \varepsilon > 0,\ (\mu(\{\omega \in \Omega : \exists N \in \mathbb{N}\ \forall k \in \mathbb{N}\ (J_{k,\varepsilon}\{X_n\}(\omega) < N)\}) = 1))$;

(J3) $(\forall \varepsilon, \lambda > 0,\ \exists N \in \mathbb{N}\ (\mu(\{\omega \in \Omega : \forall k \in \mathbb{N}\ (J_{k,\varepsilon}\{X_n\}(\omega) < N)\}) > 1 - \lambda)$;

(J4) $(\forall \varepsilon, \lambda > 0,\ \exists N \in \mathbb{N}\ \forall k \in \mathbb{N}\ (\mu(\{\omega \in \Omega : J_{k,\varepsilon}\{X_n\}(\omega) < N\}) > 1 - \lambda)$.

As in the case of convergence, we have

$$\text{(J1)} \Rightarrow \text{(J2)}, \quad \text{(J3)} \Rightarrow \text{(J2)}, \quad \text{and} \quad \text{(J3)} \Rightarrow \text{(J4)}.$$

Furthermore, it is clear that

$$\text{(J2)} \iff \text{(B)},$$

where formulation (B) is the almost sure convergence given by

$$(\forall \varepsilon > 0, (\mu(\{\omega \in \Omega : \exists N \in \mathbb{N}\ \forall i, j \geq N\ (|X_i(\omega) - X_j(\omega)| \leq \varepsilon)\}) = 1)).$$

Therefore, Example 1.1 demonstrates that the pointwise ergodic theorem cannot satisfy formulation (J2) in the finitely additive setting and thus cannot satisfy formulation (J3). However, Theorem 2.8 implies formulation (J3), so it cannot hold in the finitely additive setting. We shall show that formulation (J4) of Theorem 2.8 does hold via the Calderón transference principle [7] for finitely additive spaces.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Theorem 1.3 will follow from our formulation (J4) of Theorem 2.8 by establishing a quantitative connection between bounds on the fluctuations and learnable rates of uniform convergence. We first introduce the notion of a *modulus of finite fluctuations*, as in [29], which quantitatively captures formulation (J4).

Definition 2.8. Let $\{X_n\}$ be a sequence of real-valued functions on Ω satisfying

$$(\forall i,\, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})).$$

A function $\phi : (0,1) \times (0,1) \to \mathbb{R}^+$ is a *modulus of finite fluctuations* for $\{X_n\}$ if

$$(\forall \varepsilon, \lambda \in (0,1), \forall k \in \mathbb{N},\ (\mu(J_{k,\varepsilon}\{X_n\} < \phi(\varepsilon,\lambda)) > 1 - \lambda)).$$

Remark 2.9. This definition is well posed because, if

$$(\forall i,\, j \in \mathbb{N}, \forall \varepsilon > 0,\ (\{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F})),$$

then

$$(\forall k \in \mathbb{N}, \forall \varepsilon, N > 0,\ (\{\omega \in \Omega : J_{k,\varepsilon}\{X_n\}(\omega) < N\} \in \mathcal{F})).$$

In [29], the author and Powell examined how quantitative notions of almost sure convergence, including those introduced here, relate to one another. For example, from a learnable rate of uniform convergence, one can obtain a rate of uniform metastability. Furthermore, the authors abstracted away from almost sure convergence and instead focused on general logical formulas of the same quantifier complexity, yielding more general quantitative results.

At the time of writing [29], the quantitative relationship between moduli of finite fluctuations and learnable rates of uniform convergence in the σ-additive setting was left as an open problem. Recently, Powell communicated to the author a partial solution to this problem by constructing a learnable rate of uniform convergence from a modulus of finite fluctuations. Powell's construction also lifts naturally to the finitely additive setting. Namely, one has the following result.

PROPOSITION 2.10. *Let $\{X_n\}$ be a sequence of real-valued functions on Ω satisfying*

$$(\forall i,\, j \in \mathbb{N}, \forall \varepsilon > 0,\ \{\omega \in \Omega : |X_i(\omega) - X_j(\omega)| \leq \varepsilon\} \in \mathcal{F}).$$

Suppose $\{X_n\}$ has a modulus of finite fluctuations ϕ. Then, $\{X_n\}$ has a learnable rate of uniform convergence given by

$$\psi(\lambda, \varepsilon) := \frac{2\,\phi(\lambda/2, \varepsilon)}{\lambda}.$$

Proof. Fix $\lambda, \varepsilon \in (0,1]$ and consider sequences $a_0 < b_0 \leq a_1 < b_1 \leq \cdots$. For $k \in \mathbb{N}$, define the event

$$B_k := \{\omega \in \Omega : J_{k,\varepsilon}\{X_n\}(\omega) \geq \phi(\lambda/2, \varepsilon)\} \in \mathcal{F}.$$

By the definition of the modulus of finite fluctuations, $\mu(B_k) \leq \lambda/2$ for all $k \in \mathbb{N}$. For $a, b \in \mathbb{N}$, set

$$C(a,b) := \{\omega \in \Omega : \exists i,\, j \in [a,b],\ |X_i(\omega) - X_j(\omega)| > \varepsilon\} \in \mathcal{F}.$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Suppose, for a contradiction, that for all $n \leq \psi(\lambda, \varepsilon)$, we have $\mu(C(a_n, b_n)) > \lambda$. Then, for all $n \leq \psi(\lambda, \varepsilon)$,

$$\mu(C(a_n, b_n) \cap B_k^c) = \mu(C(a_n, b_n)) - \mu(C(a_n, b_n) \cap B_k) > \lambda - \mu(B_k) \geq \frac{\lambda}{2},$$

where we set $k := b_{\phi(\lambda/2,\varepsilon)} + 1$.

Now, observe that

$$\sum_{n \leq \psi(\lambda,\varepsilon)} I_{C(a_n,b_n)\cap B_k^c} < \phi(\lambda/2, \varepsilon),$$

and, thus,

$$\frac{(\psi(\lambda, \varepsilon) + 1)\lambda}{2} \leq \sum_{n \leq \psi(\lambda)} \mu(C(a_n, b_n) \cap B_k^c) = \mu\left(\sum_{n \leq \psi(\lambda)} I_{C(a_n,b_n)\cap B_k^c} \right) < \phi(\lambda/2, \varepsilon),$$

which is a contradiction. Here, the last inequality follows from the monotonicity of the integral. □

Remark 2.11. The optimality of this construction follows from [29, Example 4.18]. The converse relationship (constructing a modulus of finite fluctuations from a learnable rate of uniform convergence) is currently an open problem.

3. *Proof of main results*

The proof of the main result follows by establishing a variation of the Calderón transference principle [7] for finitely additive spaces.

We adopt the same setup for transference as in [19]. Suppose that for each $N \in \mathbb{N}$, we have mappings $\mathcal{O}_N : \mathbb{R}^N \to \mathbb{R}^+$ and $C : \mathbb{R}^+ \to \mathbb{R}^+$ such that, for any $K \in \mathbb{N}$, $f : [1, 2K] \to \mathbb{R}$, and $a > 0$, we have

$$|\{k \in [1; 2K] : \mathcal{O}_K(\{\tilde{A}_n f(k)\}_{n\in[1,K]}) \geq a\}| \leq C(a) \sum_{i=1}^{2K} |f(i)|, \tag{†}$$

where

$$\tilde{A}_n f(k) := \frac{1}{n} \sum_{i=1}^{n} f(k + i \bmod 2K).$$

Our goal is to obtain such an inequality for an arbitrary finitely additive dynamical system.

THEOREM 3.1. *Let $(\Omega, \mathcal{F}, \mu)$ be a finitely additive probability space, $\tau : \Omega \to \Omega$ a measure-preserving automorphism, and $f \in L^1(\Omega, \mathcal{F}, \mu)$. If the averages*

$$A_n f(\omega) := \frac{1}{n} \sum_{i=1}^{n} f(\tau^i \omega)$$

satisfy

$$(\forall K \in \mathbb{N}, \forall a > 0,\ (\{\omega \in \Omega : \mathcal{O}_K(\{A_n f(\omega)\}_{n\in[1,K]}) \geq a\} \in \mathcal{F})), \tag{⋆}$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

then for all $a > 0$ and $K \in \mathbb{N}$,

$$\mu(\mathcal{O}_K(\{A_n f\}_{n\in[1,K]}) \geq a) \leq C(a)\|f\|_1.$$

Proof. Fix $a > 0$ and $K \in \mathbb{N}$. For each $\omega \in \Omega$, define $f_\omega : [1, 2K] \to \mathbb{R}$ by

$$f_\omega(k) := f(\tau^k(\omega)).$$

Then, for $k, n \in [1, K]$,

$$\tilde{A}_n f_\omega(k) = \frac{1}{n}\sum_{i=1}^{n} f_\omega(k+i) = \frac{1}{n}\sum_{i=1}^{n} f(\tau^i(\tau^k(\omega))) = A_n f(\tau^k(\omega)).$$

For $k \in [1, K]$, define

$$\begin{aligned} E_k &:= \{\omega \in \Omega : \mathcal{O}_K(\{A_n f(\tau^k(\omega))\}_{n\in[1,K]}) \geq a\} \\ &= \{\omega \in \Omega : \mathcal{O}_K(\{\tilde{A}_n f_\omega(k)\}_{n\in[1,K]}) \geq a\}. \end{aligned}$$

By $(\star)$, $E_0 \in \mathcal{F}$. Moreover, $\mu(E_k) = \mu(E_0)$ for all $k \in [1, K]$ since τ is measure preserving and invertible, and $E_k = (\tau^k)^{-1}(E_0)$. Hence,

$$\mu(E_0) = \frac{1}{2K}\sum_{k=1}^{2K}\mu(E_k) = \frac{1}{2K}\mu\Big(\sum_{k=1}^{2K} I_{E_k}\Big).$$

For each $\omega \in \Omega$,

$$\begin{aligned} \sum_{k=1}^{2K} I_{E_k}(\omega) = |\{k \in [1, 2K] : \mathcal{O}_K(\{\tilde{A}_n f_\omega(k)\}_{n\in[1,K]}) \geq a\}| &\leq C(a)\sum_{k=1}^{2K}|f_\omega(k)| \\ &= C(a)\sum_{k=1}^{2K}|f(\tau^k(\omega))|. \end{aligned}$$

By the monotonicity of the integral, we conclude

$$\mu(E_0) \leq \frac{C(a)}{2K}\sum_{k=1}^{2K}\mu(|f(\tau^k)|) = C(a)\|f\|_1,$$

and the result follows. □

We can now establish a version of Theorem 2.8 in the finitely additive setting.

THEOREM 3.2. *There exists a constant C such that for any finitely additive probability space $(\Omega, \mathcal{F}, \mu)$, measure-preserving automorphism $\tau : \Omega \to \Omega$, and $f \in L^1(\Omega, \mathcal{F}, \mu)$, if*

$$(\forall i, j \in \mathbb{N}, \forall \varepsilon > 0,\ \{\omega \in \Omega : |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon\} \in \mathcal{F}),$$

then for all $a, \varepsilon > 0$ and $K \in \mathbb{N}$,

$$\mu(J_{K,\varepsilon}\{A_n f\} \geq a) \leq \frac{C\|f\|_1}{\varepsilon\sqrt{a}}.$$

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

Proof. Define $\mathcal{O}_K(\{x_n\}) := \varepsilon\sqrt{J_{K,\varepsilon}\{x_n\}}$. Condition $(\star)$ is satisfied since the set in question is a finite union of finite intersections of sets of the form $\{\omega \in \Omega : |A_i f(\omega) - A_j f(\omega)| \leq \varepsilon\}$.

Applying Theorem 2.8 to the discrete space $([1, 2K], \mathcal{P}([1, 2K]), \mu_K)$ with $\mu_K(A) := |A|/2K$ and $\tau_K(\omega) = \omega + 1 \bmod 2K$ gives condition (†). The result then follows by Theorem 3.1. □

Remark 3.3. As is the case for the Calderón transference principle in the σ-additive setting, our Theorem 3.1 allows us to extend several weak-type (1, 1) inequalities on the ergodic averages (in the context of further measureability assumptions in the case where $\mathcal{F}$ is only assumed to be an algebra). This includes: finitely additive analogs of the maximal ergodic theorem, upcrossing inequalities (see, for example, [**6**, **11**]), and inequalities on the variation and oscillation seminorms (see, for example, [**12**]) to the finitely additive setting.

Proof of Theorems 1.3 and 1.4. Theorem 3.2 implies that

$$\phi(\lambda, \varepsilon) := \left(\frac{C\|f\|_1}{\varepsilon\lambda}\right)^2$$

is a modulus of finite fluctuations (cf. Definition 2.8) for the sequence $\{A_n f\}$. Moreover, by Proposition 2.10, we obtain that

$$\delta(\lambda, \varepsilon) := \frac{8}{\lambda}\left(\frac{C\|f\|_1}{\varepsilon\lambda}\right)^2$$

constitutes a learnable rate of uniform convergence (cf. Definition 2.6) for $\{A_n f\}$. Consequently, Proposition 2.6 yields that

$$\Phi(\lambda, \varepsilon, g) := \tilde{g}^{(\lceil \delta(\lambda,\varepsilon)\rceil)}(0) \quad \text{where } \tilde{g}(n) := n + g(n),$$

serves as a rate of uniform metastability (cf. Definition 2.5) for $\{A_n f\}$. This yields Theorem 1.4. Theorem 1.3 then follows from Proposition 2.5. □

Remark 3.4. In [**29**, Theorem 7.9], the asymptotically sharper bound

$$\Delta(\lambda, \varepsilon) := \left(\frac{cK}{\varepsilon\lambda}\right)^2$$

is established as a learnable rate of uniform convergence for $\{A_n f\}$ in the σ-additive setting, where $c > 0$ is a universal constant and $K := \max\{1, \|f\|_1\}$.

The methods employed in [**29**] crucially rely on countable additivity and therefore do not extend directly to the finitely additive framework. Determining whether comparably sharp learnable rates of uniform convergence can be obtained for the finitely additive version of the pointwise ergodic theorem remains an open problem, which we leave for future investigation.

Remark 3.5. There is an alternative way to obtain Theorem 1.3 without appealing to Proposition 2.10. From a modulus of finite fluctuations, one can derive what is known as a rate of *pointwise* metastability convergence (cf. [**29**, Theorem 4.16]), that is, a functional Φ satisfying

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

$$(\forall \varepsilon, \lambda \in (0,1), \forall g : \mathbb{N} \to \mathbb{N},\ \mu(\exists n \leq \Phi(\lambda, \varepsilon, g) \forall i, j \in [n, n + g(n)]\ (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda).$$

The notion of *pointwise* metastable convergence was introduced in [2], where the motivation was to develop a quantitative version of Tao's metastable dominated convergence theorem (cf. [39, Theorem A.2]). Tao's theorem shows that one can obtain rates of metastability for the convergence of the integrals in the dominated convergence theorem, which depend only on a slightly stronger variant of a rate of *pointwise* metastability for the random variables appearing in its premise.

The authors of [2] observed that if one has a rate of *uniform* metastability for the random variables in the premise of the dominated convergence theorem, then one can easily obtain a rate of metastability for the corresponding integrals. Their main result (cf. [2, Theorem 3.1]) explicitly constructs a rate of uniform metastability from a rate of pointwise metastability (which they term a quantitative version of Egorov's theorem) and uses this quantitative result to strengthen Tao's metastable dominated convergence theorem. However, the construction given in [2] is highly complex and involves a form of recursion on trees known as *bar recursion* (introduced in [37]). As a consequence, the resulting bounds exhibit an enormous blow-up in complexity. For example, as shown in [2, p. 11], even comparatively simple input rates of pointwise metastability can yield tower-exponential bounds.

It was shown in [25] that the results of [2] also hold in finitely additive probability spaces. Thus, one may use the modulus of finite fluctuations obtained in Theorem 3.2 to derive a rate of pointwise metastability, and then appeal to the finitely additive adaptation of [2] given in [25] to obtain Theorem 1.3, albeit at the cost of very weak quantitative bounds on the resulting rate of uniform metastability.

Acknowledgements. We would like to express our gratitude to Nathan Creighton, John Griesmer, Pedro Pinto, Nicholas Pischke, and Thomas Powell for their valuable comments and stimulating discussions during the preparation of this work. We are especially indebted to Thomas Powell for communicating his construction, from which Proposition 2.10 was obtained by a straightforward modification.

References

[1] R. P. Agnew and A. P. Morse. Extensions of linear functionals, with applications to limits, integrals, measures, and densities. *Ann. of Math. (2)* **39**(1) (1938), 20–30.
[2] J. Avigad, E. Dean and J. Rute. A metastable dominated convergence theorem. *J. Log. Anal.* (2012).
[3] J. Avigad, P. Gerhardy and H. Towsner. Local stability of ergodic averages. *Trans. Amer. Math. Soc.* **362**(1) (2010), 261–288.
[4] J. Avigad and J. Rute. Oscillation and the mean ergodic theorem for uniformly convex Banach spaces. *Ergod. Th. & Dynam. Sys.* **35**(4) (2015), 1009–1027.
[5] K. Bhaskara Rao and M. Bhaskara Rao. *Theory of Charges (Pure and Applied Mathematics, 109)*. Elsevier, London, 1983.
[6] E. Bishop. *Foundations of Constructive Analysis*. McGraw-Hill, New York, 1967.
[7] A. P. Calderón. Ergodic theory and translation-invariant operators. *Proc. Natl. Acad. Sci. USA* **59**(2) (1968), 349–353.
[8] R. Chen. Some finitely additive versions of the strong law of large numbers. *Israel J. Math.* **24**(3) (1976), 244–259.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press

[9] R. Chen. On almost sure convergence in a finitely additive setting. *Z. Wahrsch. Verw. Geb.* **37**(4) (1977), 341–356.
[10] L. E. Dubins and L. J. Savage. *How to Gamble If You Must: Inequalities for Stochastic Processes*. McGraw-Hill, New York, 1965.
[11] V. V. Ivanov. Oscillations of averages in the ergodic theorem. *Dokl. Akad. Nauk* **347** (1996), 736–738.
[12] R. Jones, R. Kaufman, J. Rosenblatt and M. Wierdl. Oscillation in ergodic theory. *Ergod. Th. & Dynam. Sys.* **18**(4) (1998), 889–935.
[13] A. Kachurovskii. The rate of convergence in ergodic theorems. *Russian Math. Surveys* **51**(4) (1996), 653–703.
[14] S. Kalikow and B. Weiss. Fluctuations of ergodic averages. *Illinois J. Math.* **43**(3) (1999), 480–488.
[15] R. L. Karandikar. A general principle for limit theorems in finitely additive probability. *Trans. Amer. Math. Soc.* **273**(2) (1982), 541–550.
[16] Kohlenbach, U. *Applied Proof Theory: Proof Interpretations and Their Use in Mathematics (Springer Monographs in Mathematics)*. Springer, Berlin, 2008.
[17] U. Kohlenbach and L. Leuştean. Effective metastability of Halpern iterates in CAT(0) spaces. *Adv. Math.* **231**(5) (2012), 2526–2556.
[18] Kohlenbach, U. and Sipoş, A. The finitary content of sunny nonexpansive retractions. *Commun. Contemp. Math.* **23**(1) (2021), 1950093.
[19] D. Kosz. Sharp constants in inequalities admitting the Calderón transference principle. *Ergod. Th. & Dynam. Sys.* **44**(6) (2024), 1597–1608.
[20] G. Kreisel. On the interpretation of non-finitist proofs, part I. *J. Symbolic Logic* **16** (1951), 241–267.
[21] G. Kreisel. On the interpretation of non-finitist proofs, part II: interpretation of number theory. *J. Symbolic Logic* **17** (1952), 43–58.
[22] M. Neri. A finitary Kronecker's lemma and large deviations in the strong law of large numbers on Banach spaces. *Ann. Pure Appl. Logic*, **176**(6) (2025), 103569.
[23] M. Neri. Quantitative strong laws of large numbers. *Electron. J. Probab.* **30** (2025), 1–22.
[24] M. Neri, P. Oliva and N. Pischke A systematic way of analysing proofs in probability theory. *Preprint*, 2025, arXiv:2604.08078.
[25] M. Neri and N. Pischke. Proof mining and probability theory. *Forum Math. Sigma*, **13** (2025), e187.
[26] M. Neri, N. Pischke and T. Powell. Generalized learnability of stochastic principles. *Proc. Computability in Europe (CiE 2025) (Lecture Notes in Computer Science, 15764)*. Ed. A. Beckmann, I. Oitavem and F. Manea. Springer, Cham, 2025.
[27] M. Neri, N. Pischke and T. Powell. On the asymptotic behaviour of stochastic processes, with applications to supermartingale convergence, Dvoretzky's approximation theorem, and stochastic quasi-Fejér monotonicity. *Preprint*, 2025, arXiv:2504.12922.
[28] M. Neri and T. Powell. A computational study of a class of recursive inequalities. *J. Logic Anal.* **15**(3) (2023), 1–48.
[29] M. Neri and T. Powell. On quantitative convergence for stochastic processes: crossings, fluctuations and martingales. *Trans. Amer. Math. Soc. B* **12** (2025), 974–1019.
[30] M. Neri and T. Powell. A quantitative Robbins-Siegmund theorem. *Ann. Appl. Probab.* **36**(1) (2026), 636–651.
[31] Pischke, N. and Powell, T. Asymptotic regularity of a generalised stochastic Halpern scheme. *J. Optim. Theory Appl.*, to appear.
[32] T. Powell. A finitization of Littlewood's Tauberian theorem and an application in Tauberian remainder theory. *Ann. Pure Appl. Logic*, **174**(4) (2023), 103231.
[33] R. A. Purves and W. D. Sudderth. Some finitely additive probability. *Ann. Probab.* **4** (1976), 259–276.
[34] S. Ramakrishnan. Central limit theorems in a finitely additive setting. *Illinois J. Math.* **28**(1) (1984), 139–161.
[35] S. Ramakrishnan. A finitely additive generalization of Birkhoff's ergodic theorem. *Proc. Amer. Math. Soc.* **96**(2) (1986), 299–305.
[36] E. Specker. Nicht Konstruktiv Beweisbare Sätze der Analysis. *J. Symbolic Logic* **14** (1949), 145–158.
[37] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. *Recursive Function Theory (Proceedings of Symposia in Pure Mathematics, 5)*. Ed. F. D. E. Dekker. American Mathematical Society, Providence, RI, 1962, pp. 1–27.
[38] T. Tao. Soft analysis, hard analysis, and the finite convergence principle. *Structure and Randomness: Pages from Year One of a Mathematical Blog*. Ed. T. Tao. American Mathematical Society, Providence, RI, 2008.
[39] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergod. Th. & Dynam. Sys.* **28** (2008), 657–688.

https://doi.org/10.1017/etds.2026.10322 Published online by Cambridge University Press