Electronic Colloquium on Computational Complexity, Revision 3 of Report No. 130 (2026)

# Optimal Bitwise Pseudorandom Generators for Modular Sums Modulo Every Fixed Integer

Gonen Krak[†]

**Abstract**

We study pseudorandom bit generators for Boolean linear sums modulo a fixed integer $M$. A source $X \in \{0,1\}^n$ succeeds with error $\varepsilon$ if, for every $a \in \mathbb{Z}_M^n$, the residue

$$\sum_{i=1}^{n} a_i X_i \pmod{M}$$

is $\varepsilon$-close in statistical distance to the corresponding residue under independent uniform bits. Lovett, Reingold, Trevisan, and Vadhan obtained the optimal seed length for prime-power moduli. We prove the same bound for every fixed modulus:

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The proof builds the generator in stages. It keeps two quantities separate: the prime field $\mathbb{F}_p$ containing the generator's output, and the modulus $m$ used by the test. A state $(p, m)$ means that we have a PRG on $\mathbb{F}_p^n$ that fools every modulus-$m$ phase whose coordinates contribute separately.

We use two steps to move between states. The first step, $\mathrm{LIFT}_p$, changes the modulus from $m$ to $pm$. The second step, $\mathrm{TRANSFER}_{r\to p}$, changes the alphabet from $\mathbb{F}_r$ to a smaller field $\mathbb{F}_p$.

Starting from the trivial modulus 1, we process the prime factors of $M$ until we reach the Boolean state $(2, M)$.

[†]Tel Aviv University. gonenok10@gmail.com



ISSN 1433-8092

# Contents

# 1 Introduction

A distribution over $\{0,1\}^n$ is *pseudorandom* for a class of tests if no test in the class can distinguish it from $n$ independent uniform bits. The smaller the class, the shorter the seed one may hope for, and for the simplest nontrivial class — linear tests modulo 2 — the problem was solved in the seminal work of Naor and Naor [NN93]: *small-bias generators* fool all parities with error $\varepsilon$ using a seed of length $O(\log n + \log \frac{1}{\varepsilon})$, which is optimal up to constant factors [AGHP92]. Small-bias generators have since become a basic primitive of pseudorandomness, and a natural next question is what happens when the modulus 2 is replaced by an arbitrary fixed modulus $M$.

Concretely, for a coefficient vector $a \in \mathbb{Z}_M^n$ consider the modular sum

$$S_a^{(M)}(x) = \sum_{i=1}^{n} a_i x_i \pmod{M}, \qquad x \in \{0,1\}^n.$$

A bit distribution $X$ *$\varepsilon$-fools Boolean linear sums modulo $M$* if, for every $a \in \mathbb{Z}_M^n$, the distribution of $S_a^{(M)}(X)$ is $\varepsilon$-close in statistical distance to that of $S_a^{(M)}(U_n)$. Modular sums are among the simplest tests computable in logarithmic space; they are exactly the read-once products over the cyclic group $\mathbb{Z}_M$, that is, tests of the form $\prod_i g_i^{x_i}$ with $g_i \in \mathbb{Z}_M$; and fooling them is a necessary step in several programs on pseudorandomness for group products and read-once branching programs.

Lovett, Reingold, Trevisan, and Vadhan [LRTV09] constructed generators with the optimal seed length $O_M(\log n + \log \frac{1}{\varepsilon})$ for every fixed *prime-power* modulus $M = p^k$. Their second construction works for every $M$, at seed length

$$O\left(\log n + \log \frac{M}{\varepsilon} \cdot \log\left(M \log \frac{1}{\varepsilon}\right)\right),$$

which for fixed $M$ is $O_M\big(\log n + \log \frac{1}{\varepsilon} \cdot \log\log \frac{1}{\varepsilon}\big)$: optimal in $n$, but short of optimal by a $\log\log \frac{1}{\varepsilon}$ factor. In the years since, the composite case has resisted matching the prime-power one. The recent work of Lee and Viola [LV25], which achieves optimal seed length $c_G \log(n/\varepsilon)$ for read-once group products over every $p$-group, explicitly identifies the handling of composite moduli as a challenge reminiscent of a known roadblock in circuit complexity. In particular, prior to this work no generator with optimal seed length was known even for $M = 6$, the smallest modulus with two distinct prime factors.

## 1.1 Our result

We remove the restriction to prime powers entirely.

**Theorem 1.1** (Main result; informal)**.** *For every fixed integer $M \geq 2$ there is an explicit generator that $\varepsilon$-fools all Boolean linear sums modulo $M$ with seed length*

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The formal statement is Theorem 7.3. The seed length is optimal up to the constant depending on $M$: any generator fooling linear sums modulo $M$ also fools them modulo any prime divisor of $M$ (Proposition 7.4), where the bound $\Omega(\log n + \log \frac{1}{\varepsilon})$ is standard.

The generator is built from off-the-shelf components. The only pseudorandomness primitive imported is explicit small bias over fixed prime fields [NN93], [AMN98]; everything else, including the base case, is proved here. What the construction does with those components is where the content lies.

**AI-use disclosure.** The exposition and parameter calculations were reorganized and cross-checked with the assistance of ChatGPT and Claude; the author retains responsibility for every statement and proof.

## 1.2 Related work

**Modular sums.** Naor and Naor [NN93] solved the case $M = 2$ optimally. Lovett, Reingold, Trevisan, and Vadhan [LRTV09] gave the two constructions displayed above. Gopalan, Kane, and Meka [GKM18] took a different route, constructing a single generator that fools the discrete Fourier transforms of linear functions and thereby all modular tests with modulus up to $M$ simultaneously, with seed length

$$O\left(\log \frac{Mn}{\varepsilon} \cdot \left(\log \log \frac{Mn}{\varepsilon}\right)^2\right).$$

The strength of their result is its uniformity: one generator for every modulus up to $M$, with near-logarithmic dependence on $M$. For a fixed modulus it does not improve on [LRTV09]. What Theorem 1.1 does is remove the residual $\log \log \frac{1}{\varepsilon}$ loss of the LRTV arbitrary-modulus construction, for every fixed modulus rather than for prime powers alone.

**Group products and branching programs.** Linear sums modulo $M$ are read-once products over the cyclic group $\mathbb{Z}_M$, a commutative special case of read-once permutation branching programs. For group products, Meka and Zuckerman [MZ09] gave generators over $S_3$ with polynomially small error at seed length $O(\log n)$, handling the modulus-3 part through generators for low-degree polynomials over $\mathbb{F}_3$. Lee and Viola [LV25] recently achieved optimal seed length $c_G \log(n/\varepsilon)$ for read-once group products over every $p$-group, and nearly optimal seed length over every commutative group; they point to composite moduli as the boundary of current techniques. Theorem 1.1 gives the exactly optimal seed length for every cyclic group $\mathbb{Z}_M$, including those that are not $p$-groups. By a standard character argument it should also yield the corresponding statement for every finite abelian group, with the error adjusted by a constant depending on the group; we do not pursue that here.

**The Meka–Zuckerman conjecture.** For $M = 6$, the final generator has a familiar form: it is the XOR of a small-bias distribution and a bit distribution that fools all linear sums modulo 3. This is exactly the type of generator that Meka and Zuckerman conjectured would fool width-3 read-once permutation branching programs [MZ09].

Our theorem does not prove that conjecture. Linear tests modulo 6 correspond to

$$\mathbb{Z}_6 \cong \mathbb{Z}_3 \times \mathbb{Z}_2,$$

whereas width-3 permutation programs involve

$$S_3 \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_2.$$

In $S_3$, the modulo-2 part can reverse the modulo-3 part, creating a dependence on earlier input bits that our analysis does not handle.

**Relation to degree descent over $\mathbb{F}_2$.** Our technique resembles a line of work on pseudorandom generators for low-degree polynomials [BV10, Lov09, Vio09]. In those works, Cauchy–Schwarz introduces a directional derivative, each derivative lowers the polynomial degree, and the induction ends at a small-bias base case. Lemma 4.4 uses the same analytic pattern. The algebraic step in this paper also lowers degree, but only after changing the alphabet to the prime field $\mathbb{F}_p$ and reducing one local exponent modulo $p$. Every function on $\mathbb{F}_p$ then has degree at most $p-1$, so $p-1$ finite differences make it constant modulo $p$ and reveal one factor $p$ in the modulus. The overall construction therefore descends in the modulus while the inner algebra descends in degree. The need to change the alphabet is the new feature: no analogous bounded-degree representation over the Boolean field is available for counting modulo an odd prime, as reflected by the Razborov–Smolensky method [Smo87].

**Organization.** Section 2 gives the complete recursive route before any technical machinery is introduced. Section 3 defines the intermediate test classes and the error measure. Section 4 proves LIFT using polynomial degree reduction and a one-step derivative-lifting lemma. Section 5 proves TRANSFER by combining a small-bias estimate for low-weight tests with a cancellation estimate for high-weight tests. Section 6 follows the prime factors of $M$, and Section 7 converts the final phase guarantee into statistical distance and proves optimality.

# 2 The recursive construction plan

The main theorem concerns a Boolean generator for modulus $M$. Before defining phases, derivatives, or Fourier coefficients, this section gives the entire route of the construction. Every later theorem will prove one arrow in this route.

## 2.1 Reduction direction and construction direction

Suppose $M = pm$, where $p$ is prime and $m < M$. The central reduction is

| generator for the modulus-$m$ problem $\quad\Longrightarrow\quad$ generator for the modulus-$pm$ problem. |
|---|

Thus the proof reduces a larger modulus to a smaller one. Repeatedly removing prime factors gives

$$M \longrightarrow M/p_1 \longrightarrow M/(p_1p_2) \longrightarrow \cdots \longrightarrow 1.$$

The modulus-1 problem is trivial. The generator is assembled in the reverse direction: start from modulus 1 and insert one prime factor at a time. Keeping these two directions separate prevents a common source of confusion:

**Key takeaway.** The analysis moves from larger moduli to smaller moduli, whereas the construction moves from smaller moduli to larger moduli.

## 2.2 Why a state has two parameters

To insert a factor $p$, the reduction works over the alphabet $\mathbb{F}_p$. An intermediate problem is therefore described by a pair $(p, m)$, where $p$ is the current alphabet and $m$ is the current test modulus. We say that the state $(p, m)$ is solved when we have an explicit source on $\mathbb{F}_p^n$ that fools the relevant modulus-$m$ tests with the required error and seed length. The precise tests are defined in Section 3.

The target state is $(2, M)$, because $\mathbb{F}_2^n = \{0,1\}^n$. A starting state $(p, 1)$ costs no randomness, since every modulus-1 phase is the constant function 1.

## 2.3 The two moves

**LIFT: insert one matching prime.** If state $(p, m)$ has been solved, $\mathrm{LIFT}_p$ constructs state $(p, pm)$:

$$\boxed{\mathrm{LIFT}_p : (p, m) \longrightarrow (p, pm).}$$

The alphabet stays $\mathbb{F}_p$ and the modulus gains one factor $p$. The construction adds $p-1$ independent small-bias random vectors, one for each finite-difference level. Section 4 proves the error bound.

**TRANSFER: move to the next alphabet.** If $p < r$ and state $(r, m)$ has been solved, $\mathrm{TRANSFER}_{r\to p}$ constructs state $(p, m)$:

$$\boxed{\mathrm{TRANSFER}_{r\to p} : (r, m) \longrightarrow (p, m).}$$

The modulus stays fixed. The construction applies a coordinatewise map $\mathbb{F}_r \to \mathbb{F}_p$ that reaches every element of $\mathbb{F}_p$ and adds one independent small-bias random vector. Section 5 gives the proof.

## 2.4 Processing the prime factors

Write the prime factorization with multiplicity as

$$M = p_1 p_2 \cdots p_k, \qquad p_1 \le p_2 \le \cdots \le p_k.$$

The construction processes the factors from largest to smallest. It begins at $(p_k, 1)$ and first applies $\mathrm{LIFT}_{p_k}$. Suppose the factors $p_{j+1}, \ldots, p_k$ have already been inserted, and let $m_j = p_{j+1} \cdots p_k$.

If $p_j = p_{j+1}$, the alphabet already matches the next factor, so apply

$$(p_j, m_j) \xrightarrow{\mathrm{LIFT}_{p_j}} (p_j, p_j m_j).$$

If $p_j < p_{j+1}$, first change alphabets and then insert the factor:

$$(p_{j+1}, m_j) \xrightarrow{\text{TRANSFER}_{p_{j+1}\to p_j}} (p_j, m_j) \xrightarrow{\text{LIFT}_{p_j}} (p_j, p_j m_j).$$

After all factors are inserted, the state is $(p_1, M)$. If $p_1 > 2$, one final transfer reaches $(2, M)$.

### 2.5 Example: modulus 15

For $15 = 5 \cdot 3$, the complete route is

$$\boxed{(5,1) \xrightarrow{\text{LIFT}_5} (5,5) \xrightarrow{\text{TRANSFER}_{5\to 3}} (3,5) \xrightarrow{\text{LIFT}_3} (3,15) \xrightarrow{\text{TRANSFER}_{3\to 2}} (2,15).}$$

The first and third arrows insert prime factors. The second and fourth arrows change the alphabet without changing the modulus. The route starts from a test class that is identically 1 and ends at the Boolean modulus-15 class.

**What this section establishes.**
The proof is reduced to two statements, $\text{LIFT}_p$ and $\text{TRANSFER}_{r\to p}$. Since $M$ is fixed, the number of arrows is a constant depending only on $M$. Each arrow will require a constant number of small-bias random vectors and will preserve a seed bound of the form $O_M(\log n + \log(1/\varepsilon))$.

## 3 The intermediate test classes

Section 2 described the route using states $(p, m)$. We now define exactly which tests belong to a state, what it means for a source to fool them, and the Fourier facts used later. At the end of the section, LIFT and TRANSFER will be precise mathematical claims rather than informal arrows.

### 3.1 From modular sums to phase functions

Throughout, $n$ denotes the output length and $U_n$ the uniform distribution on $\{0,1\}^n$. For random variables $X, Y$ taking values in a finite set $\Omega$, the statistical distance is

$$\text{dist}(X, Y) := \frac{1}{2} \sum_{z\in\Omega} \big|\Pr[X = z] - \Pr[Y = z]\big|.$$

Fix an integer $M \geq 2$. For a coefficient vector $a \in \mathbb{Z}_M^n$ consider the modular sum

$$S_a^{(M)}(x) := \sum_{i=1}^{n} a_i x_i \pmod{M}, \qquad x \in \{0,1\}^n.$$

A bit distribution $X$ *$\varepsilon$-fools Boolean linear sums modulo $M$* if

$$\text{dist}\big(S_a^{(M)}(X),\ S_a^{(M)}(U_n)\big) \leq \varepsilon \qquad \text{for every } a \in \mathbb{Z}_M^n.$$

Let

$$\omega_M := e^{2\pi i/M}.$$

The map $r \mapsto \omega_M^r$ translates a residue into a point on the complex unit circle, and is useful because

$$u \equiv v \pmod{M} \quad \Longleftrightarrow \quad \omega_M^u = \omega_M^v, \qquad \omega_M^{u+v} = \omega_M^u \omega_M^v.$$

For a fixed coefficient vector $a$, the associated *phase function* is

$$\boxed{\phi_a^{(M)}(x) := \omega_M^{S_a^{(M)}(x)} = \omega_M^{\sum_i a_i x_i}, \qquad x \in \mathbb{F}_2^n.}$$

**Object type: the phase is a function, not a distribution.** The map $\phi_a^{(M)} : \mathbb{F}_2^n \to \mathbb{C}$ is a general complex-valued function. Its values have modulus 1; they need not be nonnegative and do not sum to 1. It is therefore not a probability distribution. Passing from residues to phases turns modular arithmetic into multiplicative algebra, and turns balance of the residues into cancellation of a complex expectation.

The one feature of $\phi_a^{(M)}$ that drives everything below is that it *factorises across coordinates*:

$$\phi_a^{(M)}(x) = \prod_{i=1}^{n} \omega_M^{a_i x_i}.$$

Each factor depends on a single input coordinate. The class of tests we shall actually fool is obtained by keeping this product structure and relaxing everything else: the exponent $a_i x_i$ is replaced by an arbitrary function of $x_i$, and the input alphabet $\{0, 1\}$ is replaced by an arbitrary prime field.

## 3.2 The test class for an intermediate state

Section 3.1 replaced a modular sum by its phase function

$$\phi_a^{(M)}(x) = \omega_M^{\sum_i a_i x_i} = \prod_{i=1}^{n} \omega_M^{a_i x_i}.$$

This representation exposes the one structural feature that the recursive construction will preserve throughout the paper: the test is a product of contributions from the individual coordinates.

However, the special exponent $a_i x_i$ is too rigid for the intermediate stages of the proof. The construction will perform two operations on the input tests.

**First operation: change the alphabet.** In the TRANSFER step, a coordinate over one prime field is replaced by a function of a coordinate over another prime field. If

$$\tau : \mathbb{F}_r \longrightarrow \mathbb{F}_p$$

is the coordinate map, then an exponent of the form $a_i x_i$ becomes $a_i \tau(y_i)$. This still depends on only one coordinate, but it need not be linear in that coordinate.

**Second operation: compare two shifted copies of a phase.** In the LIFT step, we will compare two shifted copies of the same phase. On a single coordinate, this replaces an exponent map $q_i$ by an expression of the form

$$q_i(z+u) - q_i(z+v).$$

Again, the result is an arbitrary function of one coordinate, even if the original exponent was much simpler.

These two observations lead to a single natural choice:

| keep the coordinatewise product structure, but allow arbitrary local exponent maps. |
|---|

This enlargement is exactly what we need. It still contains the original Boolean modular phases, and it is stable under both operations used by the construction.

**Definition 3.1** (Local phase class)**.** *Let $p$ be a prime and $m \geq 1$ an integer. For a tuple $\mathbf{q} = (q_1, \ldots, q_n)$ of* local exponent maps

$$q_i : \mathbb{F}_p \longrightarrow \mathbb{Z}_m,$$

*define the associated* local phase *by*

$$\Phi_{\mathbf{q}}(x) := \omega_m^{\sum_{i=1}^n q_i(x_i)}, \qquad x \in \mathbb{F}_p^n.$$

*Equivalently,*

$$\Phi_{\mathbf{q}}(x) = \prod_{i=1}^n \omega_m^{q_i(x_i)}.$$

*The class of all such phases is denoted*

$$\boxed{\mathsf{Local}_n(p,m)}.$$

*For $m = 1$ we use the convention $\mathbb{Z}_1 = \{0\}$ and $\omega_1 = 1$, so $\mathsf{Local}_n(p,1)$ contains only the constant function* 1.

The two parameters play different roles. The prime $p$ specifies the *alphabet* of the input $x \in \mathbb{F}_p^n$, while $m$ specifies the *modulus* of the exponent and hence the roots of unity taken by the phase. No relation between $p$ and $m$ is assumed.

We now verify the three reasons this is the right test class.

**Remark 3.2** (Modular sums are local phases)**.** *Let $a \in \mathbb{Z}_M^n$. Define $q_i(0) := 0$ and $q_i(1) := a_i$. Since $q_i(x_i) = a_i x_i$ for $x_i \in \mathbb{F}_2 = \{0,1\}$, we obtain*

$$\Phi_{\mathbf{q}}(x) = \omega_M^{\sum_i a_i x_i} = \phi_a^{(M)}(x).$$

*Hence every Boolean modular phase lies in $\mathsf{Local}_n(2,M)$. Fooling $\mathsf{Local}_n(2,M)$ therefore suffices for the main theorem, and this is why the final target state of the paper is $(2,M)$.*

**Remark 3.3** (Local phases are stable when the alphabet changes)**.** *Let $\tau : \mathbb{F}_r \to \mathbb{F}_p$ be any map and let $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$. Then*

$$\Phi_{\mathbf{q}}\big(\tau(y_1), \dots, \tau(y_n)\big) = \omega_m^{\sum_i q_i(\tau(y_i))} = \omega_m^{\sum_i (q_i \circ \tau)(y_i)}.$$

*Each composition $q_i \circ \tau : \mathbb{F}_r \to \mathbb{Z}_m$ is again a valid local exponent map, so the resulting function belongs to $\mathsf{Local}_n(r, m)$. This is the closure property used by* TRANSFER*: the proof may change alphabets without leaving the family of tests being fooled.*

**Remark 3.4** (Local phases are stable under shifted comparison)**.** *Let $u, v \in \mathbb{F}_p^n$ and $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$. Then*

$$\begin{aligned}\Phi_{\mathbf{q}}(x+u)\overline{\Phi_{\mathbf{q}}(x+v)} &= \omega_m^{\sum_i q_i(x_i+u_i) - \sum_i q_i(x_i+v_i)} \\ &= \omega_m^{\sum_i \left(q_i(x_i+u_i) - q_i(x_i+v_i)\right)}.\end{aligned}$$

*Thus the comparison of two shifted copies is again a local phase, with new local exponent maps*

$$q_i'(z) := q_i(z+u_i) - q_i(z+v_i).$$

*This is the closure property used by* LIFT*. In Section 4 we will introduce this operation formally as the* multiplicative derivative *and iterate it to simplify modulus-$pm$ phases.*

**Definition 3.5** (Error against a class)**.** *For a distribution $X$ on $\mathbb{F}_p^n$ and a class $\mathcal{C}$ of functions $\mathbb{F}_p^n \to \mathbb{C}$, define*

$$\boxed{\mathrm{Err}_X(\mathcal{C}) := \sup_{f \in \mathcal{C}} \Big| \mathbb{E} f(X) - \mathbb{E} f(U_p^n) \Big|.}$$

*We say that $X$ $\varepsilon$-fools $\mathcal{C}$ when $\mathrm{Err}_X(\mathcal{C}) \le \varepsilon$.*

This quantity gives a single numerical meaning to the phrase "the source fools the state." It will be used uniformly throughout the recursion: it measures how well the best test in the class distinguishes $X$ from the uniform input. As usual, if $\mathcal{C}' \subseteq \mathcal{C}$, then $\mathrm{Err}_X(\mathcal{C}') \le \mathrm{Err}_X(\mathcal{C})$.

## 3.3 The precise meaning of a solved state

State $(p, m)$ is solved with error $\varepsilon$ if there is an explicit source $X$ on $\mathbb{F}_p^n$ such that

$$\mathrm{Err}_X\big(\mathsf{Local}_n(p, m)\big) \le \varepsilon$$

and whose seed length is

$$O_{p,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The starting state $(p, 1)$ is solved by the constant-zero source with error zero. The target state is $(2, M)$; by Remark 3.2, it contains every Boolean modular phase needed for the theorem.

With this definition, the remaining construction claims are exactly

$$\begin{aligned}\mathrm{LIFT}_p &: \text{solve } (p, m) \text{ sufficiently accurately and obtain } (p, pm), \\ \mathrm{TRANSFER}_{r \to p} &: \text{solve } (r, m) \text{ sufficiently accurately and obtain } (p, m).\end{aligned}$$

The rest of the paper proves these statements and tracks the required accuracy.

## 3.4 Fourier analysis and small-bias random vectors

Both moves are analysed by Fourier analysis over the input group $\mathbb{F}_p^n$, and the final conversion to statistical distance uses Fourier analysis over $\mathbb{Z}_M$. We record the common framework once and then specialise.

Let $H$ be a finite abelian group, written additively. A *character* of $H$ is a function $\chi : H \to \mathbb{C}$ with $\chi(x+y) = \chi(x)\chi(y)$ for all $x, y \in H$; a character converts the group operation into multiplication. Since $H$ is finite, $|\chi(x)| = 1$ and $\chi(-x) = \overline{\chi(x)}$. The characters form a group $\widehat{H}$ under pointwise multiplication, with $|\widehat{H}| = |H|$, and they are orthonormal for the normalised inner product $\langle f, g\rangle := \mathbb{E}_{x\sim U_H}[f(x)\overline{g(x)}]$. Being $|H|$ in number, they form an orthonormal basis of all functions $H \to \mathbb{C}$.

**Conceptual meaning of the Fourier transform.** The characters are multiplicative with respect to the group operation, while the Fourier expansion is linear. Thus Fourier analysis represents a function as a linear combination of multiplicative basis functions. This basis is especially useful when a problem involves addition, translation, or averaging after a random shift — which is exactly the situation here, since every construction step adds an independent random vector.

We now fix the input group. For a prime $p$ write $\omega_p = e^{2\pi i/p}$, and for $\xi, x \in \mathbb{F}_p^n$ let

$$\chi_\xi(x) := \omega_p^{\langle \xi, x\rangle}, \qquad \langle \xi, x\rangle = \sum_{i=1}^{n} \xi_i x_i \pmod p.$$

These are all the characters of $\mathbb{F}_p^n$. For $f : \mathbb{F}_p^n \to \mathbb{C}$ define

$$\widehat{f}(\xi) := \mathbb{E}_{x\sim U_p^n}\left[f(x)\overline{\chi_\xi(x)}\right], \qquad \text{so that} \qquad f(x) = \sum_{\xi\in\mathbb{F}_p^n} \widehat{f}(\xi)\chi_\xi(x),$$

and Parseval's identity holds:

$$\boxed{\sum_{\xi\in\mathbb{F}_p^n} |\widehat{f}(\xi)|^2 = \mathbb{E}_{x\sim U_p^n}|f(x)|^2.}$$

In particular, if $|f| = 1$ everywhere then $\sum_\xi |\widehat{f}(\xi)|^2 = 1$, a normalisation used at several points below. Note also that $\widehat{f}(0) = \mathbb{E}f(U_p^n)$: the trivial Fourier coefficient of a test is its uniform expectation, which is precisely the quantity that Definition 3.5 compares against.

**Definition 3.6** (Small bias)**.** *A distribution $B$ on $\mathbb{F}_p^n$ is* $\delta$-biased *if*

$$|\mathbb{E}\chi_\xi(B)| \le \delta \qquad \textit{for every } \xi \neq 0.$$

**Imported primitive: small bias over a fixed prime field.** For every fixed prime $p$, every $n$, and every $0 < \delta \le 1$ there is an explicit $\delta$-biased distribution on $\mathbb{F}_p^n$ with seed length

$$O_p\left(\log n + \log\frac{1}{\delta}\right)$$

[NN93], [AGHP92], [AMN98]. This is the only pseudorandomness primitive the upper bound imports; everything else is proved here.

Two consequences of the definitions will be used in both moves. First, adding an independent random vector multiplies each character expectation: if $B$ is independent of $X$ then, by multiplicativity of characters,

$$\mathbb{E}\chi_\xi(X+B) = \mathbb{E}\chi_\xi(X)\cdot\mathbb{E}\chi_\xi(B),$$

so in particular $X+B$ is $\delta$-biased whenever $B$ is, irrespective of $X$. Second, and equivalently, if $h(x) := \mathbb{E}_B[f(x+B)]$ denotes the smoothing of $f$ by $B$, then

$$\boxed{\widehat{h}(\xi) = \widehat{f}(\xi)\cdot\mathbb{E}\chi_\xi(B).}$$

Smoothing by a small-bias random vector thus attenuates every nontrivial Fourier coefficient by a factor of at most $\delta$, and leaves the trivial coefficient untouched.

**Summary of Section 3.** The definitions introduced in this section fit together as follows:

$\phi_a^{(M)}$ ← encode a modular sum as a multiplicative phase, so the test factorises by coordinates;

$\mathsf{Local}_n(p,m)$ ← enlarge the test class while preserving that coordinatewise structure, so it survives both LIFT and TRANSFER;

$\mathrm{Err}_X\big(\mathsf{Local}_n(p,m)\big)$ ← measure how well a source fools that class, giving a numerical meaning to a solved state;

$(p,m)$ ← the intermediate goal "fool $\mathsf{Local}_n(p,m)$ over $\mathbb{F}_p^n$," which the recursion moves between;

$\widehat{f}$, $\delta$-bias ← the Fourier language and smoothing primitive, which power the later LIFT and TRANSFER proofs.

In short: Section 3.1 defines the tests, Section 3.3 defines what it means to fool them at an intermediate state, and Section 3.4 records the Fourier tools that will drive the two construction moves.

# 4 The LIFT reduction

The state $(p,m)$ is assumed to be solved, and the next goal is the state $(p,pm)$. The alphabet remains $\mathbb{F}_p$; only the modulus gains a factor $p$. This section proves that move. The proof has two parts: repeated phase derivatives simplify a modulus-$pm$ test to a modulus-$m$ test, and one small-bias random vector reverses one derivative step.

**Proof strategy.** For the target class $\mathsf{Local}_n(p,pm)$, take $p-1$ derivatives. On each coordinate, reducing the exponent modulo $p$ gives a polynomial of degree at most $p-1$, and every nonzero finite difference lowers its degree. After $p-1$ differences the exponent is constant modulo $p$, so the differentiated phase is a constant complex factor times a member of $\mathsf{Local}_n(p,m)$. Starting from a source that fools this final derivative class, add $p-1$ independent small-bias random vectors, one for each derivative level, until the original class is recovered.

## 4.1 Phase derivatives

Let $f : \mathbb{F}_p^n \to \mathbb{C}$ have absolute value 1 everywhere. For shifts $u, v \in \mathbb{F}_p^n$, define its *multiplicative derivative* by

$$D_{u,v}f(x) := f(x+u)\overline{f(x+v)}. \tag{4.1}$$

If $f(x) = \omega_m^{P(x)}$, then

$$D_{u,v}f(x) = \omega_m^{P(x+u)-P(x+v)}.$$

Thus the multiplicative derivative of a phase is the ordinary finite difference of its exponent. We use a ratio of the shifted phases rather than their ordinary difference because the ratio is again a phase.

For a class $\mathcal{C}$ of unit-modulus functions, let

$$\partial\mathcal{C} := \{D_{u,v}f : f \in \mathcal{C},\ u, v \in \mathbb{F}_p^n\},$$

and define $\partial^0\mathcal{C} = \mathcal{C}$ and $\partial^{j+1}\mathcal{C} = \partial(\partial^j\mathcal{C})$.

A derivative of a local phase is still a local phase. Indeed, if

$$\Phi_{\mathbf{q}}(x) = \omega_m^{\sum_i q_i(x_i)},$$

then

$$D_{u,v}\Phi_{\mathbf{q}}(x) = \omega_m^{\sum_i (q_i(x_i+u_i)-q_i(x_i+v_i))}.$$

The new local exponent map in coordinate $i$ is

$$q_i'(z) = q_i(z+u_i) - q_i(z+v_i).$$

Therefore repeated derivatives preserve the coordinatewise form of the test.

## 4.2 Why $p-1$ derivatives remove one factor $p$

The algebraic point is easiest to see after reducing an exponent modulo $p$. We use the following standard fact.

**Lemma 4.1** (Polynomial representation of functions on $\mathbb{F}_p$)**.** *Every function $g : \mathbb{F}_p \to \mathbb{F}_p$ is represented on $\mathbb{F}_p$ by a polynomial $G \in \mathbb{F}_p[z]$ of degree at most $p-1$.*

*Proof.* For each $a \in \mathbb{F}_p$, let

$$L_a(z) = \prod_{b \in \mathbb{F}_p \setminus \{a\}} \frac{z-b}{a-b}.$$

Then $L_a(a) = 1$ and $L_a(b) = 0$ for $b \neq a$. Hence

$$G(z) = \sum_{a \in \mathbb{F}_p} g(a) L_a(z)$$

has degree at most $p-1$ and satisfies $G(a) = g(a)$ for every $a \in \mathbb{F}_p$. □

**Lemma 4.2** (A nonzero finite difference lowers the degree)**.** *Let $G \in \mathbb{F}_p[z]$ have degree $d$ with $1 \le d \le p-1$, and let $u, v \in \mathbb{F}_p$ with $u \neq v$. Then*

$$G(z+u) - G(z+v)$$

*has degree $d-1$. If $u = v$, the difference is the zero polynomial.*

*Proof.* Write the leading term of $G$ as $a_d z^d$. The coefficient of $z^{d-1}$ in $G(z+u) - G(z+v)$ is

$$a_d d(u-v).$$

It is nonzero in $\mathbb{F}_p$ because $a_d \neq 0$, $1 \le d \le p-1$, and $u \neq v$. All terms of degree $d$ cancel, so the resulting degree is exactly $d-1$. The case $u = v$ is immediate. □

**Proposition 4.3** (Prime-factor descent)**.** *Let $p$ be prime and $m \ge 1$. Every function in $\partial^{p-1}\mathsf{Local}_n(p, pm)$ has the form*

$$\lambda\Psi, \qquad |\lambda| = 1, \qquad \Psi \in \mathsf{Local}_n(p, m).$$

*Consequently, any source that $\alpha$-fools $\mathsf{Local}_n(p, m)$ also $\alpha$-fools $\partial^{p-1}\mathsf{Local}_n(p, pm)$.*

*Proof.* Take a phase $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, pm)$ and apply $p-1$ derivatives, using arbitrary vector shift pairs. Fix one coordinate $i$. Reducing the original local exponent map

$$q_i : \mathbb{F}_p \to \mathbb{Z}_{pm}$$

modulo $p$ gives a function $\overline{q}_i : \mathbb{F}_p \to \mathbb{F}_p$. By Lemma 4.1, represent it by a polynomial of degree at most $p-1$.

Each derivative replaces the polynomial by a finite difference. If the two coordinate shifts are equal, the new polynomial is zero. Otherwise, Lemma 4.2 lowers its degree by one whenever it is nonconstant. After $p-1$ derivatives, the resulting exponent map is therefore constant modulo $p$. Thus there are $c_i \in \mathbb{Z}_{pm}$ and a map $r_i : \mathbb{F}_p \to \mathbb{Z}_m$ such that the final exponent in coordinate $i$ is

$$c_i + p r_i(z) \qquad \text{in } \mathbb{Z}_{pm}.$$

Indeed, all final values lie in the same residue class modulo $p$, and that residue class has the form $c_i + p\mathbb{Z}_m$ inside $\mathbb{Z}_{pm}$.

Let $D$ denote the chosen sequence of $p-1$ derivatives. Applying this argument to every coordinate gives

$$\begin{aligned} D\Phi_{\mathbf{q}}(x) &= \omega_{pm}^{\sum_i (c_i + p r_i(x_i))} \\ &= \underbrace{\omega_{pm}^{\sum_i c_i}}_{\lambda}\, \omega_m^{\sum_i r_i(x_i)}. \end{aligned}$$

The second factor is a member $\Psi$ of $\mathsf{Local}_n(p, m)$, and the first is a constant complex number of absolute value 1.

Finally, multiplying a test by such a constant does not change its distinguishing error:

$$\left|\mathbb{E}[\lambda\Psi(X)] - \mathbb{E}[\lambda\Psi(U_p^n)]\right| = \left|\mathbb{E}\Psi(X) - \mathbb{E}\Psi(U_p^n)\right|.$$

This proves the final assertion. □

**Key takeaway.** The descent uses $p-1$, rather than $p$, derivatives. The price is only a constant complex factor, which is harmless because the proof compares expectations in absolute value.

## 4.3 One small-bias vector lifts one derivative level

The previous subsection moves downward among test classes. We now prove the opposite movement for the source: if a source fools the derivative class, then adding one small-bias random vector makes it fool the original class.

**Lemma 4.4** (Derivative lifting)**.** *Let $\mathcal{C}$ be a class of unit-modulus functions on $\mathbb{F}_p^n$. Suppose $X$ is a distribution on $\mathbb{F}_p^n$ satisfying*

$$\mathrm{Err}_X(\partial\mathcal{C}) \le \alpha,$$

*and let $B$ be an independent $\delta$-biased distribution on $\mathbb{F}_p^n$. Then*

$$\boxed{\mathrm{Err}_{X+B}(\mathcal{C}) \le \sqrt{3\alpha + \delta^2}.}$$

*Proof.* Fix $f \in \mathcal{C}$. Write

$$U \sim U_p^n, \qquad W = X + B, \qquad \mu = \mathbb{E}f(U),$$

and define the smoothed function

$$h(x) = \mathbb{E}_B f(x + B).$$

The error for this test is

$$\Delta := \mathbb{E}f(W) - \mu = \mathbb{E}_X h(X) - \mu.$$

We prove $|\Delta|^2 \le 3\alpha + \delta^2$.

First, $W$ still $\alpha$-fools $\partial\mathcal{C}$. Indeed, if $g \in \partial\mathcal{C}$ and $b \in \mathbb{F}_p^n$, then $g_b(x) := g(x + b)$ also belongs to $\partial\mathcal{C}$: for $g = D_{u,v}f$,

$$g_b(x) = D_{u+b,v+b}f(x).$$

For fixed $b$, the hypothesis gives

$$|\mathbb{E}g(X + b) - \mathbb{E}g(U)| = |\mathbb{E}g_b(X) - \mathbb{E}g_b(U)| \le \alpha,$$

where we used that $U + b$ is uniform. Averaging over $b = B$ gives

$$|\mathbb{E}g(W) - \mathbb{E}g(U)| \le \alpha \qquad \text{for every } g \in \partial\mathcal{C}. \tag{4.2}$$

We next record a correlation identity. Let $V$ be uniform on $\mathbb{F}_p^n$ and independent of everything else. For each fixed value of $W$, the random variable $W + V$ is uniform, so

$$\mathbb{E}_V\left[f(W)\overline{f(W + V)}\right] = f(W)\overline{\mu}.$$

The same identity with $W$ replaced by $U$ gives $\mu\overline{\mu}$. Subtracting and using $f(x)\overline{f(x + v)} = D_{0,v}f(x)$ yields

$$\Delta\overline{\mu} = \mathbb{E}_V\left[\mathbb{E}D_{0,V}f(W) - \mathbb{E}D_{0,V}f(U)\right].$$

Every function $D_{0,v}f$ lies in $\partial\mathcal{C}$, so (4.2) implies

$$|\Delta\overline{\mu}| \le \alpha. \tag{4.3}$$

Now center the smoothed function before applying Cauchy–Schwarz:

$$|\Delta|^2 = |\mathbb{E}_X[h(X) - \mu]|^2 \le \mathbb{E}_X|h(X) - \mu|^2. \tag{4.4}$$

Expanding the square gives

$$\mathbb{E}_X|h(X) - \mu|^2 = \mathbb{E}_X|h(X)|^2 - |\mu|^2 - 2\,\mathrm{Re}(\overline{\mu}\Delta). \tag{4.5}$$

Let $B'$ be an independent copy of $B$. Since $|f| = 1$,

$$|h(x)|^2 = \mathbb{E}_{B,B'}\left[f(x+B)\overline{f(x+B')}\right] = \mathbb{E}_{B,B'}D_{B,B'}f(x).$$

For every fixed $b, b'$, the function $D_{b,b'}f$ belongs to $\partial\mathcal{C}$. The hypothesis on $X$, followed by averaging over $B, B'$, therefore gives

$$\mathbb{E}_X|h(X)|^2 \le \mathbb{E}_U|h(U)|^2 + \alpha. \tag{4.6}$$

It remains to bound the uniform energy of $h$. The smoothing identity from Section 3.4 gives

$$\widehat{h}(\xi) = \widehat{f}(\xi)\mathbb{E}\chi_\xi(B).$$

By Parseval, the small-bias condition, and $|f| = 1$,

$$\begin{aligned}\mathbb{E}_U|h(U)|^2 &= \sum_{\xi\in\mathbb{F}_p^n} |\widehat{f}(\xi)|^2\,|\mathbb{E}\chi_\xi(B)|^2 \\ &\le |\widehat{f}(0)|^2 + \delta^2\sum_{\xi\neq 0}|\widehat{f}(\xi)|^2 \\ &\le |\mu|^2 + \delta^2.\end{aligned}$$

Here $\widehat{f}(0) = \mu$, and Parseval gives $\sum_\xi |\widehat{f}(\xi)|^2 = 1$.

Combining (4.4), (4.5), (4.6), and (4.3), we obtain

$$\begin{aligned}|\Delta|^2 &\le \delta^2 + \alpha - 2\,\mathrm{Re}(\overline{\mu}\Delta) \\ &\le \delta^2 + \alpha + 2|\overline{\mu}\Delta| \\ &\le \delta^2 + 3\alpha.\end{aligned}$$

Taking square roots proves the lemma. □

## 4.4 The LIFT theorem

**Theorem 4.5** (LIFT). *Let $p$ be prime, $m \ge 1$, and $0 < \varepsilon \le 1$. Set*

$$\theta := \left(\frac{\varepsilon}{4}\right)^{2^{p-1}}.$$

*Let $X_0$ be a distribution on $\mathbb{F}_p^n$ satisfying*

$$\mathrm{Err}_{X_0}(\mathsf{Local}_n(p,m)) \le \theta,$$

*and let $B_1, \dots, B_{p-1}$ be independent $\theta$-biased distributions on $\mathbb{F}_p^n$, also independent of $X_0$. Define*

$$X_j = X_{j-1} + B_j, \qquad 1 \le j \le p-1.$$

*Then*

$$\boxed{\mathrm{Err}_{X_{p-1}}(\mathsf{Local}_n(p, pm)) \le \varepsilon.}$$

*Moreover, if $X_0$ has seed length $O_{p,m}(\log n + \log(1/\theta))$, then $X_{p-1}$ has seed length*

$$O_{p,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

*Proof.* Define the derivative hierarchy

$$\mathcal{C}_j := \partial^{p-1-j}\mathsf{Local}_n(p, pm), \qquad 0 \le j \le p-1.$$

Thus $\mathcal{C}_{p-1} = \mathsf{Local}_n(p, pm)$ and $\partial \mathcal{C}_j = \mathcal{C}_{j-1}$ for $j \ge 1$. By Proposition 4.3, the source $X_0$ fools $\mathcal{C}_0$ with error at most $\theta$.

Let

$$e_j = \mathrm{Err}_{X_j}(\mathcal{C}_j).$$

Lemma 4.4 gives

$$e_j \le \sqrt{3e_{j-1} + \theta^2} \qquad (1 \le j \le p-1). \tag{4.7}$$

To solve this recursion, define

$$E_0 = \theta, \qquad E_j = \sqrt{3E_{j-1} + \theta^2}.$$

Inductively, $e_j \le E_j$. Also $E_j \ge \theta$ for every $j$, and $0 < \theta \le 1$, so $\theta^2 \le E_{j-1}$. Therefore

$$E_j \le \sqrt{4E_{j-1}} = 2\sqrt{E_{j-1}}.$$

A second induction now gives

$$E_j \le 4\theta^{1/2^j}.$$

For $j = 0$ this is immediate. If it holds for $j-1$, then

$$E_j \le 2\sqrt{E_{j-1}} \le 2\sqrt{4\theta^{1/2^{j-1}}} = 4\theta^{1/2^j}.$$

At the final level,

$$e_{p-1} \le 4\theta^{1/2^{p-1}} = 4 \cdot \frac{\varepsilon}{4} = \varepsilon.$$

For the seed length,

$$\log \frac{1}{\theta} = 2^{p-1} \log \frac{4}{\varepsilon} = O_p\left(1 + \log \frac{1}{\varepsilon}\right).$$

The construction adds $p-1$ independent small-bias vectors, a number depending only on the fixed prime $p$. Their seed lengths and the seed length of $X_0$ therefore sum to $O_{p,m}(\log n + \log(1/\varepsilon))$. □

**What this section establishes.**
The operation $\mathrm{LIFT}_p$ is proved:

$$\boxed{(p, m) \xrightarrow{\mathrm{LIFT}_p} (p, pm).}$$

It uses $p - 1$ small-bias random vectors. The algebra lowers the test class by $p - 1$ derivative levels, and Lemma 4.4 climbs those levels back one at a time.

# 5 The TRANSFER reduction

The operation $\mathrm{LIFT}_p$ can insert a prime factor only when the source alphabet is $\mathbb{F}_p$. After one prime factor has been processed, the construction may need to move to a smaller prime alphabet before inserting the next one. This section proves that changing the alphabet preserves the modulus and the optimal seed-length form.

Fix primes $p < r$, a modulus $m \geq 2$, and a map

$$\sigma : \mathbb{F}_r \longrightarrow \mathbb{F}_p$$

that reaches every element of $\mathbb{F}_p$. Starting from a source $Y$ on $\mathbb{F}_r^n$, define

$$\boxed{X = \sigma(Y) + B,}$$

where $\sigma$ is applied separately to every coordinate and $B$ is an independent small-bias random vector on $\mathbb{F}_p^n$.

For a fixed value $B = \beta$ and a target test $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$, the function

$$y \longmapsto \Phi_{\mathbf{q}}(\sigma(y) + \beta) = \omega_m^{\sum_i q_i(\sigma(y_i)+\beta_i)}$$

belongs to $\mathsf{Local}_n(r, m)$. This is exactly the closure under coordinatewise maps established in Section 3.2. Hence the guarantee for $Y$ can be applied after the alphabet change.

The map $\sigma$ does not, however, send a uniform element of $\mathbb{F}_r$ to a uniform element of $\mathbb{F}_p$. The proof repairs this discrepancy in two ways: small bias handles tests that depend on few coordinates, while cancellation across coordinates handles tests that depend on many coordinates.

## 5.1 Active coordinates and the low-weight estimate

A coordinate $i$ is *active* for $\mathbf{q}$ when $q_i : \mathbb{F}_p \to \mathbb{Z}_m$ is nonconstant. Let

$$w = \mathrm{wt}(\mathbf{q})$$

be the number of active coordinates. Inactive coordinates contribute only a constant root of unity, so $w$ is the number of coordinates on which the test genuinely depends.

**Lemma 5.1** (Low-weight estimate)**.** *Let $Z$ be any distribution on $\mathbb{F}_p^n$, let $B$ be an independent $\delta$-biased distribution, and set $X = Z + B$. For every $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$ of weight $w$,*

$$\boxed{\left|\mathbb{E}\Phi_{\mathbf{q}}(X) - \mathbb{E}\Phi_{\mathbf{q}}(U_p^n)\right| \leq \delta p^{w/2}.}$$

*Proof.* For every nonzero character $\chi_\xi$ of $\mathbb{F}_p^n$, independence gives

$$\mathbb{E}\chi_\xi(X) = \mathbb{E}\chi_\xi(Z)\,\mathbb{E}\chi_\xi(B),$$

so $|\mathbb{E}\chi_\xi(X)| \le \delta$. Expanding the test in the character basis,

$$\left|\mathbb{E}\Phi_{\mathbf{q}}(X) - \mathbb{E}\Phi_{\mathbf{q}}(U_p^n)\right| \le \delta \sum_{\xi \neq 0} |\widehat{\Phi_{\mathbf{q}}}(\xi)|.$$

A Fourier coefficient of a function on $\mathbb{F}_p^n$ is labeled by a vector

$$\xi = (\xi_1, \dots, \xi_n) \in \mathbb{F}_p^n.$$

We call $\xi$ the *Fourier index*. It labels the character

$$\chi_\xi(x) = \omega_p^{\langle \xi, x\rangle} = \omega_p^{\sum_{i=1}^n \xi_i x_i},$$

and the Fourier coefficient at this index is

$$\widehat{\Phi_{\mathbf{q}}}(\xi) = \mathbb{E}_{x \in \mathbb{F}_p^n}\left[\Phi_{\mathbf{q}}(x)\overline{\chi_\xi(x)}\right].$$

Up to an overall constant factor, the test depends only on its $w$ active coordinates. To see the consequence for its Fourier transform, consider an inactive coordinate $i$. The corresponding local exponent map is constant, so for some constant $c_i$ the contribution of this coordinate to the test is always $c_i$. Its contribution to the Fourier coefficient is therefore

$$c_i\,\mathbb{E}_{z \in \mathbb{F}_p} \omega_p^{-\xi_i z}.$$

This average equals $c_i$ when $\xi_i = 0$, and it equals $0$ when $\xi_i \neq 0$, because a nontrivial character has average zero over $\mathbb{F}_p$. Hence a Fourier coefficient can be nonzero only if

$$\xi_i = 0$$

at every inactive coordinate. Only the $w$ active coordinates may have arbitrary Fourier-index entries, and each of them has $p$ possible values. Therefore the Fourier transform has at most

$$p^w$$

nonzero coefficients.

Since $|\Phi_{\mathbf{q}}| = 1$, Parseval gives

$$\sum_{\xi} |\widehat{\Phi_{\mathbf{q}}}(\xi)|^2 = 1.$$

Cauchy–Schwarz now yields

$$\sum_{\xi} |\widehat{\Phi_{\mathbf{q}}}(\xi)| \le \sqrt{p^w}\left(\sum_{\xi} |\widehat{\Phi_{\mathbf{q}}}(\xi)|^2\right)^{1/2} = p^{w/2}.$$

Substituting this bound proves the lemma. □

## 5.2 Cancellation from one active coordinate

We next give one explicit constant that works for both distributions appearing in the high-weight argument.

**Lemma 5.2** (Cancellation when every value has noticeable probability)**.** *Let $p < r$ be primes and $m \geq 2$. Define*

$$\boxed{\kappa_{r,m} := \sqrt{1 - \frac{4\sin^2(\pi/m)}{r^2}}.}$$

*Then $0 < \kappa_{r,m} < 1$. Moreover, if $\nu$ is any distribution on $\mathbb{F}_p$ satisfying*

$$\nu(a) \geq \frac{1}{r} \qquad \textit{for every } a \in \mathbb{F}_p,$$

*and $q : \mathbb{F}_p \to \mathbb{Z}_m$ is nonconstant, then*

$$\left|\mathbb{E}_{A\sim\nu}\omega_m^{q(A)}\right| \leq \kappa_{r,m}.$$

*Proof.* Let $A, A'$ be independent samples from $\nu$, and set $Z = \omega_m^{q(A)}$ and $Z' = \omega_m^{q(A')}$. Since $|Z| = |Z'| = 1$,

$$\mathbb{E}|Z - Z'|^2 = 2 - 2\,\mathrm{Re}\,\mathbb{E}[Z\overline{Z'}] = 2 - 2|\mathbb{E}Z|^2.$$

Equivalently,

$$1 - |\mathbb{E}Z|^2 = \frac{1}{2}\mathbb{E}|Z - Z'|^2.$$

Because $q$ is nonconstant, choose $a, b \in \mathbb{F}_p$ with $q(a) \neq q(b)$. Two distinct $m$-th roots of unity are at distance at least $2\sin(\pi/m)$. Keeping only the two ordered events $(A, A') = (a, b)$ and $(b, a)$ in the nonnegative expectation above gives

$$\begin{aligned} 1 - |\mathbb{E}Z|^2 &\geq \nu(a)\nu(b)\left|\omega_m^{q(a)} - \omega_m^{q(b)}\right|^2 \\ &\geq \frac{4\sin^2(\pi/m)}{r^2}. \end{aligned}$$

Rearranging proves $|\mathbb{E}Z| \leq \kappa_{r,m}$. Since $r \geq 3$, the expression inside the square root lies strictly between $0$ and $1$. □

The lemma applies to both one-coordinate distributions needed below. Under $U_p$, every point has probability $1/p \geq 1/r$. Under $\sigma(U_r) + \beta$, every point has probability at least $1/r$, because $\sigma$ reaches every element of $\mathbb{F}_p$.

## 5.3 The TRANSFER theorem

**Theorem 5.3** (TRANSFER)**.** *Let $p < r$ be primes, let $m \geq 2$, and let $0 < \varepsilon \leq 1$. Fix a map $\sigma : \mathbb{F}_r \to \mathbb{F}_p$ that reaches every element of $\mathbb{F}_p$, and set*

$$\kappa = \kappa_{r,m}, \qquad W = \left\lceil \frac{\log(4/\varepsilon)}{\log(1/\kappa)} \right\rceil, \qquad \delta = \frac{\varepsilon}{2p^{W/2}}.$$

*Suppose $Y$ is a distribution on $\mathbb{F}_r^n$ such that*

$$\mathrm{Err}_Y(\mathsf{Local}_n(r,m)) \le \frac{\varepsilon}{2},$$

*and let $B$ be an independent $\delta$-biased distribution on $\mathbb{F}_p^n$. Then*

$$X = \sigma(Y) + B$$

*satisfies*

$$\boxed{\mathrm{Err}_X(\mathsf{Local}_n(p,m)) \le \varepsilon.}$$

*If $Y$ can be sampled with seed length $O_{r,m}(\log n + \log(1/\varepsilon))$, then $X$ can be sampled with seed length*

$$O_{p,r,m}\left(\log n + \log\frac{1}{\varepsilon}\right).$$

*Proof.* Fix $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p,m)$ and let $w = \mathrm{wt}(\mathbf{q})$. Lemma 5.1, applied with $Z = \sigma(Y)$, gives

$$\left|\mathbb{E}\Phi_{\mathbf{q}}(X) - \mathbb{E}\Phi_{\mathbf{q}}(U_p^n)\right| \le \delta p^{w/2}. \tag{5.1}$$

For the second estimate, condition on $B = \beta$. The function

$$y \longmapsto \Phi_{\mathbf{q}}(\sigma(y) + \beta)$$

belongs to $\mathsf{Local}_n(r,m)$, so the hypothesis on $Y$ gives

$$\left|\mathbb{E}_Y\Phi_{\mathbf{q}}(\sigma(Y)+\beta) - \mathbb{E}_{U_r^n}\Phi_{\mathbf{q}}(\sigma(U_r^n)+\beta)\right| \le \frac{\varepsilon}{2}.$$

Under $U_r^n$, the coordinates are independent. If coordinate $i$ is active, then $q_i(\sigma(\cdot)+\beta_i)$ is nonconstant because $\sigma$ reaches every element of $\mathbb{F}_p$. Lemma 5.2 therefore bounds its one-coordinate expectation by $\kappa$. Inactive coordinates contribute factors of absolute value $1$. Hence

$$\left|\mathbb{E}_{U_r^n}\Phi_{\mathbf{q}}(\sigma(U_r^n)+\beta)\right| \le \kappa^w.$$

Averaging over $\beta = B$ gives

$$\left|\mathbb{E}\Phi_{\mathbf{q}}(X)\right| \le \frac{\varepsilon}{2} + \kappa^w.$$

The same factorization under $U_p^n$, again using Lemma 5.2, gives

$$\left|\mathbb{E}\Phi_{\mathbf{q}}(U_p^n)\right| \le \kappa^w.$$

Therefore

$$\left|\mathbb{E}\Phi_{\mathbf{q}}(X) - \mathbb{E}\Phi_{\mathbf{q}}(U_p^n)\right| \le \frac{\varepsilon}{2} + 2\kappa^w. \tag{5.2}$$

We now choose the better of (5.1) and (5.2). If $w \le W$, then

$$\delta p^{w/2} \le \delta p^{W/2} = \frac{\varepsilon}{2}.$$

If $w > W$, then $0 < \kappa < 1$ and the definition of $W$ imply

$$\kappa^w \le \kappa^W \le \frac{\varepsilon}{4}.$$

The second estimate is therefore at most

$$\frac{\varepsilon}{2} + 2 \cdot \frac{\varepsilon}{4} = \varepsilon.$$

Thus every target test has error at most $\varepsilon$.

Finally, $\kappa < 1$ is a constant depending only on $r$ and $m$, so

$$W = O_{r,m}\left(1 + \log \frac{1}{\varepsilon}\right).$$

Consequently,

$$\log \frac{1}{\delta} = \log \frac{2}{\varepsilon} + \frac{W}{2} \log p = O_{p,r,m}\left(1 + \log \frac{1}{\varepsilon}\right).$$

One explicit $\delta$-biased vector therefore costs $O_{p,r,m}(\log n + \log(1/\varepsilon))$ seed bits. Adding its seed to the seed of $Y$ proves the claimed bound. □

**What this section establishes.**
The operation $\mathrm{TRANSFER}_{r\to p}$ is proved:

$$\boxed{(r, m) \xrightarrow{\mathrm{TRANSFER}_{r\to p}} (p, m), \qquad p < r.}$$

The small-bias estimate increases with the number of active coordinates, while the cancellation estimate decreases with it. The threshold $W$ guarantees that every test is covered by one of the two estimates.

# 6 Assembling the generator for an arbitrary modulus

The two construction moves are now complete. This section follows the route from Section 2: start with the trivial modulus 1, insert the prime factors of $M$ from largest to smallest, and finish over the Boolean alphabet.

Call a state $(p, m)$ *solvable* if, for every $n$ and every $0 < \varepsilon \le 1$, there is an explicit source $X$ on $\mathbb{F}_p^n$ satisfying

$$\mathrm{Err}_X(\mathsf{Local}_n(p, m)) \le \varepsilon$$

with seed length

$$O_{p,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The state $(p, 1)$ is solvable for every prime $p$: the test class contains only the constant function 1, so the constant-zero source has error 0 and uses no seed.

Theorems 4.5 and 5.3 give two rules:

$$(p, m) \text{ solvable} \Longrightarrow (p, pm) \text{ solvable},$$
$$(r, m) \text{ solvable} \Longrightarrow (p, m) \text{ solvable} \qquad (p < r).$$

For the first rule, a target error $\varepsilon$ requires the previous state only at accuracy

$$\left(\frac{\varepsilon}{4}\right)^{2^{p-1}},$$

whose reciprocal has logarithm $O_p(1 + \log(1/\varepsilon))$. For the second rule, a target error $\varepsilon$ requires the previous state at accuracy $\varepsilon/2$, and the added small-bias vector also has logarithmic seed length. Thus both rules preserve solvability.

**Theorem 6.1** (Assembly over the prime factors)**.** *For every fixed integer $M \geq 2$, every $n$, and every $0 < \varepsilon \leq 1$, there is an explicit source $G$ on $\mathbb{F}_2^n = \{0,1\}^n$ such that*

$$\boxed{\mathrm{Err}_G(\mathsf{Local}_n(2, M)) \leq \varepsilon}$$

*and whose seed length is*

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

*Proof.* Write the prime factorization of $M$, with multiplicity, as

$$M = p_1 p_2 \cdots p_k, \qquad p_1 \leq p_2 \leq \cdots \leq p_k.$$

Begin with the solvable state $(p_k, 1)$ and apply $\mathrm{LIFT}_{p_k}$, obtaining $(p_k, p_k)$.

Now process $j = k-1, k-2, \ldots, 1$. Suppose the already inserted suffix of the modulus is

$$m_j = p_{j+1} \cdots p_k$$

and the current state is $(p_{j+1}, m_j)$. If $p_j < p_{j+1}$, first apply $\mathrm{TRANSFER}_{p_{j+1} \to p_j}$ to obtain $(p_j, m_j)$. If $p_j = p_{j+1}$, this transfer is unnecessary because the alphabet already matches. In either case, apply $\mathrm{LIFT}_{p_j}$ to reach

$$(p_j, p_j m_j).$$

After all factors have been inserted, the state is $(p_1, M)$. If $p_1 > 2$, one final application of $\mathrm{TRANSFER}_{p_1 \to 2}$ reaches $(2, M)$; if $p_1 = 2$, no final transfer is needed.

The number of moves depends only on the fixed modulus $M$, and each move preserves solvability. Hence the final state has the stated error and seed length. □

**Why the prime factors are processed in this order.** The operation $\mathrm{LIFT}_p$ inserts the factor $p$ only when the current alphabet is $\mathbb{F}_p$. The operation $\mathrm{TRANSFER}_{r \to p}$ moves only from a larger prime alphabet to a smaller one. Therefore the alphabet can only decrease, which forces the construction to process the prime factors from largest to smallest.

# 7 From phase estimates to modular-sum distributions

Theorem 6.1 gives a Boolean source whose expectation is close to uniform for every modulus-$M$ local phase. The main theorem is stated in terms of statistical distance between residue distributions. This final section converts the phase estimates into that distributional statement.

## 7.1 The phase guarantee supplied by the construction

**Theorem 7.1** (Phase-error generator)**.** *For every fixed $M \geq 2$, every $n$, and every $0 < \gamma \leq 1$, there is an explicit source $G$ on $\{0,1\}^n$ with seed length $O_M(\log n + \log(1/\gamma))$ such that, for every $a \in \mathbb{Z}_M^n$,*

$$\boxed{\left|\mathbb{E}\omega_M^{\sum_i a_i G_i} - \mathbb{E}\omega_M^{\sum_i a_i U_i}\right| \leq \gamma,}$$

*where $U$ is uniform on $\{0,1\}^n$.*

*Proof.* Apply Theorem 6.1 with error $\gamma$. For every $a \in \mathbb{Z}_M^n$, define $q_i(0) = 0$ and $q_i(1) = a_i$. Then

$$\omega_M^{\sum_i a_i x_i} = \omega_M^{\sum_i q_i(x_i)}$$

is a member of $\mathsf{Local}_n(2, M)$, so the displayed estimate is exactly the error guarantee from Theorem 6.1. □

## 7.2 Fourier inversion on $\mathbb{Z}_M$

**Lemma 7.2** (From phase error to statistical distance)**.** *Let $X$ be a bit distribution such that, for every $b \in \mathbb{Z}_M^n$,*

$$\left|\mathbb{E}\omega_M^{\sum_i b_i X_i} - \mathbb{E}\omega_M^{\sum_i b_i U_i}\right| \leq \gamma.$$

*Then, for every $a \in \mathbb{Z}_M^n$,*

$$\boxed{\operatorname{dist}\left(S_a^{(M)}(X), S_a^{(M)}(U)\right) \leq \frac{M-1}{2}\,\gamma.}$$

*Proof.* Fix $a$, and let $P$ and $Q$ be the probability mass functions of $S_a^{(M)}(X)$ and $S_a^{(M)}(U)$ on $\mathbb{Z}_M$. For $t \in \mathbb{Z}_M$, define

$$\widehat{P}(t) = \sum_{z \in \mathbb{Z}_M} P(z)\omega_M^{tz}, \qquad \widehat{Q}(t) = \sum_{z \in \mathbb{Z}_M} Q(z)\omega_M^{tz}.$$

Then

$$\widehat{P}(t) = \mathbb{E}\omega_M^{tS_a^{(M)}(X)} = \mathbb{E}\omega_M^{\sum_i (ta_i)X_i},$$

and similarly for $Q$. At $t = 0$ both coefficients equal 1. For $t \neq 0$, the vector $ta$ is another member of $\mathbb{Z}_M^n$, so the hypothesis gives

$$|\widehat{P}(t) - \widehat{Q}(t)| \leq \gamma.$$

Fourier inversion on $\mathbb{Z}_M$ gives, for every residue $z$,

$$P(z) - Q(z) = \frac{1}{M} \sum_{t \in \mathbb{Z}_M} (\widehat{P}(t) - \widehat{Q}(t)) \omega_M^{-tz}.$$

The term $t = 0$ is zero, and the triangle inequality therefore gives

$$|P(z) - Q(z)| \le \frac{1}{M} \sum_{t \neq 0} |\widehat{P}(t) - \widehat{Q}(t)| \le \frac{M-1}{M} \gamma.$$

Summing this bound over the $M$ residues,

$$\operatorname{dist}(P, Q) = \frac{1}{2} \sum_{z} |P(z) - Q(z)| \le \frac{M-1}{2} \gamma.$$

□

**Theorem 7.3** (Main theorem)**.** *For every fixed integer $M \ge 2$, every $n \ge 1$, and every $0 < \varepsilon \le \frac{1}{2}$, there is an explicit generator*

$$G_{M,n,\varepsilon} : \{0,1\}^s \longrightarrow \{0,1\}^n, \qquad s = O_M\left(\log n + \log \frac{1}{\varepsilon}\right),$$

*such that, for every $a \in \mathbb{Z}_M^n$,*

$$\operatorname{dist}\big(S_a^{(M)}(G_{M,n,\varepsilon}), S_a^{(M)}(U_n)\big) \le \varepsilon.$$

*Proof.* Set

$$\gamma = \frac{2\varepsilon}{M-1}.$$

Since $0 < \varepsilon \le 1/2$ and $M \ge 2$, we have $0 < \gamma \le 1$. Let $G$ be the source from Theorem 7.1 for this value of $\gamma$. Its seed length is

$$O_M\left(\log n + \log \frac{1}{\gamma}\right) = O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

Lemma 7.2 gives, for every $a \in \mathbb{Z}_M^n$,

$$\operatorname{dist}\big(S_a^{(M)}(G), S_a^{(M)}(U_n)\big) \le \frac{M-1}{2} \gamma = \varepsilon.$$

□

## 7.3 Explicitness and optimality

The construction uses a number of LIFT and TRANSFER steps depending only on $M$. Every alphabet map is a fixed lookup table, every added random vector comes from an explicit small-bias construction over a fixed prime field, and all operations are coordinatewise. Hence the output is computable in time polynomial in $n$.

The dependence on both $n$ and $\varepsilon$ is optimal up to the constant depending on $M$.

**Proposition 7.4** (Reduction to a prime divisor)**.** *Let $q$ be a prime dividing $M$. Every bit distribution that $\varepsilon$-fools Boolean linear sums modulo $M$ also $\varepsilon$-fools Boolean linear sums modulo $q$.*

*Proof.* Given $b \in \mathbb{Z}_q^n$, define

$$a_i = \frac{M}{q} b_i \pmod{M}.$$

For every Boolean input $x$,

$$S_a^{(M)}(x) = \frac{M}{q} S_b^{(q)}(x) \pmod{M}.$$

The map

$$\iota : \mathbb{Z}_q \longrightarrow \mathbb{Z}_M, \qquad \iota(z) = \frac{M}{q} z$$

is one-to-one. Applying a one-to-one map to two distributions preserves their statistical distance. Therefore the distance between the two modulo-$q$ sums equals the distance between the corresponding modulo-$M$ sums, which is at most $\varepsilon$ by hypothesis. □

Consequently, any generator for modulus $M$ also fools linear sums modulo a fixed prime divisor of $M$. The standard lower bound for that prime-modulus problem is

$$\Omega\left(\log n + \log \frac{1}{\varepsilon}\right)$$

[AGHP92], [LRTV09]. Hence the seed length in Theorem 7.3 is optimal up to a constant depending on $M$.

# 8 References


[AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta, "Simple Constructions of Almost $k$-wise Independent Random Variables," *Random Structures & Algorithms*, 3(3):289–304, 1992.

[AMN98] Y. Azar, R. Motwani, and J. Naor, "Approximating Probability Distributions Using Small Sample Spaces," *Combinatorica*, 18(2):151–171, 1998. Supplies explicit small-bias distributions over non-binary alphabets, used here for every prime alphabet $p > 2$.

[BV10] A. Bogdanov and E. Viola, "Pseudorandom Bits for Polynomials," *SIAM Journal on Computing*, 39(6):2464–2486, 2010. Preliminary version in *FOCS 2007*.

[GKM18] P. Gopalan, D. M. Kane, and R. Meka, "Pseudorandomness via the Discrete Fourier Transform," *SIAM Journal on Computing*, 47(6):2451–2487, 2018. Preliminary version in *FOCS 2015*.

[LRTV09] S. Lovett, O. Reingold, L. Trevisan, and S. Vadhan, "Pseudorandom Bit Generators That Fool Modular Sums," *RANDOM 2009*. Their optimal-order construction covers every fixed prime-power modulus; the present paper imports nothing from it.

[Lov09] S. Lovett, "Unconditional Pseudorandom Generators for Low-Degree Polynomials," *Theory of Computing*, 5:69–82, 2009. Preliminary version in *STOC 2008*.

[LV25] C. H. Lee and E. Viola, "Pseudorandom Bits for Non-Commutative Programs," *40th Computational Complexity Conference (CCC 2025)*, 9:1–9:22.

[MZ09] R. Meka and D. Zuckerman, "Small-Bias Spaces for Group Products," *RANDOM 2009*.

[NN93] J. Naor and M. Naor, "Small-Bias Probability Spaces: Efficient Constructions and Applications," *SIAM Journal on Computing*, 22(4):838–856, 1993. Explicit small-bias distributions are the only pseudorandomness primitive imported by the upper bound.

[Smo87] R. Smolensky, "Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity," *19th Annual ACM Symposium on Theory of Computing (STOC 1987)*, 77–82.

[Vio09] E. Viola, "The Sum of $d$ Small-Bias Generators Fools Polynomials of Degree $d$," *Computational Complexity*, 18(2):209–217, 2009. Preliminary version in *CCC 2008*.




ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il