2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS) | 978-1-6654-5519-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/FOCS54457.2022.00045

Fooling polynomials using invariant theory*

Harm Derksen
Department of Mathematics
Northeastern University
Boston, USA
ha.derksen@northeastern.edu

Emanuele Viola
Khoury College of Computer Sciences
Northeastern University
Boston, USA
viola@ccs.neu.edu

***Abstract*—We revisit the problem of constructing explicit pseudorandom generators that fool with error ϵ degree-d polynomials in n variables over the field $\mathbb{F}_q$, in the case of large q. Previous constructions either have seed length $\geq 2^d \log q$, and thus are only non-trivial when $d < \log n$, or else rely on a seminal reduction by Bogdanov (STOC 2005). This reduction yields seed length not less than $d^4 \log n + \log q$ and requires fields of size $q \geq d^6/\epsilon^2$; and explicit generators meeting such bounds are known.**

Departing from Bogdanov's reduction, we develop an algebraic analogue of the Bogdanov-Viola paradigm (FOCS 2007, SICOMP 2010) of summing generators for degree-one polynomials. Whereas previous analyses of the paradigm are restricted to degree $d < \log n$, we give a new analysis which handles large degrees. A main new idea is to show that the construction preserves indecomposability of polynomials. Apparently for the first time in the area, the proof uses invariant theory.

Our approach in particular yields several new pseudorandom generators. In particular, for large enough fields we obtain seed length $O(d \log n + \log q)$ which is optimal up to constant factors. We also construct generators for fields of size as small as $O(d^4)$. Further reducing the field size requires a significant change in techniques: Most or all generators for large-degree polynomials rely on Weil bounds; but such bounds are only applicable when $q > d^4$.

***Index Terms*—pseudorandom generator, polynomial, invariant theory, algebraic geometry, sum**

A *pseudorandom generator* for degree-d polynomials over the field $\mathbb{F}_q$ in n variables with error ϵ is an explicit map $P : S \to \mathbb{F}_q^n$ that "ϵ-fools" any such polynomial g, that is, the distributions $g(U)$ and $g(P(U))$ have *statistical distance* (or *error*) at most ϵ. Here U denotes the uniform distribution over the appropriate domain ($\mathbb{F}_q^n$ in the first occurrence and S in the second). The *seed length* of P is $\log_2 |S|$. The minimum possible seed length is $\Omega(d \log(n/d) + \log q + \log 1/\epsilon)$, at least when $d < n^{0.99}$ and q is prime [1], [2]. Explicit constructions of generators (i.e., upper bounds on the seed length) have been intensely studied for at least 30 years. Two main lines of work exist. The first applies to any field [1], [3]–[8]. The last paper gives seed length $O(\log n + 2^d \log q/\epsilon) \cdot d$ which is the best available for small fields such as $\mathbb{F}_2$. The corresponding generators are obtained within the Bogdanov-Viola paradigm [1]: to fool polynomials of degree d, sum $\ell \geq d$ independent copies of generators for degree-one polynomials. While the parameters given by the analysis in [8] are non-trivial only for $d \leq \log n$, it is unknown whether the paradigm also works for larger degrees. If it did it would yield a breakthrough in complexity theory. For example, it would imply generators for small constant-depth circuits with parity gates, thanks to a well-known approximation due to Razborov [9].

The second lines of works applies only to fields of large size $q \gg d$, but can handle much larger degrees. Here Bogdanov's seminal paper [10] laid a paradigm that reduces constructing pseudorandom generators to constructing *hitting-set* generators for polynomials, an easier task. Bogdanov's paper was followed by a series of better and better constructions of hitting-set generators by Lu [11], Cohen and Ta-Shma [12], and Guruswami and Xing [13]; see also [14] for earlier related work by Klivans and Spielman. Optimal hitting-set constructions are now known; in combination with Bogdanov's reduction they yield the following pseudorandom generators.

Theorem 1. *[[10]+ [13]+([11] or [14])] There exist explicit pseudorandom generator that fool degree-d polynomials in n variables over $\mathbb{F}_q$ with seed length $O(d^4 \log n + \log q)$, provided $q \geq O(d^6/\epsilon^2)$.*

The notation $O(.)$ and $\Omega(.)$ denotes absolute constants. To connect with previous expressions for the seed length, note that adding a $\log 1/\epsilon$ term to the seed length in Theorem 1 does not change it since $q \geq 1/\epsilon$.

The parameters in Theorem 1 are essentially the best one can achieve using the reduction in [10], as we now explain. That reduction proceeds by showing that restricting a polynomial g onto a "good" plane preserves its output distribution with high probability. Once a good plane is found, one can then just pick a uniform element from the plane, which only costs two field elements. To find a good plane, [10] relies on results by Kaltofen [15] showing that (the coefficients of) planes that are bad for g are zeroes of a low-degree polynomial K_g. One can then use a hitting set to find a good plane. A bottleneck in this reduction is that the degree of K_g is at least d^4. So one needs a hitting-set generator for polynomials of degree at least d^4, resulting in the d^4 factor in the final seed length. The same loss arises in earlier work dealing with polynomials over complex numbers, see [15] for discussion. Over fields of large characteristic the degree can be improved from $O(d^4)$ to $O(d^2)$, which is known to be optimal, see [16]. Thus, this approach does not yield seed length less than $d^2 \log n$. For related reasons, the reduction in [10] requires the field size to be at least d^6.

Harm Derksen is partially supported by NSF grant DMS 2147769. Emanuele Viola is supported by NSF grant CCF-2114116.

978-1-6654-5519-0/22/$31.00 ©2022 IEEE
DOI 10.1109/FOCS54457.2022.00045

Constructions of pseudorandom generators in the two lines of research above have followed different paradigms. By contrast, we shall prove that the [1] paradigm works also for large-degree polynomials, at least as long as the field is large enough. This in particular yields pseudorandom generators with improved parameters, stated next.

Theorem 1. *There are explicit pseudorandom generators that fool with error ϵ degree-d polynomials in n variables over $\mathbb{F}_q$ with seed length $O(d \cdot m \cdot \log(dk+dm) + \log q)$, provided that $q \geq O(dk)^4/\epsilon^2$, for any integers m and k such that $\binom{m+k-2}{m-1} \geq n$.*

In particular we can have either

(1) seed length $O(d\log(dn) + \log q)$ provided that $q \geq O(d^4 n^{0.001})/\epsilon^2$, or

(2) seed length $O(d\log n \cdot \log(d\log n) + \log q)$ provided that $q \geq O(d\log n)^4/\epsilon^2$.

Item (1) achieves optimal seed length up to constant factors, when $d < n^{0.99}$. In particular it improves on the $\Omega(d^4 \log n)$ seed lengths of previous constructions. The field size improves on the $\Omega(d^6/\epsilon^2)$ field size of previous constructions (Theorem 1) when say $d > n^{0.001}$. This item is obtained by suitably setting $m = O(1)$ and $k = n^{\Omega(1)}$.

Item (2) achieves optimal seed length up to the lower-order factor $\log(d\log n)$. The field size improves on previous constructions for $d \geq \omega(\log^2 n)$. This item is obtained by setting $m = O(\log n)$ and $k = O(\log n)$.

We also obtain pseudorandom generators with the same seed length as previous constructions, but that only require $q \geq O(d^4)$, see Theorem 5. This improves on the $\Omega(d^6)$ field size of previous constructions. Further reducing the field size will require a significant change in techniques: Most or all generators for large-degree polynomials rely on Weil bounds, cf. Fact 12 or [17, Page 92]; but such bounds are only applicable when $q > d^4$.

Proof overview: A central concept in our proof, which was apparently not used before in the pseudorandomness literature, is that of *indecomposability*.

Definition 2. *A polynomial g over a field $\mathbb{F}$ is* indecomposable *if it cannot be written as $c \circ h$ where c is a univariate polynomial of degree ≥ 2 and both c and h are over $\mathbb{F}$.*

Let g be a polynomial we aim to fool. We begin by writing $g = c(h)$ where c is a univariate polynomial of maximal degree. We observe that the polynomial h is indecomposable, for else the degree of c is not maximal. A main technical contribution (discussed more below) is a universal (i.e., independent from g) construction of polynomials $f_1, f_2, \ldots, f_n$ that (i) are on few variables, (ii) have low degree, and (iii) *preserve indecomposability:* if $h(f_1, f_2, \ldots, f_n)$ is decomposable, then so is $h(x_1, x_2, \ldots, x_n)$. As observed above, the latter is not decomposable; hence the former is not decomposable either. We then prove (Lemma 9 in Section II) that the output distribution of indecomposable polynomials is close to uniform. This proof combines several results in algebraic geometry, including Weil's bound and results about reducibility of shifts of indecomposable polynomials.

Putting the above together we conclude that the f_i fool g because

$$\begin{aligned} g(U) &= c(h(U)) \approx c(U) \\ &\approx c(h(f_1, f_2, \ldots, f_n))(U) = g(f_1, f_2, \ldots, f_n)(U). \end{aligned}$$

Hence we have reduced the problem of fooling g to that of fooling g composed with the f_i. The gain is that by (i) we have reduced the number of variables. The main cost is an increase in degree, but this increase is small by (ii). Overall we obtain the following result, which is a main technical contribution of this work.

Theorem 2. *For every positive integers n, d, k and field $\mathbb{F}_q$:*

There is an explicit family of degree-k polynomials $f_1, f_2, \ldots, f_n$ over $\mathbb{F}_q$ in $(d+1)m$ variables such that for any polynomial g over $\mathbb{F}_q$ of degree d in n variables the statistical distance between $g(U)$ and $g(f_1, f_2, \ldots, f_n)(U)$ is $O(d^2k^2/\sqrt{q})$, for any m and k as in Theorem 1.

If we plug uniform values for the variables of the f_i we obtain pseudorandom generators with seed length as in Theorem 1 except that the factor $\log(dk+dm)$ is replaced with $\log q$. This is sufficient to prove the theorem when q is polynomial in dn. If q is larger, for example $q \geq 2^d$, it is not sufficient, and we need to improve the dependence on q from multiplicative to additive. To achieve this we combine Theorem 2 with another pseudorandom generator which we construct (Theorem 5). The latter generator combines Bogdanov's template [10] discussed earlier with some of our proof ideas. Compared with [10] and subsequent works, this generator has two main differences. First, we give a variant of Bogdanov's reduction of pseudorandom to hitting-set generators, again relying on preserving indecomposability. This allows us to improve the dependence on the field size. Note however that one can already obtain non-trivial generators over fields of size $O(d^4)$ from Theorem 2 (suitably set $k = O(1)$ and $m = n^{\Omega(1)}$). Second, we need to hit polynomials whose degree is larger than the number of variables, whereas in most previous work the degree is smaller. We note that such a hitting set can be obtained by combining [11], [13].

The construction of the f_i and its analysis using invariant theory: Let $M_1, M_2, \ldots$ be an enumeration of distinct monomials of degree k in m variables (in some cases we need some mild conditions on these monomials, discussed below). We take ℓ copies of the variables, and define $f_i := M_i^{[1]} + M_i^{[2]} + \ldots + M_i^{[\ell]}$ where $M_i^{[j]}$ is the monomial M_i where the variables are taken from copy j. Hence the construction is simple and very explicit.

The proof that the f_i preserve indecomposability uses *invariant theory*, apparently for the first time in this area, and proceeds as follows. Consider the polynomial $G := g(f_1, f_2, \ldots, f_n)$. First, note that G is *invariant* under permutation of the copies of variables (simply because the f_i are). Now assume that G can be decomposed as $G = c(H)$ for some

univariate polynomial c. We show that H must be invariant as well. Next, we show that the f_i are a basis for the invariant polynomials; this allows us to write $H = h(f_1, f_2, \ldots, f_s)$ for some low-degree polynomial h, where note a priori s could be much larger than n. Hence we obtained

$$g(f_1, f_2, \ldots, f_n) = c(h(f_1, f_2, \ldots, f_s)).$$

Finally, we show that this implies $s = n$ and $g(x_1, x_2, \ldots, x_n) = c(h(x_1, x_2, \ldots, x_n))$ as desired.

Three results on preserving indecomposability: We give three formal versions of the analysis in the previous paragraph.

The first version (Theorem 3 in Section I) has the easiest proof, requires fields of characteristic $> dk$, and takes $\ell > dk$ copies of variables. This version suffices to obtain generators with seed length $\tilde{O}(d \log^2 n) + O(\log q)$ over such fields, where $\tilde{O}(x)$ stands for $x \log^{O(1)} x$. Using the construction recursively, one can improve the seed length to $\tilde{O}(d \log n) + O(\log q)$, thus matching Item (2) in Theorem 1 up to lower-order factors for large characteristic, and in particular for prime fields. However these ideas do not suffice to obtain the optimal seed length in Item (1), for example. For this first version we can take any distinct monomials. This version also allows us to draw a close analogy with the Bogdanov-Viola paradigm [1]: We note that one can replace the M_i with any set of polynomials N_i of the same degree that fool degree-one polynomials. To verify this we can write the N_i as linear combinations of the M_i and use that the linear maps are full rank since the N_i fool degree-one polynomials.

The second version (Theorem 4 in Section IV) has a slightly more complicated proof, but requires only characteristic $> d$ and more importantly takes only $\ell = d + 1$ copies. This essentially matches the number $\ell = d$ of copies in [1], [8]. For this we need a certain mild condition on the monomials. This version suffices to prove Theorem 1 for fields of characteristic $> d$, and in particular for prime fields.

The third version is omitted and is the most complicated, but works over any characteristic, and again takes only $\ell = d + 1$ copies. Here we need to avoid obvious counterexamples; for example over $\mathbb{F}_2$ we cannot take $M_1 = x^2$ because $g = x$ is trivially indecomposable but $g(f_1) = (x^{[1]})^2 + (x^{[2]})^2 + \cdots + (x^{[\ell]})^2 = (x^{[1]} + x^{[2]} + \cdots + x^{[\ell]})^2$ is decomposable. It turns out that it suffices to take any M_i that are indecomposable. This version can be used to prove Theorem 2 as stated, for fields of any characteristic. Besides this, the results in this section allow us to preserve indecomposability over any field, even small. The only restriction on the field size then comes from Weil's bound (cf. Fact 12).

Open problems: A natural goal is to reduce the field size in Item (1) in Theorem 1 to $O(d^4)$. This would yield a single generator that improves on all those in this paper. The current bounds on the field size arise from applying Weil's bound to polynomials of degree dk rather than d. However, these polynomials of degree dk have a special structure as they arise from the composition of an arbitrary polynomial of degree d with the M^Σ's. It is conceivable that Weil's bound can be improved for such composed polynomials, perhaps to obtain bounds similar to those for degree-d polynomials. We raise this as an open problem.

I. Preserving indecomposability

In this section we give a first construction of polynomials that preserve indecomposability. We state the main theorem next after some notation. Then we proceed with the proof which involves several intermediate claims.

Let $\mathbb{F}_q$ be a field of characteristic p and let $R = \mathbb{F}_q[x_1, x_2, \ldots, x_m]$ be the polynomial ring in m variables. We define $R^{\otimes \ell} = \mathbb{F}_q[\{x_j^{[i]}\}]$ as the polynomial ring in the variables $x_j^{[i]}$ with $1 \le i \le \ell$ and $1 \le j \le m$. We can arrange the $\ell \cdot m$ variables in a matrix

$$X = \begin{pmatrix} x_1^{[1]} & x_2^{[1]} & \cdots & x_m^{[1]} \\ x_1^{[2]} & x_2^{[2]} & \cdots & x_m^{[2]} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{[\ell]} & x_2^{[\ell]} & \cdots & x_m^{[\ell]} \end{pmatrix} \tag{1}$$

Definition 3. *A* monomial *is a product of powers of variables (with leading coefficient* 1*). For a monomial* $M = M(x_1, x_2, \ldots, x_m) \in R = \mathbb{F}_q[x_1, x_2, \ldots, x_m]$ *we define* $M^{[i]} = M(x_1^{[i]}, x_2^{[i]}, \ldots, x_m^{[i]})$ *and* $M^\Sigma = \sum_{i=1}^{\ell} M^{[i]}$.

Theorem 3. *Suppose that* $M_1, M_2, \ldots, M_r \in R$ *are distinct non-constant monomials of degree* $\le k$*, and let* $g(x_1, x_2, \ldots, x_r)$ *be a non-constant polynomial of degree* d*. Let* $G := g(M_1^\Sigma, M_2^\Sigma, \ldots, M_r^\Sigma)$ *and assume that* $p \ge dk + 1$ *and* $\ell \ge \max\{5, dk + 1\}$*. If* G *is decomposable then* g *is decomposable.*

We remark that $\ell = d$ does not suffice for example for $d = 1$ and $k = 2$: take $M_1 = x_1^2$.

The rest of this section is devoted to proving the theorem. We view the *symmetric group* S_ℓ of permutations on ℓ elements as acting on $R^{\otimes \ell}$ by $\sigma(x_j^{[i]}) = x_j^{[\sigma(i)]}$ for all i, j. So the action of S_ℓ permutes the rows in X.

Definition 4. *For a monomial* M *in* $R^{\otimes \ell}$*, the* diversity *of* M *is the smallest number* d *such that the variables in* M *come from* d *rows in* X*. For a nonzero polynomial* $f \in R^{\otimes \ell}$*, the diversity* $\mathrm{div}(f)$ *is the largest diversity over all monomials appearing in* f.

For a subgroup G of S_ℓ and a polynomial $g \in R^{\otimes \ell}$ we say that g is G-invariant if $\sigma g = g$ for all $\sigma \in G$. Note that M^Σ is invariant under the action of S_ℓ and that $\mathrm{div}(M^\Sigma) = 1$ when M is not constant. In general we have the following proposition.

Proposition 5. *Suppose that* $f \in R^{\otimes \ell}$ *is an* S_ℓ*-invariant polynomial with* $\mathrm{div}(f) = d$ *and* $p > d$*. Then* f *can be written as a polynomial of degree* d *in the* M^Σ*'s.*

Proof. The *orbit sum* of a monomial $M := M_1^{[1]} M_2^{[2]} \cdots M_\ell^{[\ell]}$ where the M_j are in R is the sum of all monomials in the S_ℓ orbit $\{M_1^{[\pi 1]} M_2^{[\pi 2]} \cdots M_\ell^{[\pi \ell]} : \pi \in S_\ell\}$ of M. We note that any S_ℓ-invariant polynomial f can be written as a linear combination of orbit sums of monomials. Using this fact we

now prove the proposition by induction on $d = \mathrm{div}(f)$. If $d = 1$ then the orbit sums above are orbit sums of monomials that only involve one set of variables, so they are of the form M^{Σ}.

Now suppose $d > 1$. Without loss of generality, we may assume that f does not have monomials of diversity $< d$. Consider the orbit sum of a monomial $M_1^{[i_1]} M_2^{[i_2]} \cdots M_d^{[i_d]}$ where $M_1, M_2, \ldots, M_d \in \mathbb{F}_q[x_1, x_2, \ldots, x_m]$ are non-constant monomials. If the M_i are all distinct, then this orbit sum can be written as the sum

$$\sum_{i_1, i_2, \ldots, i_d} M_1^{[i_1]} M_2^{[i_2]} \cdots M_d^{[i_d]} \tag{2}$$

over all $(i_1, i_2, \ldots, i_d) \in \{1, 2, \ldots, \ell\}^d$ with $i_1, i_2, \ldots, i_d$ distinct.

If however some of the M_j's coincide, then in the sum (2) some of the monomials $M_1^{[i_1]} M_2^{[i_2]} \cdots M_d^{[i_d]}$ are summed more than once. (For example, if $d = \ell = 2$ and $M_1 = M_2 = x_1$ then the orbit sum of $M = M_1^{[1]} M_2^{[2]}$ has size 1, whereas in the sum above the same monomial would appear twice.) In the worst case, when all M_j's are the same, the same monomial is summed $d!$ times. This is not a problem, because the characteristic p of $\mathbb{F}_q$ is $> d$, so we can still write the orbit sum as the sum (2) multiplied by a non-zero field element.

So we can write f as a linear combination of sums (2) (for various choices of the M_i). Consider one such sum S. Note that the polynomial

$$S - M_1^{\Sigma} M_2^{\Sigma} \cdots M_d^{\Sigma}$$

has diversity $< d$. By the induction hypothesis, $S - M_1^{\Sigma} M_2^{\Sigma} \cdots M_d^{\Sigma}$ can be written as a polynomial of degree $< d$ in the M^{Σ}'s. So f can be written as a polynomial of degree d in the M^{Σ}'s. □

Let A_ℓ be the *alternating subgroup* of S_ℓ.

Lemma 6. *If a polynomial $f \in R^{\otimes \ell}$ is A_ℓ-invariant and $\deg(f) \leq \ell - 2$ then f is S_ℓ-invariant.*

Proof. First, assume that f is an A_ℓ orbit sum, i.e., there is a monomial N such that f is the sum of all elements in the set $\{\sigma \cdot N \mid \sigma \in A_\ell\}$. Because $\deg(N) \leq \ell - 2$, there exist two rows in (1), say i and j, such that N does not contain any variables from those rows. Then we have $(i\ j) \cdot N = N$, and since f was already A_ℓ-invariant we conclude that f is S_ℓ-invariant.

If f is arbitrary, we use the general fact that for any group G, if a polynomial is G-invariant, then f can be written as the sum of orbit sums polynomials. Hence we can apply the argument above to each orbit sum, and conclude the general case as well. □

Lemma 7. *If $f \in R^{\otimes \ell}$ is S_ℓ-invariant, $\deg(f) \leq \ell - 1$, $\ell \geq 5$ and $u \in R^{\otimes \ell}$ divides f, then u is S_ℓ-invariant.*

Proof. We can factor $f = f_1 f_2 \cdots f_s$ where f is irreducible. Factorization in the polynomial ring into irreducible factor is unique up to permuting factors and multiplying factors with nonzero constant scalars. From $f = \pi(f) = \pi(f_1)\pi(f_2)\cdots\pi(f_s)$ follows that for every i there exists a j and a nonzero constant $c \in \mathbb{F}_q^{\times} = \mathbb{F}_q - \{0\}$ such that $\pi(f_i) = c f_j$. In other words $\pi(L_i) = L_j$ where L_i is the span of f_i. Let $\mathcal{L} = \{L_1, L_2, \ldots, L_s\}$. Note that the set $\mathcal{L}$ may have less than s elements, because some factors may be the same up to a nonzero constant. Then S_ℓ acts on $\mathcal{L}$. Let $H_i \subseteq S_\ell$ be the stabilizer subgroup of L_i, that is, $\pi \in H_i$ if and only if $\pi(L_i) = L_i$. By the *orbit-stabilizer* theorem, the index $|S_\ell|/|H_i|$ of H_i in S_ℓ equals the size of the orbit of L_i. The latter is $\leq |\mathcal{L}| \leq s$. Moreover, $s \leq \deg(f) \leq \ell - 1$, where the second inequality is by assumption. Hence, the index of H_i is $< \ell$. It is known that the only proper subgroup of S_ℓ of index $< \ell$ is A_ℓ, see e.g. [18, p. 84]. So it follows that $H_i = A_\ell$ or S_ℓ. This proves that $\pi(L_i) = L_i$ for all $\pi \in A_\ell$.

We now argue that in fact even $\pi(f_i) = f_i$ for all i and all $\pi \in A_\ell$. Fix i. From $\pi(L_i) = L_i$ for all $\pi \in A_\ell$ we know that for every $\pi \in A_\ell$ there exists a (unique) element $\chi_i(\pi) \in \mathbb{F}_q - \{0\}$ such that $\pi(f_i) = \chi_i(\pi) f_i$. Notice that $\chi_i : A_\ell \to \mathbb{F}_q^{\times}$ is a group homomorphism. Let K be its kernel. The kernel of any group homomorphism is a normal subgroup, so K is a normal subgroup of A_ℓ. On the other hand, A_ℓ is simple for $\ell \geq 5$, that is, it has no non-trivial normal subgroups. So either $K = A_\ell$ or $K = \{1\}$. We can exclude the latter possibility because it would imply that A_ℓ is commutative, which is not true. (We would have $\pi \cdot \pi' = \chi_i^{-1}\chi_i(\pi \cdot \pi') = \chi_i^{-1}(\chi_i(\pi) \cdot \chi_i(\pi')) = \chi_i^{-1}(\chi_i(\pi') \cdot \chi_i(\pi)) = \chi_i^{-1}\chi_i(\pi' \cdot \pi) = \pi' \cdot \pi$using that $\mathbb{F}_q^{\times}$ is commutative.) Hence $K = A_\ell$ and $\pi(f_i) = \chi_i(\pi) f_i = f_i$ for all $\pi \in A_\ell$.

Therefore, $f_1, f_2, \ldots, f_s$ are A_ℓ-invariant. If $s = 1$, then $f_t = f$ is S_ℓ-invariant. If $s > 1$, then $\deg(f_i) \leq \ell - 2$ for all i, and f_i is S_ℓ-invariant by lemma 6. Up to a constant, u is a product of the f_i's, so u is S_ℓ-invariant. □

Proposition 8. *Suppose that $M_1, M_2, \ldots, M_r \in \mathbb{F}_q[x_1, x_2, \ldots, x_m]$ are distinct non-constant monomials, $g(x_1, x_2, \ldots, x_r)$ is a polynomial of degree $d \leq \ell$ and $p > d$. If $g(M_1^{\Sigma}, M_2^{\Sigma}, \ldots, M_r^{\Sigma}) = 0$, then $g = 0$.*

Proof. Consider a monomial of maximal degree d in g, say $x_{i_1} x_{i_2} \cdots x_{i_d}$ with $i_1 \leq i_2 \leq \cdots \leq i_d$. Then the monomial $M_{i_1}^{[1]} M_{i_2}^{[2]} \cdots M_{i_d}^{[d]}$ appears in $M_{i_1}^{\Sigma} M_{i_2}^{\Sigma} \cdots M_{i_d}^{\Sigma}$. Here we use the assumption on the characteristic, needed for example if $i_1 = i_2 = \ldots = i_d$. Also, if $j_1 \leq j_2 \leq \cdots \leq j_d$ and $(i_1, i_2, \ldots, i_d) \neq (j_1, j_2, \ldots, j_d)$, then $M_{i_1}^{[1]} M_{i_2}^{[2]} \cdots M_{i_d}^{[d]}$ does not appear in $M_{j_1}^{\Sigma} M_{j_2}^{\Sigma} \cdots M_{j_d}^{\Sigma}$. Also, $M_{i_1}^{[1]} M_{i_2}^{[2]} \cdots M_{i_d}^{[d]}$ does not appear in $M_{j_1}^{\Sigma} M_{j_2}^{\Sigma} \cdots M_{j_{d'}}^{\Sigma}$ if $d' < d$ since the latter has diversity $\leq d'$ while the former has diversity d.

This shows that the monomial $M_{i_1}^{[1]} M_{i_2}^{[2]} \cdots M_{i_d}^{[d]}$ appears in $g(M_1^{\Sigma}, M_2^{\Sigma}, \ldots, M_r^{\Sigma})$. In particular, $g(M_1^{\Sigma}, M_2^{\Sigma}, \ldots, M_r^{\Sigma}) \neq 0$. □

We can now prove the main theorem of this section.

Proof of Theorem 3. Suppose that G can be decomposed as $G = c(H)$ for some $H \in R^{\otimes \ell}$ and univariate polynomial

$c \in \mathbb{F}_q[x]$ of degree $e \geq 1$. Note that G has degree $\leq dk$. Let $\alpha \in \overline{\mathbb{F}}_q$ be a root of $c(x)$. Then $x - \alpha$ divides $c(x)$, and so $H - \alpha$ divides $c(H)$. Then $H - \alpha$, and hence H, is S_ℓ-invariant by Lemma 7, using that $\ell \geq dk + 1$. Note that if $\alpha \in \overline{\mathbb{F}}_q$ does not lie in $\mathbb{F}_q$, then we have to apply Lemma 7 after replacing $\mathbb{F}_q$ with a finite field extension of $\mathbb{F}_q$ that contains α.

From the degree bounds on $G = c(H)$ and c it follows that H has degree $\leq dk/e$. In particular, $\mathrm{div}(H) \leq dk/e$. By Proposition 5 we can write H as a polynomial of degree $\leq dk/e$ in all M^Σ's, say $H = h(M_1^\Sigma, M_2^\Sigma, \ldots, M_s^\Sigma)$ for some s. Note that s may be larger than r.

If we set $u(x_1, x_2, \ldots, x_s) = g(x_1, x_2, \ldots, x_r) - c(h(x_1, x_2, \ldots, x_s))$, then we have

$$u(M_1^\Sigma, M_2^\Sigma, \ldots, M_s^\Sigma) = 0.$$

Proposition 8 implies that $u = 0$. So $g(x_1, x_2, \ldots, x_r) = c(h(x_1, x_2, \ldots, x_s))$. So $h(x_1, x_2, \ldots, x_s) = h(x_1, x_2, \ldots, x_r)$ only depends on $x_1, x_2, \ldots, x_r$ and the degree of h is $\leq d/e$. □

II. Indecomposability implies equidistribution

In this section we prove the following lemma.

Lemma 9. *Let h be a polynomial of degree d in n variables over $\mathbb{F}_q$. If h is indecomposable then $h(U)$ is $O(d^2/\sqrt{q})$-close to uniform over $\mathbb{F}_q$.*

For the proof we need several facts from the algebraic-geometry literature.

Fact 10. *[19] Let h be a polynomial of degree d in n variables over an algebraically-closed field K. Suppose that h is indecomposable. Then the number of $\lambda \in K$ such that $h - \lambda$ is reducible in K is at most d.*

[19] generalizes several previous works; we refer to [19] for the history of this type of results. Our polynomials are over $\mathbb{F}_q$ which is not algebraically closed. However the following fact allows us to bypass this apparent obstacle. If K is a field the notation $\overline{K}$ denotes its *algebraic closure*.

Fact 11. *[20, Theorem 4.2.] If a polynomial is indecomposable over $\mathbb{F}_q$ then it is also indecomposable over $\overline{\mathbb{F}_q}$.*

Finally, we use the following version of Weil's bound.

Fact 12. *[10, Proposition 2.6] Let h be a non-constant polynomial of degree d in n variables over $\mathbb{F}_q$ that cannot be reduced in $\overline{\mathbb{F}_q}$. Then $|\mathbb{P}[h(U) = 0] - 1/q| \leq O(d^2 q^{-3/2})$, assuming $q > 5d^4$.*

Proof of Lemma 9. By Fact 11 h is also indecomposable over $\overline{\mathbb{F}_q}$. By Fact 10, $h - \lambda$ is not reducible in $\overline{\mathbb{F}_q}$ except for at most d values of $\lambda \in \mathbb{F}_q$. For each value λ for which it is not reducible, Fact 12 yields $|\mathbb{P}[h(U) = \lambda] - 1/q| \leq O(d^2 q^{-3/2})$. Note we can assume $q > 5d^4$ for else the conclusion of the lemma holds. For any other value of λ, by Schwartz-Zippel, $|\mathbb{P}[h(U) = \lambda] - 1/q| \leq d/q$. Combining these facts, the statistical distance between $h(U)$ and uniform is at most $O(d^2/\sqrt{q}) + d^2/q = O(d^2/\sqrt{q})$. □

III. Toy pseudorandom generators with what we have so far

In this section we derive "toy" pseudorandom generators with the results of the previous two sections, over fields of characteristic $> dk$. Define $f_i := M_i^\Sigma$ as in the introduction. The generator simply picks ℓm uniform values for the variables of the f_i and outputs $(f_1, f_2, \ldots, f_n)(U)$. The analysis goes as follows. Let g be a polynomial of degree d that we aim to fool. Let c be a univariate polynomial of maximal degree such that $g(x_1, x_2, \ldots, x_n) = c(h(x_1, x_2, \ldots, x_n))$. In particular we have $g(f_1, f_2, \ldots, f_n) = c(h(f_1, f_2, \ldots, f_n))$. Note that h has degree $\leq d$ and is indecomposable, for else the degree of c is not maximal. By Theorem 3, $h(f_1, f_2, \ldots, f_n)$ is indecomposable as well. By Lemma 9, $h(f_1, f_2, \ldots, f_n)(U)$ is $O(d^2k^2/\sqrt{q})$-close to uniform, and the same bound holds for $h(U)$.

Hence we obtained generators with seed length $O(\ell m \log q) = O(dkm \log q)$ and error $O(dk)^2/\sqrt{q}$. Here we just need $\binom{m+k}{m} \geq n$. For example, we can pick m and k to be $O(\log n)$. This gives seed length $O(d \log^2 n \log q)$. As mentioned earlier, one can improve the seed length to $O(d \log n \log^{O(1)} \log(dn))$ by applying the construction recursively.

IV. Improving bounds for indecomposability

In this section we improve the bounds in Theorem 3 to get the preservation of indecomposability for $\ell \geq d + 1$ instead of $\ell \geq dk + 1$. The factor-k loss in the previous argument arises when bounding the diversity of H by the degree of H, where the latter is a priori as large as dk/e, see the proof of Theorem 3. In this section we consider a more constrained set Q of monomials, defined shortly. Using this, we can recoup a factor k when bounding the diversity of a polynomial in terms of its degree, see Lemma 13.

We fix a positive integer k and let $Q \subseteq R = \mathbb{F}_q[x_1, x_2, \ldots, x_m]$ be the subring spanned by all monomials of the form $x_1^{a_1} x_2^{a_2} \cdots x_m^{a_m} \in R$ where $a_1 + a_2 + \cdots + a_{m-1} = (k-1)a_m$. Note that the degree of a polynomial in Q is ka_m which is always divisible by k. Let $Q^{\otimes \ell} \subset R^{\otimes \ell}$ be the subring spanned by all monomials $M_1^{[1]} M_2^{[2]} \cdots M_\ell^{[\ell]}$ where $M_1, M_2, \ldots, M_\ell \in Q$.

We modify Theorem 3 by only considering monomials M^Σ where M is a monomial in the subring $Q \subset R$ rather than in R. By doing so, as we mentioned, we improve the parameters as follows.

Theorem 4. *Suppose that $M_1, M_2, \ldots, M_r \in Q$ are distinct non-constant monomials of degree k, and let $g(x_1, x_2, \ldots, x_r)$ be a non-constant polynomial of degree d. Let $G := g(M_1^\Sigma, M_2^\Sigma, \ldots, M_r^\Sigma)$ and assume that $p \geq d + 1$ and $\ell \geq \max\{5, d+1\}$. If G is decomposable then g is decomposable.*

The rest of this section is devoted to the proof of this theorem. The proof follows the same outline of the proof of the corresponding Theorem 3 in Section I, but some of the steps are more involved.

First, as mentioned above, we give a tighter connection between diversity and degree for polynomials in $Q^{\otimes \ell}$.

Lemma 13. *If $f \in Q^{\otimes \ell}$ has degree $\leq dk$, then $\mathrm{div}(f) \leq d$.*

Proof. The polynomial f is by definition a linear combination of monomials of the form $M_1^{[1]} M_2^{[2]} \cdots M_\ell^{[\ell]}$ where M_i is a monomial in Q of total degree $\leq dk$. If $M_i \neq 1$, then the degree of M_i is at least k. So $M_i \neq 1$ for at most d distinct indices i. This proves that $\mathrm{div}(f) \leq d$. □

We also modify Proposition 5 as follows.

Proposition 14. *Suppose that $f \in Q^{\otimes \ell}$ is an S_ℓ-invariant polynomial with $\mathrm{div}(f) = d$ and $p > d$. Then f can be written as a polynomial of degree d in the M^Σ's, where M ranges over monomials in Q.*

Proof. We follow the proof of Proposition 5 and note that all the monomials that appear can be chosen in Q and $Q^{\otimes \ell}$ instead of R and $R^{\otimes \ell}$ respectively. □

One difficulty that we face when generalizing the other statements in Section I such as Lemma 7 is that of arguing that the polynomials we encounter lie in $Q^{\otimes \ell}$ instead of $R^{\otimes \ell}$. For this purpose it is convenient to introduce the ring homomorphism

$$\Phi : R^{\otimes \ell} \to R^{\otimes \ell}[t^{[1]}, \ldots, t^{[\ell]}, (t^{[1]})^{-1}, \ldots, (t^{[\ell]})^{-1}]$$

with $\Phi(x_j^{[i]}) = t^{[i]} x_j^{[i]}$ for $j < m$, $\Phi(x_m^{[i]}) = (t^{[i]})^{1-k} x_m^{[i]}$ and $\Phi(\alpha) = \alpha$ for $\alpha \in \mathbb{F}_q$. Note that the image of Φ is a *Laurent polynomial*, that is, the exponents of the variables $t^{[i]}$ may be negative. If we work over the algebraically closed field $\overline{\mathbb{F}}_q$, then Φ corresponds to an action of the ℓ-dimensional torus group $T = (\overline{\mathbb{F}}_q^\times)^\ell$ on the ring $R^{\otimes \ell}$. This motivates the terminology that follows.

A polynomial $f \in R^{\otimes \ell}$ is *T-invariant* when $\Phi(f) = f$, i.e., the variables $t^{[i]}$ cancel out. Note that if $M \in Q$ then $M^{[i]}$ is T-invariant. More generally, $Q^{\otimes \ell}$ is the ring of all T-invariants. A polynomial $f \in R^{\otimes \ell}$ is called *T-semi-invariant* if $\Phi(f) = (\prod_{i=1}^{\ell} (t^{[i]})^{a^{[i]}}) f$ for some $a = (a^{[1]}, a^{[2]}, \ldots, a^{[\ell]}) \in \mathbb{Z}^\ell$, called *weight*. The monomials in $R^{\otimes \ell}$ are all T-semi-invariant.

As in Lemma 7, we need to argue about factors of invariant polynomials. We begin with the following lemma which will help us argue that these factors lie in $Q^{\otimes \ell}$ (as opposed to $R^{\otimes \ell}$).

Lemma 15. *Suppose that $u, f \in R^{\otimes \ell}$ and u divides f. If f is T-semi-invariant, then so is u.*

We shall only use this for T-invariant f, but the proof is the same.

Proof. Note that f is T-semi-invariant if and only if $\Phi(f)$, as a Laurent polynomial in $t^{[1]}, \ldots, t^{[\ell]}$, consists of a single monomial. If $f = uv$ then note $\Phi(f) = \Phi(u)\Phi(v)$. By assumption, $\Phi(f)$ consists of a single monomial as a Laurent polynomial. Then the same is also true for $\Phi(u)$ and $\Phi(v)$. Here we are using the general fact that if, say, $\Phi(u)$ has more than one term, then the product $\Phi(u)\Phi(v)$ has more than one term. To see this, consider the lexicographic order on monomials, and note that the product of the smallest monomial in $\Phi(u)$ with the smallest monomial in $\Phi(v)$ cannot be obtained by multiplying any other two monomials, and the same holds for the product of the largest monomials. Hence the product has at least two monomials. □

We modify Lemma 7 to the following statement:

Lemma 16. *If $f \in Q^{\otimes \ell}$ is S_ℓ-invariant, $\deg(f) \leq k\ell - 1$, $\ell \geq 5$ and $u \in R^{\otimes \ell}$ divides f, then u lies in $Q^{\otimes \ell}$ and is S_ℓ-invariant.*

Proof. We modify the proof of Lemma 7. We started with a factorization $f = f_1 f_2 \ldots f_s$ where $f_1, f_2, \ldots, f_s$ are irreducible. Since $f \in Q^{\otimes \ell}$ it is T-invariant, and therefore T-semi-invariant. We defined $L_i = \mathbb{F}_q f_i$ and considered the action of S_ℓ on $\mathcal{L} = \{L_1, L_2, \ldots, L_s\}$. As before H_i is the stabilizer of L_i. By Lemma 15, $f_1, f_2, \ldots, f_s$ are also semi-invariant. Let us fix some j, and let $(a^{[1]}, a^{[2]}, \ldots, a^{[\ell]})$ be the weight of the semi-invariant f_j.

We prove that $a^{[i]} \geq 0$ for all i. First recall that a monomial in Q is of the form $x_1^{a_1} x_2^{a_2} \cdots x_m^{a_m}$ with $a_1 + a_2 + \cdots + a_m = ka_m$. This means that the total degree of a polynomial $f \in Q$ in the variables $x_1, x_2, \ldots, x_m$ is exactly k times the degree of f as a polynomial in the variable x_m with coefficients in $\mathbb{F}_q[x_1, \ldots, x_{m-1}]$. Similarly, if $f \in Q^{\otimes \ell}$ then the total degree $\deg(f)$ of f is k times the degree $\deg_m(f)$ of f in the variables $x_m^{[1]}, x_m^{[2]}, \ldots, x_m^{[\ell]}$. Therefore for $f \in Q^{\otimes \ell}$ we have $\deg_m(f) = \deg(f)/k < \ell$.

Now suppose towards a contradiction that $a^{[i]} < 0$ for some i. Then $x_m^{[i]}$ must appear in f_j, so $\deg_m(f_j) \geq 1$. The orbit of L_j has $|S_\ell|/|H_j|$ elements, which correspond to as many irreducible factors of f that are distinct and have degree ≥ 1 with respect to $\deg_m$. This implies that $|S_\ell|/|H_j| \leq \deg_m(f) < \ell$. As in the proof of Lemma 7, this implies that $H_j = A_\ell$ or $H_j = S_\ell$.

For a permutation σ, $\sigma(f_j)$ is semi-invariant. Its weight is $\sigma(a) = (a^{[\sigma(1)]}, \ldots, a^{[\sigma(\ell)]})$. For $\sigma \in H_j$, $\sigma(f_j)$ and f_j are the same up to a constant, so $\sigma(a) = a$ for all $\sigma \in H_j$. If $H_j = S_\ell$ then $a^{[1]} = a^{[2]} = \cdots = a^{[\ell]} < 0$. The same holds if $H_j = A_\ell$ for $\ell \geq 3$ as we can again map i to any other value via $\sigma \in A_\ell$. This implies that all the monomials in f_j contain $\prod_{i=1}^m x_m^{[i]}$ and $\deg_m(f_j) \geq \ell$, contradicting the bound above.

We proved that $a^{[i]} \geq 0$ for all i, so $\Phi(f_j)$ lies in the polynomial ring $R^{\otimes \ell}[t^{[1]}, \ldots, t^{[\ell]}]$ for all j. From $\prod_{j=1}^s \Phi(f_j) = \Phi(f) = f \in R^{\otimes \ell}$ it follows that $\Phi(f_j) \in R^{\otimes \ell}$ for all j. This implies that $f_j \in Q^{\otimes \ell}$ for all j.

There remains to argue that the f_j are S_ℓ-invariant. As in the proof of Lemma 7, $\sigma(L_i) = L_i$ for all $\sigma \in A_\ell$ implies that $\sigma(f_i) = f_i$ for all $\sigma \in A_\ell$. Hence, the f_i are A_ℓ-invariant.

If $s = 1$, then $f = f_1$ is S_ℓ-invariant. Otherwise, $\deg_m(f_j) \leq \ell - 2$. Since $f_j \in Q^{\otimes \ell}$ we get $\deg(f_j) = k \deg_m(f_j) \leq (\ell - 2)k$. By Lemma 13, $\mathrm{div}(f_j) \leq \ell - 2$. Using Lemma 6 we conclude that f_j is S_ℓ-invariant. □

Proof of Theorem 4. Suppose that G can be decomposed as $G = c(H)$ for some $H \in R^{\otimes \ell}$ and univariate polynomial $c \in \mathbb{F}_q[x]$ of degree $e \geq 1$. We claim that in fact $H \in Q^{\otimes \ell}$. To verify this, note that from $G = c(H)$ it follows $\Phi(G) = c(\Phi(H))$. Since $\Phi(G)$ is T-invariant, i.e., constant in the variables $t^{[1]}, t^{[2]}, \ldots, t^{[\ell]}$, so is $\Phi(H)$ and $H \in Q^{\otimes \ell}$.

Note that G has degree $\leq dk$. Let $\alpha \in \overline{\mathbb{F}}_q$ be a root of $c(x)$. Then $x - \alpha$ divides $c(x)$, and so $H - \alpha$ divides $c(H)$. Because $\deg(H - \alpha) \leq dk < k\ell$, $H - \alpha$ lies in $Q^{\otimes \ell}$ and is S_ℓ-invariant by Lemma 16. (Possibly, we may have to replace $\mathbb{F}_q$ by a finite field extension.) It follows that $H \in Q^{\otimes \ell}$ and is S_ℓ-invariant.

From the degree bounds on $G = c(H)$ and c it follows that H has degree $\leq dk/e$. By Lemma 13, we get $\mathrm{div}(H) \leq d/e < \ell$. By Proposition 14 we can write H as a polynomial of degree $\leq d/e$ in all M^Σ's with $M \in Q$, say $H = h(M_1^\Sigma, M_2^\Sigma, \ldots, M_r^\Sigma, \ldots, M_s^\Sigma)$. If we set $u(x_1, x_2, \ldots, x_s) = g(x_1, x_2, \ldots, x_\ell) - c(h(x_1, x_2, \ldots, x_s))$, then we have

$$u(M_1^\Sigma, M_2^\Sigma, \ldots, M_s^\Sigma) = 0.$$

and the degree of u is $\leq d$. Proposition 8 implies that $u = 0$. So $g(x_1, x_2, \ldots, x_r) = c(h(x_1, x_2, \ldots, x_s))$. So $h(x_1, x_2, \ldots, x_s) = h(x_1, x_2, \ldots, x_r)$ only depends on $x_1, x_2, \ldots, x_r$ and the degree of h is $\leq d/e$. □

V. Bogdanov-style generators

In this section we prove the following theorem.

Theorem 5. *There are explicit pseudorandom generators that fool with error ϵ degree-d polynomials in n variables over $\mathbb{F}_q$, provided $q \geq O(d^4/\epsilon^2)$, with seed length either*
(1) $O(n \log(d+n) + \log q)$ or
(2) $O(d^4 \log n + \log q)$.

First we refine Bogdanov's reduction of pseudorandom generators to hitting-set generators. An explicit map $H : S \to \mathbb{F}_q^n$ is a *δ-hitting-set* generator for degree-d polynomials in n variables over $\mathbb{F}_q$ if for any such polynomial f, if $f \neq 0$ then $\mathbb{P}[f(H(U)) = 0] \leq \delta$. The seed length of H is $\log_2 |S|$.

We obtain the following refinement of Bogdanov's reduction:

Lemma 17. *Suppose there exists a δ-hitting-set generator with seed length s for polynomials of degree $3d^4$ in $2n$ variables over $\mathbb{F}_q$. Then there exists a pseudorandom generator for polynomials of degree d in n variables over $\mathbb{F}_q$ with seed length $2s + 2\log q$ and error $O(\delta + d^2/\sqrt{q})$.*

[10, Theorem 3.1] proves the same but with error $O(\sqrt{\delta} d + d^2/\sqrt{q} + d^6/q)$. To prove Lemma 17 first we use the following result to relate indecomposability and *irreducibility*.

Fact 18. *[21, Lemma 7] Let $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ be a non-constant polynomial. Then f is indecomposable over $\overline{\mathbb{F}}$ iff $f - y$ is irreducible in $\overline{\mathbb{F}(y)}[x_1, x_2, \ldots, x_n]$.*

Here $\overline{\mathbb{F}(y)}$ is the algebraic closure of the function field $\mathbb{F}(y)$, where y is a variable.

We also need the following fact, mentioned already in [10] when $\mathbb{E} = \overline{\mathbb{F}}$.

Fact 19. *Let $\mathbb{F} \subseteq \mathbb{E}$ be a field extension. Let H be a δ-hitting-set generator for degree-d polynomials over $\mathbb{F}$. Then H is also a δ-hitting-set generator for polynomials over $\mathbb{E}$.*

This fact follows because $\mathbb{E}$ is a vector space over $\mathbb{F}$.

Proof of Lemma 17. Let g be a polynomial that we aim to fool. As in Section III, write $g = c(h)$ where c is a univariate polynomial of maximal degree, and h is indecomposable. It suffices to preserve the output distribution of h, which by Lemma 9 is close to uniform. We relate indecomposability to irreducibility via Fact 18, inspired by the proof of Theorem 8 in [21], then reason as in [10], using Theorem 5 in [15].

By Fact 11, h is indecomposable over $\overline{\mathbb{F}_q}$ as well. Hence we can apply Fact 18 to conclude that $h - y$ is irreducible in $\mathbb{E}[x_1, x_2, \ldots, x_n]$ where $\mathbb{E} := \overline{\mathbb{F}_q(y)}$. We now use Theorem 5 in [15] over the field $\mathbb{E}$. For $v_{1..n} \in \mathbb{E}^n$ and $w_{2..n}, z_{2..n} \in \mathbb{E}^{n-1}$ define the following bivariate restriction of h:

$$h|_{v,w,z}[s,t] := h(s + v_1, w_2 s + z_2 t + v_2 \ldots, w_n s + z_n t + v_n).$$

Theorem 5 in [15] shows that $h|_{v,w,z}$ is absolutely irreducible except when v, w are zeroes of a polynomial of degree $O(d^2)$ over $\mathbb{E}$, or z is the zero of a polynomial of degree $O(d^4)$ over $\mathbb{E}$ (where the latter polynomial may depend on the first).

For our generator, we pick (v, w) with a δ-hitting-set generator for polynomials of degree $O(d^2)$ and z with an independent δ-hitting-set generator with error ϵ for polynomials of degree $O(d^4)$. For the variables s and t we plug uniform values in $\mathbb{F}_q$.

By Fact 19, these hitting-set generators are also δ-hitting-set generators polynomials over $\mathbb{E}$. Hence, $h|_{v,w,z}$ is absolutely irreducible over $\mathbb{E}[s,t]$ with probability $\geq 1 - O(\delta)$. Then from Fact 18 we obtain that $h|_{v,w,z}$ is indecomposable with at least the same probability over the choice of v, w, z from the hitting-set generators. Whenever it is indecomposable, by Lemma 9 its output distribution is $O(d^2)/\sqrt{q}$-close to uniform. □

To prove Theorem 5 there remains to construct δ-hitting-set generators. Such for polynomials of degree d in n variables are known with optimal seed length $O(d \log n + \log 1/\delta)$, provided $q \geq O(d/\delta)$ [13]. In particular, for polynomials of degree d^4 in $O(n)$ variables we can set $\delta := \epsilon d^2/\sqrt{q}$ and have seed length $O(d^4 \log n + \log q)$, provided $q \geq O(d^4/(\epsilon d^2/\sqrt{q}))$. The last provision is equivalent to $q \geq O(d^4/\epsilon^2)$, which we can always assume for else the theorem is trivial. This gives Item (2) in Theorem 5.

Over fields of characteristic $\geq O(d^2)$ the d^4 factor can be improved to d^2 using Corollary 8 in [16] – and that is the best possible, see Corollary 7 and the surrounding discussion in the same paper.

For Item (1) in Theorem 5 we need a different hitting-set generator, stated next.

Lemma 20. *[Implicit in [11], [13]] There is an explicit δ-hitting-set generator with seed length $O(n \log(n+d) + \log 1/\delta)$*

for polynomials of degree d in n variables over $\mathbb{F}_q$, provided $q \geq O(d/\delta)$.

This should be compared to the Schwartz-Zippel lemma, which yields a δ-hitting-set generator with seed length $n \log(d/\delta)$ provided $q \geq d/\delta$. As the above lemma is not stated in those works we quickly sketch how it follows from [11], [13]. Lu [11] (Theorem 1) gives a δ-hitting-set generator for polynomials with s terms with seed length $O(\log(sd/\delta))$ provided $q \geq d^{1.01}/\delta$. (Lu's proof focuses on constant δ, but as noted there and in [13] one can also obtain the stated parameters.) A degree-d polynomial in n variables has $s \leq \binom{n+d}{n}$ monomials. Hence we obtain seed length $O(n \log(n+d) + \log 1/\delta)$. Guruswami and Xing [13] use multiplication-friendly codes to bring down the bound on the field size to $q \geq O(d/\delta)$.

To prove Item (1) in Theorem 5, use the δ-hitting-set generator in Lemma 20 for polynomials of degree d^4 in $O(n)$ variables, setting $\delta := \epsilon d^2/\sqrt{q}$.

VI. Proof of main results for fields of characteristic $> d$

In this section we prove our main results, Theorem 1 and Theorem 2, in the case of fields of characteristic $> d$ (for example, prime fields).

Proof of Theorem 2. Let Q and $M_1, M_2, \ldots$ be as in Theorem 4. The number of distinct monomials in Q is at least the number of positive integers $a_1, a_2, \ldots, a_{m-1}$ with sum equal to $k-1$ (corresponding to the setting $a_m = 1$ in Section IV). This number is $\binom{m-1+k-1}{m-1}$, which is $\geq n$ by assumption. Define $f_i := M_i^{\Sigma}$. The analysis is the same as in Section III. □

Proof of Theorem 1. From Theorem 2 we reduce our task to that of fooling polynomials with degree $d' := dk$ in $n' := \ell m = (d+1)m$ variables, up to an error $O(d'^2/\sqrt{q})$. This error is $\leq \epsilon$ by our assumption that $q \geq O(dk)^4/\epsilon^2$.

Item (1) in Theorem 5 shows how to fool such polynomials with seed length $O(n' \log(d'+n') + \log q)$ and error β, provided $q \geq O(d'^4/\beta^2)$. This allows us to set $\beta := O(d'^2/\sqrt{q})$ and the provision is true. Again by our assumption that $q \geq O(dk)^4/\epsilon^2$, we have $\beta = O(\epsilon)$. Hence the combined error from the two steps is $O(\epsilon)$. The final seed length is $O(dm \log(dk+dm) + \log q)$, as desired. □

References

[1] A. Bogdanov and E. Viola, "Pseudorandom bits for polynomials," *SIAM J. on Computing*, vol. 39, no. 6, pp. 2464–2486, 2010.

[2] N. Alon, I. Ben-Eliezer, and M. Krivelevich, "Small sample spaces cannot fool low degree polynomials," in *12th Workshop on Randomization and Computation (RANDOM)*. Springer, 2008, pp. 266–275.

[3] J. Naor and M. Naor, "Small-bias probability spaces: efficient constructions and applications," *SIAM J. on Computing*, vol. 22, no. 4, pp. 838–856, 1993.

[4] N. Alon, O. Goldreich, J. Håstad, and R. Peralta, "Simple constructions of almost k-wise independent random variables," *Random Structures & Algorithms*, vol. 3, no. 3, pp. 289–304, 1992.

[5] M. Luby, B. Veličković, and A. Wigderson, "Deterministic approximate counting of depth-2 circuits," in *2nd Israeli Symposium on Theoretical Computer Science (ISTCS)*, 1993, pp. 18–24.

[6] E. Viola, "Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates," *SIAM J. on Computing*, vol. 36, no. 5, pp. 1387–1403, 2007.

[7] S. Lovett, "Unconditional pseudorandom generators for low degree polynomials," *Theory of Computing*, vol. 5, no. 1, pp. 69–82, 2009. [Online]. Available: http://www.theoryofcomputing.org/articles/v005a003

[8] E. Viola, "The sum of d small-bias generators fools polynomials of degree d," *Computational Complexity*, vol. 18, no. 2, pp. 209–217, 2009.

[9] A. Razborov, "Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function," *Akademiya Nauk SSSR. Matematicheskie Zametki*, vol. 41, no. 4, pp. 598–607, 1987, english translation in Mathematical Notes of the Academy of Sci. of the USSR, 41(4):333-338, 1987.

[10] A. Bogdanov, "Pseudorandom generators for low degree polynomials," in *ACM Symp. on the Theory of Computing (STOC)*, 2005, pp. 21–30.

[11] C. Lu, "Hitting set generators for sparse polynomials over any finite fields," in *IEEE Conf. on Computational Complexity (CCC)*. IEEE Computer Society, 2012, pp. 280–286. [Online]. Available: https://doi.org/10.1109/CCC.2012.20

[12] G. Cohen and A. Ta-Shma, "Pseudorandom generators for low degree polynomials from algebraic geometry codes," *Electron. Colloquium Comput. Complex.*, p. 155, 2013. [Online]. Available: https://eccc.weizmann.ac.il/report/2013/155

[13] V. Guruswami and C. Xing, "Hitting sets for low-degree polynomials with optimal density," in *IEEE Conf. on Computational Complexity (CCC)*. IEEE Computer Society, 2014, pp. 161–168. [Online]. Available: https://doi.org/10.1109/CCC.2014.24

[14] A. R. Klivans and D. A. Spielman, "Randomness efficient identity testing of multivariate polynomials," in *ACM Symp. on the Theory of Computing (STOC)*, J. S. Vitter, P. G. Spirakis, and M. Yannakakis, Eds. ACM, 2001, pp. 216–223. [Online]. Available: https://doi.org/10.1145/380752.380801

[15] E. Kaltofen, "Effective noether irreducibility forms and applications," *J. Comput. Syst. Sci.*, vol. 50, no. 2, pp. 274–295, 1995. [Online]. Available: https://doi.org/10.1006/jcss.1995.1023

[16] G. Lecerf, "Improved dense multivariate polynomial factorization algorithms," *J. Symbolic Comput.*, vol. 42, no. 4, pp. 477–494, 2007. [Online]. Available: https://doi.org/10.1016/j.jsc.2007.01.003

[17] W. Schmidt, *Equations Over Finite Fields: An Elementary Approach*. Kendrick Press, 2004.

[18] A. Clark, *Elements of Abstract Algebra*, ser. Dover Books on Mathematics Series. Dover Publications, 1984. [Online]. Available: https://books.google.com/books?id=bj1kOY8gOfcC

[19] S. Najib, "Une généralisation de l'inégalité de Stein-Lorenzini," *Journal of Algebra*, vol. 292, no. 2, pp. 566–573, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021869305004060

[20] A. Bodin, P. Dèbes, and S. Najib, "Indecomposable polynomials and their spectrum," *Acta Arith.*, vol. 139, no. 1, pp. 79–100, 2009. [Online]. Available: https://doi.org/10.4064/aa139-1-7

[21] G. Chèze and S. Najib, "Indecomposability of polynomials via Jacobian matrix," *J. Algebra*, vol. 324, no. 1, pp. 1–11, 2010. [Online]. Available: https://doi.org/10.1016/j.jalgebra.2010.01.007