2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS) | 978-1-6654-5519-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/FOCS54457.2022.00028

Fast Multivariate Multipoint Evaluation Over All Finite Fields

Vishwas Bhargava
Dept. of Computer Science
Rutgers University
Piscataway, NJ, USA
vishwas1384@gmail.com

Sumanta Ghosh
Dept. of Computing and Mathematical Sciences
Caltech
Pasadena, California, USA
besusumanta@gmail.com

Zeyu Guo
Dept. of Computer Science
UT Austin
Austin, Texas, USA
zguotcs@gmail.com

Mrinal Kumar
Dept. of Computer Science & Engineering
IIT Bombay
Mumbai, India
mrinal@cse.iitb.ac.in

Chris Umans
Dept. of Computing and Mathematical Sciences
Caltech
Pasadena, California, USA
umans@cs.caltech.edu

***Abstract*—Multivariate multipoint evaluation is the problem of evaluating a multivariate polynomial, given as a coefficient vector, simultaneously at multiple evaluation points. In this work, we show that there exists a deterministic algorithm for multivariate multipoint evaluation over any finite field $\mathbb{F}$ that outputs the evaluations of an m-variate polynomial of degree less than d in each variable at N points in time $(d^m + N)^{1+o(1)} \cdot \mathrm{poly}(m, d, \log|\mathbb{F}|)$ for all $m \in \mathbb{N}$ and all sufficiently large $d \in \mathbb{N}$.**

A previous work of Kedlaya and Umans (FOCS 2008, SICOMP 2011) achieved the same time complexity when the number of variables m is at most $d^{o(1)}$ and had left the problem of removing this condition as an open problem. A recent work of Bhargava, Ghosh, Kumar and Mohapatra (STOC 2022) answered this question when the underlying field is not *too* large and has characteristic less than $d^{o(1)}$. In this work, we remove this constraint on the number of variables over all finite fields, thereby answering the question of Kedlaya and Umans over all finite fields.

Our algorithm relies on a non-trivial combination of ideas from three seemingly different previously known algorithms for multivariate multipoint evaluation, namely the algorithms of Kedlaya and Umans, that of Björklund, Kaski and Williams (IPEC 2017, Algorithmica 2019), and that of Bhargava, Ghosh, Kumar and Mohapatra, together with a result of Bombieri and Vinogradov from analytic number theory about the distribution of primes in an arithmetic progression.

We also present a second algorithm for multivariate multipoint evaluation that is completely elementary and in particular, avoids the use of the Bombieri–Vinogradov Theorem. However, it requires a mild assumption that the field size is bounded by an exponential-tower in d of bounded *height*.

Vishwas Bhargava's research is supported in part by the Simons Collaboration on Algorithms and Geometry and NSF grant CCF-1909683. Zeyu Guo is supported by a Simons Investigator Award (#409864, David Zuckerman). Mrinal Kumar is supported by a seed grant from IIT Bombay.

***Index Terms*—polynomial evaluation, multivariate multipoint evaluation, finite fields**

I. Introduction

We study the problem of multivariate multipoint evaluation: given an m-variate polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of degree less than d in each variable, and N points $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \in \mathbb{F}^m$, output $f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_N)$. Here $\mathbb{F}$ is the underlying field. The input polynomial $f(\mathbf{x})$ is given by its coefficient vector. Therefore, the overall input can be represented by a list of $(d^m + mN)$ elements in $\mathbb{F}$. A trivial algorithm for this problem is to evaluate $f(\mathbf{x})$ at each $\mathbf{a}_i$ separately. Since evaluating $f(\mathbf{x})$ at each $\mathbf{a}_i$ takes $d^m \cdot \mathrm{poly}(d, m)$ operations over $\mathbb{F}$, this algorithm needs $Nd^m \cdot \mathrm{poly}(d, m)$ $\mathbb{F}$-operations in total. For $N = \Theta(d^m)$, the time complexity of this algorithm is quadratic with respect to the input size. Therefore, a natural algorithmic question here is to seek faster algorithms for this problem. Of particular interest would be to have an algorithm for this problem whose time complexity is *nearly linear*, more specifically $(d^m + N)^{1+o(1)}$ (multiplied by lower-order $\mathrm{poly}(d, n, \log|\mathbb{F}|)$ terms), with respect to the input size.

In addition to its innate appeal as a fundamental and natural question in computational algebra, fast algorithms for multivariate multipoint evaluation are closely related to fast algorithms for other important algebraic problems such as polynomial factorization and modular composition. For a detailed discussion on these connections, we refer to the work of Kedlaya and Umans [1]. In a recent work, Bhargava, Ghosh, Kumar and Mohapatra [2] used the special structure of their algorithm for multivariate multipoint evaluation to show an upper bound on the rigidity of Vandermonde matrices and very efficient

978-1-6654-5519-0/22/$31.00 ©2022 IEEE
DOI 10.1109/FOCS54457.2022.00028

algebraic data structures for the polynomial evaluation problem over finite fields.

For the setting of univariate polynomials, Borodin and Moenck [3] showed that the multipoint evaluation can be solved in nearly linear time. Their algorithm is short, simple and elementary, and proceeds via an application of the Fast Fourier Transform (FFT). However, this approach does not seem to extend when the number of variables exceeds one; in fact, even when the number of variables is two. However, for the multivariate case, when the input points form a product set, one can naturally extend the ideas in Borodin and Moenck [3] to get a nearly linear time algorithm for this problem. But, when the input points are arbitrary, getting a sub-quadratic algorithm for multipoint evaluation seems to be significantly more difficult. In fact, about three decades after Borodin and Moenck's work, Nüsken and Ziegler [4] proved that multipoint evaluation can be solved in most $O(d^{\omega_2/2+1})$ operations for $m = 2$ and $N = d^2$, where ω_2 is the exponent for multiplying a $d \times d$ and a $d \times d^2$ matrix. The work [4] extends to general m and gives an algorithm for multipoint evaluation that performs $O(d^{\omega_2/2\cdot(m-1)+1})$ field operations.

Two significant milestones in this line of work are the results of Umans [5] and Kedlaya and Umans [1]. Umans [5] gave a nearly linear time (that is, $(d^m + N)^{1+o(1)} \cdot \mathrm{poly}(m, d, \log|\mathbb{F}|)$-time) *algebraic* algorithm for this problem over finite fields, provided that the characteristic of the field and the number of variables are at most $d^{o(1)}$. Later, Kedlaya and Umans [1] gave a nearly linear time *non-algebraic* algorithm for *all* finite fields, but they also need $m = d^{o(1)}$. In a recent work, Bhargava, Ghosh, Kumar and Mohapatra [2] improve the result of Umans [5] by removing the restriction on m over finite fields whose characteristics are small and sizes are not too large. More specifically, they gave a nearly linear time algebraic algorithm for multivariate multipoint evaluation, provided that the characteristic of the field is $d^{o(1)}$ and the size of the field is at most $(\exp(\exp(\cdots(\exp(d)))))$, where the height of this tower of exponentials is fixed. Another closely related result is a recent work of Björklund, Kaski and Williams [6] who (among other results) gave an algorithm for multivariate multipoint evaluation, but their time complexity depends polynomially on the field size (and not polynomially on the logarithm of the field size), and instead of d^m, their time complexity is nearly linear in D^m where D is the total degree of the polynomial. Nevertheless, their results play a crucial role in proving the results of this paper and we will discuss them in more detail in Section II-B.

Thus, from the context of previous work, a very natural and interesting open question is to design an algorithm for multivariate multipoint evaluation that runs in nearly linear time and works for all finite fields and all ranges of the number of variables. Indeed, Kedlaya and Umans [1] mention this as an open problem.

In this work, we answer this question by giving two different algorithms for multivariate multipoint evaluation over finite fields. While our first algorithm works over all finite fields, the second algorithm still requires that the field size is not too large in terms of d. We now state our results and discuss the pros and the cons of the two algorithms and compare them to the algorithms known in prior work. Both our algorithms happen to be non-algebraic, i.e. we need more than just arithmetic operations over the underlying field.

A. Our Results

We state our main result as follows.

Theorem I.1. *There is a deterministic algorithm that given the coefficient vector of an m-variate polynomial $f(\mathbf{x})$ of degree less than d in each variable over a finite field $\mathbb{F}$ and N points $\mathbf{a}_1, \ldots, \mathbf{a}_N \in \mathbb{F}^m$, outputs $f(\mathbf{a}_1), \ldots, f(\mathbf{a}_N)$ in time $(d^m + N)^{1+o(1)}\,\mathrm{poly}(m, d, \log|\mathbb{F}|)$ for all $m \in \mathbb{N}$ and all sufficiently large $d \in \mathbb{N}$.*

Remark. *Throughout this paper, when we say d is* sufficiently large*, it means $d = \omega(1)$.* ◊

The proof of the above theorem crucially relies on a deep result from analytic number theory, known as the Bombieri–Vinogradov Theorem [7], [8], related to the distribution of primes in arithmetic progressions.

We also give a different algorithm for multivariate multipoint evaluation that avoids the Bombieri–Vinogradov Theorem and is completely elementary, but it requires the finite field to be not too large: at most $(\exp(\exp(\cdots(\exp(d)))))$, where the height of this tower of exponentials is fixed. In other words, it removes the restriction on the characteristic of the field in the work of Bhargava *et.al.* [2], but via a non-algebraic algorithm.

We remark that neither of our algorithms is algebraic, and in particular, we crucially rely on working with the bit representation of the inputs. To obtain an algebraic algorithm for multivariate multipoint evaluation, for large m, and over all finite fields is a fundamental algebraic problem that continues to remain open.

II. An Overview of the Proofs

At a high level, our algorithms rely on ideas from three of the recent prior works on multivariate multipoint evaluation, namely that of Kedlaya and Umans [1], that of Björklund, Kaski and Williams [6], and a recent work of Bhargava, Ghosh, Kumar and Mohapatra [2]. We start by giving a brief outline of these.

We start with some necessary notation. Let $\mathbb{F}$ be a finite field and let $f \in \mathbb{F}[\mathbf{x}]$ be an m-variate polynomial of degree less than d in each variable, and let $\{\mathbf{a}_i : i \in$

$[N]\}$ be a set of N inputs in $\mathbb{F}^m$. Our goal is to evaluate f on each $\mathbf{a}_i$. For simplicity, we focus on the case when the underlying field $\mathbb{F}$ is a prime field, i.e. $\mathbb{F} = \mathbb{F}_p$ for some prime p. The case of extension fields is handled in a very similar manner, with a few technicalities.

A starting observation is that multivariate multipoint evaluation has a nearly linear time algorithm (over all fields) when the set of evaluation points forms a product set (see Lemma III.7), and more generally when the set of evaluation points is *close* to a product set. At a high level, each of the algorithms in [1], [2], [6] proceeds via a very efficient reduction from multivariate multipoint evaluation over an arbitrary set of points to multivariate multipoint evaluation over product sets. However, despite this common high level structure, the details of the reductions involved are fairly different in each of the three algorithms, thereby giving these algorithms their features, both desirable and undesirable. We now elaborate a bit more on these reductions.

A. The Algorithm of Kedlaya and Umans

To solve the problem efficiently over a finite field $\mathbb{F}$, Kedlaya and Umans [1] first reduce an instance of the multivariate multipoint evaluation problem over $\mathbb{F}$ to an instance of the same problem over a ring of the form $\mathbb{Z}/r\mathbb{Z}$. Then they use their efficient algorithm for multivariate multipoint evaluation problem over $\mathbb{Z}/r\mathbb{Z}$ to solve it. Finally, from the evaluations over $\mathbb{Z}/r\mathbb{Z}$, they recover the original evaluations over $\mathbb{F}$. In the following, we describe only their algorithm over $\mathbb{Z}/r\mathbb{Z}$.

The algorithm over $\mathbb{Z}/r\mathbb{Z}$: In their algorithm, Kedlaya and Umans [1] start by *lifting* their problem instance over $\mathbb{Z}/r\mathbb{Z}$ to an instance over integers. They do this by just viewing $\mathbb{Z}/r\mathbb{Z}$ as the set of integers $\{0, 1, \ldots, r-1\}$ and this naturally maps a polynomial $f(\mathbf{x})$ over $\mathbb{Z}/r\mathbb{Z}$ to a polynomial $F(\mathbf{x})$ with coefficients in $\mathbb{Z}$. Similarly, this also gives a natural map from an input point $\mathbf{a} \in (\mathbb{Z}/r\mathbb{Z})^m$ to a point $\tilde{\mathbf{a}} \in \mathbb{Z}^m$. Clearly, for every $\mathbf{a} \in (\mathbb{Z}/r\mathbb{Z})^m$ and polynomial f, $f(\mathbf{a}) = F(\tilde{\mathbf{a}}) \mod r$. Thus, it suffices to solve this lifted instance over integers. Yet another property of this lifted instance is that the integer $F(\tilde{\mathbf{a}})$ is a non-negative integer of magnitude less than $M = d^m(r-1)^{dm}$ since each coefficient of F and each coordinate of $\tilde{\mathbf{a}}$ are in $\{0, 1, \ldots, r-1\}$, and the total degree of F is less than or equal to $(d-1)m$. Thus, to compute $F(\tilde{\mathbf{a}})$, it suffices to compute $F(\tilde{\mathbf{a}}) \mod M$. Kedlaya and Umans now proceed by finding distinct small primes $p_1, p_2, \ldots, p_k$ such that $\prod_{i\in[k]} p_i > M$, evaluating the polynomial $f_j(\mathbf{x}) = F \mod p_j$ at the point $\mathbf{b}_j = \tilde{\mathbf{a}} \mod p_j$ and [1] then combining the values $f_1(\mathbf{b}_1), f_2(\mathbf{b}_2), \ldots, f_k(\mathbf{b}_k)$ using the Chinese Remainder Theorem. The correctness follows from the observation that for every $j \in [k]$, $f_j(\mathbf{b}_j) = F(\tilde{\mathbf{a}}) \mod p_j$. The advantage of this *multi-modular* reduction is that if the primes p_j are very small (for instance, if all these primes are close to d), then the set of evaluation points of interest, that were initially *scattered* sparsely in $\mathbb{F}_p^n$ are now mapped to points that are packed densely in the space $\mathbb{F}_{p_j}^m$, which is a product set. Thus, we can use the simple multidimensional FFT to evaluate f_j on all of $\mathbb{F}_{p_j}^m$ for every j, and then combine the outcome using the Chinese Remainder Theorem. For $m < d^{o(1)}$, this indeed gives a nearly linear time algorithm for multivariate multipoint evaluation. This constraint on the number of variables m is due to a term of the form $(dm)^m$ in the final running time of the algorithm which is nearly linear in the input size only if m is small. This $(dm)^m$ essentially appears because the product of primes $p_1, p_2, \ldots, p_k$ chosen in this reduction must exceed M, and hence, the largest of these primes p_k must be $\Omega(\log M) = \Omega(dm \log r)$, and thus evaluating a polynomial f_k on all of $\mathbb{F}_{p_k}^m$ requires at least $p_k^m = \Omega(d^m m^m)$ time. Recursive application of this process leads to smaller primes but the improved dependence is on the $\log r$ factor and this $(dm)^m$ factor continues to persist in the eventual bound on the running time. Thus, one approach towards a faster algorithm for multipoint evaluation over $\mathbb{Z}/r\mathbb{Z}$ would be to replace this step of evaluating f_j on all of $\mathbb{F}_{p_j}^m$ in [1] with a faster subroutine, in particular, something that runs in nearly linear time in the input size even for large m.

Our first algorithm in this paper does precisely this. In order to obtain this gain, it crucially relies on ideas in an algorithm of Björklund, Kaski and Williams [6] which we discuss in Section II-B and a very careful choice of primes to do Chinese Remaindering with, in the multimodular reduction discussed above. Together, these steps lead to an improvement in running time and give us an algorithm that runs in nearly linear time even when the number of variables is large.

For our second algorithm, we introduce a slightly different modification in the framework of Kedlaya and Umans. Instead of working modulo small primes as in [1], which as discussed above, forces us to pick primes as large as dm, we work modulo powers of distinct primes in the multimodular reduction step. Thus, it seems conceivable that we can now work with much smaller primes than in the original algorithm, since instead of having the condition that the product of these primes is larger than M as in [1], we now need that the product of powers of these primes is larger than M. However, we still need efficient algorithms for multivariate multipoint evaluation over rings of the form $\mathbb{Z}/p^k\mathbb{Z}$ for small primes p and large $k \in \mathbb{N}$. To handle this subproblem, we extend the derivative-based techniques used in the algorithm of

[1] In other words, f_j is obtained from F by reducing each of its coefficients modulo p_j and $\mathbf{b}_j$ is obtained by reducing each of the coordinates of $\tilde{\mathbf{a}}$ modulo p_j.

Bhargava et al. [2] for fields of small characteristic so that they work over rings of the form $\mathbb{Z}/p^k\mathbb{Z}$ for small primes p and large $k \in \mathbb{N}$.

The advantage of this strategy over our first algorithm is that this gives us a completely elementary algorithm, and the disadvantage is that for this algorithm to run in nearly linear time, as desirable, the underlying ring $\mathbb{Z}/r\mathbb{Z}$ needs to be somewhat small. This issue also affects the original algorithm of Bhargava et al. [2] and seems somewhat inherent to this style of an argument.

B. The Algorithm of Björklund, Kaski and Williams

In a nutshell, the algorithm of Björklund et al. [6] proceeds via constructing a set $K \subseteq \mathbb{F}_p^m$ such that

- The size of K is not too large and K is (*close* to) a product set.
- For every $\mathbf{a} \in \mathbb{F}_p^m$, there is a curve $C_{\mathbf{a}}$ of low degree (in fact, a low degree univariate polynomial map) that passes through the point $\mathbf{a}$ and intersects the set K on at least p points [2].

These sets K can be thought of as a natural higher degree analog of Kakeya sets over finite fields from discrete geometry. Indeed, Björklund et al. refer to the set K as high degree Kakeya sets, where the degree of the set is defined to be the maximum over the degrees of the curves $C_{\mathbf{a}}$ over all $\mathbf{a} \in \mathbb{F}_p^m$.

Given such a Kakeya set K, Björklund et al.proceed by evaluating f on all points in K fast, using the multidimensional FFT algorithm. This is the preprocessing phase of the algorithm. Then, for an arbitrary point $\mathbf{a} \in \mathbb{F}_p^m$, they compute $f(\mathbf{a})$ by considering the univariate polynomial $R(y)$ obtained by taking the restriction f on the curve $C_{\mathbf{a}}$. From the properties of the set K, we know the curve $C_{\mathbf{a}}$ intersects the set K on at least p points. Thus, if the degree of $R \leq \deg(f) \cdot \deg(C_{\mathbf{a}})$ is less than p, then we can recover the polynomial R from the evaluations of f on K computed in the preprocessing step and using univariate polynomial interpolation. The quantitative bounds for this approach are therefore crucially determined by the size of the set K and the degree of the curve $C_{\mathbf{a}}$.

Björklund et al. showed that for every $u \in \mathbb{N}$ such that $u+1$ divides $p-1$, there is a Kakeya set K of degree u of size at most $((p-1)/(u+1)+1)^{m+1}$. This divisibility condition ensures the existence of a multiplicative subgroup of $\mathbb{F}_p^*$ of size $(p-1)/(u+1)$ and set K is based on this subgroup. Thus, if $\tilde{d}$ denotes $(p-1)/(u+1)$, then we can evaluate the polynomial f on the K in time $\tilde{d}^m$, which is nearly linear in the input size if $\tilde{d} \leq d^{1+o(1)}$. However, note that in this case, u is around $p/\tilde{d}$, and hence, the degree of the restriction R of f on a curve of degree u has total degree $udm = pm \cdot \frac{d}{\tilde{d}}$. Thus, if $pm \cdot \frac{d}{\tilde{d}} > p$, we cannot hope to recover R from its evaluations on just p points. To address this issue, we combine the above strategy in [6] with an idea in [2] where instead of evaluating just f on K, we evaluate all its (Hasse) derivatives of order at most $m \cdot \frac{d}{\tilde{d}}$ on K in the preprocessing phase. There are at most $\binom{m+m\cdot\frac{d}{\tilde{d}}}{m}$ such derivatives and this leads to an additional multiplicative factor of $\binom{m+m\cdot\frac{d}{\tilde{d}}}{m}$ in the final running time, but if $\tilde{d}$ is not too small compared to d, for instance, $\tilde{d} = \Theta(d)$, this binomial coefficient is at most $\exp(O(m))$ which is $d^{o(m)}$ for all growing d. Thus, with this stronger guarantee in the preprocessing step, we are guaranteed to have higher multiplicity information available to us in the local computation step. So, we can now hope to uniquely recover a univariate polynomial of degree higher than p from this information (via Hermite interpolation). However, since the degree of the univariates we have here is larger than p, this Hermite interpolation step runs in time polynomially bounded in the underlying field size p and not just polynomially bounded in $\log p$ as would have been desirable.

To summarise, if there exists an $u \in \mathbb{N}$ such that $(p-1)/(u+1) = \tilde{d}$, where $\tilde{d}$ is close to d, e.g. $\tilde{d} = \Theta(d)$, then we have an algorithm for evaluating m-variate polynomials of degree less than d in each variable on any N points in $\mathbb{F}_p^m$ in time $\mathrm{poly}(p,d,m) \cdot (d^m + N)^{1+o(1)}$. Thus, this is nearly linear time, when the field size p is not too large.

Having discussed these prior results, we are now ready to give an outline of our algorithms. We start with the first algorithm.

C. The First Algorithm

As discussed earlier in this section, the plan for our algorithm is to somehow replace the multidimensional FFT step in the algorithm of Kedlaya and Umans [1] (over rings of the form $\mathbb{Z}/r\mathbb{Z}$) with the Kakeya-set-based algorithm above over a field $\mathbb{F}_{p_j}$. However, in order to effectively use the Kakeya-set-based algorithm outlined in the previous section to obtain nearly linear time algorithms for multipoint evaluation, we need to ensure two properties.

- The underlying field size p_j is small. For instance, we would need $p_j = (d^m + N)^{o(1)}$ for a nearly linear time algorithm.
- There exists $u \in \mathbb{N}$ such that $u+1$ divides $p_j - 1$ and $(p_j - 1)/(u+1)$ is an integer close to d.

In fact, instead of the second condition here, it suffices if there is a small $t \in \mathbb{N}$ such that there exists a $u \in \mathbb{N}$ such that $u+1$ divides $p_j^t - 1$ and $(p_j^t - 1)/(r+1) = d^{1+o(1)}$, since we can always view the problem over $\mathbb{F}_p$ as a problem over an extension of $\mathbb{F}_p$. However, we need the

[2]This notion of a curve passing through a point here is slightly different to that in other related works like [2]. However, for the sake of simplicity, we gloss over this technical detail right now.

degree of the extension to be small in order to get useful final quantitative bounds.

The first condition about the primes p_j being small does not appear too difficult to ensure in isolation and in particular, is also true for the algorithm of Kedlaya and Umans. However, the second divisibility condition seems trickier to guarantee even with the flexibility of working over low degree extensions of $\mathbb{F}_{p_j}$ as outlined earlier in this section. In particular, it is not clear to us if for every pair d, p_j, there always exists small t such that $p_j^u - 1$ has a divisor in the vicinity of d.

Getting around these technical difficulties is the main technical content of our algorithm. In a nutshell, we proceed by following the multimodular reduction step of Kedlaya and Umans, but via a careful choice of primes $p_1, p_2, \ldots, p_k$ (as opposed to picking a sufficiently large number of small primes as in [1]). This careful choice preserves the fact that these primes are all small (at most $\mathrm{poly}(d, m, \log p)$) and additionally guarantees that the divisibility condition needed to invoke the Kakeya-set-based framework of [6]. More formally, we choose $p_1, p_2, \ldots, p_k$ so that they are all at most $\mathrm{poly}(d, m, \log p)$, their product exceeds $M = d^m(p-1)^{dm}$ and there exists a $\tilde{d} \in [0.8d, d]$ such that for every $j \in [k]$, $\tilde{d}$ divides $p_j - 1$. Thus, we can use the Kakeya-set-based framework outlined in Section II-B, with the parameter u_j to be set equal to $(p_j - 1)/\tilde{d} - 1$. This satisfies both the conditions highlighted earlier, and the final running time of this algorithm does indeed turn out to be nearly linear in the input size. The details can be found in Section V. Once we have this algorithm for multipoint evaluation over the rings of the form $\mathbb{Z}/r\mathbb{Z}$, we use exactly the same strategy as Kedlaya and Umans did to solve this problem over all finite fields. For details, see the full version of this paper.

Thus, if we can find distinct primes $p_1, p_2, \ldots, p_k$ with the properties outlined above, we would be done. However, it is not immediately clear how to do find such a set of numbers efficiently, or whether such a collection of primes and the parameter $\tilde{d}$ should even exist. The appearance of the parameter $\tilde{d} = \Theta(d)$ is also slightly mysterious. For instance, it would be aesthetically nice if $\tilde{d}$ would have been equal to d. Perhaps surprisingly, we do not know how to even show the existence of primes $p_1, p_2, \ldots, p_k$ satisfying the desired properties with $\tilde{d} = d$! We now outline our approach to finding such primes and the parameter $\tilde{d}$. However, for a start, let us attempt to do this with $\tilde{d} = d$ and try to understand the issues that arise.

The intuition on showing the existence of such primes follows from the observation that if d divides $p_j - 1$ for each $j \in [k]$ then, each of the primes $p_1, p_2, \ldots, p_k$ lies in the arithmetic progression (AP) $A_d = (1, 1 + d, 1 + 2d, \ldots)$. It follows from a classical theorem of Dirichlet (see Chapter 5 in [9] for more details) that this arithmetic progression A_d indeed contains an infinite number of primes for every $d \in \mathbb{N}$. Thus, if we take k to be sufficiently large, then there exist primes $p_1, p_2, \ldots, p_k$ each congruent to 1 modulo d such that their product is greater than $M = d^m(p-1)^{dm}$. However, it is not enough for our application. We also need to show that these primes are not too large, e.g. each $p_i \leq \mathrm{poly}(d, m, \log p)$, and that they can be found efficiently. For this, it would be sufficient to show that not only does the arithmetic progression A_d contains an infinite number of primes, but the set of primes in A_d is also a sufficiently dense subset of A_d. The prime number theorem gives such a statement for the progression A_1, i.e. for the set of natural numbers and here, a similar statement for arbitrary arithmetic progressions is needed. An unconditional bound on the density of primes in an arithmetic progression A_d is given by the well-known Siegel-Walfisz theorem [10], [11] which implies a lower bound on the number of primes less than x in the AP A_d for all $x \geq 0$ with $x > 2^{d^\varepsilon}$ for any constant ε. However, this estimate does not appear to be sufficient for us, since for the algorithm, we need the magnitude of these primes to be at most $\mathrm{poly}(d, m, \log p)$ and not exponentially growing in d, and it is not clear if such a guarantee can be obtained directly from this theorem. An improved lower bound on the density of primes in arithmetic progressions is known under the Generalized Riemann Hypothesis, and this would have been sufficient for our applications, except for the fact that the result would be conditional. For the unconditional result in this paper, we rely on the following theorem of Bombieri and Vinogradov, which gives an improved lower bound on the density of primes in an AP on average. For $x > 0, t \in \mathbb{N}$, let $\pi(x, t)$ be the number of primes less than x in the AP starting at 1 and with common difference t, $\pi(x)$ denote the number of primes less than x, and $\phi : \mathbb{N} \to \mathbb{N}$ be the Euler Totient function. Various versions of this theorem can be found in literature, for instance, [7], [8], Theorem 18.1 in [9]. Here we rely on the bound in equation 1.1. in [12].

Theorem II.1 (Bombieri–Vinogradov)**.** *For any fixed $a > 0$, there exist constants $c = c(a)$ and $b = b(a)$ such that for all sufficiently large $x > 0$, $\sum_{t \leq d} \left| \pi(x, t) - \frac{\pi(x)}{\phi(t)} \right| \leq cx(\log x)^{-a}$, where $d \leq x^{1/2}(\log x)^{-b}$.*

Thus, if x is sufficiently large compared to d, e.g. $x = d^3$, this theorem can be viewed as saying that on average (over $t \in \mathbb{N}, t \leq d$), an AP with common difference t contains at least $\frac{\pi(x)}{\phi(t)} - cxd^{-1}(\log x)^{-a}$ primes less than x. Clearly, $\phi(t) \leq t \leq d$ and $\pi(x) = \Theta(x/\log x)$ by the prime number theorem. Thus, if we

take $a > 1$, the number of primes less than x is at least $\Omega(\pi(x)/d)$. For our final argument, we combine this average-case statement about the density of primes in an AP with a standard application of Markov's inequality to deduce that there exists a $\tilde{d} \in [0.8d, d]$ such that the AP with common difference $\tilde{d}$ has at least $\Omega(\pi(x)/\tilde{d})$ many primes less than x. By choosing x to be a sufficiently large polynomial in $d, m, \log p$, we get precisely what we want: sufficiently many primes $p_1, p_2, \ldots, p_k$, each at most $\mathrm{poly}(d, m, \log p)$ in absolute value such that their product exceeds M and they are all congruent to 1 modulo $\tilde{d}$, for $\tilde{d} = \Theta(d)$. This application of Markov's inequality is precisely why we have to settle for working with the quantity $\tilde{d}$ and not d itself.

D. The Second Algorithm

In this section, we give a brief overview of our second algorithm. It implies that Theorem I.1 holds as long as the size of the finite field is bounded by $(\exp(\exp(\cdots(\exp(d)))))$, where the height of this tower of exponentials is fixed via an elementary algorithm. In particular, this algorithm does not rely on the Bombieri–Vinogradov theorem necessary for the first algorithm.

For simplicity, we only explain our algorithm over rings of the form $\mathbb{Z}/r\mathbb{Z}$, or $\mathbb{Z}/r^s\mathbb{Z}$ for some $s \leq m$. This covers the case of prime finite fields $\mathbb{F}_p$ by choosing $r = p$ and $s = 1$. The general case of arbitrary finite fields (and certain extension rings of $\mathbb{Z}/r\mathbb{Z}$) is addressed in the full version of this paper.

The algorithm over $\mathbb{Z}/r\mathbb{Z}$: Recall that Kedlaya and Umans [1] use multimodular reduction together with the Chinese Remainder Theorem to reduce the multivariate multipoint evaluation problem over $\mathbb{Z}/r\mathbb{Z}$ to that over $\mathbb{F}_{p_j}$ for a collection of small primes p_j. As discussed in Section II-A, for the Chinese Remainder Theorem, the primes p_j need to be chosen such that $\prod_{i\in[k]} p_i > M := d^m(r-1)^{dm}$. The problem here is that, as the primes p_j are distinct, the largest prime would have order $O(\log M) = O(dm \log r)$. The $\log r$ factor can be further reduced by repeating the multimodular reduction. However, the dm factor persists. As a consequence, the time complexity of the Kedlaya–Umans algorithm has a factor $(dm)^m$, which is nearly linear in d^m only when $m = d^{o(1)}$.

In our algorithm, we introduce the new idea of using the *prime powers* p_j^m as the moduli for Chinese remaindering instead of the primes p_j. That is, we compute the evaluations over the rings $\mathbb{Z}/p_j^m\mathbb{Z}$ and then combine them via Chinese Remainder Theorem to obtain the evaluations over the integers. Assuming this can be done, then we only need to choose the primes p_j such that $\prod_{i\in[k]} p_i^m > M$. So the largest prime may have order $O(\frac{1}{m}\log M) = O(d \log r)$, which is independent of m.

Now, to make this idea work, we need a fast algorithm for multivariate multipoint evaluation over $\mathbb{Z}/p_j^m\mathbb{Z}$, for small primes p_j. In particular, if we have an algorithm over $\mathbb{Z}/p_j^m\mathbb{Z}$ that runs in time $(p_j^m + N)^{1+o(1)}$, then, overall, we have an algorithm that runs in time $(d^m(\log r)^m + N)^{1+o(1)}$. Note that this has already enabled us to get rid of the m^m factor in the running time as in [1]. So, up to the factor of $(\log r)^m$ in the running time, we seem to have made some progress and we soon elaborate further on how to reduce this $(\log r)^m$ factor further.

But first, we note that naively evaluating the polynomial at all points in $(\mathbb{Z}/p_j^m\mathbb{Z})^m$ would be extremely inefficient, as the size of $(\mathbb{Z}/p_j^m\mathbb{Z})^m$ is exponential in m^2. So, we need a significantly faster algorithm for multivariate multipoint evaluation over $\mathbb{Z}/p_j^m\mathbb{Z}$ to have any hope of making this strategy work.

In their algorithm, Kedlaya and Umans [1] deal with the $(\log r)^m$ factor by recursively applying the multimodular reduction a few times. So, to reduce the $(\log r)^m$ in the discussion above, we could also try to do something similar. We already see that one application of the reduction reduces the modulus r to p_j^m for a collection of primes p_j, where $\prod_{i\in[k]} p_i > d(r-1)^d$. Fix a prime p_j and suppose we want to apply the multimodular reduction again. We may lift the instance over $\mathbb{Z}/p_j^m\mathbb{Z}$ to an instance over the integers, and then reduce it modulo $p_i'^m$ for a collection of primes p_i'. The problem here is that, if we simply lift the evaluation points from $(\mathbb{Z}/p_j^m\mathbb{Z})^m$ to $\{0, 1, \ldots, p_j^m - 1\}^m$, we would have an upper bound $M' = d^m(p_j^m - 1)^{dm}$ for the evaluations over the integers, which is too large for us. The primes p_i' would have to satisfy $\prod_i p_i' > M'^{1/m} = d(p_j^m - 1)^d$, and then the order of the largest prime must depend (at least polynomially) on m.

We address the above two challenges, namely that of obtaining a fast multipoint evaluation algorithm over $\mathbb{Z}/p_j^m\mathbb{Z}$ that does not require evaluating on all of $\mathbb{Z}/p_j^m\mathbb{Z}^m$ and that of reducing the factor $(\log r)^m$ using the following observation: over $\mathbb{Z}/r^s\mathbb{Z}$, the evaluation of an m-variate polynomial $f(\mathbf{x})$ at a point $\mathbf{a} \in (\mathbb{Z}/r^s\mathbb{Z})^m$ can be derived from the evaluations of the Hasse derivatives of $f(\mathbf{x})$ of sufficiently high order at another point $\mathbf{b} \in (\mathbb{Z}/r^s\mathbb{Z})^m$, provided that the coordinates of $\mathbf{a} - \mathbf{b}$ are all multiples of r. Intuitively, this means if $\mathbf{a}$ and $\mathbf{b}$ are "close enough," then we can learn the evaluation of $f(\mathbf{x})$ at $\mathbf{a}$ from the evaluations at $\mathbf{b}$ of all the Hasse derivatives of f of sufficiently high order.

Formally, for all $\mathbf{e} \in \mathbb{N}^m$, let $\overline{\partial}_{\mathbf{e}}(f) \in (\mathbb{Z}/r^s\mathbb{Z})[\mathbf{x}]$ be the Hasse derivative of $f(\mathbf{x})$ with respect to $\mathbf{x}^{\mathbf{e}}$. For $\mathbf{a}, \mathbf{b} \in (\mathbb{Z}/r^s\mathbb{Z})^m$, we get from Taylor's expansion of $f(\mathbf{x})$ at $\mathbf{b}$ that $f(\mathbf{a}) = \sum_{\mathbf{e}\in\mathbb{N}^m} \overline{\partial}_{\mathbf{e}}(f)(\mathbf{b})(\mathbf{a}-\mathbf{b})^{\mathbf{e}}$. Suppose the coordinates of $\mathbf{a} - \mathbf{b}$ are all multiples of r. In this case, observe that $(\mathbf{a}-\mathbf{b})^{\mathbf{e}} = 0$ in $\mathbb{Z}/r^s\mathbb{Z}$ for all

$\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 \geq s$. Hence,

$$f(\mathbf{a}) = \sum_{\mathbf{e}\in\mathbb{N}^m : |\mathbf{e}|_1 < s} \overline{\partial}_{\mathbf{e}}(f)(\mathbf{b})(\mathbf{a}-\mathbf{b})^{\mathbf{e}}. \tag{II.2}$$

So we may compute $f(\mathbf{a})$ from the evaluations of Hasse derivatives $(\overline{\partial}_{\mathbf{e}}(f)(\mathbf{b}))_{\mathbf{e}\in\mathbb{N}^m : |\mathbf{e}|_1 < s}$.

We apply this idea to resolve the above two issues. First, in a base case of the recursive algorithm, instead of evaluating $f(\mathbf{x})$ at all points in $(\mathbb{Z}/p_j^m\mathbb{Z})^m$, we evaluate the Hasse derivatives $\overline{\partial}_{\mathbf{e}}(f)$ at the points in S^m using a fast evaluation algorithm for product sets, where S is the subset of $\mathbb{Z}/p_j^m\mathbb{Z}$ represented by $\{0, 1, \ldots, p_j - 1\}$. Note that for any $\mathbf{a} \in (\mathbb{Z}/p_j^m\mathbb{Z})^m$, we may find $\mathbf{b} \in S^m$ such that the coordinates of $\mathbf{a} - \mathbf{b}$ are multiples of p_j. Then $f(\mathbf{a})$ can be computed from $\overline{\partial}_{\mathbf{e}}(f)(\mathbf{b})$ using (II.2). The advantage of this is that the size of S^m is only p_j^m, which is much smaller than the size $p_j^{m^2}$ of the whole set $(\mathbb{Z}/p_j^m\mathbb{Z})^m$.

Similarly, when applying the multimodular reduction over a ring $\mathbb{Z}/p_j^m\mathbb{Z}$, the idea above allows us to use a small yet non-exact lift of each evaluation point $\mathbf{a}_i$. Namely, suppose $\tilde{\mathbf{a}}_i \in \mathbb{Z}^m$ is the unique lift of $\mathbf{a}_i \in (\mathbb{Z}/p_j^m\mathbb{Z})^m$ with coordinates in $\{0, 1, \ldots, p_j^m - 1\}$. We compute $\tilde{\mathbf{a}}'_i \in \{0, 1, \ldots, p_j - 1\}^m$ whose coordinates are obtained by reducing the corresponding coordinates of $\tilde{\mathbf{a}}_i$ modulo p_j. Then $\tilde{\mathbf{a}}'_i$ is a lift of some $\mathbf{a}'_i \in (\mathbb{Z}/p_j^m\mathbb{Z})^m$ such that the coordinates of $\mathbf{a}_i - \mathbf{a}'_i$ are all multiples of p_j. We compute the evaluation $\overline{\partial}_{\mathbf{e}}(f)(\mathbf{a}'_i)$ at the point $\mathbf{a}'_i$ (instead of $\mathbf{a}$), and then $f(\mathbf{a}_i)$ can be computed from $\overline{\partial}_{\mathbf{e}}(f)(\mathbf{a}'_i)$ using (II.2). The advantage of evaluating at $\mathbf{a}'_i$ instead of $\mathbf{a}_i$ is that the coordinates of its lift $\tilde{\mathbf{a}}'_i$ are bounded by $p_j - 1$ instead of $p_j^m - 1$. This translates into a better bound for the primes that we choose in multimodular reduction, thereby resolving the second issue.

Finally, at each level of the recursive algorithm, we need to evaluate not only $f(\mathbf{x})$, but also the Hasse derivatives $\overline{\partial}_{\mathbf{e}}(f)$ of order less than m. In addition, we need to solve the subproblem for each prime p_j. This means the number of subproblems blows up by a factor of $2^{O(m)} \cdot O(d \log r)$ each time. However, as we assume the original r (= the field size when r is prime) is reasonably bounded in terms of d, it takes only a constant number of rounds to reduce r to $d^{1+o(1)}$. So the total blow-up is reasonably controlled, and we obtain a nearly linear time algorithm when d is sufficiently large. For details, see Section VI.

a) Comparison with the first algorithm: Compared to our first algorithm, which uses the ideas of generalized Kakeya sets and the Bombieri–Vinogradov theorem, our second algorithm uses a different idea, namely the Chinese Remainder Theorem with prime powers as the moduli. At a high level, this may be seen as an analogue of the "method of multiplicities" applied to the ring $\mathbb{Z}$ and polynomial rings over $\mathbb{Z}$. To see this, note that for a univariate polynomial $f(x)$ over a field, knowing the evaluations of all (Hasse) derivatives $f^{(i)}(x)$ of order $< s$ at a point a is equivalent to knowing the remainder of f modulo the power $(x-a)^s$. So from an ideal-theoretic point of view, the idea of applying the Chinese Remainder Theorem to learn an integer from its remainders modulo prime powers is analogous to applying Hermite interpolation to learn a univariate polynomial from the evaluations of its Hasse derivatives, the latter playing a crucial role in [2].

III. Preliminaries

Define $\mathbb{N} = \{0, 1, \ldots\}$, $\mathbb{N}^+ = \{1, 2, \ldots\}$, $[n] = \{1, 2, \ldots, n\}$, and $[\![n]\!] = \{0, 1, \ldots, n-1\}$. The cardinality of a set S is denoted by $|S|$.

All rings in this paper are commutative rings with unity. For univariate polynomials $f(x), g(x)$ over a ring R such that $g(x)$ is monic of positive degree, there exist unique $h(x), r(x) \in R[x]$ such that $f(x) = g(x)h(x) + r(x)$ and $\deg(r) < \deg(g)$ [13, Theorem 1.1]. Define $f(x) \bmod g(x) := r(x)$, which can be computed using polynomially many R-operations via long division.

By $\mathbf{x}$ and $\mathbf{z}$, we denote the variable tuples $(x_1, \ldots, x_m)$ and $(z_1, \ldots, z_m)$, respectively. For any $\mathbf{e} = (e_1, \ldots, e_m) \in \mathbb{N}^m$, $\mathbf{x}^{\mathbf{e}}$ denotes the monomial $\prod_{i=1}^m x_i^{e_i}$. By $|\mathbf{e}|_1$, we denote the sum $e_1 + \cdots + e_m$.

For every positive integer k, $k!$ denotes $\prod_{i=1}^k i$. For $k = 0$, $k!$ is defined as 1. For two non-negative integers i and k with $k \geq i$, $\binom{k}{i}$ denotes $\frac{k!}{i!(k-i)!}$. For $k < i$, $\binom{k}{i} = 0$. For $\mathbf{a} = (a_1, \ldots, a_m), \mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{N}^m$, $\binom{\mathbf{a}}{\mathbf{b}} = \prod_{i=1}^m \binom{a_i}{b_i}$.

All logarithms in this paper are with respect to base 2. For a non-negative integer c, $\log^{\circ c}(n)$ denotes the c-times composition of the logarithm function with itself. For example, $\log^{\circ 2}(n) = \log\log(n)$. We denote by $\log^\star(n)$ the smallest non-negative integer c such that $\log^{\circ c}(n) \leq 1$.

We need the following number-theoretic result.

Lemma III.1 ([1, Lemma 2.4]). *For all $N \geq 2$, the product of the primes $p \leq 16 \log N$ is greater than N.*

A. Chinese Remainder Theorem

For our algorithms, we crucially use the Chinese Remainder Theorem. For completeness, we formally state the version we use and refer to Chapter 10 of [14] for a proof.

Theorem III.2 (Chinese Remainder Theorem). *Let $n_1, n_2, \ldots, n_t$ be pairwise relatively prime natural numbers greater than or equal to 2 and let $u_1, u_2, \ldots, u_t$ be arbitrary natural numbers such that for every $i \in [t]$, $u_i \leq n_i - 1$. Then, there is a unique $v \in \mathbb{N}$ with*

$v < \prod_{i=1}^{t} n_i$ *such that for every* $i \in [t]$, $v \equiv u_i \pmod{n_i}$.

Moreover, there is a deterministic algorithm, that when given $n_1, n_2, \ldots, n_t$ *and* $u_1, u_2, \ldots, u_t$ *as input, outputs* v *in time at most* $\mathrm{poly}(\sum_{i\in[t]} \log n_i)$*, i.e., in time polynomial in the input size.*

B. Hasse Derivatives

In this section, we briefly discuss the notion of Hasse derivatives that plays a crucial role in our results.

Definition III.3 (Hasse derivative)**.** *Let* $f(\mathbf{x})$ *be an* m*-variate polynomial over a commutative ring* R*. Let* $\mathbf{e} = (e_1, \ldots, e_m) \in \mathbb{N}^m$*. Then, the Hasse derivative of* f *with respect to the monomial* $\mathbf{x}^{\mathbf{e}}$ *is the coefficient of* $\mathbf{z}^{\mathbf{e}}$ *in the polynomial* $f(\mathbf{x}+\mathbf{z}) \in (R[\mathbf{x}])[\mathbf{z}]$. ◊

Notations: Suppose that $f(\mathbf{x})$ is an m-variate polynomial over a commutative ring R. For $\mathbf{a} \in \mathbb{N}^m$, denote by $\overline{\partial}_{\mathbf{a}}(f)$ the Hasse derivative of $f(\mathbf{x})$ with respect to the monomial $\mathbf{x}^{\mathbf{a}}$. For any non-negative integer k, define

$$\overline{\partial}^{\leq k}(f) := \left\{\overline{\partial}_{\mathbf{a}}(f) \mid \mathbf{a} \in \mathbb{N}^m \text{ s.t. } |\mathbf{a}|_1 \leq k\right\}, \text{ and}$$
$$\overline{\partial}^{<k}(f) := \{\overline{\partial}_{\mathbf{a}}(f) \mid \mathbf{a} \in \mathbb{N}^m \text{ s.t. } |\mathbf{a}|_1 < k\}.$$

For a univariate polynomial $h(t)$ over $\mathbb{F}$ and a non-negative integer k, denote by $h^{(k)}(t)$ the Hasse derivative of $h(t)$ with respect to the monomial t^k, that is, $\mathrm{coeff}_{z^k}(h(t+z))$.

The following lemma states that Hasse derivatives of polynomials can be computed efficiently. We defer its proof to the full version of this paper.

Lemma III.4. *Let* R *be either a finite field or a ring of the form* $\mathbb{Z}/r\mathbb{Z}$*. There exists an algorithm that given an* m*-variate polynomial* $f(\mathbf{x})$ *of individual degree less than* d *over* R *and* $\mathbf{e} \in \mathbb{N}^m$ *with* $|\mathbf{e}|_1 \leq dm$*, computes* $\overline{\partial}_{\mathbf{e}}(f)$ *in time* $O(d^m) \cdot \mathrm{poly}(m, d, \log|R|)$.

We now state a lemma that gives an algorithm for fast evaluation of all the Hasse derivatives of $h(t) = f(\mathbf{g}(t))$ over a finite field $\mathbb{F}_q$.

Lemma III.5. *Let* $f(\mathbf{x})$ *be an* m*-variate, individual degree less than* d *polynomial over a finite field* $\mathbb{F}_q$ *and* $\mathbf{g}(t) = (g_1, g_2, \ldots, g_m)$ *where* $g_i \in \mathbb{F}_q[t]$ *with degree bounded by* r*. Then, given access to evaluations of* $\overline{\partial}^{\leq 2m}(f)$ *on* $\mathbb{F}_q$*, there exists an algorithm that computes the evaluations of all* $\leq 2m$ *order Hasse derivatives of the polynomial* $h(t) = f(\mathbf{g}(t))$ *at all points in* $\mathbb{F}_q$ *in time* $\Theta(1)^m \cdot \mathrm{poly}(q, r, d, m)$.

The proof of the above lemma (and its promised algorithm) follows directly from Algorithm 4 in [2] and its correctness, thus is skipped here. The only change is that Algorithm 4 looked at $\leq m$-th order Hasse derivatives, and here we look at $\leq 2m$-th order Hasse derivatives. It is an easy exercise to see that the analysis of the algorithm in [2] extends as it is to this case.

C. Hermite Interpolation

The following lemma gives a stronger version of univariate polynomial interpolation, known as Hermite interpolation. To interpolate a univariate polynomial of degree d, we need its evaluations at $d+1$ distinct points. However, for Hermite interpolation, the number of evaluation points can be less than d, provided that evaluations of Hasse derivatives of the polynomial are available up to a certain order.

Lemma III.6 (Hermite interpolation)**.** *Let* R *be either a finite field or a ring of the form* $\mathbb{Z}/r\mathbb{Z}$*. Let* $f(x)$ *be a univariate polynomial over* R *and* $e_1, \ldots, e_\ell$ *be positive integers such that* $d := e_1 + \cdots + e_\ell$ *is greater than* $\deg(f)$*. Let* $a_1, a_2, \ldots, a_\ell \in R$ *such that for distinct* $i, j \in [\ell]$*,* $a_i - a_j$ *has multiplicative inverse in* R*. For all* $i \in [\ell]$ *and* $j \in [\![e_j]\!]$*, let* $\beta_{ij} = f^{(j)}(a_i)$*. Then given* (a_i, β_{ij}) *for all* $i \in [\ell]$ *and* $j \in [\![e_j]\!]$*,* $f(x)$ *can be computed in time* $\mathrm{poly}(d, \log|R|)$*. Equivalently, given* $(a_i, f(x) \bmod (x-a_i)^{e_i})$ *for all* $i \in [\ell]$*,* $f(x)$ *can be computed in time* $\mathrm{poly}(d, \log|R|)$.

We refer to the full version of the paper for the proof.

D. Fast Multivariate Multipoint Evaluation for Product Sets

The following lemma states that multivariate multipoint evaluation can be solved very efficiently if the set of evaluation points is a product set. We defer its proof to the full version of this paper.

Lemma III.7. *Let* R *be either a finite field or a ring of the form* $\mathbb{Z}/r\mathbb{Z}$*. There exists an algorithm that given an* m*-variate polynomial* $f(\mathbf{x})$ *of individual degree less than* d *over* R *and a finite subset* S *of* R*, outputs the evaluations* $f(\mathbf{a})$ *for all* $\mathbf{a} \in S^m$ *in time* $O(d^m + |S|^m) \cdot \mathrm{poly}(m, d, \log|R|)$.

IV. The Necessary Building Blocks

In this section, we set up some of the necessary building blocks for our algorithm. Due to space constraints, the proofs have been skipped here and can be found in the full version of the paper [15].

A. Primes in an Arithmetic Progression

The first ingredient we need is the existence of *sufficiently many* primes in the arithmetic progression $A_d = \{1, 1+d, 1+2d, \ldots\}$ that are not too large. When d is small, and x tends to infinity, a well-known result of Dirichlet (Theorem 5.5 in [9]) shows that the density of primes less than x in the arithmetic progression A_d tends to $\Theta(\frac{x}{\phi(d)\log x})$, where ϕ is the Euler totient function. However, for our application, we will need x and d to

be *close* to each other and hence it becomes important to carefully look at the error term in the prime counting function for the progression A_d.

While we do not know how to show such a statement, we end up working with a weaker statement that turns out to be sufficient for our application. This weaker statement that we use follows (immediately) from a deep result of Bombieri and Vinogradov that we state now. But first, we need some notation. For any $x \geq 0$, we denote by $\pi(x)$ the number of primes less than or equal to x. For $x \geq 0$ and $t \in \mathbb{N}$, we also use $\pi(x,t)$ to denote the number of primes less than or equal to x in the arithmetic progression $A_t = \{1, 1+t, 1+2t, \ldots, \}$

We are now ready to state the theorem of Bombieri and Vinogradov that we use. Various versions of the theorem can be found in literature, for instance, [7], [8], Theorem 18.1 in [9]. Here we rely on the bound in Equation 1.1 in [12].

Theorem IV.1 (Bombieri–Vinogradov)**.** *For any fixed $a > 0$, there exist constants $c = c(a)$ and $b = b(a)$ such that for all sufficiently large $x > 0$, $\sum_{t \leq Q} \left|\pi(x,t) - \frac{\pi(x)}{\phi(t)}\right| \leq cx(\log x)^{-a}$, where $Q \leq x^{1/2}(\log x)^{-b}$.*

Semantically, Theorem IV.1 says that on average (over $t \leq Q$), the quantity $\left|\pi(x,t) - \frac{\pi(x)}{\phi(t)}\right|$ is bounded by $(cx(\log x)^{-a})$. For our application, we would require a similar statement in the worst-case choice of t. This, however, is not known unconditionally when t is large compared to x[3] (which will turn out to be the case here), unless we assume the Generalized Riemann Hypothesis. Thankfully, it turns out that we have some wriggle room, and we can in fact work with the average-case statement above (up to some small loss in the parameters). More formally, we need the following immediate consequence of Theorem IV.1.

Lemma IV.2. *For any fixed $a > 1$, there exist constants $c = c(a)$ and $b = b(a)$ such that for all sufficiently large $x > 0$, $Q \leq x^{1/2}(\log x)^{-b}$ and $\delta > 1$, there is a $t_0 \in \mathbb{N}$ with $Q(1 - 2/\delta) \leq t_0 \leq Q$ and $\pi(x, t_0) \geq \frac{x}{4Q \log x}$.*

We now state the following consequence of this lemma that will be directly useful for us in the Chinese Remaindering step of our algorithm.

Lemma IV.3. *Let D, M be natural numbers and let D be sufficiently large. Then, there exists a natural number $\tilde{D} \in [0.8D, D]$ such that there are distinct primes $p_1, p_2, \ldots, p_k$ in the arithmetic progression $A_{\tilde{D}} = (1, 1+\tilde{D}, 1+2\tilde{D}, \ldots,)$ with the following properties.*

1) $k \leq D^2(\log M)^3$
2) *For every $i \in [k]$, $p_i \leq (D \log M)^3$*
3) $\prod_{i=1}^{k} p_i > M$

Moreover, there is a deterministic algorithm that on input D, M outputs $p_1, \ldots, p_k, \tilde{D}$ in time $\mathrm{poly}(D, \log M)$.

[3]More specifically, we would like x and t to be polynomially related to each other.

B. Explicit Kakeya Sets of Higher Degree

We start with the definition of Kakeya sets of high degree.

Definition IV.4 ([6])**.** *Let $\mathbb{F}$ be a finite field and let $u, m \in \mathbb{N}$. A set $K \subseteq \mathbb{F}^m$ is said to be a Kakeya set of degree u in $\mathbb{F}^m$ if there exist functions $g_0, g_1, \ldots, g_{u-1} : \mathbb{F}^m \to \mathbb{F}^m$ such that for every $\mathbf{a} \in \mathbb{F}^m$, the set of points*

$$\{g_0(\mathbf{a}) + g_1(\mathbf{a}) \cdot \tau + \cdots + g_{u-1}(\mathbf{a}) \cdot \tau^{u-1} + \mathbf{a} \cdot \tau^u : \tau \in \mathbb{F}\}$$

is a subset of K. ◊

For ease of notation, we denote the curve

$$\{g_0(\mathbf{a}) + g_1(\mathbf{a}) \cdot y + \cdots + g_{u-1}(\mathbf{a}) \cdot y^{u-1} + \mathbf{a} \cdot y^u : y \in \mathbb{F}\}$$

of degree u by $G_{\mathbf{a}}(y)$.

In their work [6], Björklund, Kaski and Williams gave an explicit construction of Kakeya sets of degree u of non-trivially small size, provided that the degree u and the field size $\mathbb{F}$ satisfy an appropriate divisibility condition. This construction will be crucial for our algorithm.

Theorem IV.5 (Explicit Kakeya sets of degree u [6])**.** *Let $\mathbb{F}$ be a finite field of size q, and let $u \in \mathbb{N}$ be such that $u+1$ divides $q-1$. Then, for every $m \in \mathbb{N}$, there is a Kakeya set K of degree u in $\mathbb{F}^m$ of size at most $\left(\frac{q-1}{u+1} + 1\right)^{m+1}$.*

Moreover, this set K is a union of at most q product sets in $\mathbb{F}^m$ and there is a deterministic algorithm that on input $u, m, \mathbb{F}$, outputs K and the associated functions $g_0, g_1, \ldots, g_{u-1}$ in time $O(q|K|)$.

Using the property that the set K in Theorem IV.5 is a union of product sets and that for product sets we have nearly linear algorithms for multipoint evaluation using Lemma III.7, we get the following.

Lemma IV.6. *Let $\mathbb{F}$ be a finite field of size q, $u \in \mathbb{N}$ be such that $u+1$ divides $q-1$, and $m \in \mathbb{N}$ be a natural number. Let K be the Kakeya set of degree u given by Theorem IV.5 over $\mathbb{F}^m$ and let $f(\mathbf{x})$ be a polynomial of degree less than d in each variable with coefficients in $\mathbb{F}$. Then there is a deterministic algorithm that takes as input the set K and the coefficient vector of f and outputs the evaluation of f at every point in K in time $O(|K| + d^m) \cdot \mathrm{poly}(m, d, q)$.*

C. Fast Multipoint Evaluation over Nice Finite Fields

Theorem IV.7. *Let $\mathbb{F}$ be a finite field of size q and let $d, \tilde{d}, m \in \mathbb{N}$ be such that $\tilde{d} \in [0.8d, d]$ and $\tilde{d} - 1$ divides $q - 1$. Then there is an algorithm that given a*

homogeneous m-variate polynomial in $\mathbb{F}[\mathbf{x}]$ of degree less than d in every variable and a set of N input points in $\mathbb{F}^m$, outputs the evaluation of this polynomial on these inputs in time $(d^m + N) \cdot \Theta(1)^m \cdot \mathrm{poly}(q, m, d)$.

Proof. Let f be the input polynomial and $\mathbf{a}_1, \ldots, \mathbf{a}_N \in \mathbb{F}^m$ be the input points of interest.

At a high level, the algorithm here is similar in structure to that in [2]. We first evaluate the polynomial on an appropriate product set $\mathcal{P}$ in nearly linear time using Lemma III.7 in the preprocessing phase. Next, in the local computation step, we look at the restriction of f on a curve $C_{\mathbf{a}}$ *through* any point $\mathbf{a} \in \mathbb{F}^m$ of interest. Based on the construction of the aforementioned product set $\mathcal{P}$, we will guarantee that there is a curve $C_{\mathbf{a}}$ through $\mathbf{a}$ such that the intersection of $C_{\mathbf{a}}$ with the set $\mathcal{P}$ is sufficiently large, so that the univariate polynomial obtained by restricting f to $C_{\mathbf{a}}$ can be uniquely *decoded* using the evaluation of f on $\mathcal{P}$. We then use this decoded polynomial to obtain $f(\mathbf{a})$.

Despite this high-level similarity, there are some technical differences between the algorithm here and that in [2]. Primarily, these differences arise due to the fact that unlike the setting in [2], we are no longer working over fields of small characteristic. So, the construction of the set $\mathcal{P}$ is different here and is based on the ideas in [6]. We now specify the details, starting with the description of the algorithm.

a) The algorithm:

1) From the coefficient vector of f, compute each of Hasse derivatives of f of order at most $2m$.
2) Using Theorem IV.5, we construct a Kakeya set K of degree $u = (q-1)/(\tilde{d}-1) - 1$. As is necessary, $u+1$ divides $q-1$. Note that, $|K| \leq \tilde{d}^{(m+1)}$.
3) For every Hasse derivative $\tilde{f}$ of f of order at most $2m$, evaluate $\tilde{f}$ on K using Lemma IV.6.
4) For every $i \in [N]$:
   a) We consider the univariate polynomial $R_i(y)$ obtained by the restriction of f on the curve $G_{\mathbf{a}_i}(y)$. This is a univariate polynomial of degree at most $(d-1)m \cdot (q-1)/(\tilde{d}-1) < 2m(q-1)$. Using Lemma III.5, compute the evaluation of $R_i(y)$ and all its $\leq 2m$ order Hasse derivatives on $\mathbb{F}$.
   b) Since degree of R_i is less than $2m(q-1)$, and we have the evaluation of R_i and all its derivatives of order at most $2m$ on q points, we can recover R_i uniquely from this information. In particular, we use Lemma III.6 to recover $R_i(y)$.
   c) We output $f(\mathbf{a}_i)$ to be equal to the coefficient of $y^{\deg(f)\cdot u}$ in $R_i(y)$.

Due to space constraints, we skip the proofs of correctness and the analysis of the running time of the algorithm here, and defer these details to the full version of the paper [15].

□

V. The First Algorithm over $\mathbb{Z}/r\mathbb{Z}$

With the necessary background in place, we are now ready to describe our first algorithm for fast multivariate multipoint evaluation over rings of the form $\mathbb{Z}/r\mathbb{Z}$. This already handles the case of prime fields, and contains most of our main ideas.

The case of extension rings as well as many of the proofs can be found in the full version of this paper [15].

A. The Description of the Algorithm

Algorithm 1 The First Algorithm over $\mathbb{Z}/r\mathbb{Z}$

Algorithm MME-A$(f, \mathbf{a}_1, \ldots, \mathbf{a}_N, r)$

where f is an m-variate *homogeneous* polynomial over $\mathbb{Z}/r\mathbb{Z}$ of individual degree less than d and $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \in (\mathbb{Z}/r\mathbb{Z})^m$ are evaluation points.

1) Let $F \in \mathbb{Z}[\mathbf{x}]$ be the m-variate homogeneous polynomial of individual degree less than d obtained from f by replacing each of its coefficients with its natural lift in the set $[\![r]\!]$ of integers.
2) For every $i \in [N]$, let $\tilde{\mathbf{a}}_i \in [\![r]\!]^m$ be the lift of $\mathbf{a}_i \in (\mathbb{Z}/r\mathbb{Z})^m$ to the integers.
3) Let $M = d^m r^{dm}$. We invoke Lemma IV.3 with parameters d and M and obtain a natural number $\tilde{d} \in [0.8d, d]$ and primes $p_1, p_2, \ldots, p_k$ where $k \leq d^2(\log M)^3$, each $p_i \leq d^3(\log M)^3$ and is congruent to 1 modulo $\tilde{d}$, and $\prod_{i\in[k]} p_i > M$.
4) For $j \in [k]$, let $f_j(\mathbf{x}) \in \mathbb{F}_{p_i}[\mathbf{x}]$ be the m-variate homogeneous polynomial of individual degree less than d obtained by reducing each of the coefficients of F modulo the prime p_j. Similarly, for every $i \in [N]$, let $\mathbf{a}_{i,j} \in \mathbb{F}_{p_j}^m$ be obtained by reducing each of the coordinates of $\tilde{\mathbf{a}}_i$ modulo p_j.
5) For every $j \in [k]$, invoke the algorithm in Theorem IV.7 for the polynomial f_j, input points $\{\mathbf{a}_{i,j} : i \in [N]\}$ and parameters $d, \tilde{d}$ as above, and get $f_j(\mathbf{a}_{i,j})$ for all $j \in [k]$ and $i \in [N]$. Note that each f_j is a homogeneous polynomial, and from the guarantees of Lemma IV.3, $\tilde{d}$ is in the range $[0.8d, d]$ and $\tilde{d} - 1$ divides $p_i - 1$ as needed by Theorem IV.7.
6) For every $i \in [N]$, use the Chinese Remainder Theorem (Theorem III.2) to compute $F(\tilde{\mathbf{a}}_i)$ from $\{f_j(\mathbf{a}_{i,j}) : j \in [k]\}$.
7) For every $i \in [N]$, output $f(\mathbf{a}_i) = F(\tilde{\mathbf{a}}_i) \mod r$.

We summarize the correctness and the time complexity of the algorithms in the following theorem, and refer to the full version of the paper [15] for the proof.

Theorem V.1. *Let $f(\mathbf{x})$ be a homogeneous m-variate polynomial over $\mathbb{Z}/r\mathbb{Z}$ of individual degree less than d. Let $\mathbf{a}_1, \ldots, \mathbf{a}_N$ be N points from $(\mathbb{Z}/r\mathbb{Z})^m$. Then, given $(f, \mathbf{a}_1, \ldots, \mathbf{a}_N, r)$ as the input to Algorithm 1, it computes $f(\mathbf{a}_i)$ for all $i \in [N]$ in time $(d^m+N)\cdot\Theta(1)^m \cdot \mathrm{poly}(m, d, \log r)$.*

VI. The Second Algorithm over Rings of the Form $\mathbb{Z}/r\mathbb{Z}$

The main result of this section is the following theorem.

Theorem VI.1. *Over $\mathbb{Z}/r\mathbb{Z}$, for all $m \in \mathbb{N}$ and sufficiently large $d \in \mathbb{N}$, there exists a deterministic algorithm that outputs the evaluation of an m-variate polynomial of degree less than d in each variable on N points in time $(d^m+N)^{1+o(1)}\cdot\mathrm{poly}(m, d, \log r)$ provided that $\log^{\circ c} r \leq d^{o(1)}$ for some fixed constant $c \in \mathbb{N}$.*

We need the following lemma. It gives a way of computing the evaluation of $f(\mathbf{x})$ over a ring R at a point $\mathbf{a}$ from the evaluations of Hasse derivatives of $f(\mathbf{x})$ at another point $\mathbf{b}$, provided that the coordinates of $\mathbf{a}-\mathbf{b}$ are in a nilpotent ideal of R.

Lemma VI.2. *Let $f(\mathbf{x})$ be an m-variate polynomial over a commutative ring R. Let I be an ideal of R and s be a positive integer such that $I^s = 0$. Let $\mathbf{a} = (a_1, \ldots, a_m), \mathbf{b} = (b_1, \ldots, b_m) \in R^m$ such that $a_i \equiv b_i \pmod{I}$ for $i \in [m]$. Then*

$$f(\mathbf{a}) = \sum_{\mathbf{e}\in\mathbb{N}^m : |\mathbf{e}|_1 < s} \overline{\partial}_{\mathbf{e}}(f)(\mathbf{b}) \cdot (\mathbf{a}-\mathbf{b})^{\mathbf{e}}.$$

A. A Basic Algorithm

We first describe a basic algorithm, MME-Product-Set, that evaluates a polynomial $f(\mathbf{x}) \in (\mathbb{Z}/r^s\mathbb{Z})[\mathbf{x}]$ at N points in $(\mathbb{Z}/r^s\mathbb{Z})^m$ simultaneously.

Algorithm 2 Basic Algorithm

Algorithm MME-Product-Set$(f, \mathbf{a}_1, \ldots, \mathbf{a}_N, r, s)$

where $f(\mathbf{x})$ is an m-variate polynomial over $\mathbb{Z}/r^s\mathbb{Z}$ of individual degree at most $d-1$, $\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_N$ are evaluation points in $(\mathbb{Z}/r^s\mathbb{Z})^m$, and $s \in [m]$.

1) For all $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$, use Lemma III.4 to compute $f_{\mathbf{e}}(\mathbf{x}) := \overline{\partial}_{\mathbf{e}}(f)(\mathbf{x})$.
2) For all $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$, use Lemma III.7 to compute $f_{\mathbf{e}}(\mathbf{a})$ for $\mathbf{a} \in [\![r]\!]^m$, where $[\![r]\!]$ is identified with a subset of $\mathbb{Z}/r^s\mathbb{Z}$ via $i \mapsto i + r^s\mathbb{Z}$.
3) For all $i \in [N]$, compute $\bar{\mathbf{a}}_i \in [\![r]\!]^m \subseteq (\mathbb{Z}/r^s\mathbb{Z})^m$ such that the coordinates of $\bar{\mathbf{a}}_i$ are the remainders of the corresponding coordinates of $\mathbf{a}_i$ modulo r.
4) For all $i \in [N]$, compute and output

$$f(\mathbf{a}_i) = \sum_{\mathbf{e}\in\mathbb{N}^m : |\mathbf{e}|_1 < s} f_{\mathbf{e}}(\bar{\mathbf{a}}_i) \cdot (\mathbf{a}_i - \bar{\mathbf{a}}_i)^{\mathbf{e}}. \quad \text{(VI.3)}$$

Lemma VI.4. *Given the input $(f, \mathbf{a}_1, \ldots, \mathbf{a}_N, r, s)$, the algorithm* MME-Product-Set *computes $f(\mathbf{a}_i)$ for all $i \in [N]$ in time $O(\binom{m+s-1}{s-1}(d^m + r^m + N)) \cdot \mathrm{poly}(m, d, \log r)$.*

B. The Description of the Algorithm

We describe the second algorithm MME-B now.

Algorithm 3 The Second Algorithm over $\mathbb{Z}/r^s\mathbb{Z}$

Algorithm MME-B$(f, \mathbf{a}_1, \ldots, \mathbf{a}_N, r, s, t)$

where $f(\mathbf{x})$ is an m-variate polynomial over $\mathbb{Z}/r^s\mathbb{Z}$ of individual degree at most $d-1$, $\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_N$ are evaluation points in $(\mathbb{Z}/r^s\mathbb{Z})^m$, $s \in [m]$, and $t \geq 0$ is the depth of the reduction tree.

1) If $t = 0$, invoke MME-PRODUCT-SET with input $(f(\mathbf{x}), \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N, r, s)$ to compute $f(\mathbf{a}_i)$ for $i \in [N]$, and return.
2) For all $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$, use Lemma III.4 to compute $f_{\mathbf{e}}(\mathbf{x}) := \overline{\partial}_{\mathbf{e}}(f)(\mathbf{x})$, and then compute a lift $\tilde{f}_{\mathbf{e}}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$ of $f_{\mathbf{e}}(\mathbf{x})$ with coefficients in $[\![r^s]\!]$.
3) For all $i \in [N]$, compute $\tilde{\mathbf{a}}_i \in [\![r]\!]^m$ such that the coordinates of $\tilde{\mathbf{a}}_i$ are the remainders of the corresponding coordinates of $\mathbf{a}_i$ modulo r, and compute $\bar{\mathbf{a}}_i := \tilde{\mathbf{a}}_i \bmod r^s \in (\mathbb{Z}/r^s\mathbb{Z})^m$.
4) Let $M := d(r-1)^d$. Find primes $p_1 < p_2 < \cdots < p_k \leq 16 \log M$ such that $\prod_{j=1}^k p_j > M$.
5) For all $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$ and $j \in [k]$, compute $f_{\mathbf{e},j}(\mathbf{x}) := \tilde{f}_{\mathbf{e}}(\mathbf{x}) \bmod p_j^m \in (\mathbb{Z}/p_j^m\mathbb{Z})[\mathbf{x}]$.
6) For all $i \in [N]$ and $j \in [k]$, compute $\mathbf{a}_{i,j} := \tilde{\mathbf{a}}_i \bmod p_j^m \in (\mathbb{Z}/p_j^m\mathbb{Z})^m$.
7) For $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$ and $j \in [k]$, invoke MME-B on input $(f_{\mathbf{e},j}, \mathbf{a}_{1,j}, \ldots, \mathbf{a}_{N,j}, p_j, m, t-1)$ to compute $f_{\mathbf{e},j}(\mathbf{a}_{i,j})$ for $i \in [N]$.
8) For all $\mathbf{e} \in \mathbb{N}^m$ with $|\mathbf{e}|_1 < s$ and $i \in [N]$, use the Chinese Remainder Theorem (Theorem III.2) to compute $\tilde{f}_{\mathbf{e}}(\tilde{\mathbf{a}}_i)$ as the unique $Q_i \in \left[\!\!\left[\prod_{j=1}^k p_j^m\right]\!\!\right]$ such that $Q_i \bmod p_j^m = f_{\mathbf{e},j}(\mathbf{a}_{i,j})$ for $j \in [k]$, and then compute $f_{\mathbf{e}}(\bar{\mathbf{a}}_i) = \tilde{f}_{\mathbf{e}}(\tilde{\mathbf{a}}_i) \bmod r^s \in \mathbb{Z}/r^s\mathbb{Z}$.
9) For all $i \in [N]$, compute and output

$$f(\mathbf{a}_i) = \sum_{\mathbf{e} \in \mathbb{N}^m : |\mathbf{e}|_1 < s} f_{\mathbf{e}}(\bar{\mathbf{a}}_i) \cdot (\mathbf{a}_i - \bar{\mathbf{a}}_i)^{\mathbf{e}}. \quad \text{(VI.5)}$$

We refer to the full version of the paper [15] for the correctness and time complexity of the algorithm.

ACKNOWLEDGMENT

Mrinal is thankful to Swastik Kopparty for introducing him to the question of multipoint evaluation and the work of Kedlaya–Umans [1] and to Prahladh Harsha and Ramprasad Saptharishi for many helpful discussions.

REFERENCES

[1] K. Kedlaya and C. Umans, "Fast polynomial factorization and modular composition," *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1767–1802, 2011. [Online]. Available: https://doi.org/10.1137/08073408X

[2] V. Bhargava, S. Ghosh, M. Kumar, and C. K. Mohapatra, "Fast, algebraic multivariate multipoint evaluation in small characteristic and applications," *arXiv preprint arXiv:2111.07572*, 2021, to appear in STOC 2022. [Online]. Available: https://arxiv.org/abs/2111.07572

[3] A. Borodin and R. Moenck, "Fast modular transforms," *Journal of Computer and System Sciences*, vol. 8, no. 3, pp. 366–386, 1974. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022000074800292

[4] M. Nüsken and M. Ziegler, "Fast multipoint evaluation of bivariate polynomials," in *Algorithms – ESA 2004*, S. Albers and T. Radzik, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 544–555.

[5] C. Umans, "Fast polynomial factorization and modular composition in small characteristic," in *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, C. Dwork, Ed. ACM, 2008, pp. 481–490. [Online]. Available: https://doi.org/10.1145/1374376.1374445

[6] A. Björklund, P. Kaski, and R. Williams, "Generalized kakeya sets for polynomial evaluation and faster computation of fermionants," *Algorithmica*, vol. 81, no. 10, pp. 4010–4028, 2019. [Online]. Available: https://doi.org/10.1007/s00453-018-0513-7

[7] E. Bombieri, "On the large sieve," *Mathematika*, vol. 12, pp. 201–225, 1965.

[8] A. I. Vinogradov, "The density hypothesis for the dirichlet l-series," *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 29, pp. 903–934, 1965.

[9] K. Kedlaya, "Lecture notes for the course 'Analytic Number Theory'," 2015. [Online]. Available: https://kskedlaya.org/papers/ant-overall.pdf

[10] C. Siegel, "Über die classenzahl quadratischer zahlkörper," *Acta Arithmetica*, vol. 1, no. 1, pp. 83–86, 1935. [Online]. Available: https://eudml.org/doc/205054

[11] A. Walfisz, "Zur additiven Zahlentheorie. II." *Mathematische Zeitschrift*, vol. 40, no. 1, pp. 592–607, 1936.

[12] J. Maynard, "Primes in arithmetic progressions to large moduli I: Fixed residue classes," *arXiv preprint arXiv:2006.06572*, 2020. [Online]. Available: https://arxiv.org/abs/2006.06572

[13] S. Lang, *Algebra*, 3rd ed. Springer-Verlag, New York Inc., 2002.

[14] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, 3rd ed. Cambridge University Press, 2013.

[15] V. Bhargava, S. Ghosh, Z. Guo, M. Kumar, and C. Umans, "Fast multivariate multipoint evaluation over all finite fields," 2022. [Online]. Available: https://arxiv.org/abs/2205.00342

[16] M. Agrawal, N. Kayal, and N. Saxena, "Primes is in P," *Annals of Mathematics*, vol. 160, no. 2, pp. 781–793, 2004.

[17] V. Shoup, *A Computational Introduction to Number Theory and Algebra*, 2nd ed. New York: Cambridge University Press, 2008, available from https://shoup.net/ntb/.