entropy

MDPI

Article

On the Algorithm Complexity of Generating Discrete Uniform Distribution from a Biased Coin

Mengqi Zhang [1,*], Guangqiang Teng [2] and Xiaoyu Lei [3]

[1] School of Economics, Peking University, Beijing 100871, China
[2] School of Mathematics, Harbin Institute of Technology, Harbin 150001, China; tenggq@alu.hit.edu.cn
[3] Department of Statistics, University of Wisconsin-Madison, 1205 University Ave, Madison, WI 53706, USA; xlei35@wisc.edu
* Correspondence: 2406395240@pku.edu.cn

Abstract

Lei proposed an algorithm Algorithm $\mathcal{A}_3$ in 2023 to generate an exact discrete uniform distribution from an unknown biased Bernoulli source. The present paper does not claim a new extraction algorithm. Its contributions are analytical: first, we provide a Fourier-analytic proof of the uniformity mechanism based on roots of unity and coefficient extraction; second, we derive explicit acceptance-probability and expected-runtime bounds, with a rigorous treatment of composite moduli; third, we show how independent accepted $\mathcal{A}_3$ digits yield a continuous Uniform$(0,1)$ limit through a base-n expansion and quantify the cost of finite-digit simulation.

Keywords: random number generation; biased coin; discrete uniform distribution; number theory; characteristic function; algorithm complexity; Monte Carlo methods; fairness in algorithms

1. Introduction

1.1. The Quest for Perfect Randomness in Imperfect Worlds

New theory and large experiments show that coin tossing is not perfectly fair in the strict physical sense. The dynamical model from [1] predicts a small precession in flight that makes the coin slightly more likely to land on the same face it started with. A large experiment led by the University of Amsterdam collected 350,757 flips and found the probability of landing on the starting face is about 50.8% with a 95% confidence interval of [0.506, 0.509], and this result closely matches the approximately 51% theoretical prediction while the overall heads versus tails frequencies remain symmetric with heads at roughly 50.0% and a 95% confidence interval of [0.498, 0.502]. No systematic bias exists toward heads or tails but a subtle starting-orientation effect does exist, meaning raw coin tossing is not ideal for those demanding exact fairness.

This mild but detectable bias shows the value of the method from [2] published in Electronic Communications in Probability in 2023. Even when the base bit source, which is the coin, has an unknown bias, one can generate exactly discrete-uniform random numbers without parameters and with proof. The algorithm uses n tosses to form the head count N and a rank-sum R and it outputs R mod n whenever $\gcd(N,n)=1$; it adds an independent random cyclic shift if the flips are independent but not identically distributed, and this process yields an exactly uniform value on $\{0,\ldots,n-1\}$. Even with a starting-orientation bias in physical tossing, this number-theoretic–probabilistic "uniformization" delivers

Check for updates

Academic Editor: Boris Ryabko

Received: 21 June 2026
Revised: 14 July 2026
Accepted: 16 July 2026
Published: 17 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

strictly fair n-ary outcomes that are useful for instructional labs, low-cost randomness acquisition and any application that needs verifiable fairness.

1.2. Historical Context and Theoretical Foundations

The present problem is related to broader literature on extracting unbiased randomness from biased sources, beginning with von Neumann's extractor [3] and including refinements by Elias [4], Knuth-Yao [5], and Peres [6]. These methods optimize different efficiency or source-assumption criteria.

The mathematical problem of making fair random bits from biased coins has a long history. It started with von Neumann's important 1951 work [3]. For a coin with unknown bias p, von Neumann suggested a simple algorithm. Flip the coin twice. Use the first flip as the result if the outcomes are different. Discard the two flips and try again if they are the same. This simple process gets exactly one fair bit from two biased flips. It has a variable cost and the expected runtime is $1/(2p(1-p))$ flips per output bit.

Dijkstra [7] extended this framework to prime numbers. He created an algorithm that generates a uniform distribution on p elements from p coin flips. This method uses the algebraic structure of cyclic groups and lexicographic ranking within orbits made by cyclic shifts. The primality of p ensures each orbit has exactly p distinct configurations and this allows uniform sampling.

Lei [2] simplified this method by using rank sums modulo p instead of complex orbit analysis, replacing the latter with a straightforward modular computation. His simplification laid the groundwork for generalizing the method to any integer n, using a number-theoretic approach that combines ideas from probability theory and algebra. Lei's algorithm $\mathcal{A}_3$ is distinctive in that it produces exact n-ary uniform variables directly from a Bernoulli source through a number-theoretic acceptance condition. Our contribution is to analyze this mechanism using Fourier coefficients and to quantify the acceptance cost.

1.3. Applications in Modern Data Science

The algorithm's importance extends beyond theoretical interest to practical applications across modern data science:

Statistical computing: The generation of non-uniform deviates via inversion, acceptance-rejection, or MCMC fundamentally relies on a source of i.i.d. uniform draws. By supplying such draws from an arbitrarily biased Bernoulli stream, $\mathcal{A}_3$ immunises these procedures against source-level imperfections. In particular, Metropolis-Hastings samplers require proposal steps with exact uniformity to preserve detailed balance; the proposed method provides a parameter-free guarantee that is otherwise difficult to achieve without knowledge of the physical bias.

Machine Learning and Randomized Algorithms:Modern machine learning uses more and more randomness. It uses randomness for initialization (neural network weights), regularization (dropout), and optimization (stochastic gradient descent). This algorithm has provable fairness. It makes sure that these random parts do not add unintended biases. Ensemble methods like random forests use bootstrap sampling and random feature selection. This algorithm ensures that each data point and feature is selected with equal probability.

Reinforcement Learning: In Reinforcement learning (RL) and bandit algorithms, exploration strategies such as ϵ-greedy or Thompson sampling algorithms need uniform sampling to explore state spaces and K-arms; see [8]. A biased generator can cause systematic under-exploration of some action trajectories. MCMC inference uses an acceptance ratio. This ratio needs to compare a computed likelihood with a uniform variable $u \sim \mathcal{U}[0,1]$. The bias in u breaks the detailed balance condition. This means the chain converges to the

wrong stationary distribution. We extend $\mathcal{A}_3$ to the continuous domain (Section 5) to fix this verification gap.

Cryptography and Security: Cryptographic protocols need high-quality randomness for key generation, nonce creation and probabilistic encryption. Cryptographic applications usually need stronger security guarantees than physical coin tossing alone can provide. This algorithm gives a valuable post-processing step. It helps distill entropy from imperfect sources in randomness extractors [9].

A concrete use case is finite experimental randomization from a simple biased physical or hardware Bernoulli source. Suppose a laboratory device produces binary events with unknown but stable bias p, and the experiment requires exact assignment among n treatment arms. Estimating p and then correcting the source introduces calibration error, whereas $\mathcal{A}_3$ gives an exact uniform n-ary output under the i.i.d. Bernoulli model. This is the regime in which the algorithm is preferable: exact finite-sample n-ary uniformization is required, the source bias is unknown, and cryptographic security is not the primary objective.

1.4. Contributions and Paper Overview

In summary, the contribution of this paper is as follows: (1) We provide a characteristic-function proof of Lei's algorithm $\mathcal{A}_3$ [2] using roots of unity and coefficient extraction. (2) We prove explicit acceptance-probability and expected-runtime bounds, including the composite-modulus case. (3) We use independent accepted $\mathcal{A}_3$ digits to construct a continuous Uniform$(0,1)$ limit and quantify finite-digit runtime and approximation error.

The specific contributions are as follows:

1. A Novel Characteristic Function Proof (Section 3): We use a rigorous analytic proof with Characteristic Functions, which are Fourier analysis on $\mathbb{Z}_n$, instead of the combinatorial group-action proof from the original paper. This proof is more robust and directly links the number-theoretic condition $\gcd(N,n)=1$ to the vanishing of Fourier coefficients at non-zero frequencies.
2. Complexity Analysis (Section 4): We establish rigorous bounds on the algorithm's expected runtime. We show the expected number of coin flips is $O(n)$ when n is prime and $O(n \log\log n)$ when n is composite. We find explicit constants with Euler's totient function $\phi(n)$ and give detailed asymptotic expansions using analytic number theory tools.
3. Extension to Continuous Uniform Distribution (Section 5): We show $\mathcal{A}_3$ can be used to generate continuous uniform random variables on $[0,1]$ with base-n expansion. We prove convergence with characteristic function arguments and give error bounds for practical use.

The paper is organized as follows: Section 2 formally presents $\mathcal{A}_3$ and establishes notation. Section 3 provides the characteristic function proof of uniformity. Section 4 analyzes computational complexity with detailed asymptotic expansions. Section 5 extends this method to the generation of continuous uniform distributions. Section 6 concludes with discussion and future directions.

2. Preliminaries and Algorithm Formulation

We operate within a standard probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Let $\mathbb{N} = \{1,2,\dots\}$ and $\mathbb{N}_0 = \{0,1,\dots\}$.

2.1. Model and Sufficient Statistics

Assumption 1 (Bernoulli Source)**.** *Let $X_0, X_1, \dots$ be a sequence of independent and identically distributed (i.i.d.) random variables taking values in $\{0,1\}$. We denote the success probability by*

$p = P(X_i = 1)$ and the failure probability by $q = P(X_i = 0) = 1 - p$. We assume $p \in (0,1)$ is unknown.

We denote a sequence of n flips as a random vector $\mathbf{X}^n = (X_0, X_1, \ldots, X_{n-1}) \in \{0,1\}^n$.

The algorithm relies on two statistics of the sample $\mathbf{X}^n$: the Hamming weight (head count) and a weighted positional sum (rank sum).

Definition 1 (Head Count)**.** *For a vector $\mathbf{X}^n$, the head count is defined as:*

$$N(\mathbf{X}^n) = \sum_{i=0}^{n-1} X_i. \tag{1}$$

Definition 2 (Rank Sum)**.** *The rank sum is defined as the sum of indices where heads occur:*

$$R(\mathbf{X}^n) = \sum_{i=0}^{n-1} i \cdot X_i. \tag{2}$$

Definition 3 (Coprimality)**.** *Let $\gcd(x, y)$ denote the greatest common divisor of integers x and y. We say N and n are coprime if $\gcd(N, n) = 1$. The set of integers in $\{0, \ldots, n-1\}$ coprime to n forms the multiplicative group of integers modulo n, denoted by $\mathbb{Z}_n^\times$. The size of this group is given by Euler's totient function $\phi(n)$.*

2.2. Algorithm $\mathcal{A}_3$

The core algorithm proposed in [2] is a rejection sampling procedure. It generates batches of size n and accepts a batch only if the head count N is coprime to the batch size n (see Algorithm 1).

Algorithm 1. Discrete Uniform Generation

Require: Sequence of biased coin flips $X_i \sim \text{Ber}(p)$. Target size $n \in \mathbb{N}, n \geq 2$.
Ensure: Random variable Y uniformly distributed on $\{0, 1, \ldots, n-1\}$.

```
Initialize valid ← false.
while not valid do
    Sample n independent flips X^n = (X_0, ..., X_{n-1}).
    Calculate N = Σ_{i=0}^{n-1} X_i.
    if gcd(N, n) == 1 then
        Calculate R = Σ_{i=0}^{n-1} i · X_i.
        Set Y = R (mod n).
        valid ← true.
    else
        Discard X^n and restart.
    end if
end while
return Y.
```

2.3. Illustrative Examples

Example 1 ($n = 2$: Connection to von Neumann's Algorithm)**.** *For $n = 2$, $\mathcal{A}_3$ reduces to a variant of von Neumann's original procedure. Consider all possible outcomes:*

- $(0,0)$*:* $N = 0$, $\gcd(0,2) = 2 > 1$, *reject*
- $(1,1)$*:* $N = 2$, $\gcd(2,2) = 2 > 1$, *reject*
- $(0,1)$*:* $N = 1$, $\gcd(1,2) = 1$, $R = 0 \cdot 0 + 1 \cdot 1 = 1$, *output* $1 \mod 2 = 1$
- $(1,0)$*:* $N = 1$, $\gcd(1,2) = 1$, $R = 0 \cdot 1 + 1 \cdot 0 = 0$, *output* $0 \mod 2 = 0$

The algorithm accepts exactly when the two flips differ, producing outputs 0 for (H, T) and 1 for (T, H), matching von Neumann's procedure.

Example 2 ($n = 6$: Composite Case). *Consider $n = 6$ and the outcome $X^6 = (1, 0, 1, 0, 1, 0)$:*

- *Head count:* $N = 1 + 0 + 1 + 0 + 1 + 0 = 3$
- $\gcd(3, 6) = 3 > 1$, *so reject*

Now consider $X^6 = (1, 1, 0, 0, 0, 1)$:

- *Head count:* $N = 1 + 1 + 0 + 0 + 0 + 1 = 3$
- $\gcd(3, 6) = 3 > 1$, *reject again*

Finally, consider $X^6 = (1, 0, 0, 1, 0, 0)$:

- *Head count:* $N = 1 + 0 + 0 + 1 + 0 + 0 = 2$
- $\gcd(2, 6) = 2 > 1$, *reject*

These examples illustrate that many outcomes are rejected when n has multiple prime factors.

3. Characteristic Function Proof of Uniformity

The algorithm admits a natural algebraic interpretation. Consider the random variable $Y_i \in \{\overline{0}, \overline{1}\}$ where $\overline{0}$ represents tails and $\overline{1}$ represents heads. Then, the n flips correspond to a random vector $(Y_0, \ldots, Y_{n-1})$ in $\mathbb{Z}_n^n$. The head count condition $\gcd(N, n) = 1$ ensures that the sum $\sum_{i=0}^{n-1} Y_i$ generates the entire group $\mathbb{Z}_n$, while the rank sum condition $R \mod n$ provides a linear combination that, under this coprimality condition, yields uniform distribution.

Theorem 1. *Let $X_0, \ldots, X_{n-1}$ be identically distributed Bernoulli variables with the probability of success $p \in (0, 1)$ and the probability of failure $q = 1 - p$. Let $N = \sum_{i=0}^{n-1} X_i$ and $R = \sum_{i=0}^{n-1} iX_i$. For any integer c such that $\gcd(c, n) = 1$, the conditional distribution of $R \pmod n$ given $N = c$ is uniform on $\{0, \ldots, n-1\}$.*

3.1. Refined Proof: Group Action Method

Proof. Let $\Omega = \{0, 1\}^n$ be the sample space of outcomes $\mathbf{x} = (x_0, \ldots, x_{n-1})$. Consider the cyclic shift operator $T : \Omega \to \Omega$ defined by

$$T(x_0, \ldots, x_{n-1}) = (x_1, \ldots, x_{n-1}, x_0).$$

Since the random variables $X_0, \ldots, X_{n-1}$ are i.i.d., the probability measure P on Ω is invariant under T; that is, $P(\mathbf{X} = \mathbf{x}) = P(\mathbf{X} = T\mathbf{x})$ for all $\mathbf{x} \in \Omega$.

We examine the behavior of the statistics N and R under this transformation. The total count is invariant:

$$N(T\mathbf{x}) = \sum_{i=0}^{n-1} x_i = N(\mathbf{x}).$$

For the rank sum R, observing that indices shift by $-1 \pmod n$, we compute:

$$\begin{aligned} R(T\mathbf{x}) &= \sum_{i=0}^{n-2} i x_{i+1} + (n-1)x_0 \\ &= \sum_{j=1}^{n-1} (j-1)x_j + (n-1)x_0 = \sum_{j=1}^{n-1} j x_j - \sum_{j=1}^{n-1} x_j + (n-1)x_0 \\ &= \sum_{j=0}^{n-1} j x_j - (N(\mathbf{x}) - x_0) + n x_0 - x_0 \\ &= R(\mathbf{x}) - N(\mathbf{x}) + n\, x_0 \\ &\equiv R(\mathbf{x}) - N(\mathbf{x}) \pmod n. \end{aligned}$$

By induction, for any $k \in \mathbb{Z}_{\geq 0}$, the k-th iteration of the shift satisfies

$$R(T^k \mathbf{x}) \equiv R(\mathbf{x}) - kN(\mathbf{x}) \pmod{n}.$$

Consider the set of outcomes $S_c = \{\mathbf{x} \in \Omega : N(\mathbf{x}) = c\}$ for a fixed count c satisfying $\gcd(c, n) = 1$. The operator T partitions S_c into disjoint orbits. For any $\mathbf{x} \in S_c$, the orbit is $\mathcal{O}_\mathbf{x} = \{\mathbf{x}, T\mathbf{x}, \ldots, T^{n-1}\mathbf{x}\}$.

Because $\gcd(c, n) = 1$, the map $k \mapsto -kc \pmod{n}$ is a bijection on $\mathbb{Z}_n$. Consequently, the set of values $\{R(\mathbf{y}) \pmod{n} : \mathbf{y} \in \mathcal{O}_\mathbf{x}\}$ is exactly $\{R(\mathbf{x}) - kc \pmod{n}\}_{k=0}^{n-1}$, which is a permutation of $\{0, \ldots, n-1\}$.

Since the measure is invariant under T, every element in a given orbit $\mathcal{O}_\mathbf{x}$ has equal probability. Thus, conditional on falling into a specific orbit, $R \pmod{n}$ is uniform. Since S_c is a disjoint union of such orbits, $R \pmod{n}$ is uniform on S_c. □

This algebraic perspective connects the algorithm to deeper number-theoretic structures and motivates the characteristic function approach developed. We now use roots of unity to characterize the discrete uniform distribution on the set $\{0, 1, \ldots, n-1\}$.

Lemma 1. *Let Z be a random variable supported on $\{0, 1, \ldots, n-1\}$ and write*

$$\phi_Z(t) \;=\; \mathbb{E}\Big[e^{itZ}\Big], \qquad t \in \mathbb{R}.$$

Let $\omega := e^{2\pi i/n}$ and define the discrete Fourier coefficients

$$\widehat{p}(m) \;:=\; \mathbb{E}\Big[\omega^{mZ}\Big] \;=\; \phi_Z\big(\tfrac{2\pi m}{n}\big), \qquad m = 0, 1, \ldots, n-1.$$

Then the following are equivalent:

1. *Z is uniformly distributed on $\{0, 1, \ldots, n-1\}$.*
2. *$\widehat{p}(0) = 1$ and $\widehat{p}(m) = 0$ for all $m = 1, \ldots, n-1$; equivalently*

$$\phi_Z\big(\tfrac{2\pi m}{n}\big) = \begin{cases} 1, & m = 0, \\ 0, & m = 1, \ldots, n-1. \end{cases}$$

Proof. *(1 ⇒ 2).* If Z is uniform on $\{0, 1, \ldots, n-1\}$, then $P(Z = k) = 1/n$ for all k and

$$\widehat{p}(m) = \frac{1}{n} \sum_{k=0}^{n-1} \omega^{mk}.$$

For $m = 0$ this is 1. For $m = 1, \ldots, n-1$ we use the geometric series:

$$\sum_{k=0}^{n-1} \omega^{mk} = \frac{1 - (\omega^m)^n}{1 - \omega^m} = \frac{1 - 1}{1 - \omega^m} = 0,$$

so $\widehat{p}(m) = 0$. Because $\widehat{p}(m) = \phi_Z(2\pi m/n)$, the stated values of ϕ_Z follow.

(2 ⇒ 1). Let $p_k = P(Z = k)$, $k = 0, \ldots, n-1$, and consider the discrete Fourier transform (DFT)

$$\widehat{p}(m) = \sum_{k=0}^{n-1} p_k \, \omega^{mk}, \qquad m = 0, \ldots, n-1.$$

On the finite abelian group $\mathbb{Z}_n$, the DFT is invertible with the inversion formula

$$p_k = \frac{1}{n} \sum_{m=0}^{n-1} \widehat{p}(m) \, \omega^{-mk}, \qquad k = 0, \ldots, n-1, \tag{3}$$

which follows from the orthogonality of nth roots of unity:

$$\frac{1}{n}\sum_{m=0}^{n-1}\omega^{m(k-\ell)} = \begin{cases} 1, & k=\ell, \\ 0, & k\neq\ell. \end{cases}$$

Now apply (3) with the assumption $\widehat{p}(0)=1$ and $\widehat{p}(m)=0$ for $m\geq 1$:

$$p_k = \frac{1}{n}\Big(\widehat{p}(0)\,\omega^0 + \sum_{m=1}^{n-1}\widehat{p}(m)\,\omega^{-mk}\Big) = \frac{1}{n}\,.$$

Hence $p_k = 1/n$ for all k, i.e., Z is uniform on $\{0,1,\ldots,n-1\}$. □

While the group-action proof is combinatorially elegant, the characteristic-function approach reveals directly how $\gcd(N,n)=1$ forces the Fourier coefficients to vanish. The argument proceeds as follows.

3.2. New Proof: Characteristic Function Method

Proof. Let $\omega = e^{2\pi i/n}$. To establish uniformity on $\mathbb{Z}_n$, it suffices to show that the conditional characteristic function of R vanishes at all non-zero frequencies. Specifically, we must show that for all $k\in\{1,\ldots,n-1\}$:

$$\mathbb{E}[\omega^{kR} \mid N=c] = 0.$$

Since $gcd(c,n)=1$ implies $1\leq c\leq n-1$ and $p\in(0,1)$, we have $P(N=c) = \binom{n}{c}p^c(1-p)^c > 0$. Hence the conditional expectation is well-defined by

$$\frac{\mathbb{E}[\omega^{kR}\mathbf{1}_{\{N=c\}}]}{P(N=c)}.$$

We analyze the numerator. Consider the joint characteristic function (CF) of (R,N), defined for $z\in\mathbb{C}$ as:

$$\Psi_k(z) = \mathbb{E}\Big[\omega^{kR}z^N\Big] = \mathbb{E}\left[\prod_{j=0}^{n-1}\omega^{kjX_j}z^{X_j}\right].$$

Since $|\omega^{kR}z^N| \leq |z|^N$ and $\mathbb{E}[|z|^N] = (1-p+p|z|)^n < \infty$ for $|z|\leq 1$, the series converges absolutely and coefficient extraction is valid.

By independence, the expectation factors:

$$\Psi_k(z) = \prod_{j=0}^{n-1}\mathbb{E}\Big[(\omega^{kj}z)^{X_j}\Big] = \prod_{j=0}^{n-1}(q+pz\omega^{kj}).$$

where $p = P(X_j=1)$ and $q=1-p$.

Let $d_k = \gcd(k,n)$ and $m_k = n/d_k$. The sequence of exponents $\{kj \pmod n\}_{j=0}^{n-1}$ traverses the values $\{0,d_k,2d_k,\ldots,(m_k-1)d_k\}$, each repeated d_k times. With $\eta=\omega^{d_k}$, a primitive m_k-th root of unity,

$$\Psi_k(z) = \left\{\prod_{r=0}^{m_k-1}(q+pz\eta^r)\right\}^{d_k} = \{q^{m_k}-(-pz)^{m_k}\}^{d_k}.$$

Therefore $\Psi_k(z)$ is a polynomial in z^{m_k}. Consequently $[z^c]\Psi_k(z) = 0$ unless $m_k \mid c$. If $1 \le k \le n-1$, then $m_k > 1$ and $m_k \mid n$; since $\gcd(c,n) = 1$, $m_k \nmid c$. Hence $[z^c]\Psi_k(z) = 0$, and the nonzero Fourier coefficients vanish. We have:

$$\mathbb{E}[\omega^{kR}\mathbf{1}_{\{N=c\}}] = [z^c]\Psi_k(z).$$

where the notation $[z^c]\Psi_k(z)$ is the coefficient of the term z^c in the series expansion of $\Psi_k(z)$. Since

$$\mathbb{E}[\omega^{kR}\mathbf{1}_{\{N=c\}}] = \sum_r \sum_m P(R=r, N=m)\cdot\omega^{kr}\cdot\mathbf{1}_{\{m=c\}} = \sum_r P(R=r, N=c)\cdot\omega^{kr}\cdot 1$$

Given the hypothesis $\gcd(c,n) = 1$, c cannot be a multiple of $m_k = n/\gcd(k,n)$ unless $m_k = 1$. If $k \in \{1,\dots,n-1\}$, then $\gcd(k,n) < n$, implying $m_k > 1$. Since m_k is a divisor of n greater than 1, and $\gcd(c,n) = 1$, it follows that $m_k \nmid c$. Therefore, $[z^c]\Psi_k(z) = 0$ for all $k \not\equiv 0 \pmod{n}$.

The conditional characteristic function is thus:

$$\mathbb{E}[\omega^{kR} \mid N = c] = \begin{cases} 1 & \text{if } k \equiv 0 \pmod{n}, \\ 0 & \text{if } k \not\equiv 0 \pmod{n}. \end{cases}$$

By the uniqueness of the CF, R is uniformly distributed on $\{0,\dots,n-1\}$ conditional on $N = c$. □

4. Complexity Analysis of Algorithm $\mathcal{A}_3$

We analyze the sample complexity of Algorithm $\mathcal{A}_3$, defined as the expected number of coin flips required to generate a single output sample.

Theorem 2 (Expected Complexity)**.** *Let $p \in (0,1)$ be the success probability of the biased coin. Let T_n be the random variable representing the total number of coin flips performed by Algorithm $\mathcal{A}_3$ to generate one output in $\{0,\dots,n-1\}$.*

The expected number of flips satisfies the following asymptotic bound as $n \to \infty$:

$$\mathbb{E}[T_n] = O(n \log\log n).$$

Furthermore, if n is prime, the complexity is strictly linear:

$$\mathbb{E}[T_n] = n + o(1).$$

Proof. Algorithm $\mathcal{A}_3$ operates in independent trials. In each trial, a batch of n coin flips is generated. Let $\mathbf{X}^{(n)} = (X_0,\dots,X_{n-1})$ be the sequence of outcomes in a single batch, and let $N = \sum_{i=0}^{n-1} X_i$ be the number of heads. The batch is accepted if and only if $\gcd(N,n) = 1$.

Let α_n denote the probability of acceptance in a single trial:

$$\alpha_n = P(\gcd(N,n) = 1) = \sum_{\substack{0 \le k \le n \\ \gcd(k,n)=1}} \binom{n}{k} p^k (1-p)^{n-k}.$$

Since trials are independent, the number of trials M required to obtain a valid output follows a Geometric distribution with parameter α_n. The total number of flips is $T_n = n \cdot M$. By the properties of the Geometric distribution ($\mathbb{E}[M] = 1/\alpha_n$), the expected number of flips is:

$$\mathbb{E}[T_n] = \frac{n}{\alpha_n}.$$

Case 1: n is Prime

If n is prime, $\gcd(k, n) = 1$ for all $k \in \{1, \ldots, n-1\}$. The condition fails only if $N = 0$ or $N = n$.

$$\alpha_n = 1 - P(N = 0) - P(N = n) = 1 - (1-p)^n - p^n.$$

For any fixed $p \in (0, 1)$, the terms $(1-p)^n$ and p^n decay exponentially to 0. Thus, $\lim_{n\to\infty} \alpha_n = 1$.

$$\mathbb{E}[T_n] \approx \frac{n}{1} = O(n).$$

Case 2: n is Composite (General Case)

We estimate α_n for composite n. As $n \to \infty$, the distribution of the sum N modulo d approaches a uniform distribution on $\{0, \ldots, d-1\}$ when d is any fixed divisor of n.

The probability that N is coprime to n behaves asymptotically like the density of integers coprime to n in the range $[0, n-1]$ and this density is given by Euler's totient function ratio $\frac{\phi(n)}{n}$:

$$\alpha_n = \frac{\phi(n)}{n} + r_n, \qquad |r_n| \le C_p \frac{\tau(n)}{\sqrt{n}}.$$

Thus, for fixed $p \in (0, 1)$ and sufficiently large n, $\alpha_n \ge \frac{1}{2}\phi(n)/n$. This is the only estimate needed for the complexity upper bound.

The relevant lemmas and proofs can be found in Appendix A. Substituting this into the expectation:

$$\mathbb{E}[T_n] = \frac{n}{\alpha_n} \le \frac{2n}{\phi(n)/n} = \frac{2n^2}{\phi(n)}.$$

By Landau's theorem (see Theorem 328 of [10]), $n/\phi(n) = O(\log\log n)$, hence

$$\mathbb{E}[T_n] = O(n \log\log n).$$

The prime case remains sharper. If n is prime, the only rejected head counts are 0 and n, so

$$\alpha_n = 1 - q^n - p^n, \qquad \mathbb{E}[T_n] = \frac{n}{1 - q^n - p^n} = n + O(n(p^n + q^n)) = n + o(1).$$

□

According to Theorem 2, when n is prime, the expected complexity of Algorithm $\mathcal{A}_3$ achieves a strict linear bound $\mathbb{E}[T_n] = n + o(1)$. This implies that the algorithm's acceptance probability α_n is approximately $1/n$ (as derived from $T_n = n/\alpha_n$), thereby achieving an optimal linear time cost for generating a single output.

For the composite case, our analysis using Fourier methods and number theory demonstrates that the acceptance probability satisfies $\alpha_n = \frac{\phi(n)}{n} + O_p\left(\frac{\tau(n)}{\sqrt{n}}\right)$, where $\phi(n)$ is Euler's totient function and $\tau(n)$ is the divisor function. By leveraging the properties $\tau(n) = n^{o(1)}$ and $n/\phi(n) = O(\log\log n)$, we establish the lower bound $\alpha_n \ge \frac{1}{2} \cdot \frac{\phi(n)}{n}$. This directly leads to the expected complexity bound of $O(n \log\log n)$.

This refined analysis not only explains the root cause of the complexity difference between prime and composite cases but also rigorously demonstrates, through a proven lower bound on the acceptance probability, that the $O(n \log\log n)$ complexity for composite n is indeed tight.

Remark 1. *The appearance of $\phi(n)/n$ has a simple interpretation. A batch is accepted precisely when its head count lies in a residue class coprime to n. Among the n residues modulo n, exactly $\phi(n)$*

are coprime to n. Thus $\phi(n)/n$ is the ideal acceptance density. The Moebius–Fourier calculation quantifies the deviation of the binomial head-count distribution from this ideal residue density.

5. Continuous Complexity Analysis of Algorithm $\mathcal{A}_3$

Having established the strict uniformity of the algorithm through a purely analytic characteristic function proof and derived tight asymptotic bounds on its expected runtime, we proceed to extend the method to the continuous uniform distribution in this section. We utilize the discrete extraction method from Theorem 3 to construct a random variable U on the continuous interval $[0, 1]$.

Definition 4 (Base-n Digit Generator)**.** *Let $\mathcal{S}$ be a stream of i.i.d. Bernoulli trials. We parse $\mathcal{S}$ into disjoint blocks of length n. For each block j, we compute the total sum N_j and rank sum R_j.*

- *If $\gcd(N_j, n) = 1$, we accept the block and output the digit $D_k = R_j \pmod n$.*
- *If $\gcd(N_j, n) \neq 1$, we discard the block.*

Let $D_1, D_2, \ldots$ be the sequence of successfully generated digits. By Theorem 3 and the independence of the blocks, $\{D_k\}_{k\geq 1}$ is a sequence of i.i.d. variables uniformly distributed on $\{0, 1, \ldots, n-1\}$.

Base-n expansions are not unique on the countable set of n-adic rationals, where a terminating expansion can also be written with an eventually all-$(n-1)$ tail. We adopt the canonical convention that excludes the eventually all-$(n-1)$ representation. The digit sequence generated by independent accepted A3 blocks is almost surely not eventually constant equal to $n-1$, so the ambiguity has probability zero and does not affect the distributional conclusion.

Theorem 3 (Continuous Uniformity)**.** *Let U be the random variable defined by the base-n expansion using the digits generated above:*

$$U = \sum_{k=1}^{\infty} \frac{D_k}{n^k}.$$

Then U follows a continuous uniform distribution on $[0, 1]$, i.e., $U \sim Uniform[0, 1]$.

Proof. We use the method of Characteristic Functions. The characteristic function of the target distribution Uniform$[0, 1]$ is:

$$\phi_{\text{target}}(t) = \int_0^1 e^{itx}\, dx = \frac{e^{it} - 1}{it}.$$

We must show that $\phi_U(t) = \mathbb{E}[e^{itU}]$ is equal to $\phi_{\text{target}}(t)$.

Case 1: Characteristic Function of a Single Digit

Since D_k is uniform on $\{0, \ldots, n-1\}$, its characteristic function is:

$$\mathbb{E}[e^{i\xi D_k}] = \frac{1}{n} \sum_{j=0}^{n-1} e^{i\xi j} = \frac{1}{n} \frac{e^{in\xi} - 1}{e^{i\xi} - 1}.$$

Case 2: Characteristic Function of the Partial Sum

Let $U_M = \sum_{k=1}^{M} D_k n^{-k}$ be the partial sum of the first M digits. Since the digits are independent, the characteristic function of the sum is the product of the characteristic functions of the individual terms. The k-th term is $X_k = D_k n^{-k}$.

$$\phi_{U_M}(t) = \mathbb{E}\left[\exp\left(it\sum_{k=1}^{M}\frac{D_k}{n^k}\right)\right] = \prod_{k=1}^{M}\mathbb{E}\left[e^{i(t/n^k)D_k}\right].$$

Substituting $\xi = t/n^k$ into the single digit formula:

$$\mathbb{E}\left[e^{i(t/n^k)D_k}\right] = \frac{1}{n}\frac{e^{in(t/n^k)}-1}{e^{i(t/n^k)}-1} = \frac{1}{n}\frac{e^{it/n^{k-1}}-1}{e^{it/n^k}-1}.$$

The preceding calculation assumes $t \neq 0$ because the displayed telescoping quotient contains $e^{it/n^M} - 1$ in the denominator. The point $t = 0$ is immediate: for every M, $\phi_{U_M}(0) = 1$, and the characteristic function of Uniform$[0,1]$ also equals 1 at zero. We therefore define $(e^{it}-1)/(it)$ at $t = 0$ by continuity. Hence the characteristic functions converge for all $t \in \mathbb{R}$.

The product over $k = 1$ to M telescopes:

$$\begin{aligned}\phi_{U_M}(t) &= \left(\frac{1}{n}\frac{e^{it}-1}{e^{it/n}-1}\right)\cdot\left(\frac{1}{n}\frac{e^{it/n}-1}{e^{it/n^2}-1}\right)\cdots\left(\frac{1}{n}\frac{e^{it/n^{M-1}}-1}{e^{it/n^M}-1}\right)\\ &= \frac{1}{n^M}\frac{e^{it}-1}{e^{it/n^M}-1}.\end{aligned}$$

Case 3: Convergence to the Continuum

We take the limit as $M \to \infty$. Consider the denominator term $e^{it/n^M} - 1$. Using the Taylor expansion $e^x \approx 1 + x$ for small x:

$$e^{it/n^M} - 1 \sim \frac{it}{n^M} \quad \text{as } M \to \infty.$$

Substituting this back into the expression for $\phi_{U_M}(t)$:

$$\lim_{M\to\infty}\phi_{U_M}(t) = \lim_{M\to\infty}\frac{1}{n^M}\frac{e^{it}-1}{(it/n^M)} = \frac{e^{it}-1}{it}.$$

The limit matches $\phi_{\text{target}}(t)$. By the Lévy Continuity Theorem, U_M converges in distribution to U, and U is uniformly distributed on $[0,1]$. □

For finite-precision simulation, the generator outputs $U_M = \sum_{k=1}^{M} D_k n^{-k}$. Since each digit requires one accepted A3 block and each block has acceptance probability α_n, the expected number of biased flips is Mn/α_n. Hence the finite-precision runtime is $Mn(1+o(1))$ for prime n and $O(Mn\log\log n)$ for general n. Under the canonical coupling with the infinite expansion, $|U - U_M| \leq n^{-M}$.

In fact, let $(D_k)_{k\geq 1}$ be the i.i.d. digits generated by accepted A3 blocks. Define

$$U = \sum_{k=1}^{\infty} D_k n^{-k}, \qquad U_M = \sum_{k=1}^{M} D_k n^{-k}.$$

Then $U \sim$ Uniform$[0,1]$ and

$$0 \leq U - U_M \leq \sum_{k=M+1}^{\infty}(n-1)n^{-k} = n^{-M}.$$

The expected number of biased coin flips required to generate U_M is

$$\mathbb{E}[T_{n,M}] = M\frac{n}{\alpha_n},$$

which is $Mn(1+o(1))$ for prime n and $O(Mn\log\log n)$ for general n.

The algorithm does not require storing the entire block. During the scan of a block, N and R mod n can be updated incrementally. Computing $\gcd(N,n)$ by the Euclidean algorithm adds only logarithmic arithmetic overhead. Therefore the dominant cost remains the generation of Bernoulli trials, while the modular arithmetic and gcd check are standard low-order implementation costs.

6. Conclusions

This paper has provided a comprehensive theoretical treatment of Lei's algorithm $\mathcal{A}_3$ for extracting exact discrete uniform distributions from biased Bernoulli trials, and has extended its reach to continuous uniform generation. Three primary contributions were presented. First, in terms of algorithm validity verification, we abandon the traditional group-theoretic proof framework and propose a novel analytic proof based on characteristic functions and roots of unity. Second, regarding sample complexity analysis, we derive the first explicit asymptotic bounds for Algorithm $\mathcal{A}_3$ by combining properties of the binomial distribution, Euler's totient function $\phi(n)$, and classical results from analytic number theory. Third, in the extension to continuous uniform distributions, we construct a Base-n Digit Generator using the discrete uniform digits output by $\mathcal{A}_3$.

From a methodological standpoint, the characteristic-function perspective not only simplifies the proof but also suggests possible generalizations to other group structures. The complexity analysis, albeit asymptotic, confirms that the rejection penalty for composite moduli is mild—only a logarithmic factor—and is unlikely to hinder practical deployment.

Future research directions can focus on three aspects: first, optimizing the rejection sampling strategy to reduce the discard rate of composite n, especially for large n with multiple prime factors; second, extending the algorithm to non-i.i.d. Bernoulli sources (e.g., sources with weak dependence) while maintaining uniformity guarantees; third, exploring the application of $\mathcal{A}_3$ in privacy-preserving machine learning, such as ensuring fairness in federated learning systems where local hardware biases may compromise global model equity.

Author Contributions: Conceptualization, M.Z. and X.L.; methodology, M.Z.; software, M.Z. and G.T.; validation, M.Z. and G.T.; formal analysis, M.Z. and G.T.; investigation, M.Z.; resources, M.Z.; data curation, M.Z. and G.T.; writing—original draft preparation, M.Z.; writing—review and editing, M.Z. and G.T.; visualization, NONE; supervision, X.L.; project administration, M.Z.; funding acquisition, NONE. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Definition A1 (Möbius function)**.** *The* Möbius function $\mu:\mathbb{N}\to\{-1,0,1\}$ *is defined by*

$$\mu(1)=1,\qquad \mu(d)=\begin{cases}0, & \text{if } d \text{ is divisible by } p^2 \text{ for some prime } p,\\ (-1)^k, & \text{if } d \text{ is a product of } k \text{ distinct primes.}\end{cases}$$

Equivalently, $\mu(d)=0$ *if* d *is not square-free, and if* d *is square-free then* $\mu(d)=(-1)^{\omega(d)}$*, where* $\omega(d)$ *denotes the number of distinct prime divisors of* d*.*

Lemma A1 (Acceptance probability via Möbius inversion and Fourier bounds)**.** *Let $n \geq 2$, and let $N \sim \mathrm{Bin}(n, p)$ with $p \in (0,1)$ and $q := 1 - p$. Define the acceptance probability*

$$\alpha_n \;:=\; P\big(\gcd(N,n) = 1\big).$$

Let $\mu(\cdot)$ be the Möbius function, $\varphi(\cdot)$ Euler's totient, and $\tau(n) := \#\{d : d \mid n\}$ the divisor-counting function. Then

$$\alpha_n \;=\; \frac{\varphi(n)}{n} \;+\; r_n, \qquad \text{with} \qquad |r_n| \;\leq\; \frac{\sqrt{\pi}}{4\sqrt{pq}} \cdot \frac{\tau(n)}{\sqrt{n}}.$$

In particular, for all sufficiently large n (depending on p),

$$\alpha_n \;\geq\; \frac{1}{2} \cdot \frac{\varphi(n)}{n}.$$

Proof. Step 1: Möbius inversion for the coprimality indicator.

For integers $k, n \geq 1$ one has the standard identity

$$\mathbf{1}\{\gcd(k,n) = 1\} \;=\; \sum_{d \mid \gcd(k,n)} \mu(d) \;=\; \sum_{d \mid n} \mu(d)\, \mathbf{1}\{d \mid k\}.$$

This identity is also valid for $k = 0$. It is important to note the handling of $gcd(0, n) = n$. Taking expectation with $k = N$ gives

$$\alpha_n \;=\; \mathbb{E}\big[\mathbf{1}\{\gcd(N,n) = 1\}\big] \;=\; \sum_{d \mid n} \mu(d)\, P(d \mid N). \tag{A1}$$

Step 2: Exact Fourier formula for $P(d \mid N)$.

Fix $d \geq 2$. Let $\zeta_d := e^{2\pi i/d}$. By the discrete Fourier inversion on $\mathbb{Z}_d$,

$$\mathbf{1}\{d \mid N\} \;=\; \frac{1}{d} \sum_{m=0}^{d-1} \zeta_d^{mN}.$$

Taking expectation yields

$$P(d \mid N) \;=\; \frac{1}{d} \sum_{m=0}^{d-1} \mathbb{E}[\zeta_d^{mN}] \;=\; \frac{1}{d} \sum_{m=0}^{d-1} \big(q + p\,\zeta_d^m\big)^n, \tag{A2}$$

since $\mathbb{E}[z^N] = (q + pz)^n$ for $N \sim \mathrm{Bin}(n, p)$.

From (A2),

$$P(d \mid N) - \frac{1}{d} \;=\; \frac{1}{d} \sum_{m=1}^{d-1} \big(q + p\,\zeta_d^m\big)^n,$$

so by the triangle inequality,

$$\left| P(d \mid N) - \frac{1}{d} \right| \;\leq\; \frac{1}{d} \sum_{m=1}^{d-1} \left| q + p\,\zeta_d^m \right|^n. \tag{A3}$$

Step 3: A uniform bound on $\left|q + pe^{i\theta}\right|^n$.

For $\theta \in \mathbb{R}$,

$$|q + pe^{i\theta}|^2 = q^2 + p^2 + 2pq\cos\theta = 1 - 2pq(1 - \cos\theta).$$

https://doi.org/10.3390/e28070817

Using $1 - x \le e^{-x}$ and the inequality $1 - \cos\theta \ge 2\theta^2/\pi^2$ for $\theta \in [0, \pi]$, we obtain for $\theta \in [0, \pi]$,

$$|q + pe^{i\theta}|^2 \le \exp\Big(-2pq(1 - \cos\theta)\Big) \le \exp\Big(-4pq\,\theta^2/\pi^2\Big),$$

hence

$$|q + pe^{i\theta}| \le \exp\Big(-2pq\,\theta^2/\pi^2\Big), \qquad \theta \in [0, \pi]. \tag{A4}$$

Now in (A3), pair m with $d - m$ to reduce to angles in $[0, \pi]$:

$$\sum_{m=1}^{d-1} \left|q + p\zeta_d^m\right|^n = 2 \sum_{m=1}^{\lfloor (d-1)/2 \rfloor} \left|q + pe^{2\pi i m/d}\right|^n,$$

and for $1 \le m \le \lfloor (d-1)/2 \rfloor$, $\theta_m := 2\pi m/d \in (0, \pi]$. Applying (A4) yields

$$\left|q + pe^{2\pi i m/d}\right|^n \le \exp\Big(-2pq\,n\,\theta_m^2/\pi^2\Big) = \exp\Big(-2pq\,n\,(2\pi m/d)^2/\pi^2\Big) = \exp\Big(-8pq\,n\,m^2/d^2\Big).$$

Therefore,

$$\sum_{m=1}^{d-1} \left|q + p\zeta_d^m\right|^n \le 2 \sum_{m=1}^{\infty} \exp\Big(-8pq\,n\,m^2/d^2\Big) \le 2 \int_0^{\infty} \exp\Big(-8pq\,n\,x^2/d^2\Big)\,dx.$$

Evaluating the Gaussian integral gives

$$2 \int_0^{\infty} e^{-8pq\,n\,x^2/d^2}\,dx = 2 \cdot \frac{d}{2} \sqrt{\frac{\pi}{8pq\,n}} = d \cdot \frac{\sqrt{\pi}}{\sqrt{8pq\,n}}.$$

Plugging this into (A3) yields the uniform bound independent of d

$$\left|P(d \mid N) - \frac{1}{d}\right| \le \frac{1}{d} \cdot d \cdot \frac{\sqrt{\pi}}{\sqrt{8pq\,n}} = \frac{\sqrt{\pi}}{\sqrt{8pq}} \cdot \frac{1}{\sqrt{n}} \le \frac{\sqrt{\pi}}{4\sqrt{pq}} \cdot \frac{1}{\sqrt{n}}. \tag{A5}$$

Step 4: Summation over divisors of n.

Decompose (A1) as

$$\alpha_n = \sum_{d|n} \frac{\mu(d)}{d} + \sum_{d|n} \mu(d) \Big(P(d \mid N) - \frac{1}{d}\Big).$$

The first sum equals $\varphi(n)/n$ (a standard identity). For the remainder, use $|\mu(d)| \le 1$ and (A5):

$$|r_n| \le \sum_{d|n} \left|P(d \mid N) - \frac{1}{d}\right| \le \tau(n) \cdot \frac{\sqrt{\pi}}{4\sqrt{pq}} \cdot \frac{1}{\sqrt{n}}.$$

This proves the stated error bound. Finally, since $\tau(n) = n^{o(1)}$, the bound implies $|r_n| = o(1)$ as $n \to \infty$, and in particular $|r_n| \le \frac{1}{2}\varphi(n)/n$ for all sufficiently large n (depending on p), which yields $\alpha_n \ge \frac{1}{2}\,\varphi(n)/n$. □

References

1. Diaconis, P.; Holmes, S.; Montgomery, R. Dynamical bias in the coin toss. *SIAM Rev.* **2007**, *49*, 211–235. [CrossRef]
2. Lei, X. Generating discrete uniform distribution from a biased coin using number-theoretic method. *Electron. Commun. Probab.* **2023**, *28*, 1–6. [CrossRef]
3. von Neumann, J. Various techniques used in connection with random digits. *Natl. Bur. Stand. Appl. Math. Ser.* **1951**, *12*, 36–38.
4. Elias, P. Universally acceptable information sources. *IEEE Trans. Inf. Theory* **1962**, *8*, 213–219.
5. Knuth, D.E.; Yao, A.C. The complexity of nonuniform random number generation. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS 1976)*; IEEE: New York, NY, USA, 1976; pp. 323–333.

6. Peres, Y. Iterating von Neumann's procedure for extracting random bits. *Ann. Stat.* **1992**, *20*, 590–597. [CrossRef]
7. Dijkstra, E.W. Making a fair roulette from a possibly biased coin. *Inf. Process. Lett.* **1990**, *36*, 193. [CrossRef]
8. Zhou, P.; Wei, H.; Zhang, H. Selective Reviews of Bandit Problems in AI via a Statistical View. *Mathematics* **2025**, *13*, 665. [CrossRef]
9. Li, X.; Ruan, F.; Wang, H.; Long, Q.; Su, W.J. A statistical framework of watermarks for large language models: Pivot, detection efficiency and optimal rules. *Ann. Stat.* **2025**, *53*, 322–351. [CrossRef] [PubMed]
10. Hardy, G.H.; Wright, E.M. *An Introduction to the Theory of Numbers*; Oxford University Press: Oxford, UK, 1979.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.