



Article

Random Finite Sum Sets and Product Sets in Subsets of the Natural Numbers

Sukrit Chakraborty^{1,*†}, Sourav Kanti Patra^{2†}

¹Department of Mathematics, Achhruram Memorial College, Jhalda, Purulia 723202, West Bengal, India

²Department of Mathematics, Kishori Sinha Mahila College, Q976+WXP, Aurangabad, Bihar 824101, India

*Corresponding author: Sukrit Chakraborty, sukritpapai@gmail.com

†The authors contributed equally to this work.

Abstract

We investigate the occurrence of additive and multiplicative structures in random subsets of the natural numbers. Specifically, for a Bernoulli random subset of \mathbb{N} where each integer is included independently with probability $p \in (0, 1)$, we prove that almost surely such a set contains finite sumsets (FS-sets) and finite product sets (FP-sets) of every finite length. We also establish that almost surely the random set contains configurations of the form $\{x, y, x+y, xy\}$. This result is presented as a probabilistic existence statement for additive–multiplicative configurations in Bernoulli random subsets of \mathbb{N} . The probabilistic setting considered here is fundamentally different from the partition-regular framework associated with the classical Hindman conjecture. Finally, we discuss a probabilistic limit theorem for sums indexed by structured sequences.

Keywords

Binomial model, Finite sumsets, Finite product sets, Random subsets, Probabilistic combinatorics, Central limit theorem

Mathematics Subject Classification (2020): Primary 05D10, 60C05; Secondary 11B75, 60F05

Article History

Received: 28 January 2026

Revised: 8 June 2026

Accepted: 12 June 2026

Available Online: 1 July 2026

Copyright

© 2026 by the authors. This article is published by the Cultech Publishing Sdn. Bhd. under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>

1. Introduction

Ramsey theory studies the emergence of highly structured configurations inside large combinatorial systems. A natural question is whether similar structures persist in probabilistic settings. Motivated by random analogues of classical results such as van der Waerden's theorem and Szemerédi's theorem, we investigate finite sumsets and finite product sets inside Bernoulli random subsets of \mathbb{N} . Our results show that, with probability one, such random subsets contain finite sumsets and finite product sets of every finite length. We also investigate related additive-multiplicative configurations and certain probabilistic limit phenomena associated with structured sequences.

The study of additive and multiplicative structures in subsets of natural numbers has a long and rich history, ranging from classical results in combinatorics to modern developments in probabilistic and additive number theory [1-4]. A central theme is understanding how structured configurations, such as arithmetic progressions, finite sumsets, and product sets, appear either deterministically or in a random setting [5,6].

Hindman's theorem [7-9] is a cornerstone of additive Ramsey theory. It asserts that for any finite coloring of the natural numbers, there exists an infinite sequence whose finite sumset is monochromatic. This profound result guarantees the existence of highly structured sets under arbitrary colorings, highlighting the inevitability of additive combinatorial structure. Extensions and related results were further developed by Bergelson and Hindman [10,11], exploring partition-regular and polynomial configurations. Analogous multiplicative phenomena have been considered in the context of product sets, although the literature is less extensive.

In probabilistic combinatorics, random subsets of \mathbb{N} provide a natural framework to study the typical presence of additive and multiplicative structures. Classical works, such as those by Erdős and Rényi [12], established that almost every random subset contains various prescribed finite patterns, including arithmetic progressions. More recent probabilistic and sparse random set results were developed by Conlon and Gowers [13] and by Rödl and Ruciński [14], with further related works, although explicit formulations concerning finite sumsets (FS-sets) and finite product sets (FP-sets) in random subsets remain largely absent from the literature.

In this work, we establish the following results, rigorously proving that almost every random subset of the natural numbers contains finite FS-sets and FP-sets of every finite length. These results establish the almost sure existence of finite additive and multiplicative configurations inside Bernoulli random subsets of \mathbb{N} :

Theorem A (Random Finite Sumsets): In a random subset of \mathbb{N} formed by including each integer independently with probability $p \in (0,1)$, almost surely for every $L \geq 1$ there exists a finite sumset of length L contained in the set (see Theorem 1 for precise statement and proof).

Theorem B (Random Finite Product Sets): Similarly, in the same random model, almost surely for every $L \geq 1$ there exists a finite product set of length L contained in the set (see Theorem 2 for precise statement and proof).

The main contributions of this paper are probabilistic existence results for finite sumsets (FS-sets) and finite product sets (FP-sets) in Bernoulli random subsets of \mathbb{N} . The proofs rely on a standard but effective probabilistic method: constructing infinitely many pairwise disjoint realizations of a fixed combinatorial pattern and applying the second Borel–Cantelli lemma.

While the underlying probabilistic framework is classical, explicit formulations for FS-set and FP-set configurations in Bernoulli random subsets do not appear to be systematically recorded in the literature. Our goal is therefore to present clean and self-contained probabilistic statements for these configurations together with transparent proofs that may be useful for future investigations.

We also include several additional observations concerning additive–multiplicative patterns and probabilistic limit behavior. These results should be viewed as complementary to the main FS/FP existence results rather than as fundamentally new probabilistic methods. One motivating question considered in this paper is to determine whether random sub-sets of the natural numbers, generated under a Bernoulli model, almost surely contain rich additive and multiplicative structures such as finite sumsets (FS-sets) and finite product sets (FP-sets) of arbitrary finite length.

While deterministic Ramsey-theoretic results guarantee the existence of such structures under coloring assumptions, it is not immediately clear whether analogous configurations persist in a probabilistic setting. This motivates a systematic investigation of the typical behavior of additive and multiplicative configurations in random subsets of \mathbb{N} .

Research Gap

Although many probabilistic results are known for arithmetic and combinatorial patterns in random sets, explicit formulations concerning finite sumsets (FS-sets) and finite product sets (FP-sets) in Bernoulli random subsets of \mathbb{N} do not appear to be systematically developed in the literature. In particular, there seems to be no direct probabilistic statement guaranteeing the almost sure existence of FS-sets and FP-sets of arbitrary finite length.

The present work addresses this gap by providing explicit and self-contained probabilistic existence results for these configurations.

Novelty and Contributions. The main contributions of this paper can be summarized as follows:

- (1) We establish that Bernoulli random subsets of \mathbb{N} almost surely contain finite sumsets (FS-sets) and finite product sets (FP-sets) of arbitrary finite length.
- (2) We provide explicit and constructive probabilistic proofs based on disjoint pattern embeddings and the second Borel–Cantelli lemma.
- (3) We demonstrate the existence of mixed additive–multiplicative configurations of the form $\{x, y, x+y, xy\}$ in random subsets.
- (4) We introduce a probabilistic limit theorem for sums indexed by structured sequences.

The remainder of the paper is organized as follows. In Section 2, we introduce the preliminaries, including the definitions of random subsets of \mathbb{N} , finite sumsets (FS-sets), and finite product sets (FP-sets), along with motivating remarks. Section 3 presents the main theorems, establishing the almost sure existence of FS-sets and FP-sets of arbitrary finite length in random subsets, accompanied by detailed proofs. In addition to these results, we show that a Bernoulli random subset of \mathbb{N} almost surely contains a configuration of the form $\{x, y, x+y, xy\}$. This should be interpreted as a probabilistic existence result for a mixed additive–multiplicative pattern. We emphasize that this does not address the classical Hindman conjecture, which concerns monochromatic configurations under arbitrary finite colorings. In Section 4, we establish an auxiliary probabilistic limit theorem concerning sums indexed by structured sequences. In Section 5, we discuss immediate corollaries and structural consequences of the main results, such as the appearance of arbitrarily long arithmetic progressions and infinitely many disjoint FS-sets and FP-sets. Finally, Section 6 concludes with a summary of our findings and outlines several open problems and directions for future research.

2. Preliminaries

We briefly introduce the notation and definitions used in the paper. We assume familiarity with basic probability theory and combinatorial number theory.

2.1 Random Subsets of \mathbb{N}

Random subsets of natural numbers form a natural probabilistic analogue of deterministic combinatorial problems. They allow us to study the typical behavior of additive and multiplicative structures rather than worst-case configurations.

Definition 1 (Random Subset of \mathbb{N}). Fix a probability $p \in (0, 1)$. A random subset $A \subset \mathbb{N}$ is constructed by including each $n \in \mathbb{N}$ independently with probability p .

This model is often called the binomial model and is denoted by N_p .

Remark 1. This construction ensures independence of membership events. It provides a rich probabilistic framework where one can ask questions like: What is the probability that a certain combinatorial pattern appears in A ? For example, almost surely, certain additive and multiplicative configurations will appear infinitely often, as we shall see.

2.2 Finite Sumsets (FS-Sets)

Additive combinatorics centers on understanding how sums of elements from a set behave. FS-sets encode all possible nonempty sums generated by a finite sequence of integers, serving as building blocks for more complex additive structures.

Definition 2 (Finite Sumset). Given integers $x_1, x_2, \dots, x_L \in \mathbb{N}$, the finite sumset generated by these numbers is

$$FS(x_1, \dots, x_L) := \left\{ \sum_{i \in F} x_i : \emptyset \neq F \subset \{1, \dots, L\} \right\}.$$

Example 1. If $x_1=1$ and $x_2=3$, then

$$FS(1, 3) = \{1, 3, 4\},$$

corresponding to sums of singletons and the sum of both elements.

Remark 2. FS-sets capture the essence of additive structure: finding an FS-set of length L inside a set A guarantees that A contains a rich collection of additive relations among L elements.

2.3 Finite Product Sets (FP-Sets)

Similarly, multiplicative combinatorics focuses on products of elements. FP-sets encode all possible nonempty products from a finite sequence, forming a multiplicative analogue of FS-sets.

Definition 3. (Finite Product Set). Given integers $x_1, x_2, \dots, x_L \in \mathbb{N}$, the finite product set is

$$\text{FP}(x_1, \dots, x_L) := \left\{ \prod_{i \in F} x_i : \emptyset \neq F \subset \{1, \dots, L\} \right\}.$$

Example 2. If $x_1=2$ and $x_2=5$, then

$$\text{FP}(2,5) = \{2,5,10\}.$$

Remark 3. Just as FS-sets reveal additive patterns, FP-sets reveal multiplicative structure. Studying FP-sets in random subsets of \mathbb{N} helps uncover typical multiplicative configurations in probabilistic settings.

3. Main Results and Their Proofs

In this section, we present the main contributions of this paper, establishing that random subsets of the natural numbers almost surely contain finite sumsets and product sets of every finite length. We provide constructive proofs demonstrating the almost sure existence of these additive and multiplicative structures. The proofs rely on combinatorial constructions, independence arguments, and the Borel–Cantelli lemma, illustrating that even in a probabilistic setting, highly structured configurations are unavoidable.

Theorem 1 (Random Finite Sumsets). Fix $p \in (0, 1)$. Let $A \subset \mathbb{N}$ be distributed as Np . Then with probability 1, the following holds:

For every integer $L \geq 1$ there exist integers $x_1 < \dots < x_L$ such that

$$\text{FS}(x_1, \dots, x_L) \subset A.$$

In other words, almost surely A contains finite FS-sets of arbitrarily large finite length.

Proof.

Fix an arbitrary $L \geq 1$. Set $R := 2^L - 1$ and consider the pattern

$$P := \{1, 2, 4, \dots, 2^{L-1}\}.$$

Note that for this choice of P , all subset sums are distinct. Indeed, binary expansion ensures that

$$\text{FS}(P) = \{1, 2, \dots, 2^L - 1\}$$

contains exactly $2^L - 1$ distinct elements. The nonempty subset-sums of P are exactly $\{1, 2, \dots, R\}$, since binary expansions represent all integers from 1 up to $2^L - 1$. Thus,

$$S := \text{FS}(P) = \{1, 2, \dots, R\}.$$

We dilate P by powers of $R + 1$. For each $j \geq 0$ define

$$P_j := (R+1)^j P = \{(R+1)^j \cdot 1, (R+1)^j \cdot 2, (R+1)^j \cdot 4, \dots, (R+1)^j \cdot 2^{L-1}\},$$

and write its set of subset-sums as

$$S_j := \text{FS}(P_j) = (R+1)^j S = \{(R+1)^j s : s \in S\}.$$

Equivalently,

$$S_j = \left\{ (R+1)^j \sum_{i \in F} i : F \subseteq P, F \neq \emptyset \right\}.$$

Each S_j lies in the interval

$$[(R+1)^j \cdot 1, (R+1)^j \cdot R],$$

and since $(R+1)^{j+1} > (R+1)^j R$ the intervals for different j are pairwise disjoint. Hence, the sets S_j are pairwise disjoint.

Now define the event

$$E_j := \{S_j \subset A\}.$$

Each S_j consists of exactly R distinct integers, and inclusion of each integer into A is independent with probability p . Therefore,

$$\mathbb{P}(E_j) = p^R > 0.$$

Because the sets S_j are disjoint for distinct j , the events $(E_j)_{j \geq 0}$ are independent. Since

$$\sum_{j=0}^{\infty} P(E_j) = \sum_{j=0}^{\infty} p^R = \infty,$$

the second Borel–Cantelli lemma implies that with probability 1 infinitely many of the events E_j occur.

Each occurrence of E_j gives $\text{FS}(P_j) = S_j \subset A$, so with probability 1 there exist integers $x_1 < \dots < x_L$ (namely the elements of P_j ordered) whose finite sumset lies entirely in A . Since the event that A contains an FS-set of length L has probability 1 for each fixed L , and since a countable intersection of probability-1 events also has probability 1, it follows that almost surely A contains FS-sets of every finite length.

Theorem 2 (Random Finite Product Sets). Fix $p \in (0, 1)$. Let $A \subset \mathbb{N}$ be distributed as N_p . Then with probability 1, the following holds:

For every integer $L \geq 1$ there exist integers $x_1 < \dots < x_L$ such that

$$\text{FP}(x_1, \dots, x_L) \subset A.$$

In other words, almost surely A contains finite FP-sets of arbitrarily large finite length.

Proof. Fix an arbitrary $L \geq 1$ and let $m := 2^L - 1$. Let $(q_j)_{j \geq 1}$ be an increasing sequence of distinct prime numbers.

For each $j \geq 1$, define

$$x_{j,i} := q_j^{2^{i-1}}, \quad i=1, \dots, L.$$

Then for any nonempty subset $F \subset \{1, \dots, L\}$ we have

$$\prod_{i \in F} x_{j,i} = \prod_{i \in F} q_j^{2^{i-1}} = q_j^{\sum_{i \in F} 2^{i-1}}.$$

The exponents $\sum_{i \in F} 2^{i-1}$ range over all integers from 1 to $2^L - 1 = m$. Thus,

$$\text{FP}(x_{j,1}, \dots, x_{j,L}) = \{q_j^1, q_j^2, \dots, q_j^m\}.$$

Define,

$$\text{FP}(x_{j,1}, \dots, x_{j,L}) = \{q_j^1, q_j^2, \dots, q_j^m\}.$$

Since, the q_j are distinct primes, the sets S_j are pairwise disjoint. (Indeed, powers of different primes are distinct integers.) Define the event

$$E_j := \{S_j \subset A\}.$$

The set S_j has m elements, and the probability that they all lie in A is

$$P(E_j) = p^m > 0.$$

Again, the sets S_j are disjoint, so the events $(E_j)_{j \geq 1}$ are independent. Since,

$$\sum_{j=1}^{\infty} P(E_j) = \sum_{j=1}^{\infty} p^m = \infty,$$

the second Borel–Cantelli lemma implies that with probability 1 infinitely many of the events E_j occur.

Consequently, each E_j guarantees that $\text{FP}(x_{j,1}, \dots, x_{j,L}) \subset A$. Thus with probability 1, A contains an FP-set of length L . Since the event that A contains an FS-set of length L has probability 1 for each fixed L , and since a countable intersection of probability-1 events also has probability 1, it follows that almost surely A contains FP-sets of every finite length.

Theorems 1 and 2 show that additive and multiplicative configurations persist robustly under Bernoulli randomization. Although the constructions are probabilistic, the resulting FS- and FP-patterns resemble finite-dimensional analogues of the structured configurations that arise in deterministic Ramsey theory. In particular, these results demonstrate that random subsets of \mathbb{N} typically retain substantial additive and multiplicative combinatorial richness despite the independence of the underlying random process.

Recent research in Ramsey theory has investigated the occurrence of exponential patterns. In particular, answering a conjecture of Sisto [15], J. Sahasrabudhe established the following result.

Theorem 3 (Sahasrabudhe–Schur Theorem [16, Theorem 2]). For any finite colouring of \mathbb{N} , there exists a monochromatic pattern of the form $\{x, y, x \cdot y, x^y\}$.

Motivated by this deterministic result, we now present a probabilistic analogue which shows that the same exponential pattern almost surely appears inside a random subset of \mathbb{N} .

Theorem 4 (Random subset version). Fix $p \in (0, 1)$. Let $A \subset \mathbb{N}$ be the Bernoulli random subset in which each $n \in \mathbb{N}$ is included independently with probability p . Then with probability 1 the set A contains infinitely many quadruples of the form $\{x, y, x \cdot y, x^y\}$. In particular, there exists at least one such quadruple with all four elements in A .

Proof. Choose two disjoint infinite sequences of distinct primes $(u_j)_{j \geq 1}$ and $(v_j)_{j \geq 1}$ (for instance, partition the odd primes into two infinite subsequences). For each j consider the quadruple

$$T_j := \{u_j, v_j, u_j \cdot v_j, u_j^{v_j}\}.$$

where u_j and v_j are distinct prime numbers, $u_j v_j$ denotes their product, and $u_j^{v_j}$ denotes u_j raised to the power v_j . Because the u_j are distinct primes and the v_j are distinct integers, all elements of distinct quadruples T_j are distinct. Thus the quadruples T_j are pairwise disjoint, and the events

$$E_j := \{T_j \subset A\}$$

are independent. Each E_j has probability

$$P(E_j) = p^4 > 0.$$

Since $\sum_{j=1}^{\infty} P(E_j) = \sum_{j=1}^{\infty} p^4 = +\infty$, the (second) Borel–Cantelli lemma implies that with probability 1 infinitely many of the E_j occur. Each occurrence yields a quadruple

$$\{x, y, x \cdot y, x^y\}$$

contained in A .

Remark 4. We record some observations regarding the construction and its generalizations:

(1) The key to the Borel–Cantelli argument is producing infinitely many *pairwise disjoint* instances of the pattern, each occurring with the same positive probability; independence then yields almost sure existence of infinitely many such quadruples. The construction using distinct primes as bases (and distinct exponents) ensures disjointness and distinctness of the power values.

(2) This method applies to many related patterns $P(x, y)$, as long as one can produce infinitely many pairwise disjoint realizations of the pattern, each with a fixed finite number of entries. For example, patterns such as $\{x, y, x + y\}$ or $\{x, y, x^y\}$ can be handled in the same way.

(3) Moreover, one can strengthen the theorem to say that for any fixed finite family of finite patterns (each involving finitely many integers), a Bernoulli(p) random subset almost surely contains infinitely many disjoint copies of each pattern, provided suitable disjoint embeddings exist.

Additive-Multiplicative Configurations in Random Sub-Sets

Motivated by the prominence of this additive–multiplicative configuration, we establish the following probabilistic existence theorem. We note that the proof strategy for the following theorem differs substantially from the approach used in the earlier version of the manuscript. The revised argument is more direct and relies on an explicit construction of pairwise disjoint additive–multiplicative configurations together with the second Borel–Cantelli lemma.

The configuration $x, x+y, x \cdot y$ is a natural additive–multiplicative pattern that has received considerable attention in Ramsey theory. In this paper, however, we do not investigate partition-regularity or coloring problems. Instead, our focus is entirely probabilistic: we establish that configurations of the form $\{x, x+y, x \cdot y\}$ occur almost surely in Bernoulli random subsets of \mathbb{N} . Thus, the result should be viewed as a probabilistic existence theorem for a mixed additive–multiplicative configuration rather than as a statement concerning classical Ramsey-theoretic partition regularity.

Theorem 5 (Existence of additive–multiplicative quadruples in random sets). Let $p \in (0, 1)$, and let $A \subset \mathbb{N}$ be the Bernoulli random subset in which each integer $n \in \mathbb{N}$ is included independently with probability p . Then, with probability 1, the set A contains infinitely many pairwise disjoint quadruples of the form

$$\{x, x+y, x \cdot y, y\},$$

where the four numbers in each quadruple are pairwise distinct. In particular, with probability 1, there exists at least one such quadruple contained entirely in A .

Proof. We construct an infinite family of pairwise disjoint candidate quadruples, each occurring with the same

positive probability.

Define an increasing sequence of integers $(m_j)_{j \geq 1}$ recursively by

$$m_1 := 1, \quad m_{j+1} := 2m_j + 2 \quad (j \geq 1).$$

In particular, $m_j \geq 1$ for every j , and the sequence is strictly increasing.

For each $j \geq 1$, set

$$\begin{aligned} x_j &:= 2^{m_j}, \\ y_j &:= 2^{m_j + 1} = 2x_j. \end{aligned}$$

Consider the associated quadruple

$$Q_j := \{x_j, y_j, x_j + y_j, x_j y_j\}.$$

Since,

$$x_j + y_j = 2^{m_j} + 2^{m_j + 1} = 3 \cdot 2^{m_j}, \quad x_j y_j = 2^{m_j} \cdot 2^{m_j + 1} = 2^{2m_j + 1}$$

we obtain,

$$Q_j = \{2^{m_j}, 2^{m_j + 1}, 3 \cdot 2^{m_j}, 2^{2m_j + 1}\}.$$

We first check that the four elements in Q_j are distinct. Indeed,

$$2^{m_j} < 2^{m_j + 1} < 3 \cdot 2^{m_j},$$

because $3 > 2$, and

$$3 \cdot 2^{m_j} < 2^{2m_j + 1}$$

since this inequality is equivalent to

$$3 < 2^{m_j + 1},$$

which holds for every $m_j \geq 1$.

Thus each Q_j consists of four distinct natural numbers.

We now show that $Q_j \cap Q_k = \emptyset$ whenever $j \neq k$. It is enough to prove that

$$\max Q_j < \min Q_{j+1} \quad \text{for every } j \geq 1.$$

Indeed, since

$$\max Q_j = x_j y_j = 2^{2m_j + 1} \quad \text{and} \quad \min Q_{j+1} = x_{j+1} = 2^{m_{j+1}},$$

it suffices to verify that

$$2^{2m_j + 1} < 2^{m_{j+1}}.$$

Because $m_{j+1} = 2m_j + 2$, we have

$$m_{j+1} = 2m_j + 2 > 2m_j + 1.$$

Hence

$$2^{2m_j + 1} < 2^{m_{j+1}}.$$

Therefore

$$\max Q_j < \min Q_{j+1},$$

and so the sets Q_j are pairwise disjoint.

For each $j \geq 1$, define the event

$$E_j := \{Q_j \subset A\}.$$

Since Q_j has exactly four elements and membership in A is independent across integers, we have

$$P(E_j) = p^4.$$

Moreover, because the sets Q_j are pairwise disjoint, the events $(E_j)_{j \geq 1}$ are independent: each E_j depends

only on the membership indicators of the four integers in Q_j , and these indicators are independent across different j .

Since

$$\sum_{j=1}^{\infty} P(E_j) = \sum_{j=1}^{\infty} p^4 = +\infty,$$

and the events E_j are independent, the second Borel–Cantelli lemma implies that

$$P(E_j \text{ infinitely often}) = 1.$$

Thus, with probability 1, infinitely many of the quadruples Q_j are contained in A .

Consequently, almost surely A contains infinitely many pairwise disjoint quadruples of the form

$$\{x, y, x+y, xy\},$$

with all four elements distinct. In particular, almost surely A contains at least one such quadruple.

The above theorem establishes a probabilistic existence result for a specific additive–multiplicative configuration inside a Bernoulli random subset of \mathbb{N} . This should not be interpreted as a resolution, even in a randomized sense, of the Hindman conjecture, which concerns the existence of monochromatic configurations under arbitrary finite colorings of \mathbb{N} . The probabilistic setting considered here is fundamentally different.

4. A Probabilistic Limit Theorem for Structured Sums

The following result is of a different nature from the preceding sections. While Theorems 1–5 establish almost sure existence of additive and multiplicative configurations in random subsets, the result below concerns probabilistic limit behavior associated with sums indexed by Hindman-type sequences. Although it is not required for the FS/FP existence theorems, it is included primarily as an auxiliary probabilistic observation concerning random sums indexed by structured sequences.

Theorem 6 (Hindman + CLT: probabilistic finite-sums and finite-products). Let \mathbb{N} be finitely colored and we have a sequence $x = (x_i)_{i \geq 1} \subset \mathbb{N}$ such that $\text{FS}(x)$ is monochromatic.

$$S_k = \sum_{j=1}^k \varepsilon_{n_j} y_j, \quad \mu_k = E[S_k] = p \sum_{j=1}^k y_j, \quad \sigma_k^2 = \text{Var}(S_k) = p(1-p) \sum_{j=1}^k y_j^2$$

Assume $\sum_{j=1}^{\infty} y_j^2 = \infty$ and write j_k for an index achieving the maximum at stage k ,

$$y_{j_k}^2 = \max_{1 \leq j \leq k} y_j^2$$

Set the residual squared sum

$$R_k := \sum_{\substack{1 \leq j \leq k \\ j \neq j_k}} y_j^2$$

Then with probability 1, the following statements are true. The following two regimes arise under separate hypotheses and should be interpreted independently.

(1) Dominated single-term regime. If,

$$\frac{y_{j_k}^2}{\sigma_k^2} = \frac{y_{j_k}^2}{y_{j_k}^2 + R_k} \xrightarrow{k \rightarrow \infty} 1,$$

then the variance is asymptotically carried by the single largest summand. In this case

$$\frac{S_k - \mu_k}{\sigma_k} \xrightarrow{d} \frac{B - p}{\sqrt{p(1-p)}},$$

where $B \sim \text{Bernoulli}(p)$. Equivalently, the limit law is the two-point distribution taking values $\frac{1-p}{\sqrt{p(1-p)}}$ with probability p and $\frac{-p}{\sqrt{p(1-p)}}$ with probability $1-p$.

In particular no nondegenerate Gaussian limit can hold for the full sums under this normalisation.

(2) Trimmed CLT regime. Suppose instead that there exists a deterministic choice of indices j_k (for instance the

argmax indices) such that the trimmed variance

$$(\sigma_k^{\text{trim}})^2 := p(1-p) R_k \xrightarrow{k \rightarrow \infty} \infty$$

and the largest remaining squared term is negligible relative to the residual, i.e.

$$\frac{\max_{1 \leq j \leq k, j \neq j_k} y_j^2}{R_k} \xrightarrow{k \rightarrow \infty} 0.$$

Then the centered trimmed sums

$$S_k^{\text{trim}} := \sum_{\substack{1 \leq j \leq k \\ j \neq j_k}} \varepsilon_{n_j} y_j$$

satisfy the Lindeberg condition and therefore

$$\frac{S_k^{\text{trim}} - \mathbb{E}[S_k^{\text{trim}}]}{\sigma_k^{\text{trim}}} \xrightarrow{d} \mathcal{N}(0,1).$$

Moreover, if the removed (dominating) term is negligible relative to the trimmed variance in the sense that

$$\frac{y_{j_k}^2}{R_k} \xrightarrow{k \rightarrow \infty} 0,$$

then reinserting it does not change the Gaussian limit and one obtains

$$\frac{S_k - \mu_k}{\sigma_k^{\text{trim}}} \xrightarrow{d} \mathcal{N}(0,1).$$

If instead $y_{j_k}^2/R_k \not\rightarrow 0$ then reinsertion typically changes the limit and one does not obtain a Gaussian law with the trimmed normalisation.

In all cases every atom (point-mass value) of S_k (respectively of any trimmed sum) is a finite sum of elements from $\{y_1, \dots, y_k\}$ and hence lies in $\text{FS}(x)$.

Proof.

We give a detailed, self-contained verification of the two regimes and the asserted limit laws. Fix k . Let j_k be an index attaining the maximum square among $\{y_1^2, \dots, y_k^2\}$. Write,

$$\sigma_k^2 = p(1-p) \sum_{j=1}^k y_j^2 = p(1-p)(y_{j_k}^2 + R_k).$$

For $j=1, \dots, k$ define the centred summands

$$Z_{k,j} := \varepsilon_{n_j} y_j - p y_j,$$

so that $S_k - \mu_k = \sum_{j=1}^k Z_{k,j}$. Each $Z_{k,j}$ is independent, mean zero and takes only two values:

$$Z_{k,j} = \begin{cases} (1-p)y_j & \text{with probability } p, \\ -p y_j & \text{with probability } 1-p, \end{cases}$$

hence $\text{Var}(Z_{k,j}) = p(1-p)y_j^2$. We will analyse the contribution of the maximal summand Z_{k,j_k} and of the residual sum

$$R_k' := \sum_{\substack{1 \leq j \leq k \\ j \neq j_k}} Z_{k,j},$$

whose variance equals $p(1-p)R_k$.

Proof of (A): dominated single-term limit. Assume

$$\frac{y_{j_k}^2}{\sigma_k^2} = \frac{y_{j_k}^2}{y_{j_k}^2 + R_k} \xrightarrow{k \rightarrow \infty} 1.$$

Equivalently $R_k/y_{j_k}^2 \rightarrow 0$, hence $\text{Var}(R_k') = p(1-p)R_k = o(p(1-p)y_{j_k}^2) = o(\sigma_k^2)$. Therefore

$$\frac{R_k' - E[R_k']}{\sigma_k} \xrightarrow{p} 0,$$

since the variance of the left-hand side tends to 0. More precisely,

$$\text{Var} \left(\frac{R_k' - E[R_k']}{\sigma_k} \right) = \frac{p(1-p)R_k}{\sigma_k^2} \xrightarrow{k \rightarrow \infty} 0,$$

so, the random variable converges to zero in L^2 and hence in probability.

Consequently, the asymptotic law of $(S_k - \mu_k)/\sigma_k$ is the same as that of the single-term contribution

$$\frac{Z_{k,j_k}}{\sigma_k} = \frac{\varepsilon_{n_{j_k}} y_{j_k} - P y_{j_k}}{\sigma_k}.$$

But by the domination assumption $\sigma_k \sim \sqrt{p(1-p)} y_{j_k}$, so

$$\frac{Z_{k,j_k}}{\sigma_k} \xrightarrow{d} \frac{B-p}{\sqrt{p(1-p)}},$$

where $B \sim \text{Bernoulli}(p)$. That proves the claimed two-point limit law in regime (A). In particular this limit is not Gaussian (unless degenerate), so a nondegenerate CLT for the full sums under the normalisation σ_k cannot hold.

Proof of (B): trimmed CLT. Assume now that after removing the maximal index j_k the residual squared sum R_k tends to infinity and the largest remaining squared term is negligible, i.e.

$$(\sigma_k^{\text{trim}})^2 := p(1-p)R_k \rightarrow \infty \quad \text{and} \quad \frac{\max_{1 \leq j \leq k, j \neq j_k} y_j^2}{R_k} \xrightarrow{k \rightarrow \infty} 0.$$

We shall verify the Lindeberg condition for the triangular array formed by the centred trimmed summands $\{Z_{k,j} : 1 \leq j \leq k, j \neq j_k\}$ with normalisation σ_k^{trim} . Fix any $\delta > 0$. Note that for each j , $|Z_{k,j}| \leq \max\{p, 1-p\} y_j \leq y_j$.

Hence,

$$\mathbf{1}_{\{|Z_{k,j}| > \delta \sigma_k^{\text{trim}}\}} \leq \mathbf{1}_{\{y_j > \delta \sigma_k^{\text{trim}}\}}.$$

Therefore,

$$\begin{aligned} \frac{1}{(\sigma_k^{\text{trim}})^2} \sum_{\substack{1 \leq j \leq k \\ j \neq j_k}} E \left[Z_{k,j}^2 \mathbf{1}_{\{|Z_{k,j}| > \delta \sigma_k^{\text{trim}}\}} \right] &\leq \frac{1}{(\sigma_k^{\text{trim}})^2} \sum_{\substack{1 \leq j \leq k \\ j \neq j_k, y_j > \delta \sigma_k^{\text{trim}}}} E \left[Z_{k,j}^2 \right] \\ &= \frac{p(1-p)}{p(1-p)R_k} \sum_{\substack{1 \leq j \leq k \\ j \neq j_k, y_j > \delta \sigma_k^{\text{trim}}}} y_j^2 \\ &\leq \frac{\max_{1 \leq j \leq k, j \neq j_k} y_j^2}{R_k} \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Thus, the Lindeberg condition for the trimmed triangular array holds. By the Lindeberg--Feller Central Limit Theorem,

$$\frac{S_k^{\text{trim}} - E[S_k^{\text{trim}}]}{\sigma_k^{\text{trim}}} = \frac{1}{\sigma_k^{\text{trim}}} \sum_{\substack{1 \leq j \leq k \\ j \neq j_k}} Z_{k,j} \xrightarrow{d} \text{N}(0,1),$$

as claimed.

Reinsertion of the removed summand. Finally, we check the asserted effect of reinserting the removed maximal term. Write,

$$S_k - \mu_k = Z_{k,j_k} + (S_k^{\text{trim}} - E[S_k^{\text{trim}}]).$$

If the removed term is negligible relative to the trimmed variance, namely $y_{j_k}^2/R_k \rightarrow 0$, then

$$\frac{Z_{k,j_k}}{\sigma_k^{\text{trim}}} = \frac{\varepsilon_{n_{j_k}} y_{j_k} - P y_{j_k}}{\sigma_k^{\text{trim}}}$$

has variance

$$\text{Var} \left(\frac{Z_{k,j_k}}{\sigma_k^{\text{trim}}} \right) = \frac{p(1-p)y_{j_k}^2}{p(1-p)R_k} = \frac{y_{j_k}^2}{R_k} \xrightarrow{k \rightarrow \infty} 0,$$

hence $Z_{k,j_k}/\sigma_k^{\text{trim}} \xrightarrow{P} 0$. Combining this with the CLT for the trimmed sums and Slutsky's theorem yields

$$\frac{S_k - \mu_k}{\sigma_k^{\text{trim}}} = \frac{S_k^{\text{trim}} - E[S_k^{\text{trim}}]}{\sigma_k^{\text{trim}}} + \frac{Z_{k,j_k}}{\sigma_k^{\text{trim}}} \xrightarrow{d} N(0,1).$$

Thus, reinsertion does not affect the Gaussian limit when $y_{j_k}^2/R_k \rightarrow 0$. If instead $y_{j_k}^2/R_k \not\rightarrow 0$, then reinsertion typically changes the normalisation needed and in general one does not obtain a Gaussian limit with the trimmed normalisation; behavior in that case can range from mixed (sum of Gaussian limit plus an independent two-point mass in the limit) to complete domination by the single-term law as in regime (A), depending on the precise rates.

It is obvious that, for any fixed k , every possible value (atom) of S_k is a finite sum of elements drawn from $\{y_1, \dots, y_k\}$. Since these elements are a subsequence of the original Hindman sequence x , their finite sums lie in $FS(x)$. The same remark applies to any trimmed sum. Finally, we note that every value of S_k is a finite sum of elements from $\{y_1, \dots, y_k\}$ and hence lies in $FS(x)$ by definition. This observation does not rely on Hindman's theorem itself, but only on the structure of finite sums.

5. Immediate Consequences

The following observations are immediate consequences of the constructions used in the proofs of Theorems 1 and 2 and are included only for completeness.

Corollary 1 (Infinitely many disjoint FS- and FP-sets of fixed length). Fix $L \geq 1$. Then almost surely, A contains infinitely many disjoint FS- and FP-sets of length L .

Proof.

Proof of infinitely many disjoint FS-sets of fixed length: In the proof of Theorem 1, the translated patterns P_j are pairwise disjoint. Each event $E_j := \{FS(P_j) \subset A\}$ is independent with positive probability. By the Borel--Cantelli lemma, infinitely many E_j occur almost surely, yielding infinitely many disjoint FS-sets of length L .

Proof of infinitely many disjoint FP-sets of fixed length: In Theorem 2, the FP-sets are constructed using distinct primes q_j , so the sets $S_j := FP(x_{j,1}, \dots, x_{j,L})$ are disjoint. Each event $E_j := \{S_j \subset A\}$ is independent with positive probability. By the Borel--Cantelli lemma, infinitely many E_j occur almost surely, giving infinitely many disjoint FP-sets of length L .

Corollary 2 (Arbitrarily large minimal elements in FS and FP-sets). Almost surely, for any $M > 0$ and $L \geq 1$, A contains a FS-set and a FP-set of length L all of whose elements are greater than M .

Proof.

Here we present the proof for the product case. We now briefly indicate the argument for the sum case. In the construction of Theorem 1, the elements of P_j grow like $(R+1)^j$. By choosing j sufficiently large so that $(R+1)^j > M$, all elements of $FS(P_j)$ exceed M . Since infinitely many such j occur almost surely, the conclusion follows.

Choose j sufficiently large so that the prime $q_j > M$. Then the corresponding FP-set $S_j = \{q_j^1, q_j^2, \dots, q_j^{L-1}\} \subset A$ will have all elements larger than M . Since infinitely many E_j occur almost surely (by Borel--Cantelli), infinitely many such FP-sets exist.

6. Conclusion and Open Problems

We have rigorously established that almost every random subset of \mathbb{N} contains finite sumsets and product sets of arbitrary finite length. These results show that simple additive and multiplicative configurations appear almost surely in Bernoulli random subsets of \mathbb{N} .

Several natural questions remain open:

(1) While we guarantee arbitrarily large finite FS-sets and FP-sets, the existence of infinite FS-sets (additive IP-sets) or infinite FP-sets (multiplicative IP-sets) in a random subset N_p remains open. Developing probabilistic analogues of infinite finite-sums phenomena remains an intriguing direction.

(2) For random subsets of $[n]$, determining the precise threshold probability $p = p(n)$ at which FS-sets or FP-sets of given lengths appear with high probability is a natural extension, akin to results in random van der Waerden theory.

- (3) Investigating finite sumsets and product sets in random subsets of other algebraic structures, such as \mathbb{Z}^d or finite fields, may reveal richer combinatorial phenomena.
- (4) Can one obtain quantitative Berry–Esseen bounds for the rate of convergence in Theorem 6, specialized to Hindman subsequences?
- (5) Is it possible to enforce additional arithmetic structure (such as prescribed residues) on the Hindman sequence while still satisfying the variance growth needed for the CLT?
- (6) What probabilistic limit theorems hold for infinite partitions, where Hindman’s theorem no longer applies in its classical form?

The methods used in this paper are intentionally elementary and rely primarily on independence arguments together with the second Borel–Cantelli lemma. We hope that the explicit FS/FP formulations presented here may serve as a useful starting point for more refined probabilistic investigations of Ramsey-theoretic structures.

Conflict of Interest

The authors declare no conflict of interest.

Generative AI Statement

The authors declare that no Generative AI was used in the creation of this manuscript.

References

- [1] Nathanson MB. Additive number theory. New York: Springer; 1996. DOI: 10.1017/9781108775267.013
- [2] Tao T, Vu VH. Additive combinatorics. Cambridge University Press; 2006. DOI: 10.1017/CBO9780511755149
- [3] Alon N, Spencer JH. The probabilistic method. John Wiley & Sons; 2016. Available at: http://lib.ysu.am/disciplines_bk/39cbf4832349c9024453be49f58db93e.pdf (accessed on 20 May 2025).
- [4] Bollobás B. Random graphs. In Modern graph theory 2011 May 28 (pp. 215-252). New York, NY: Springer New York. DOI: 10.1007/978-1-4612-0619-4_7
- [5] Gowers WT. A new proof of Szemerédi's theorem. Geometric & Functional Analysis GAFA, 200, 11(3), 465-588. DOI:10.1007/s00039-001-0332-9
- [6] Komlós J, Simonovits M. Szemerédi's Regularity Lemma and its applications in graph theory. Center for Discrete Mathematics & Theoretical Computer Science, 1995. Available at: <https://dl.acm.org/doi/abs/10.5555/868009> (accessed on 20 May 2025).
- [7] Hindman N. Finite sums from sequences within cells of a partition of \mathbb{N} . Journal of Combinatorial Theory, Series A. 1974, 17(1), 1-1. DOI: 10.1016/0097-3165(74)90023-5
- [8] Hindman N. Partitions and sums and products of integers. Transactions of the American Mathematical Society. 1979, 247, 227-245. DOI: 10.1090/S0002-9947-1979-0517693-4
- [9] Baumgartner JE. A short proof of Hindman's theorem. Journal of Combinatorial Theory, Series A. 1974, 17(3), 384-386. DOI: 10.1016/0097-3165(74)90103-4
- [10] Bergelson V, Hindman N. Partition regular structures contained in large sets are abundant. Journal of Combinatorial Theory, Series A. 2001, 93(1), 18-36. DOI: 10.1006/jcta.2000.3061
- [11] Bergelson V, Leibman A. Sets of large values of correlation functions for polynomial cubic configurations. Ergodic Theory and Dynamical Systems. 2018, 38(2), 499-522. DOI: 10.1017/etds.2016.49
- [12] Erdős P, Rényi A. On the strength of connectedness of a random graph. Acta Mathematica Hungarica. 1961, 12(1-2), 261-267. DOI: 10.1007/bf02066689
- [13] Conlon D, Gowers WT. Combinatorial theorems in sparse random sets. Annals of Mathematics. 2016, 184, 367-454. DOI: 10.4007/annals.2016.184.2.2
- [14] Rödl V, Ruciński A. Threshold functions for Ramsey properties. Journal of the American Mathematical Society. 1995, 8(4), 917-942. DOI: 10.1090/S0894-0347-1995-1276825-6
- [15] Sisto A. Exponential triples. The Electronic Journal of Combinatorics, 2011, 18(1). DOI: 10.37236/634
- [16] Sahasrabudhe J. Exponential patterns in arithmetic Ramsey theory. Acta Arithmetica, 2018, 182(1), 13-42. DOI:10.4064/aa8603-9-2017