# COMPLEXITY OF COUNTING POINTS ON CURVES, AND THE FACTOR $P_1(T)$ OF THE ZETA FUNCTION OF SURFACES

#### DIPTAJIT ROY, NITIN SAXENA <sup>(D)</sup>, AND MADHAVAN VENKATESH

ABSTRACT. The article concerns the computational complexity of a fundamental problem in number theory: counting points on curves and surfaces over finite fields. There is no subexponential-time algorithm known and it is unclear if it can be NP-hard.

Given a curve (say, f(x, y) = 0 of degree d over field  $\mathbb{F}_q$ ), we present the first efficient Arthur-Merlin protocol to certify its point-count, its Jacobian group structure, and its Hasse-Weil zeta function. We place this problem in AM  $\cap$  coAM, while the previous best was BQP (Kedlaya'06). We extend this result to a smooth projective *surface* (say, dimension 2 in  $\mathbb{P}^4$  and degree D) to certify the factor  $P_1(T)$ , corresponding to the *first* Betti number, of the zeta function; the previous best was  $\mathbb{P}^{\#\mathbb{P}}$  by using the counting oracle. Famously, the complex reciprocal roots of  $P_1(T)$  have norm  $\sqrt{q}$  (Deligne's proof of the Weil-Riemann Hypothesis, 1974), and it tells us all about the Picard variety of the surface. We give the first algorithm to compute  $P_1(T)$  that is poly(log q)-time if the degree D of the input surface is *fixed*; and in quantum poly( $D \log q$ )-time in general.

Our technique in the curve case, is to sample hash functions using the Weil and Riemann-Roch bounds, to certify the group order of its Jacobian. For *higher* dimension varieties, we first reduce to the case of a surface, which is fibred as a Lefschetz pencil of hyperplane sections over  $\mathbb{P}^1$ . The formalism of vanishing cycles, and the inherent *big monodromy*, enable us to prove an effective version of Deligne's 'theoreme du pgcd' using the hard-Lefschetz theorem and an equidistribution result due to Katz. These reduce our investigations to that of computing the zeta function of a curve, defined over a finite field extension  $\mathbb{F}_Q/\mathbb{F}_q$  of poly-bounded degree. This explicitization of the theory yields the first nontrivial upper bounds on the computational complexity.

#### CONTENTS

1. Introduction	1
2. Zeta function of curves	6
3. Surfaces: Vanishing cycles, monodromy, and equidistribution	11
4. $P_1(T)$ for smooth projective varieties	16
5. Conclusion	21
Acknowledgements	22
References	22

#### 1. INTRODUCTION

Since antiquity mathematicians have studied 'simple' equations to find, and count, the roots; unearthing powerful theories. A classical family is the projective curve C (in  $\mathbb{P}^2$ ):  $a_0x_0^d + a_1x_1^d + a_2x_2^d \equiv 0 \mod p$ , and the projective surface S (in  $\mathbb{P}^3$ ):  $a_0x_0^d + a_1x_1^d + a_2x_2^d + a_3x_3^d \equiv 0 \mod p$ , for a prime p, and numbers  $a_i$ 's, d. One would like to count the roots, denoted  $|C(\mathbb{F}_p)|$  resp.  $|S(\mathbb{F}_p)|$ , in time polynomial in  $\log p$  and d. We can trivially estimate the counts to be p resp.  $p^2$ , but how good are these estimates? This has been studied, for various cases, at least since the times of Gauß (1800s) [Pie03], Jacobi [JBW<sup>+</sup>84], Lebesgue, Hardy & Littlewood, Davenport & Hasse; till the modern formulation of Weil-Riemann hypothesis of a zeta function was given by Weil [Wei49, Wei48a]. It uses the topological and geometric properties of a variety to reflect on its arithmetic properties.

<sup>2020</sup> Mathematics Subject Classification. Primary 11Y16 11G25 68Q25 14G15.

Key words and phrases. Zeta function, Weil conjectures, ell-adic, étale, cohomology, Frobenius, Jacobian variety, Lefschetz pencils, vanishing cycles, equidistribution, monodromy, Arthur-Merlin protocol, quantum.

Specifically, over characteristic zero, one can associate to a smooth variety its singular and de Rham cohomology groups. In this setting, there have been algorithmic results on computing these topological invariants e.g., the number of irreducible components in [BS10], and more general cohomology computations using real algebraic geometry [BPR06, Sch07] and algebraic de Rham cohomology [OT99, Sch19].

We are interested in an arithmetic analogue of this line of work. This study has found numerous applications in modern computing; especially motivated by the example of curves  $[CFA^+05]$  and surfaces [Ber20]. In particular, the genus of a curve and its number of rational points can be read off from its zeta function, and selecting a curve with optimised such parameters is a natural question that crops up in the theory of algebraic-geometric (AG) codes [TV13]. For a higher dimensional variety, the first cohomology encodes information about its *Picard* variety, a natural abelian variety parametrising codimension one subvarieties modulo an equivalence relation; that plays the analogue of the Jacobian of a curve.

In this work we will clarify the complexity of the curve case in a significant way, and we will take the first steps in the surface case. The topological invariants involved in the latter are much harder as they demand the most abstruse cohomology theory [FK13]. In particular, we provide the first explicit, computational results on the Picard variety of higher dimensional varieties.

Let  $X_0$  be a smooth, projective variety of dimension n over the finite field  $\mathbb{F}_q$  of characteristic p > 0. Denote by X the base-change to the algebraic closure  $\overline{\mathbb{F}}_q$ . To encode the number of its points over all finite field extensions, the zeta function of X is defined as

$$Z(X/\mathbb{F}_q, T) := \exp\left(\sum_{j=1}^{\infty} \#X(\mathbb{F}_{q^j})\frac{T^j}{j}\right) \in \mathbb{Z}[[T]].$$

It is an exponential of the generating function of the *point-counts*. The result is seen as a formal power series in T. Let  $\ell$  be a prime distinct from p. By the foundational work of Grothendieck et al.[G<sup>+</sup>77] on  $\ell$ -adic cohomology, it is known that the zeta function can be written as a rational function:

(1.1) 
$$Z(X/\mathbb{F}_q, T) = \frac{P_1(T)P_3(T)\cdots P_{2n-1}(T)}{P_0(T)P_2(T)\cdots P_{2n}(T)} \in \mathbb{Q}(T),$$

where  $P_i(T) = \det(1 - TF_q^* \mid H^i(X, \mathbb{Q}_\ell))$  is the characteristic polynomial<sup>1</sup> of the map  $F_q^*$  induced on the cohomology by the geometric Frobenius. Further, the zeta function satisfies the functional equation

$$Z(X/\mathbb{F}_q, 1/(q^n T)) = \pm q^{n \cdot \chi/2} \cdot T^{\chi} \cdot Z(X/\mathbb{F}_q, T)$$

where  $\chi = \sum_{i=0}^{2n} (-1)^i \cdot \dim \operatorname{H}^i(X, \mathbb{Q}_\ell)$  is the  $\ell$ -adic Euler-Poincaré characteristic of X. Denote  $\beta_i := \dim \operatorname{H}^i(X, \mathbb{Q}_\ell)$ , also called the *i*th Betti number. As a result of Deligne's proof [Del74] of the Weil-Riemann hypothesis, we have  $P_i(T) = \prod_{j=1}^{\beta_i} (1 - \alpha_{i,j}T) \in \mathbb{Z}[T]$ , with  $\alpha_{i,j}$  being algebraic numbers such that  $|\iota(\alpha_{i,j})| = q^{i/2}$  for any embedding  $\iota : \mathbb{Q}(\alpha_{i,j}) \to \mathbb{C}$ . In particular, it follows that the  $P_i(T)$  are independent of  $\ell$ .

The complexity of computing the zeta function of a variety over a finite field is a natural question, being the generalisation of the ancient problem of counting the number of congruent solutions of a given polynomial equation modulo a prime p. Let  $X \subset \mathbb{P}^N$  be a smooth, projective variety of dimension n and degree D, presented as the zero set of homogeneous polynomials  $f_1, \ldots, f_m$  each of total degree  $\leq d$ . The dimension n of the variety (and that of its embedding space, N) is considered fixed. This is because the Betti numbers of a variety, and hence also the degree of its zeta function, are *exponential* in N.

<sup>&</sup>lt;sup>1</sup>or *reversed* characteristic polynomial, according to another convention

So, in practice, one seeks algorithms to compute  $Z(X/\mathbb{F}_q, T)$  efficiently in only two parameters, namely  $\log q$  and D. Such an algorithm which is polynomial-time in both is unknown, despite being a well-studied problem in the intersection of mathematics and computer science. A special case of a question of Serre [Ser16, Preface] asks the following (paraphrased), which fundamentally motivates our work.

**Question** (Compute). Let  $\mathcal{X}_0$  be a (fixed) smooth, projective variety over  $\mathbb{Q}$ . Is there an algorithm which, given a prime p of good reduction of  $\mathcal{X}_0$ ; computes the point-count of the reduction,  $\#X(\mathbb{F}_p)$ , in time polynomial in  $\log p$ ?

We obtain the first polynomial-time (in log q) algorithm to compute  $P_1(T)$  for smooth varieties (of dimension  $\geq 2$ ) of fixed degree D, extending a line of work that goes back to elliptic curves [Sch85] and abelian varieties [Pil90]. Consequently, for a surface X, we can now compute all  $P_i$ 's except  $P_2(T)$ ; thus, computing  $Z(X/\mathbb{F}_q, T) \cdot P_2(T)$ .

One notices that even for the simple-to-present hyperelliptic curves,  $y^2 = f(x) \mod p$ , that are quite useful in cryptography [CFA<sup>+</sup>05], there is no fast algorithm known to compute the zeta function, in time polynomial in both log p and deg(f). So, one wonders if an 'easier' verification question (see [LPP03, Question 15]) should be answered first:

Question (Certify). Given a variety X, a rational function Q(T) and some 'data', is there a polynomial-time algorithm to *verify* that Q(T) is the zeta function of X? In other words, is the zeta function computation problem in NP, or in coNP? More generally, given input polynomials  $Q_i(T) \in \mathbb{Z}[T]$ , for all *i*, is *verifying* 

$$Q_i(T) \stackrel{!}{=} P_i(T)$$
 in NP  $\cap$  coNP ?

In this work, we take a major step towards answering the above question about verifying the zeta function. Unfortunately, our protocol does not translate into a practical algorithm. But, we do show that the problem of computing zeta function of a smooth projective curve (with D, log q variables) is *unlikely* to be NP-hard, or has 'intermediate' complexity (in the sense of [AB09, §8.2.4]).

Further, generalising work of Kedlaya [Ked06] (which was restricted to curves), we obtain the first quantum algorithm for computing  $P_1(T)$  that is polynomial-time in log q and D.

1.1. **Prior work.** It is possible to interpret (1.1) via a trace formula in a suitable *Weil cohomology* theory. Examples include  $\ell$ -adic cohomology, for primes  $\ell$  distinct from the characteristic, developed by Grothendieck [G<sup>+</sup>77]; and rigid cohomology, an extension of crystalline cohomology due to Berthelot [Ber86]. In general, algorithms for computing the zeta function can be classified broadly into two distinct families,  $\ell$ -adic resp. *p*-adic, usually based on the nature of the cohomology theory being employed. The progenitor of the  $\ell$ -adic class of algorithms is the work of Schoof [Sch85], who gave an algorithm to compute the zeta function of an elliptic curve over  $\mathbb{F}_q$  with complexity polynomial in log q. This method was generalised by Pila [Pil90] to curves (of genus g), and abelian varieties, with improvements for some special cases due to Huang-Ierardi [HI98] and Adleman-Huang [AH01]. The complexity of these algorithms, while polynomial in log q is exponential or worse in g. A common theme is the realisation of the étale cohomology  $H^1(X, \mu_\ell)$  as the  $\ell$ -torsion Pic<sup>0</sup>(X)[ $\ell$ ] in the Picard scheme. This has, so far, limited their application to varieties where this realisation can be made explicit, namely curves and abelian varieties. There has been work showing the computability of étale cohomology in higher degrees as well [MO15], but it has not proven amenable to complexity analysis yet.

On the other hand, *p*-adic methods encompass a more diverse range of algorithms. Some early examples are Satoh's algorithm for elliptic curves [Sat00] using canonical lifts and Kedlaya's algorithm for hyperelliptic curves [Ked01] using Monsky-Washnitzer cohomology (and extensions thereof [DV06, CDV06]). Lauder-Wan [LW06], inspired by work of Dwork on the rationality of the zeta function [Dwo60], proposed a more general algorithm capable of handling arbitrary varieties.

Lauder [Lau04] also developed an algorithm for hypersurfaces based on *p*-adic deformation theory. More recently, there is the 'non-cohomological' work of Harvey [Har15], who devised an algorithm based on a novel trace formula. The complexity of these algorithms, while polynomial in the degree D of the variety, is exponential in log p. A common theme is that they involve a *p*-adic lift of the Frobenius, which necessitates working with O(p) monomials in the basis for the respective *p*-adic cohomology theory.

1.2. A detour to basic complexity notions [AB09]. Since the zeta function is defined via an infinite sum of point-counts, the problem of computing  $P_1(T)$  of a smooth projective variety could potentially be *un*computable! A lot of work has been done to pinpoint the *complexity* of this problem [MO15]; but a complete solution is unknown even in the case of smooth projective curves.

This paper is motivated by the class of *Interactive Protocol*, where the verification algorithm (called *Arthur*) is allowed to have a number of interactions with the oracle (*Merlin*). In the Arthur-Merlin class, denoted by AM, we assume that Arthur has access to Merlin only once throughout the computation. Arthur is allowed to use randomisation in the verification algorithm (thus, it is like a randomised NP protocol). Problems lying in AM  $\cap$  coAM class are 'unlikely' to be NP-hard (else, the polynomial-hierarchy collapses, see [BHZ87]); optimistically, we can even conjecture them to have quasipolynomial-time algorithms. Many famous problems are known to be in AM  $\cap$  coAM - e.g. integer factoring, discrete logarithm, graph isomorphism, algebra isomorphism, and algebraic dependence (see [KS06, GSS19] and the references therein). A major byproduct of this work is to conclude that *computing*  $P_1(T)$  *is unlikely to be* NP-hard, as we show it to be in AM  $\cap$  coAM.

Another popular complexity class is that of *quantum* polynomial-time, denoted BQP. It is not clear how it compares with the complexity classes we defined earlier, except the trivial comparison of BPP  $\subseteq$  BQP. Many famous problems are known to be in BQP — e.g. integer factoring, discrete logarithm, zeta function of curves, and the hidden-subgroup problem of abelian groups (see [Ked06, NC01]). It is unknown if there is any NP-hard problem contained in BQP, or if BQP  $\subseteq$  NP  $\cup$  coNP. Similarly, BQP and AM are (currently) incomparable classes. Both of them are solvable using the *counting class* #P as an oracle (e.g. the problem of counting satisfying assignments).

## 1.3. Main results: Certify or Compute.

**Certification.** For a smooth, projective, geometrically irreducible *curve*  $C \subset \mathbb{P}^N$  of genus g > 0, the zeta function has the form

$$Z(C/\mathbb{F}_q, T) = \frac{P_1(C/\mathbb{F}_q, T)}{(1-T)(1-qT)},$$

where  $P_1(C/\mathbb{F}_q, T) \in \mathbb{Z}[T]$  is of degree 2g, with constant term 1. Somewhat surprisingly, we will not only verify  $P_1(C/\mathbb{F}_q, T)$  but also the abelian group structure of the Jacobian variety over the base field. It addresses a question of Kedlaya [Ked06, §9] on verifying the order of the Jacobian as a black-box group.

**Theorem 1.1** (Zeta & Jacobian). Given an input polynomial  $P(T) \in \mathbb{Z}[T]$ , deciding whether P(T) is the numerator polynomial of the zeta function of the smooth, projective curve C, given as above (with variable  $g \log q$ ), is in AM  $\cap$  coAM. Moreover, given a finite Abelian group G (via additive generators), the verification problem

$$G \stackrel{\cdot}{\simeq} \operatorname{Jac}(C)(\mathbb{F}_q) \quad is \ in \quad \operatorname{AM} \cap \operatorname{coAM}.$$

The above protocol reduces to the verification of a few group orders  $\mathcal{N}_j := \# \operatorname{Jac}(C)(\mathbb{F}_{q^j})$  of the Jacobian of C, which entails the verification of independence for a set of generators. The well-known "mod- $\ell$  pairing"-based arguments do not give a protocol immediately; as, the order  $\ell \mid \mathcal{N}_j$  of a generator can be very large. In which case, it can require an exponential degree extension  $\mathbb{F}_Q/\mathbb{F}_q$  for the Tate pairing to be non-degenerate on  $\operatorname{Jac}(\mathbb{F}_Q)[\ell]$  [FR94].

We now assume the input is a smooth, projective variety  $X_0 \subset \mathbb{P}^N$  of dimension<sup>2</sup>  $n \geq 1$  and degree D, over the finite field  $\mathbb{F}_q$ , presented as a system of m homogeneous polynomials  $f_1, \ldots, f_m$  of degree  $\leq d^3$ . Further, we assume  $X_0$  is obtained via good reduction of a smooth projective variety  $\mathcal{X}_0$  over a number field K at a prime  $\mathfrak{p}$ . As we are interested in the regime of varying the characteristic, we assume accordingly that we have a globally defined *smooth* model over characteristic zero (i.e., a number field). Write  $X := X_0 \times \overline{\mathbb{F}}_q$ . We have the following certification result.

**Theorem 1.2** (Certify  $P_1$ ). Given  $Q_1(T) \in \mathbb{Z}[T]$ , deciding whether  $Q_1(T) = P_1(X/\mathbb{F}_q, T)$ , for  $X_0$  given as above (with variable  $D \log q$ ), is in AM  $\cap$  coAM.

The technical heart of the results in this work lies in the proof of Theorem 4.7, an effective version of Deligne's 'théorème du pgcd' (from the celebrated work [Del80]). This allows us to reduce the computation of  $P_1(T)$ <sup>4</sup> for X to the computation of the zeta function of smooth curves obtained by taking successive hyperplane sections of X, while the result for curves is proved in Theorem 1.1. **Algorithmic results.** We also give the first quantum polynomial-time algorithm (allowing the degree D to vary) to compute  $P_1(X/\mathbb{F}_q, T)$ , by applying Kedlaya's algorithm [Ked06] for the curve case.

**Theorem 1.3** (Quantum). There exists a quantum algorithm that computes  $P_1(X/\mathbb{F}_q, T)$  in time polynomial in  $D \log q$ , for any  $X_0$  given as above.

For varieties of constant degree D, by our reduction to the case of curves and applying work of Pila [Pil90] and Huang-Ierardi [HI98], we have the following.

**Theorem 1.4** (Fixed D). There exists a classical randomised algorithm that, given  $X_0$  as above of fixed degree D, computes  $P_1(X/\mathbb{F}_q,T)$  in time polynomial in  $\log q$ .

A major obstacle to computing the above was the lack of a concise and explicit representation, in general, for the étale cohomology group  $H^1(X, \mu_{\ell})$ ; despite it being known to be isomorphic to the  $\ell$ -torsion in the Picard scheme of X. A priori, elements therein are a formal sum of codimension-1 subvarieties (modulo an equivalence relation), and it is uncertain how one may directly produce  $\ell$ -torsion elements due to the highly *non-explicit* nature of the group law. There has been a strategy laid out by Levrat [Lev22, IV.3.5, VI.4] for surfaces, under some strongly restrictive hypotheses; but the general-case complexity is either unclear or exponential-time (see also [Lev23, §5]).

*Remark.* In Theorems 1.3 and 1.4, the stated runtimes are bounded by polynomial functions of the data as claimed, whose degree and coefficients are independent of  $D \cdot \log q$  for Theorem 1.3 and  $\log q$  for Theorem 1.4.

1.4. New techniques and proof ideas. Certifying the zeta function of a smooth curve  $C/\mathbb{F}_q$  of genus g boils down to a certification of the group orders  $\#\operatorname{Jac}(C)(\mathbb{F}_{q^j})$  for  $1 \leq j \leq 2g$ . The addition law on the Jacobian can be made explicit (after reducing to a plane model) by an effective Riemann-Roch algorithm (Algorithm 1). Utilising the additive structure of  $\operatorname{Jac}(C)(\mathbb{F}_q) \simeq \mathbb{Z}/n_1 \times \ldots \times \mathbb{Z}/n_r$ (with  $n_i|n_{i+1}$  and  $r \leq 2g$ ) as an abelian group, it suffices to certify that a candidate generating-set of divisors  $(D_i)_{1\leq i\leq r}$ , with each  $D_i$  of order  $n_i$ , actually generates the full group. Using the Weil bound for the size of the Jacobian, we are able to certify, with high probability, the 'independence' of the divisors  $D_i$  (Algorithm 2). This is done by random sampling in a family of hash functions a classical technique that originated in the famous protocol of Goldwasser-Sipser [GS86] to certify the lower bound of a, possibly exponential-size, set. This addresses Theorem 1.1.

<sup>&</sup>lt;sup>2</sup>the dimension of the embedding space, N, is considered to be fixed (say, N = 2,3 for n = 1,2 respectively), as we are primarily interested in *curves* and *surfaces* in this work. The degree D and field size q are allowed to vary.

<sup>&</sup>lt;sup>3</sup>the complexity can be measured in D or d, as for N fixed, each is bounded by a polynomial function in the other. <sup>4</sup>we write  $P_1(X/\mathbb{F}_q, T)$  to specify the q-power Frobenius.

More generally for smooth, projective varieties X, the theory of étale cohomology, in particular the Kummer sequence, allows us to relate the group  $\mathrm{H}^1(X, \mathbb{Z}/\ell\mathbb{Z}) \simeq \mathrm{Pic}^0(X)[\ell]^{\vee}$  to the  $\ell$ -torsion in the Picard scheme of X. Define the  $\ell$ -adic versions of the cohomology,

$$\mathrm{H}^{1}(X,\mathbb{Z}_{\ell}) := \lim_{\leftarrow j} \mathrm{H}^{1}(X,\mathbb{Z}/\ell^{j}\mathbb{Z}) \text{ and } \mathrm{H}^{1}(X,\mathbb{Q}_{\ell}) := \mathrm{H}^{1}(X,\mathbb{Z}_{\ell}) \otimes \mathbb{Q}_{\ell}.$$

By an application of the weak-Lefschetz theorem (Theorem 4.3), we notice that to compute  $P_1(X/\mathbb{F}_q, T)$ , it is sufficient to compute  $P_1(Y/\mathbb{F}_q, T)$  where Y is a smooth projective surface obtained by successively taking smooth hyperplane sections of X. By Algorithm 3, we produce a Lefschetz pencil of hyperplane sections on Y, denoted  $(H_t)_{t\in\mathbb{P}^1}$ , with  $Y_t := H_t \cap Y$  being smooth curves, for t in an open dense subscheme  $U_0 \subseteq \mathbb{P}^1$ . Denote  $U := U_0 \times \overline{\mathbb{F}}_q$ .

This procedure gives us (implicitly) a morphism<sup>5</sup>  $\pi : \tilde{Y} \to \mathbb{P}^1$ , whose fibre at any  $t \in \mathbb{P}^1$  is  $Y_t$ . By the Leray spectral sequence, we have

$$\mathrm{H}^{1}(Y, \mathbb{Q}_{\ell}) \simeq \mathrm{H}^{1}(\tilde{Y}, \mathbb{Q}_{\ell}) \simeq \mathrm{H}^{0}(\mathbb{P}^{1}, \mathcal{F}),$$

where  $\mathcal{F} := R^1 \pi_* \mathbb{Q}_\ell$  is the étale sheaf, on the projective line, obtained by the first direct image relative to  $\pi$ . Further, by the proper base-change theorem, we have for any  $t \in \mathbb{P}^1$ , the stalk  $\mathcal{F}_t \simeq \mathrm{H}^1(Y_t, \mathbb{Q}_\ell)$ . We notice that  $\mathcal{F}|_U$  is a locally constant sheaf on U and has as a subsheaf  $\mathcal{E} \subset \mathcal{F}|_U$ , the sheaf of vanishing cycles. The sheaf  $\mathcal{E}$  is locally constant and of rank (say) 2r.

We prove an effective version (Theorem 4.7) of Deligne's 'théorème du pgcd' ("polynomial gcd theorem" from the celebrated work [Del80]). In particular, we show that there exists an extension  $\mathbb{F}_Q/\mathbb{F}_q$  of bounded degree such that we can recover (with high probability)  $P_1(Y/\mathbb{F}_Q, T)$ , merely from the curve-case polynomials  $P_1(Y_{u_i}/\mathbb{F}_Q, T)$  with  $u_i \in U(\mathbb{F}_Q)$  chosen randomly, for  $1 \leq i \leq 2$ ; by computing their gcd. The Theorems 1.2, 1.3 and 1.4 follow from this. The ingredients are as follows. The hard-Lefschetz theorem (Theorem 3.2) states

$$\mathrm{H}^{1}(Y_{u}, \mathbb{Q}_{\ell}) = \mathrm{H}^{1}(Y, \mathbb{Q}_{\ell}) \oplus \mathcal{E}_{u}$$

for  $u \in U$ . We proceed to understand the action of the Frobenius at u on  $\mathcal{E}_u$ , which for our purposes behaves as a 'random group' contribution. The sheaf  $\mathcal{E}_{\mathbb{Z}_\ell} \subset R^1 \pi_* \mathbb{Z}_\ell |_U$  of  $\ell$ -adic integral vanishing cycles on U corresponds to a representation of the étale fundamental group  $\rho : \pi_1(U_0, u) \to$  $\operatorname{GL}(2r, \mathbb{Z}_\ell)$  via its action on the stalk of  $\mathcal{E}_{\mathbb{Z}_\ell}$  at u. We next study the geometric mod- $\ell$  monodromy  $\overline{\rho}_\ell : \pi_1(U, u) \to \operatorname{GL}(2r, \mathbb{F}_\ell)$ . Methods of Hall [Hal08] imply that  $\operatorname{im}(\overline{\rho}_\ell) = \operatorname{Sp}(2r, \mathbb{F}_\ell)$ , the symplectic group, when  $\ell$  is such that the  $\operatorname{H}^i(Y, \mathbb{Z}_\ell)$  are all torsion-free. An equidistribution theorem due to Katz dictates the proportion of Frobenius elements  $F_{Q,v} \in \pi_1(U_0, u)$  for  $v \in U(\mathbb{F}_Q)$ , whose image lies in a conjugacy-stable subset of the mod- $\ell$  arithmetic monodromy group. The error term therein, and an analysis of the proportion of matrices in the group of symplectic similitudes  $\operatorname{GSp}(2r, \mathbb{F}_\ell)$ with characteristic polynomial coprime to a given one; provide the reasonable bounds for  $\ell$  and Qto obtain our computational complexity result. The underlying torsion-bounds employ the work of Kweon [Kwe21], along with our good-reduction assumption for  $\mathcal{X}_0$  at the prime  $\mathfrak{p}$ .

#### 2. Zeta function of curves

In this section, we present an AM  $\cap$  coAM protocol for certifying the zeta function of a curve  $C/\mathbb{F}_q$ . We assume the input to be a smooth, projective, absolutely irreducible curve  $C_0 \subset \mathbb{P}^N$  of genus g > 0 and degree  $\delta$ , presented as a system of homogeneous polynomials  $f_1, \ldots, f_m$  with coefficients in  $\mathbb{F}_q$  and of degree  $\leq d$ . Denote by C the base change to the algebraic closure  $\mathbb{F}_q$ . We begin with the preliminary subsections 2.1 and 2.2 consisting of standard material. The AM $\cap$ coAM protocol of Theorem 1.1 and its proof is presented in 2.3.

 $<sup>{}^{5}\</sup>tilde{Y}$  is a smooth projective surface obtained by blowing up Y along  $\Delta \cap Y$ , where  $\Delta$  is the axis of the pencil.

2.1. **Preliminaries.** A divisor D on C is a formal sum  $D = \sum_{i=1}^{r} n_i P_i$ , where  $P_i \in C(\overline{\mathbb{F}}_q)$  and  $n_i \in \mathbb{Z} \setminus \{0\}$ . The set of points  $P_i$  occurring in the sum above is called the *support* of D. The sum  $\sum_i n_i$  is called the *degree* of D.

We denote the group of divisors by Div(C) and the subgroup of degree zero divisors by  $\text{Div}^0(C)$ . Let K denote the function field of C. We have a map div :  $K^* \hookrightarrow \text{Div}^0(C)$ , sending a function to the sum of its zeros and poles. The image of this map is called the subgroup of *principal divisors*, denoted  $\text{Div}^{\text{pr}}(C)$ . We call a divisor D effective, if  $n_i \ge 0$  for all i, which we denote by  $D \ge 0$ .

**Definition 2.1.** There exists an abelian variety (of dimension g) called the *Jacobian*, denoted Jac(C), whose  $\overline{\mathbb{F}}_q$ -rational points correspond to elements of the quotient group  $\text{Div}^0(C)/\text{Div}^{\text{pr}}(C)$ .

Let D be a divisor on C. We recall the *Riemann-Roch space* of D.

$$\mathcal{L}(D) := \{ f \in K^* \mid \operatorname{div}(f) + D \ge 0 \} \cup \{ 0 \}$$

Further, denoting by  $K_C$  the canonical divisor of C, the Riemann-Roch theorem states

$$\dim \mathcal{L}(D) = \deg(D) + 1 - g + \dim \mathcal{L}(K_C - D)$$

Addition on the Jacobian is performed by using an effective Riemann-Roch theorem. However, in order to invoke algorithms [HI94, ABCL02] computing the Riemann-Roch spaces, we first reduce our curve to a *planar* model.

In particular, we seek to find a curve  $C' \subset \mathbb{P}^2$  birational to C, given by a homogeneous form F. A singular point  $P \in C'$  is said to be a *node* if it is an ordinary double point, i.e., has multiplicity two, with distinct tangents. A curve is *nodal* if all its singularities are nodes. We recall [Har13, IV.3.11].

**Lemma 2.2** (Planar model). Let  $C \subset \mathbb{P}^N$  be as above. There is a randomised algorithm that computes a nodal curve  $C' \subset \mathbb{P}^2$  and a birational morphism  $\phi : C \to C'$  that runs in time polynomial in  $g \log q$ .

Proof. We describe how to obtain an equation defining C' algorithmically. The key idea is to choose a random point  $O \in \mathbb{P}^N$ , with  $O \notin C$ , and project C onto a hyperplane from O. For generic O (lying outside any secant or tangent of C) and  $N \geq 4$ , the resulting map is an embedding. Repeating the process, we get a sequence of morphisms  $C \to \mathbb{P}^{N-1} \to \cdots \to \mathbb{P}^3$ . The locus of 'bad' projections forms a subvariety of  $\mathbb{P}^3$  of dimension at most 2, with degree bounded by a polynomial in  $\delta := \deg(C)$ . Hence, this locus can be avoided with high probability at the cost of a field extension of degree at worst poly( $\delta$ ). Therefore, generically, by [Har13, Theorem V.3.10] for  $O \in \mathbb{P}^3$ , the image of the projection of C from O onto  $\mathbb{P}^2$  has at worst nodal singularities. Denote by  $\phi : C \to \mathbb{P}^2$ the composite morphism of all projections. It is a birational morphism with  $\deg(\phi(C)) \leq \delta$ . Therefore, the polynomial F cutting out C' in  $\mathbb{P}^2$  has total degree at most  $\delta$ . Writing the linear projection  $\phi$  explicitly and computing the image of  $\Theta(\delta^2)$  many points  $P_i \in C$ , we can recover Fby a bivariate interpolation algorithm. Points on the curve can be sampled by the procedure below.

Sampling points in  $C(\mathbb{F}_q)$  (which exist after an extension) can be achieved in randomised polynomial time as follows. Consider an affine piece of C in  $\mathbb{A}^N$  (with coordinates  $(y_1, \ldots, y_N)$ ) by taking the complement of a hyperplane. Fixing a value of  $y_1$  amounts to intersecting with a hyperplane in  $\mathbb{A}^N$ , giving a finite set of points. The Weil bound (see Theorem 2.3 below) for C guarantees that with high probability, after  $2g \leq 4\delta^2$  fixings of  $y_1$  in  $\mathbb{F}_q$ , the resulting zero-dimensional system has  $\mathbb{F}_q$ -rational points. Extracting them can be done in randomised polynomial-time by using the main result of [LL91] for the  $\mathbb{F}_q$ -root-finding of a zero-dimensional N-variate system.

We conclude this subsection with a statement of the Weil-Riemann hypothesis for curves [Wei48a, Wei48b].

**Theorem 2.3** (Weil).  $|\#C(\mathbb{F}_q) - (q+1)| \le 2g\sqrt{q}$ .

2.2. Jacobian arithmetic. Recall the standard results showing that elements of  $\text{Jac}(C)(\mathbb{F}_q)$  can be presented concisely and divisor arithmetic therein can be performed efficiently. We know by [Ser88, §8] that C injects into its Jacobian, by the choice of a rational point, which we call  $\infty$ .

**Lemma 2.4** (Reduced form). Given  $D \in \text{Jac}(C)$ ,  $\exists 0 \leq i \leq g$  and a unique effective divisor E of degree g - i such that  $D = E - (g - i)(\infty)$  in  $\text{Jac}(C)(\overline{\mathbb{F}}_g)$ .

Proof. By the Riemann-Roch theorem, we have dim  $\mathcal{L}(D+g(\infty)) = 1 + \dim \mathcal{L}(K_C - D - g(\infty)) > 0$ . Iteratively, subtracting  $\infty$  from the divisor  $D + g(\infty)$ , we choose the largest  $0 \leq i \leq g$  so that  $\dim \mathcal{L}(D+(g-i)(\infty))$  is still positive. In particular, for such an i, we have  $\dim \mathcal{L}(D+(g-i)(\infty)) = 1$ . Thus, one gets a 'unique' (up to a constant) rational function f in the basis of  $\mathcal{L}(D + (g - i\infty))$ . Therefore, one obtains a unique effective divisor  $E := \operatorname{div}(f) + D + (g - i)(\infty) \geq 0$ , which is the same as saying  $D = E - (g - i)(\infty)$  in the arithmetic of  $\operatorname{Jac}(C)(\overline{\mathbb{F}}_q)$ .

We recall next a method to compute bases of Riemann-Roch spaces.

**Proposition 2.5** (Riemann-Roch basis). Let D be a divisor on a curve C of degree and support-size  $\leq \delta$ . A basis of the Riemann-Roch space  $\mathcal{L}(D)$  can be computed efficiently in  $O(\delta^{12} \log q)$  time.

Proof. After computing a plane model  $\phi : C \to C' \subset \mathbb{P}^2$  one uses [HI94, §2] to compute the Riemann-Roch space of a divisor on the normalisation of C' (which is isomorphic to C). While [HI94] requires the singular points of C' to lie over the base field (essentially to ensure an efficient resolution of singularities), this can be bypassed by using [Koz94] instead. The complexity follows from [HI94, §2.5]. This strategy was also utilised in the algorithm of [Ked06, §6] as a preprocessing step to do basic arithmetic in the class group (=Jac(C)).

Using Proposition 2.5, we can now check when a divisor of degree zero is trivial in the Jacobian. Recall that for  $D \in \text{Div}^{0}(C)$ , we have dim  $\mathcal{L}(D) = 1$  if and only if  $D \in \text{Div}^{\text{pr}}(C)$ . This implies the following.

**Lemma 2.6** (Zero test). Given  $D \in \text{Div}^{0}(C)$ , whether  $D \in \text{Div}^{\text{pr}}(C)$  is testable in polynomial time. In other words zero-tests in Jac(C) can be performed in polynomial time.

Combining Lemma 2.4, Proposition 2.5 and Lemma 2.6, one obtains a polynomial time algorithm to put a given divisor  $D \in \text{Jac}(C)(\mathbb{F}_q)$  into reduced form. Indeed by Lemma 2.4, one knows that the support of D can be chosen to be of size at most poly(g). Then, Proposition 2.5 can be applied to obtain the effective divisor E and the integer i, so that  $D = E - (g - i)\infty$  is in reduced form as an element of Jac(C).

*Remark.* The points occurring in the support of the effective divisor E associated to the reduced form of D in the above description each lie in a poly(g) extension of  $\mathbb{F}_q$ . However, one never needs to go to an extension of  $\mathbb{F}_q$  containing all of them simultaneously (which may be exponentially large in degree). The issue is handled exactly the same way in [Ked06, §6]. See also [HI94, §3] for more on this *implicit representation* of divisors used in their algorithm to do Jacobian arithmetic.

We are now ready to describe a randomised polynomial-time Algorithm 1 to compute the sum of two elements in Jac(C) in the canonical representation described above.

2.3. AM **protocol.** In this subsection, we present an AM $\cap$ coAM protocol to certify the order (and group structure) of Jac(C)( $\mathbb{F}_q$ ). We then show how to certify the zeta function of C using this. We first recall a result of Weil [Wei48a, pp.70-71] which generalises a theorem of Hasse [Has36, p.206] from elliptic curves (g = 1) to abelian varieties ( $g \ge 1$ ).

**Proposition 2.7** (Hasse-Weil interval). For an abelian variety A of dimension g over the finite field  $\mathbb{F}_q$ , the number of  $\mathbb{F}_q$ -rational points is in the following range:

$$(\sqrt{q}-1)^{2g} \leq #A(\mathbb{F}_q) \leq (\sqrt{q}+1)^{2g}.$$

### Algorithm 1 Adding two points on the Jacobian

- Input: Two divisors  $D_1 = E_1 m_1(\infty)$  and  $D_2 = E_2 m_2(\infty)$  of degree zero, with  $m_1, m_2 \leq g$  lying in the Jacobian of a smooth projective curve  $C/\mathbb{F}_q$ , presented in the reduced form as per Lemma 2.4.
- Output:  $D_3 = D_1 + D_2$  as  $D_3 = E_3 m_3(\infty)$  where  $E_3$  is effective of degree  $m_3$ .
- (Reduction loop) For each i, compute L(D<sub>1</sub> + D<sub>2</sub> + (g − i)(∞)) using Proposition 2.5, starting from i = 0. If dim L(D<sub>1</sub> + D<sub>2</sub> + (g − i)(∞)) = 1 then we get a unique effective divisor E := div(f) + D<sub>1</sub> + D<sub>2</sub> + (g − i)(∞), where the representation of div(f) can be found in randomised polynomial-time [LL91]. Choose the largest such i and set m<sub>3</sub> = g − i and E<sub>3</sub> = E.
   Output E<sub>3</sub> − m<sub>3</sub>(∞).

**Reduced gap.** Given an input curve of genus g we want to choose q so that the above gap is small enough, namely,  $((\sqrt{q}+1)/(\sqrt{q}-1))^{2g} < 2$ . In particular, we require

$$1 + \frac{2}{\sqrt{q} - 1} < 2^{1/2g} = \exp\left(\frac{\log 2}{2g}\right) = 1 + \frac{\log 2}{2g} + \frac{(\log 2)^2}{8g^2} + \dots$$

Truncating, we notice that  $q > (8g+1)^2$  suffices.

(

**Hash functions** are pseudorandom maps from large strings to small strings, in a way that minimizes *collision* as much as possible. Let  $h : \{0,1\}^n \mapsto \{0,1\}^k$ ;  $k \ll n$  be from a hash family. We require that for  $X \in \{0,1\}^n$  and a random  $Y \in \{0,1\}^k$ ,  $\Pr_{h,Y}[h(X) = Y] = 1/2^k$ . One can show that, for a random  $k \times n$  matrix A over  $\mathbb{F}_2$ , and a random vector  $b \in \{0,1\}^k$ ,  $h : X \mapsto AX + b$  satisfies this property (see [AB09, Theorem 8.15]). Using this concept, Algorithm 2 is the AM  $\cap$  coAM protocol to verify the Jacobian size, over  $\mathbb{F}_Q \supseteq \mathbb{F}_q$  assuming  $Q > (8g+1)^2$ .

**Lemma 2.8** (Probability of Algorithm 2). In Algorithm 2 (given candidate  $\mathcal{N}$ ), if  $\# \operatorname{Jac}(C)(\mathbb{F}_Q) = \mathcal{N}$ , then Arthur accepts with probability > 2/3. Else, Arthur rejects with probability > 2/3.

Proof. We adapt the protocol from [AB09, §9.4]. Let  $S \subset \{0,1\}^{2g \log Q}$  denote the set  $\operatorname{Jac}(C)(\mathbb{F}_Q)$  with the elements written as binary strings. Let  $\mathcal{G}$  be the group generated by the divisors  $D_i$ 's that Merlin provided. Suppose it has size  $\mathcal{N}$ , as Merlin claimed. In particular,  $\mathcal{G} = S$  as we have made the Hasse-Weil 'gap' small enough so that only a unique multiple of  $\mathcal{N}$  can lie in that interval. For a random  $y \in \{0,1\}^{L+1}$  and a random hash function h (chosen from a uniform distribution over matrices A and vectors b such that  $h: x \mapsto Ax + b$ ), the probability that there is an  $x \in \mathcal{G} = S$ , such that h(x) = y is

Pr[
$$\exists x \in \mathcal{G} = S, \ h(x) = y$$
]  $\geq \binom{\#S}{1} \cdot \frac{1}{2^{L+1}} - \binom{\#S}{2} \cdot \frac{1}{2^{2(L+1)}} > \frac{\#S}{2^{L+1}} - \frac{(\#S)^2}{2^{2(L+1)+1}}$   
>  $\frac{\#S}{2^{L+1}} \left(1 - \frac{\#S}{2^{L+2}}\right) \geq 0.75 \cdot \frac{\#S}{2^{L+1}}$ .

from the inclusion-exclusion-principle, and applying the inequality  $2^{L-1} < \#S = \mathcal{N} \leq 2^L$ .

Conversely, suppose  $\#S \neq \mathcal{N}$ , as Merlin bluffed (so,  $\mathcal{G} \neq S$ ). Since Arthur checked that the product of the orders of the divisors  $D_i$ 's equals  $\mathcal{N}$ , we deduce that  $\#\mathcal{G} \leq \#S/2$  (as the order of the subgroup  $\mathcal{G}$  properly divides that of the group S). So, simply by the union-bound we get

(2.2) 
$$\Pr[\exists x \in \mathcal{G}, \ h(x) = y] \le \binom{\#\mathcal{G}}{1} \cdot \frac{1}{2^{L+1}} \le 0.5 \cdot \frac{\#S}{2^{L+1}}$$

Thus, Eqns.2.1-2.2 have a noticeable difference in the probability estimate. Now, we can repeat, with Arthur choosing several (h, y) pairs, take the 'majority vote', and use the Chernoff bound [AB09, §7.4.1]. This *amplification* trick brings the probabilities above 2/3 (in Eqn.2.1) and below 1/3 (in Eqn.2.2) respectively. The number of repetitions will be inverse-polynomial in  $\#S/2^{L+1} > 1/4$ ; which is only a constant blowup in our time complexity.

## **Algorithm 2** Verifying the size and structure of the Jacobian of $C/\mathbb{F}_Q$

**Input:** A smooth projective curve  $C \subset \mathbb{P}^N$  of genus g and degree  $\delta$ , given by polynomials  $(f_i)_{1 \leq i \leq m}$ . A candidate integer  $\mathcal{N}$  lying in the Hasse-Weil interval. Set L:  $2^{L-1} < \mathcal{N} \leq 2^L$ .

- 1: Arthur: Choose a random hash function  $h : \{0,1\}^{2g \log Q} \to \{0,1\}^{L+1}$  by picking a matrix A and a vector b randomly as stated above. Pick a random  $y \in \{0,1\}^{L+1}$  and send (h, y) to Merlin as a *challenge*. Note: Arthur could send O(L) many such independently chosen pairs (h, y) to reduce the error probability exponentially. Below, we use only one pair for the simplicity of exposition.
- 2: Merlin:
  - Pick r generators  $D_i \in \text{Jac}(C)(\mathbb{F}_Q)$   $(i \in [r])$  such that

$$\operatorname{Jac}(C)(\mathbb{F}_Q) \simeq \langle D_1 \rangle \times \ldots \times \langle D_r \rangle$$

with each  $D_i$  of order  $n_i$ , with  $n_i|n_{i+1}$  and  $\prod_{i=1}^r n_i = \mathcal{N}$ . Each  $D_i$  is presented in canonical form as  $D_i = E_i - m_i(\infty)$ , with  $E_i$  effective of degree  $m_i$ . The divisors  $E_i$  in turn are presented as a sum of  $\overline{\mathbb{F}}_Q$  – rational points of C, each defined over an extension of  $\mathbb{F}_Q$  of degree at most poly(g) thanks to Lemma 2.4.

- Send a response consisting of r quadruples  $\{(c_i, D_i, n_i, P_i)\}_{1 \le i \le r}$  with the claim that the divisor  $\sum_i c_i D_i =: x$ , for  $c_i \in \mathbb{Z}/n_i\mathbb{Z}$ , satisfies h(x) = y. Every  $P_i$  is a set of pairs: each consisting of a prime factor of  $n_i$  and the corresponding exponent in its factorisation.
- 3: Arthur:
  - Check whether  $D_i$  indeed represents a point in  $\text{Jac}(C)(\mathbb{F}_Q)$ . This is done by evaluating the Frobenius  $F_Q$  on  $D_i = E_i m_i(\infty)$  and checking for invariance. If not, *Reject*.
  - Check the factorization data  $P_i$  of each  $n_i$ . Check the order  $n_i$  as follows: verify  $n_i D_i = 0$ , and for each distinct prime factor  $p_{i,j}$  of  $n_i$ , verify  $(n_i/p_{i,j})D_i \neq 0$ . Check that  $\mathcal{N} = \prod_{i=1}^r n_i$ . If a check fails, *Reject*. Calculate  $x = \sum_i c_i D_i$ .
  - Check  $h(x) = h(\sum_i c_i D_i) = y$ , if yes then *Accept*; otherwise *Reject*. All the checks can be easily performed by Arthur using: basic arithmetic, or Algorithm 1, combined with the standard trick of repeated-doubling.

*Remark.* The steps of Merlin require exponential resources (i.e. Step 2), so we do not know how to compute them in polynomial-time in practice. The purpose is to only provide a concise certificate, using which Arthur can verify the Jacobian-size efficiently and reliably (with high probability).

**Lemma 2.9** (Complexity of Algorithm 2). Arthur's verification algorithm runs in randomised polynomial-time.

*Proof.* Step 1 simply involves addition and multiplication of binary matrices of size  $poly(g \log Q)$ , so can be accomplished in  $poly(g \log q)$  time. In Step 3, since the number of prime factors of any integer n is  $O(\log n)$ , the prime factor checking computation can be performed in  $poly(\log N)$  time. Applying the Hasse-Weil bound, this is in fact  $poly(g \log q)$  time. For the Jacobian arithmetic, Arthur uses Algorithm 1 and repeated-doubling. This sums up the complexity of our protocol to  $poly(g, \log q)$ -time.

The zeta function is intimately connected to the order of the Jacobian. From [Ked06, §8]:

**Lemma 2.10** (Count to zeta function). Assume we are given  $\#\operatorname{Jac}(C)(\mathbb{F}_{q^j})$ , for every  $1 \leq j \leq \max(18, 2g)$ . Then,  $P_1(C/\mathbb{F}_q, T)$  can be reconstructed from these counts, in  $\operatorname{poly}(g \log q)$ -time.

Kedlaya [Ked06, §8] also shows the following, connecting the zeta function of a larger Frobenius to that of a *smaller* Frobenius.

**Lemma 2.11** (Base zeta function). Let primes  $m_1, m_2$  with  $m_1 < m_2$ , be such that  $m_j - 1$  is divisible by a prime greater than 2g, for  $j \in \{1, 2\}$ . Assume further that  $q^{m_1} > (8g+1)^2$ . Then,  $P_1(C/\mathbb{F}_q, T)$  can be recovered from  $P_1(C/\mathbb{F}_q^{m_j}, T)$ ,  $j \in \{1, 2\}$ , in time polynomial in  $g \log q$ .

Further, the existence of such  $m_1, m_2$  bounded by a polynomial in  $g \log q$  is guaranteed.<sup>6</sup>

Proof of Theorem 1.1. Using Algorithm 2, we can verify the structure of  $\operatorname{Jac}(C)(\mathbb{F}_Q)$  for any  $Q > (8g+1)^2$ . This implies  $P_1(C/\mathbb{F}_q,T)$  can be certified by first certifying  $P_1(C/\mathbb{F}_{q^{m_1}},T)$  and  $P_1(C/\mathbb{F}_{q^{m_2}},T)$  and next applying Lemma 2.11. Each  $P_1(C/\mathbb{F}_{q^{m_j}},T)$  can be computed, uniquely, using the counts  $\#\operatorname{Jac}(C)(\mathbb{F}_{q^{im_j}})$ , for  $1 \leq i \leq \max(18, 2g)$ , by Lemma 2.10. This completes the proof of the first part of the theorem, verifying the zeta function.

Group structure. In the second part of the theorem statement, suppose a candidate G has been provided via additive generators  $\{A_1, \ldots, A_r\}$ , with each  $A_i$  of order  $n_i$  such that G decomposes as a direct sum of the subgroups  $\langle A_i \rangle$ , where  $n_i \mid n_{i+1}$ . We need to verify whether  $\operatorname{Jac}(C)(\mathbb{F}_q) \simeq G$ . For this, Merlin first convinces Arthur of the structure of  $\operatorname{Jac}(C)(\mathbb{F}_Q)$ , and provides the additive generators for  $Q > (8g + 1)^2$ . Using this, Arthur can compute  $P_1(C/\mathbb{F}_q, T)$ , thereby obtaining the count  $\#\operatorname{Jac}(C)(\mathbb{F}_q) = P_1(C/\mathbb{F}_q, 1)$ . For the subgroup  $\operatorname{Jac}(C)(\mathbb{F}_q) \subset \operatorname{Jac}(C)(\mathbb{F}_Q)$ , Merlin presents divisors  $D_i$  with support in  $C(\mathbb{F}_Q)$ , that are candidates corresponding to each  $A_i$ . Arthur first checks whether the  $D_i$  all belong to  $\operatorname{Jac}(C)(\mathbb{F}_q)$  (by evaluating the q-Frobenius on them and verifying invariance). Next, Arthur verifies the independence of the  $D_i$  as in Algorithm 2. This provides a lower bound for #G. Comparing it with the verified count  $\#\operatorname{Jac}(C)(\mathbb{F}_q)$  certifies the structure. The proof then follows from Lemmas 2.8-2.9.

#### 3. Surfaces: Vanishing cycles, monodromy, and equidistribution

This section is devoted to the technical background necessary to prove our main theorems in the higher dimensional case. In particular, 3.1 reviews the theory of vanishing cycles on a surface, including a statement of the hard-Lefschetz theorem in this case and the general 'gcd theorem' of Deligne. Next, in 3.2, the Picard-Lefschetz formulas and the  $(\ell$ -adic and mod- $\ell$ ) monodromy of a Lefschetz pencil of hyperplane sections on a surface are discussed including torsion bounds, followed by the statement of an equidistribution result of Katz. Finally in 3.3, we briefly review symplectic groups over finite fields and deduce a probability estimate that we later use to prove an effective version of Deligne's gcd theorem.

3.1. Vanishing cycles. In this subsection, we give a brief overview of the theory of vanishing cycles associated to a surface fibred as a Lefschetz pencil over  $\mathbb{P}^1$ . Then, we discuss the 'hard-Lefschetz theorem' and some implications for the first étale cohomology. Finally, we wrap with a statement of Deligne's 'théorème du pgcd', which enables us to recover the characteristic polynomial of Frobenius, acting on the first cohomology, from its action on the cohomology of the fibres.

Let  $X_0$  be a smooth, projective, geometrically irreducible surface over the finite field  $\mathbb{F}_q$  of characteristic p > 0. Denote by X the base change to the algebraic closure. Assume we have a Lefschetz fibration  $\pi : \tilde{X} \to \mathbb{P}^1$  following Algorithm 3. As usual, we let  $Z \subset \mathbb{P}^1$  denote the set giving rise to singular fibres (nodal curves), and let U denote its complement. Let  $X_\eta$  be the generic fibre of  $\pi$ . It is a smooth curve of genus g over the function field of  $\mathbb{P}^1$ .

Let  $\ell$  be an odd prime, coprime to p. Consider the sheaf  $\mathcal{F}^{\ell} := R^1 \pi_{\star} \mu_{\ell}$  on  $\mathbb{P}^1$ . By the proper basechange theorem, we have that its stalk at a point  $u \to \mathbb{P}^1$  is the group  $\mathrm{H}^1(X_u, \mu_{\ell}) \simeq \mathrm{Pic}^0(X_u)[\ell]$ . Further, we know that  $\mathcal{F}^{\ell}|_U$  is a locally constant sheaf of rank 2g on U. We seek to understand the behaviour of  $\mathcal{F}^{\ell}$  at points  $z \in Z$ . Let  $X'_z \to X_z$  be the normalisation (which has genus g-1) of such a singular fibre, and denote by  $V_z$ , the kernel of the map  $\mathcal{F}^{\ell}_z \to \mathrm{Pic}^0(X'_z)[\ell]$ . We call  $V_z$  the group of vanishing cycles at z. We now recall a collection of results from [Mil80, V.3].

 $<sup>^{6}</sup>$ by [Har05, Theorem 1.2]

**Proposition 3.1.** With the above setup, the following are true:

- For any  $u \in \mathbb{P}^1$  there exists a cospecialisation map  $\mathcal{F}_u^\ell \to \mathcal{F}_\eta^\ell$  which is an isomorphism if and only if  $u \in U$ .
- If  $z \in Z$ , the cospecialisation map  $\mathcal{F}_z^{\ell} \to \mathcal{F}_{\eta}^{\ell}$  is an injection. In particular  $\mathcal{F}_z^{\ell} \simeq (\mathbb{Z}/\ell\mathbb{Z})^{2g-1}$ . Further,  $V_z$  is the exact annihilator of  $\mathcal{F}_z^{\ell}$  under the Weil-pairing map

$$\langle \cdot, \cdot 
angle : \mathcal{F}^{\ell}_{\eta} imes \mathcal{F}^{\ell}_{\eta} \longrightarrow \mu_{\ell}(\overline{\mathbb{F}}_q) \,.$$

•  $\mathcal{F}^{\ell}$  is tamely ramified at all  $z \in Z$ .

In particular, for  $z \in Z$ , we have  $V_z \simeq \mathbb{Z}/\ell\mathbb{Z}^{-7}$ , and we denote by  $\delta_z$ , the element that maps to 1. We may also identify  $\delta_z$  with its image under the map  $\mathcal{F}_z^{\ell} \to \mathcal{F}_\eta^{\ell}$ , and call  $\mathcal{E}^{\ell}(X_\eta)$  the subspace generated by all the  $\delta_{z_i}$  in  $\mathcal{F}_\eta^{\ell}$  for  $z_i \in Z$ . By the cospecialisation map, we refer to the corresponding subspace generated in  $\mathcal{F}_u^{\ell}$  for  $u \in U$  by  $\mathcal{E}^{\ell}(X_u)$ . By passage to the limit and tensoring, we also obtain the  $\mathbb{Q}_{\ell}$ -vector space of vanishing cycles  $\mathcal{E}(X_u)$ . Moreover, there exists a locally constant subsheaf  $\mathcal{E} \subset R^1 \pi_* \mathbb{Q}_{\ell}|_U$ , called the *sheaf of vanishing cycles* with stalk  $\mathcal{E}_u = \mathcal{E}(X_u)$  for  $u \in U$ . We now recall the 'hard-Lefschetz' theorem for surfaces, which measures precisely the discrepancy between  $\mathrm{H}^1(X_u, \mathbb{Q}_{\ell})$  and  $\mathrm{H}^1(X, \mathbb{Q}_{\ell})$ .

Theorem 3.2 (Hard-Lefschetz). We have the decomposition

$$\mathrm{H}^{1}(X_{u}, \mathbb{Q}_{\ell}) \simeq \mathrm{H}^{1}(X, \mathbb{Q}_{\ell}) \oplus \mathcal{E}_{u}$$

with respect to the symplectic pairing. In particular,  $\mathrm{H}^1(X, \mathbb{Q}_\ell) \simeq \mathcal{E}_u^{\perp}$  when viewed as a subspace of  $\mathrm{H}^1(X_u, \mathbb{Q}_\ell)$  under the weak-Lefschetz map.

The general result is a deep theorem of Deligne [Del80, 4.3.9]. The surface case is easier to handle and is done in [Kle68, 2.A.10]. We conclude this subsection with a statement of Deligne's 'théorème du pgcd' [Del80, 4.5.1].

Let  $X_0/\mathbb{F}_q$  now be a smooth, projective, geometrically irreducible variety of dimension n and let X be the base change to the algebraic closure.

**Theorem 3.3** (Le théorème du pgcd). Let  $(X_t)_{t\in\mathbb{P}^1}$  be a Lefschetz pencil of hypersurface sections of degree  $d \geq 2$  on X. Then  $P_{n-1}(X/\mathbb{F}_q, T)$  is the least common multiple of all polynomials  $f(T) \in \mathbb{C}[T]$ , satisfying the condition that for any  $t \in \mathbb{F}_{q^r}$  such that  $X_t$  is smooth, the polynomial <sup>8</sup>  $f(T)^{(r)}$  divides  $P_{n-1}(X_t/\mathbb{F}_{q^r}, T)$ .

Deligne derived the above as a consequence of his proof of the Weil-Riemann hypothesis and hard-Lefschetz theorem for  $\ell$ -adic cohomology. Theorem 3.3 has been used by Katz-Messing in [KM74] to deduce the same facts for any Weil cohomology theory. The theorem was also used by Gabber in [Gab83] to show torsion-freeness of the integral  $\ell$ -adic cohomology for smooth projective varieties for 'almost all'  $\ell$ .

3.2. Monodromy and equidistribution. In this subsection, we introduce the notion of monodromy in a Lefschetz pencil. We then recall a big mod- $\ell$  monodromy result, obtained by an adaptation of work of Hall. Finally, we state a version of Deligne's equidistribution theorem [Del80, 3.5.3] due to Katz. As before, let  $\pi : \tilde{X} \to \mathbb{P}^1$  be a Lefschetz pencil of curves on a smooth, projective surface X. We denote by  $U_0 \subset \mathbb{P}^1$  the locus parameterising smooth fibres (of genus g) and by  $U = U_0 \times \overline{\mathbb{F}}_q$ . Let  $Z = \mathbb{P}^1 \setminus U$  be the finite set parameterising the critical fibres. Write  $\mathcal{F} = R^1 \pi_* \mathbb{Q}_\ell$ and  $\mathcal{F}^\ell = R^1 \pi_* \mu_\ell$  for the respective direct-image sheaves.

<sup>8</sup> if 
$$f(T) = \prod_{j} (1 - \alpha_j T)$$
, then  $f(T)^{(r)} := \prod_{j} (1 - \alpha_j^r T)$ .

<sup>&</sup>lt;sup>7</sup>we omit Tate-twists by fixing an isomorphism  $\mathbb{Z}/\ell\mathbb{Z} \simeq \mu_{\ell}(\overline{\mathbb{F}}_q)$  and choosing a generator for the group of roots of unity.

Let  $u \in U$  be a geometric point. The arithmetic étale fundamental group (see [Mur67] for the definition)  $\pi_1(U_0, u)$  acts on  $\mathcal{F}_u^{\ell}$  and by passage to the limit, on  $\mathcal{F}_u$ . This latter representation restricted to the geometric étale fundamental group  $\pi_1(U, u)$  is called the monodromy of the pencil. Since  $\mathcal{F}$  is tamely ramified, the action of  $\pi_1(U, u)$  factors through the tame fundamental group<sup>9</sup>  $\pi_1^t(U, u)$ . By a theorem of Grothendieck [Gro57, 182-27],  $\pi_1^t(U, u)$  is generated topologically by #Z elements  $\sigma_i$  for each  $z_i \in Z$ , satisfying the relation  $\prod_{i=1}^{\#Z} \sigma_i = 1$ . The Picard-Lefschetz formulas make this action explicit. See [Mil80, Ch V, Theorem 3.14] or [FK13, III.4.3] for a proof.

**Proposition 3.4** (Picard-Lefschetz formulas). For any  $\gamma \in \mathcal{F}_{u}^{\ell}$ , we have

(3.1) 
$$\sigma_i(\gamma) = \gamma - \epsilon_i \cdot \langle \gamma, \delta_{z_i} \rangle \cdot \delta_{z_i} ,$$

where for a uniformising parameter  $\theta_i$  at  $z_i$ , we have  $\sigma_i(\theta_i^{1/\ell}) = \epsilon_i \cdot \theta_i^{1/\ell}$ .

Clearly, the monodromy action respects the symplectic pairing. By the hard-Lefschetz theorem, we know that  $\mathrm{H}^1(X_u, \mathbb{Q}_\ell) \simeq \mathrm{H}^1(X, \mathbb{Q}_\ell) \oplus \mathcal{E}_u$ , with  $\mathrm{H}^1(X, \mathbb{Q}_\ell) = \mathcal{E}_u^{\perp}$ . In particular,  $\pi_1(U, u)$  acts trivially on  $\mathrm{H}^1(X, \mathbb{Q}_\ell)$ , implying that the monodromy action factors through  $\mathrm{Sp}(\mathcal{E}_u)$ , the group of symplectic transformations of the vector space  $\mathcal{E}_u$ . We know [Del74, 5.10] that the image of  $\pi_1(U, u)$ is open and Zariski-dense in  $\mathrm{Sp}(\mathcal{E}_u)$ . Further, by the conjugacy of vanishing cycles, we also know  $\pi_1(U, u)$  acts absolutely irreducibly on  $\mathcal{E}_u$ .

One seeks a version of the above to compute the mod- $\ell$  geometric monodromy for certain equidistribution estimates coming from Theorem 3.7. Consider the torsion-free sheaf  $R^1\pi_*\mathbb{Z}_\ell$  of rank 2g on U. It has as subsheaf, the sheaf of integral  $\ell$ -adic vanishing cycles  $\mathcal{E}_{\mathbb{Z}_\ell} \subset R^1\pi_*\mathbb{Z}_\ell$  of rank, say, 2r. This in turn, corresponds to representations  $\rho : \pi_1(U_0, u) \to \operatorname{GL}(2r, \mathbb{Z}_\ell)$  and  $\overline{\rho} = \rho | \pi_1(U, u)$ . Let  $\mathcal{V} := \mathcal{E}_{\mathbb{Z}_\ell} \otimes_{\mathbb{Z}_\ell} \mathbb{F}_\ell$  be the lisse  $\mathbb{F}_\ell$ -sheaf giving rise to, respectively, the mod- $\ell$  representations  $\rho_\ell$ and  $\overline{\rho}_\ell$ . There are multiple ways to show big mod- $\ell$  monodromy for 'almost all primes  $\ell$ ' (all but finitely many), but [Hal08, §4-6] gives a method that works for *every* prime  $\ell \geq 5$  invertible on the characteristic. However, the generic rank of the local system  $\mathcal{V}$  is a priori dependent on  $\ell$ , and guaranteed to be 2r only when the cohomology groups  $\operatorname{H}^i(X, \mathbb{Z}_\ell)$  are all torsion-free. The following result appears to be known to Hall and Katz ([Hal08, pg 5] and [Hal23]); for completeness, we provide a brief proof below.

**Theorem 3.5** (Big mod- $\ell$  monodromy). If  $H^i(X, \mathbb{Z}_{\ell})$  are all torsion-free, we have

$$G := \overline{\rho}_{\ell} \left( \pi_1(U, u) \right) = \operatorname{Sp}(2r, \mathbb{F}_{\ell}).$$

*Proof.* As the groups  $\mathrm{H}^{i}(X, \mathbb{Z}_{\ell})$  are all torsion-free, we know (analogously to the situation in [Kat11, Theorem 9.2]) that the hard-Lefschetz theorem holds with  $\mathbb{F}_{\ell}$  – coefficients, i.e., for a smooth hyperplane section  $X_{u}$  obtained as a fibre of our Lefschetz pencil, we have

$$\mathrm{H}^{1}(X_{u}, \mathbb{F}_{\ell}) \simeq \mathrm{H}^{1}(X, \mathbb{F}_{\ell}) \oplus \mathcal{V}_{u}.$$

We now show that the representation  $\mathcal{V}_u$  of  $\pi_1(U, u)$  is irreducible. Indeed if  $W \subset \mathcal{V}_u$  is a stable subspace, for any  $\gamma \neq 0 \in W$ , we must have  $\langle \gamma, \delta_j \rangle \neq 0$  for some j, as otherwise the Weil pairing would be degenerate on  $\mathrm{H}^1(X_u, \mathbb{F}_\ell)$ . Therefore, by (3.1), this implies  $\sigma_j(\gamma) - \gamma = \epsilon_j \cdot \langle \gamma, \delta_j \rangle \cdot \delta_j \in W$ . As the vanishing cycles are all conjugate under the action of  $\pi_1(U, u)$  [Kat73a, 6.6], this means  $\delta_i \in W$  for all i, or  $W = \mathcal{V}_u$ .

We then invoke a theorem of Hall [Hal08, Theorem 3.1], to conclude that the image is in fact the full symplectic group, due to the irreducibility and the transvections coming from the Picard-Lefschetz formulas.

<sup>&</sup>lt;sup>9</sup>essentially classifying finite étale covers of U tamely ramified over Z.

Let  $X_0/\mathbb{F}_q$  now be obtained via good reduction from a smooth, projective, geometrically irreducible surface  $\mathcal{X}_0$  over a number field K at a prime  $\mathfrak{p}$ . We assume  $\mathcal{X}_0 \subset \mathbb{P}^N$  is of degree D > 0 and given by the vanishing of homogeneous forms  $f_1, \ldots, f_m$  each of degree  $\leq d$ . Denote  $X := X_0 \times_{\mathbb{F}_q} \overline{\mathbb{F}}_q$ and  $\mathcal{X}^{\mathrm{an}} := \mathcal{X}_0 \times_K \mathbb{C}$ , equipped with the complex analytic topology. One has the following.

**Proposition 3.6.** There exists a prime  $\ell$  with  $(4D)^4 \leq \ell \leq 2^{11}D^{N^2}$ , coprime to q, such that  $\mathrm{H}^i(X,\mathbb{Z}_\ell)$  are all torsion-free for  $0 \leq i \leq 4$ .

*Proof.* Since  $\mathcal{X}_0$  is a surface, we know, for Betti (co)homology

$$\mathrm{H}_{1}(\mathcal{X}^{\mathrm{an}},\mathbb{Z})_{\mathrm{tors}}\simeq (\pi_{1}(\mathcal{X}^{\mathrm{an}})^{\mathrm{ab}})_{\mathrm{tors}}\simeq \mathrm{H}^{2}(\mathcal{X}^{\mathrm{an}},\mathbb{Z})_{\mathrm{tors}}\simeq \mathrm{H}^{3}(\mathcal{X}^{\mathrm{an}},\mathbb{Z})_{\mathrm{tors}},$$

by Poincaré duality, the Hurewicz theorem and the universal coefficient theorem. Further, Kweon [Kwe21, Corollary 5.4] shows the following, as a consequence of bounds for torsion in the Néron-Severi group

$$\prod_{\ell \neq p} \# \mathrm{H}^{2}(X, \mathbb{Z}_{\ell})_{\mathrm{tor}} \leq 2^{D^{N^{2}} + 2N \log N} \leq 4^{D^{N^{2}}}.$$

The result follows as a consequence of analysing the growth of the primorial function [HW79, XXII]

$$2^{n/2} \le \pi^!(n) := \prod_{\ell \text{ prime}}^{\ell \le n} \ell \le 4^n$$

and applying standard comparison theorems for étale cohomology.

We close this subsection with the statement of a powerful Chebotarev-type equidistribution theorem due to Katz [KS99, Theorem 9.7.13].

Let  $U_0/\mathbb{F}_q$  be a smooth, affine, geometrically irreducible curve. Let U be the base change to the algebraic closure. Pick a geometric point  $u \to U$ , lying over a closed point  $u_0 \in U(\mathbb{F}_q)$  and denote by  $\overline{\pi}_1 := \pi_1(U, u)$  the geometric étale fundamental group. Let  $\pi_1$  denote the arithmetic fundamental group  $\pi_1(U_0, u)$ .

For any closed point  $v \in U(\mathbb{F}_q)$ , there exists an element  $F_{q,v} \in \pi_1$  well-defined up to conjugacy, called the *Frobenius element* at v. It is defined as follows. Writing  $v = \operatorname{Spec}(\mathbb{F}_q) \to U$ , we obtain an induced map of fundamental groups

$$\operatorname{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q) \to \pi_1(U_0, v) \simeq \pi_1.$$

The element  $F_{q,v} \in \pi_1$  is simply the image in  $\pi_1$  of the frobenius element in  $\text{Gal}(\mathbb{F}_q/\mathbb{F}_q)$  under the composition of the above morphisms.

Given a map  $\rho : \pi_1 \to G$  to a finite group, and a conjugacy-stable subset  $C \subset G$ , we seek to understand the proportion of points  $v \in U(\mathbb{F}_{q^w})$  such that  $\rho(F_{q^w},v)$  lies in C. The following is [Cha97, Theorem 4.1].

**Theorem 3.7** (Katz). Assume there is a commutative diagram

$$1 \longrightarrow \overline{\pi}_{1} \longrightarrow \pi_{1} \longrightarrow \hat{\mathbb{Z}} \longrightarrow 1$$
$$\stackrel{\overline{\rho}}{\downarrow} \qquad \qquad \downarrow^{\rho} \qquad \qquad \downarrow^{1 \mapsto -\gamma}$$
$$1 \longrightarrow \overline{G} \longrightarrow G \stackrel{\mu}{\longrightarrow} \Gamma \longrightarrow 1$$

where G is a finite group,  $\Gamma$  is abelian,  $\overline{\rho}$  is surjective and tamely ramified. Let  $C \subset G$  be stable under conjugation by elements of G. Then

$$\left|\frac{\#\{v \in U(\mathbb{F}_{q^w}) \mid \rho(F_{q^w,v}) \in C\}}{\#U(\mathbb{F}_{q^w})} - \frac{\#(C \cap G^{\gamma^w})}{\#\overline{G}}\right| \leq |\chi(U)| \frac{\#G\sqrt{q^w}}{\#U(\mathbb{F}_{q^w})}$$

where  $G^{\gamma^w} = \mu^{-1}(\gamma^w)$  and  $\chi(U) = \sum_{i=0}^{1} (-1)^i \dim \mathrm{H}^i(U, \mathbb{Q}_\ell)$  is the  $\ell$ -adic Euler-Poincaré characteristic of U.

3.3. Symplectic groups over finite fields. The goal of this subsection is to obtain a probability estimate (Lemma 3.10) for use in Theorem 4.7. Let V be a vector space of rank 2r, for  $r \in \mathbb{Z}_{>0}$ , over the finite field  $\mathbb{F}_{\ell}$  of characteristic  $\ell > 0$ , equipped with a *symplectic* (i.e., alternating, nondegenerate, bilinear) pairing  $\langle \cdot, \cdot \rangle$ .

**Definition 3.8.** The group of symplectic similar similar  $\operatorname{GSp}(2r, \mathbb{F}_{\ell})$  is defined as

$$\mathrm{GSp}(2r,\mathbb{F}_{\ell}) := \{ A \in \mathrm{GL}(2r,\mathbb{F}_{\ell}) \mid \exists \ \gamma \in \mathbb{F}_{\ell}^* \text{ such that } \langle Av,Aw \rangle = \gamma \langle v,w \rangle \ \forall v,w \in V \}.$$

For  $A \in \operatorname{GSp}(2r, \mathbb{F}_{\ell})$ , the associated  $\gamma \in \mathbb{F}_{\ell}^*$  is called the *multiplicator* of A. We denote by  $\operatorname{GSp}(2r, \mathbb{F}_{\ell})^{\gamma}$  the subset of matrices with multiplicator  $\gamma$ . The matrices with multiplicator  $\gamma = 1$  form a subgroup known as the *symplectic group*, denoted  $\operatorname{Sp}(2r, \mathbb{F}_{\ell})$ . We have the following exact sequence

$$1 \longrightarrow \operatorname{Sp}(2r, \mathbb{F}_{\ell}) \longrightarrow \operatorname{GSp}(2r, \mathbb{F}_{\ell}) \xrightarrow{\operatorname{mult}} \mathbb{F}_{\ell}^* \longrightarrow 1$$

For any  $\gamma \in \mathbb{F}_{\ell}^*$ , collect the 'relevant' characteristic polynomials f in the set

$$M_r^{\gamma} := \{ f(T) = 1 + a_1 T + \ldots + a_{2r-1} T^{2r-1} + \gamma^r T^{2r} \mid a_i \in \mathbb{F}_{\ell}, \ a_{2r-i} = \gamma^{r-i} a_i, \ 0 \le i \le 2r \}.$$

We now give an estimate for the number of matrices with given characteristic polynomial f(T). See [Cha97, Theorem 3.5] for a proof.

**Lemma 3.9.** Fix  $f(T) \in M_r^{\gamma}$ . For  $\ell > 4$ , we have

$$(\ell - 3)^{2r^2} \leq \# \{ A \in \operatorname{GSp}(2r, \mathbb{F}_{\ell})^{\gamma} \mid f(T) = \det(1 - TA) \} \leq (\ell + 3)^{2r^2}.$$

We may identify  $M_r^{\gamma}$  with the points of the affine space  $\mathbb{A}_{\mathbb{F}_{\ell}}^r$  with coordinates  $(y_1, \ldots, y_r)$ , by sending a polynomial  $f(T) = 1 + \sum_{i=1}^{2r-1} a_i T^i + \gamma^r T^{2r}$  to the tuple  $(a_1, \ldots, a_r)$ .

Our goal is to obtain estimates for the proportion of characteristic polynomials that are *not* coprime to a given  $f(T) \in M_r^{\gamma}$ . Let  $W \subset \mathbb{A}_{\mathbb{F}_{\ell}}^r$  parameterise such polynomials. It is a hypersurface, given by the vanishing of  $F(y_1, \ldots, y_r)$ , described as the resultant of a formal polynomial of the type

$$g(T) = 1 + \sum_{i=1}^{r} y_i T^i + \sum_{i=1}^{r-1} \gamma^{r-i} y_i T^{2r-i} + \gamma^r T^{2r}$$

with f(T) w.r.t. T. The polynomial F is of total degree at most 4r in the  $y_i$ . The number of its rational points,  $\#W(\mathbb{F}_{\ell})$ , gives the count we need. But, by [BS86, pg 45], we have  $\#W(\mathbb{F}_{\ell}) \leq 4r\ell^{r-1}$ . Further, recalling the order formula for the symplectic group, we have

$$\ell^{2r^2}(\ell-1)^r \leq \# \operatorname{Sp}(2r, \mathbb{F}_{\ell}) = \ell^{r^2} \prod_{j=1}^r (\ell^{2j}-1) \leq \ell^{2r^2+r}$$

Therefore, combining with Lemma 3.9, the proportion of matrices in  $GSp(2r, \mathbb{F}_{\ell})^{\gamma}$  with characteristic polynomial *not* coprime to f(T) is at most

$$\frac{4r\ell^{r-1}\cdot(\ell+3)^{2r^2}}{\ell^{2r^2}(\ell-1)^r} = \frac{4r}{\ell}\left(1+\frac{1}{\ell-1}\right)^r\left(1+\frac{3}{\ell}\right)^{2r^2}$$

which is less than 1/4, for  $\ell > 16e^2r^2$ , where  $e := \exp(1)$ . We summarise what we have shown in the following.

**Lemma 3.10** (Common eigenvalue). Let  $r \in \mathbb{Z}_{>0}$  and let  $\ell > 4$  be a prime. Let f(T) be the characteristic polynomial of a matrix in  $\operatorname{GSp}(2r, \mathbb{F}_{\ell})^{\gamma}$  for some  $\gamma \in \mathbb{F}_{\ell}^*$ . Denote by  $C \subset \operatorname{GSp}(2r, \mathbb{F}_{\ell})$  the set of matrices with characteristic polynomial not coprime with f(T). Then for  $\ell > 119r^2$ ,

$$\frac{\# \left(C \cap \operatorname{GSp}(2r, \mathbb{F}_{\ell})^{\gamma}\right)}{\#\operatorname{Sp}(2r, \mathbb{F}_{\ell})} \leq 1/4 \,.$$

#### 4. $P_1(T)$ for smooth projective varieties

This section proves the rest of our main theorems. 4.1 details the reduction to the case of a surface and 4.2 delineates an algorithm to construct, with high probability, a Lefschetz pencil of hyperplane sections on a surface. In 4.3, we prove the effective gcd theorem, which forms the basis for the algorithms and the proofs of the Theorems 1.2, 1.3 and 1.4 in Section 4.4.

4.1. Reduction to smooth projective surfaces. In this subsection, we demonstrate a reduction of the problem of computing the characteristic polynomial of geometric Frobenius on the first ( $\ell$ -adic) étale cohomology of a smooth projective variety over a finite field  $\mathbb{F}_q$  of fixed dimension r > 1, to that of a smooth projective surface. This reduction is polynomial in the input data, namely the degree of the polynomials defining the variety and log q.

Let  $X_0/\mathbb{F}_q$  be a smooth, projective, geometrically irreducible variety of dimension n > 1 and degree D > 0. We suppose that it is presented as a subvariety of  $\mathbb{P}^N$ , given by a homogeneous ideal I generated by m polynomials  $f_1, \ldots, f_m$  of degree  $\leq d$  for  $d \in \mathbb{Z}_{>0}$ . Denote by X the base change to the algebraic closure. Let  $\ell$  be a prime distinct from the characteristic of the base field. We recall the following.

**Definition 4.1.** Let X be as above. A hyperplane section of X is a codimension 1 subvariety  $Y \subset X$  obtained by intersecting X with a hyperplane  $H \subset \mathbb{P}^N$ . A hyperplane H is said to intersect X transversally at  $x \in X$  if  $T_x X \not\subset H$ , i.e., H does not contain the tangent space to X at x. Equivalently, this translates to the condition that  $X \cap H$  is smooth at x. In general, H intersects X transversally if  $H \cap X$  is a smooth, irreducible subvariety of codimension 1 of X.

Denote by  $(\mathbb{P}^N)^{\vee}$  the dual projective space, parameterising hyperplanes in  $\mathbb{P}^N$ . We construct the *dual variety* to X, denoted  $\check{X} \subset (\mathbb{P}^N)^{\vee}$  as follows. Let

$$\Omega := \{ (x, H) \in X \times (\mathbb{P}^N)^{\vee} \mid x \in H, \ T_x X \subset H \}.$$

It is a closed subvariety of  $X \times (\mathbb{P}^N)^{\vee}$ . We define  $\check{X}$  to be the projection of  $\Omega$  onto its second factor. In particular,  $\check{X}$  parameterises those hyperplanes that *do not* intersect transversally with X. We now state an effective version of Bertini's theorem, that ensures the availability of smooth hyperplane sections. The following is [Bal03, Theorem 1].

**Proposition 4.2** (Effective Bertini). Let  $W \subset \mathbb{P}^N$  be a smooth, irreducible variety of dimension n and degree D, defined over  $\mathbb{F}_q$ . Let  $\mathbb{F}_Q/\mathbb{F}_q$  be an extension such that  $Q > D(D-1)^n$ . Then, there exists a hyperplane H defined over  $\mathbb{F}_Q$  that intersects transversally with W.

In the proof of the above theorem, it is shown [Bal03, Lemma 1] that  $\check{W}$  is a variety of degree at most  $D(D-1)^N \leq D^{N+1}$ . The singular locus of  $\check{W}$ , denoted  $\mathring{W}$  is a subvariety of  $(\mathbb{P}^N)^{\vee}$  of codimension at least 2 and degree (by Bézout) at most  $D^{(N+1)^2}$ .

*Remark.* The existence of smooth hypersurface sections of sufficiently large degree is given by [Poo04]. However it is unavoidable to take field extensions for our algorithmic purposes (e.g., even to ensure the existence of a rational point), so the trade-off is immaterial.

We now recall the following theorem, which is the key step in our reduction to surfaces. See [Fu11, §8.5.5] for the proof of the more general theorem, of which this is a special case.

**Theorem 4.3** (Weak-Lefschetz). Let  $Y \hookrightarrow X$  be a smooth hyperplane section. Then the induced map

$$\mathrm{H}^1(X, \mathbb{Q}_\ell) \to \mathrm{H}^1(Y, \mathbb{Q}_\ell)$$

is an isomorphism if  $n = \dim(X) > 2$  and an injection if n = 2.

With this setup, we notice that with an application of Bertini's theorem on the existence of smooth hyperplane sections, we can inductively reduce the dimension of X by intersecting with a generic hyperplane in  $\mathbb{P}^N$ . In particular, there is a chain of smooth hyperplane sections  $Y := Y_2 \subset Y_3 \subset \ldots \subset Y_{n-1} \subset X$ , where  $Y_i$  are smooth varieties of dimension *i*. Applying the weak-Lefschetz theorem, we get an isomorphism

$$\mathrm{H}^{1}(X, \mathbb{Q}_{\ell}) \to \mathrm{H}^{1}(Y, \mathbb{Q}_{\ell}),$$

compatible with the action of the respective geometric Frobenii. Writing

$$P_1(X/\mathbb{F}_q, T) := \det\left(1 - TF_q^\star \mid \mathrm{H}^1(X, \mathbb{Q}_\ell)\right)$$

and assuming Y is also defined over  $\mathbb{F}_q$ , we have  $P_1(X/\mathbb{F}_q, T) = P_1(Y/\mathbb{F}_q, T)$ . So, it suffices to compute  $P_1(Y/\mathbb{F}_q, T)$  for Y a smooth subvariety obtained from X after intersection with n-2 hyperplanes in general position.

*Remark.* We note that the ideal defining Y now is generated by the forms  $f_i, L_j$  with  $1 \le i \le m$ and  $1 \le j \le n-2$ , where the  $L_j$  are linear forms representing the generic hyperplanes in  $\mathbb{P}^N$  that we have intersected X with, to obtain Y.

4.2. Lefschetz pencils on a surface. To study the zeta function of a surface, intuitively, one wants to break it up into those of curves, each parameterized by a single variable t, and then invoke the methods of Section 2. It is not so easy because Theorem 4.3 does not give an isomorphism when X is a surface, e.g.,  $H^1(Y, \mathbb{Q}_\ell)$  can be a larger group for a generic curve Y lying on the surface X, which will make the zeta function of Y 'larger' than that of X, introducing errors called vanishing cycles (see Section 3.1).

To explore these issues, in this subsection, we introduce the classic machinery of Lefschetz pencils and describe an algorithm to fibre a given smooth projective surface  $X \subseteq \mathbb{P}^N$  of degree D over the projective line so that the fibres are curves with singularities at worst being ordinary double points. We assume X is given by m homogeneous forms  $f_1, \ldots, f_m$ , each of total degree  $\leq d \in \mathbb{Z}_{>0}$ . Denote by  $(\mathbb{P}^N)^{\vee}$ , the dual projective space.

**Definition 4.4.** Let  $X/\overline{\mathbb{F}}_q$  be as above. A *Lefschetz pencil* on X is a collection of hyperplanes  $(H_t)_{t\in\mathbb{P}^1}$  such that there exists a line  $L \simeq \mathbb{P}^1 \subset (\mathbb{P}^N)^{\vee}$ ; for e.g.,  $(\lambda, \mu) \mapsto \lambda F = \mu G$ , for linear forms F, G on  $\mathbb{P}^N$ , satisfying the following conditions

- the axis, of the pencil,  $\Delta := (F = 0) \cap (G = 0)$  in  $\mathbb{P}^N$  intersects X transversally, i.e.,  $X \cap \Delta$  is smooth of codimension 2,
- there is a dense open subset  $U \subset \mathbb{P}^1$  on which the associated intersections  $(\lambda, \mu) \to X \cap (\lambda F = \mu G)$  are smooth and geometrically irreducible for  $(\lambda, \mu) \in U$ ; and have only an ordinary double point as singularity for the finitely many  $(\lambda, \mu) \notin U$ .

It is a fundamental theorem that Lefschetz pencils exist on any smooth projective variety of dimension  $\geq 2$ , over an algebraically closed field (see [Kat73b]). Over arbitrary fields, Lefschetz pencils exist, subject to a degree  $\geq 3$  Veronese embedding.<sup>10</sup> We recall [JS12, Theorem 3].

**Proposition 4.5.** There exists a nonempty open subscheme (after possibly passing to a degree  $\geq 3$ Veronese embedding) in the Grassmannian of lines  $W_X \subset \text{Gr}(1, (\mathbb{P}^N)^{\vee})$  such that every  $L \in W_X$ defines a Lefschetz pencil for X.

Algorithmically, to construct a Lefschetz pencil, we first take a field extension to ensure the existence of a transversal hyperplane section. We saw that the dual variety  $\check{X}$  parameterises those hyperplanes that do not intersect transversally with X. Further, its singular locus  $\mathring{X}$  parameterises those hyperplanes that intersect X with singularity worse than a single ordinary double point. In other words,  $\check{X} \setminus \mathring{X}$  consists of those hyperplanes H such that  $H \cap X$  has a single node (see [Mil98,

<sup>&</sup>lt;sup>10</sup>this adds an overhead of only a polynomial in the degree D of X.

Theorem 31.2] or [Kat73b]). In light of Proposition 4.2, it suffices to randomly take two linear forms  $F, G \in (\mathbb{P}^N)^{\vee}$ . With high probability, they intersect transversally with X and the line joining them in  $(\mathbb{P}^N)^{\vee}$  intersects  $\check{X}$  at finitely many points and completely misses  $\mathring{X}$ .

# Algorithm 3 Lefschetz pencil on a surface

- Input: A smooth projective surface  $X_0/\mathbb{F}_q$  of degree D presented as a system of homogeneous polynomials of degree  $\leq d$  in the projective space  $\mathbb{P}^N$ .
- **Pre-processing:** Replace X with the degree 3 Veronese image of X in  $\mathbb{P} := \mathbb{P}^{\binom{N+3}{3}-1}$ .
- Output: Hyperplanes F and G in  $\mathbb{P}$  such that the line L through them in the dual  $(\mathbb{P})^{\vee}$ , is a Lefschetz pencil on X.
- 1: Take a field extension  $\mathbb{F}_Q/\mathbb{F}_q$  with degree bounded by a polynomial in D, such that smooth hyperplane sections exist as in Proposition 4.2.
- 2: Select two random linear forms F and G on  $\mathbb{P}$ , such that they intersect transversally with X (this is possible by Proposition 4.2).
- 3: The line L in  $(\mathbb{P})^{\vee}$  through F and G is a Lefschetz pencil on X.

## **Lemma 4.6.** Algorithm 3 succeeds with probability at least 1 - O(1/Q).

Proof. Indeed for  $Q \gg D$ , the locus of hyperplanes in  $\mathbb{P}^{\vee}$  defined over  $\mathbb{F}_Q$  that do not intersect transversally with X is given by the dual variety  $\check{X}$ , which by the Lang-Weil estimates, can be avoided with probability 1 - O(1/Q). Further, for two hyperplanes  $H_1$  and  $H_2$  that intersect transversally with X, the condition that they define a Lefschetz pencil on X is equivalent to the condition that the line through the corresponding points in  $\mathbb{P}^{\vee}$  does not intersect the singular locus  $\mathring{X}$  of  $\check{X}$ . For two randomly chosen hyperplanes, this is also ensured with probability greater than 1 - O(1/Q), again by a Lang-Weil argument.

One checks that the output is correct by computing the finite subset Z of 'bad' hyperplanes (which is possible in poly-time) and verifying that the associated fibres are indeed nodal curves. The latter can be done by blowing up at a singular point and checking that the exceptional divisor intersects the transformed curve at two points, which has a polynomial-time algorithm.

Blowing up X along  $X \cap \Delta$  gives a smooth projective surface  $\tilde{X}$  and a morphism  $\pi : \tilde{X} \to \mathbb{P}^1$ such that the fibre of a  $[\lambda : \mu] \in \mathbb{P}^1$  is the curve  $X \cap (\lambda F = \mu G)$ . Algorithmically, the locus  $\Delta \cap X$ may not all be defined over  $\mathbb{F}_q$  and going to a field extension which contains all of the points therein may be expensive. Further, computing the blowup  $\tilde{X} \to X$  and the morphism  $\pi : \tilde{X} \to \mathbb{P}^1$  may also be exponential in the input data. Fortunately, we are able to leave  $\pi$  implicit, i.e., the only knowledge we need is that the fibre of  $u \in \mathbb{P}^1$  under  $\pi$  is  $H_u \cap X$ , where  $H_u$  is the hyperplane associated to u. We describe the required construction in Algorithm 3.

Now, consider the étale sheaf  $R^{\hat{1}}\pi_{\star}\mathbb{Q}_{\ell}$  on  $\mathbb{P}^1$ . It is locally constant of rank 2g on U, where g is the genus of the generic fibre  $\tilde{X}_{\eta}$  (which is a curve over the function field of  $\mathbb{P}^1$ ), where  $\eta \to \mathbb{P}^1$  is a geometric generic point. By the proper base-change theorem, for a point  $u \in \mathbb{P}^1$ , we have  $(R^1\pi_{\star}\mathbb{Q}_{\ell})_u \simeq \mathrm{H}^1(\tilde{X}_u, \mathbb{Q}_{\ell}) = \mathrm{H}^1(X \cap H_u, \mathbb{Q}_{\ell})$ . Further, by [Mil98, Lemma 33.2], we have  $\mathrm{H}^1(\tilde{X}, \mathbb{Q}_{\ell}) \simeq \mathrm{H}^1(X, \mathbb{Q}_{\ell})$ .

We now establish bounds for the genus g of the generic smooth fibre and for the critical locus which we call  $Z := \mathbb{P}^1 \setminus U$ . Pick  $u \in U$ , the fibre  $X \cap H_u$  is a curve in  $\mathbb{P}$  of degree  $D = \deg(X)$ . By the results of [GLP83] (see also [Mac19, Theorem 3.3]), we have

$$g \le D^2 - 2D + 1.$$

Further, the number of critical points, i.e., #Z is bounded by the size of  $L \cap X$ , which by the remark following Proposition 4.2 and Bézout's theorem, is at most  $D^{N+1}$ . Denote the Betti numbers

 $\beta_i := \dim \mathrm{H}^i(X, \mathbb{Q}_\ell)$ . Clearly, we have<sup>11</sup>

$$\beta_3 = \beta_1 \le 2g \le 2D^2.$$

By [Mil80, V, Theorem 3.12], we have

$$\beta_2 = \#Z + 2\beta_1 + 2 - 4g \le 2D^{N+1}$$

4.3. An effective gcd theorem. Now, let  $X_0/\mathbb{F}_q$  be a smooth, projective, geometrically irreducible surface of degree D > 0 obtained from good reduction of a smooth, projective surface  $\mathcal{X}_0$  over a number field K at a prime  $\mathfrak{p}$ . We assume that  $X_0$  is presented as  $X_0 \subset \mathbb{P}^N$ , given by a homogeneous ideal I generated by m polynomials  $f_1, \ldots, f_m$  of degree  $\leq d$  for  $d \in \mathbb{Z}_{>0}$ . Denote by X the base change to the algebraic closure. Let  $\ell$  be an odd prime, distinct from the characteristic, chosen according to Proposition 3.6. In this subsection, we prove an effective version of Deligne's 'théorème du pgcd' [Del80, Théorème 4.5.1], that enables one to recover  $P_1(X/\mathbb{F}_q, T)$  from the zeta function of hyperplane sections of X (namely, simply by taking their gcd).

Following Algorithm 3, we may fibre X as a Lefschetz pencil of hyperplane sections  $\pi : \tilde{X} \to \mathbb{P}^1$ over  $\mathbb{F}_q$  (after possibly replacing  $\mathbb{F}_q$  by an extension of degree at most polynomial in D). Denote by  $U \subset \mathbb{P}^1$  the open subscheme<sup>12</sup> where the fibres are smooth, and Z its complement. Let g be the genus of the geometric generic fibre  $X_{\eta}$ .

Let  $u \in U(\mathbb{F}_q)$ . From the formalism of vanishing cycles and the so-called 'hard-Lefschetz theorem' [Del80, 4.3.9], we have the decomposition

$$\mathrm{H}^{1}(X_{u}, \mathbb{Q}_{\ell}) \simeq \mathrm{H}^{1}(X, \mathbb{Q}_{\ell}) \oplus \mathcal{E}_{u}$$

where  $X_u$  denotes the fibre of  $\pi$  over u, and  $\mathcal{E}_u$  is the space generated by the vanishing cycles in  $\mathrm{H}^1(X_u, \mathbb{Q}_\ell)$ . In particular, we have that

$$P_1(X_u/\mathbb{F}_q, T) = P_1(X/\mathbb{F}_q, T) \cdot P(\mathcal{E}_u/\mathbb{F}_q, T) ,$$

where  $P(\mathcal{E}_u/\mathbb{F}_q, T)$  denotes the characteristic polynomial of  $F_q^*$  acting on  $\mathcal{E}_u$ .

A theorem of Deligne (Theorem 3.3) states that  $P_1(X/\mathbb{F}_q, T)$  can be recovered from  $P_1(X_{u_i}/\mathbb{F}_{q^j}, T)$ for  $u_i \in U(\mathbb{F}_{q^j})$  over all extensions  $\mathbb{F}_{q^j}$ . We show that there is a 'small' extension, and a small number of fibres over that extension to sample, to recover  $P_1(X/\mathbb{F}_q, T)$ .

Firstly, consider the representation  $\rho : \pi_1(U_0, u) \to \operatorname{GL}(2r, \mathbb{Z}_\ell)$  of the étale fundamental group of  $U_0$  associated to the torsion-free lisse  $\mathbb{Z}_\ell$ -sheaf  $\mathcal{E}_{\mathbb{Z}_\ell} \subset R^1 \pi_* \mathbb{Z}_\ell |_U$ , of vanishing cycles. Denote by  $\overline{\rho} := \rho \mid \pi_1(U, u)$ , the restriction to the geometric fundamental group. By [Del74, 5.10], we know that the Zariski-closure of the image of  $\overline{\rho} \otimes \mathbb{Q}_\ell$  in  $\operatorname{GL}(2r, \mathbb{Q}_\ell)$  is the symplectic group  $\operatorname{Sp}(2r, \mathbb{Q}_\ell)$ . Using methods of Hall [Hal08], we deduce that the mod- $\ell$  monodromy of the family, i.e., the image of  $\overline{\rho}_\ell : \pi_1(U, u) \to \operatorname{GL}(2r, \mathbb{F}_\ell)$  is the symplectic group  $\operatorname{Sp}(2r, \mathbb{F}_\ell)$ .

Next, we note that for  $u \in U(\mathbb{F}_{q^j})$  the 'vanishing term'  $P(\mathcal{E}_u/\mathbb{F}_{q^j}, T)$  is equidistributed (mod- $\ell$ ) in the family, à la Katz (see [Cha97, Theorem 4.1] or [KS99, Theorem 9.7.13]), so can be eliminated with high probability after two samplings. This is done by first moving to a large enough extension  $\mathbb{F}_Q$  of  $\mathbb{F}_q$  (to minimise the error term coming from the aforementioned equidistribution theorem) and sampling points uniformly randomly in  $U(\mathbb{F}_Q)$ . Then the zeta functions of the associated fibres are computed and their gcd is taken. With high probability, this procedure gives  $P_1(X/\mathbb{F}_Q, T)$ , from which  $P_1(X/\mathbb{F}_q, T)$  can be easily recovered using an analogue of Lemma 2.11.

**Theorem 4.7** (Effective gcd). There exists an extension  $\mathbb{F}_Q/\mathbb{F}_q$ , with degree  $[\mathbb{F}_Q : \mathbb{F}_q]$  bounded by a polynomial in D, such that for any two distinct randomly chosen  $u_1, u_2 \in U(\mathbb{F}_Q)$ , we have with probability > 2/3

$$gcd(P_1(X_{u_1}/\mathbb{F}_Q, T), P_1(X_{u_2}/\mathbb{F}_Q, T)) = P_1(X/\mathbb{F}_Q, T).$$

<sup>&</sup>lt;sup>11</sup>by Poincaré duality

<sup>&</sup>lt;sup>12</sup>write  $U_0$  for the associated  $\mathbb{F}_q$ -scheme.

Proof. Let  $\ell \in [(4D)^4, 2^{11}D^{N^2}]$  be a prime distinct from p such that the cohomology groups  $\mathrm{H}^i(X, \mathbb{Z}_\ell)$  are all torsion-free. This is possible by Proposition 3.6. Consider the locally constant sheaf  $R^1\pi_{\star}\mathbb{Z}_{\ell}|_U$  on U. It has as subsheaf,  $\mathcal{E}_{\mathbb{Z}_\ell}$ , the sheaf of  $\mathbb{Z}_\ell$ -vanishing cycles of rank (say) 2r. Denote by  $\rho : \pi_1(U_0, u) \to \mathrm{GL}(2r, \mathbb{Z}_\ell)$  the associated  $\ell$ -adic representation and by  $\overline{\rho} = \rho | \pi_1(U, u)$ . Write  $\rho_\ell$  and  $\overline{\rho}_\ell$  respectively, for the mod- $\ell$  representations.

By the hard-Lefschetz theorem, (Theorem 3.2) we have  $2r = 2g - \beta_1$  where  $\beta_1$  is the first Betti number of X. By Theorem 3.5, we know that the sheaf  $\mathcal{E}_{\mathbb{Z}_{\ell}}$  has big mod- $\ell$  monodromy, i.e.,  $\operatorname{im}(\overline{\rho}_{\ell}) = \operatorname{Sp}(2r, \mathbb{F}_{\ell})$ . We seek to apply Theorem 3.7 to this setup. Let  $\mathbb{F}_Q/\mathbb{F}_q$  be an extension where  $Q := q^w$  and choose  $u_1 \in U(\mathbb{F}_Q)$  randomly. We estimate the number of  $v \in U(\mathbb{F}_Q)$  such that  $P(\mathcal{E}_v/\mathbb{F}_Q, T)$  is coprime to  $f(T) := P(\mathcal{E}_{u_1}/\mathbb{F}_Q, T)$ . Write  $\overline{f}(T) := f(T) \mod \ell$ .

Denote by  $C \subset \operatorname{GSp}(2r, \mathbb{F}_{\ell})$  the subset of matrices with characteristic polynomial not coprime to  $\overline{f}(T)$ . It is stable under conjugation by elements from  $\operatorname{GSp}(2r, \mathbb{F}_{\ell})$ . Applying Theorem 3.7 to C, we get

$$\frac{\#\{v \in U(\mathbb{F}_Q) \mid \rho_\ell(F_{Q,v}) \in C\}}{\#U(\mathbb{F}_Q)} \leq \frac{\#(C \cap \operatorname{GSp}(2r, \mathbb{F}_\ell)^{\gamma^w})}{\#\operatorname{Sp}(2r, \mathbb{F}_\ell)} + |\chi(U)| \frac{\#\operatorname{GSp}(2r, \mathbb{F}_\ell)\sqrt{q^w}}{\#U(\mathbb{F}_Q)}$$

By Lemma 3.10 (since  $\ell > 119r^2$ ), the first summand on the RHS is  $\leq 1/4$ . From the calculation<sup>13</sup> of the étale cohomology of U (the projective line with #Z punctures), we deduce that  $|\chi(U)| \leq \#Z \leq D^{N+1}$ . For  $q^w > 2D^{N+1}$ , we have

$$|\chi(U)| \frac{\# \operatorname{GSp}(2r, \mathbb{F}_{\ell})\sqrt{q^w}}{\# U(\mathbb{F}_Q)} \leq D^{N+1}\ell^{2g^2+g+1} \frac{\sqrt{q^w}}{q^w - D^{N+1}} \leq 2D^{N+1}(2^{11}D^{N^2})^{4D^2} \frac{\sqrt{q^w}}{q^w/2}$$

In particular, if  $Q = q^w > \Omega\left(D^{5N^2D^4}\right)$ , we have

$$\frac{\#\{v \in U(\mathbb{F}_Q) \mid \rho_\ell(F_{Q,v}) \notin C\}}{\#U(\mathbb{F}_Q)} > 2/3$$

which completes the proof.

4.4. Algorithms for  $P_1(T)$ . Let  $X_0 \subset \mathbb{P}^N$  be a smooth projective variety of dimension n > 1 and degree D over  $\mathbb{F}_q$ , obtained via good reduction from  $\mathcal{X}_0$ ; defined over a number field K, at a prime  $\mathfrak{p} \subset \mathcal{O}_K$ . An AM  $\cap$  coAM protocol for certifying  $P_1(X/\mathbb{F}_Q, T)$  for any field extension  $\mathbb{F}_Q/\mathbb{F}_q$  with

$$Q > \Omega\left(D^{5N^2D^4}\right)$$

is presented in Algorithm 4.

**Theorem 1.2 (restated).** Given  $Q_1(T) \in \mathbb{Z}[T]$ , deciding whether  $Q_1(T) = P_1(X/\mathbb{F}_q, T)$ , for  $X_0$  given as above, is in AM  $\cap$  coAM.

Proof. Let  $X_0 \subset \mathbb{P}^N$  be a smooth projective variety of dimension n > 1 and degree D, over the field  $\mathbb{F}_q$  given by homogeneous forms  $f_1, \ldots, f_m$ , each of total degree  $\leq d \in \mathbb{Z}_{>0}$ . For any extension  $\mathbb{F}_Q/\mathbb{F}_q$  such that  $Q > \Omega\left(D^{5N^2D^4}\right)$  of poly-bounded degree, we may verify  $P_1(X/\mathbb{F}_Q,T)$ using Algorithm 4. Now, choosing two field extensions  $\mathbb{F}_{Q_1}/\mathbb{F}_q$  and  $\mathbb{F}_{Q_2}/\mathbb{F}_q$  with size greater than  $\Omega\left(D^{5N^2D^4}\right)$  according to Lemma 2.11, we can recover and hence certify  $P_1(X/\mathbb{F}_q,T)$  as well.  $\Box$ 

For Theorem 1.3, we recall a theorem of Kedlaya [Ked06, Theorem 1] that enables efficient quantum computation of the zeta function of a curve.

**Theorem 4.8** (Kedlaya). Let  $C \subset \mathbb{P}^N$  be a smooth projective curve over  $\mathbb{F}_q$ , of degree D. Then, there exists a quantum algorithm that computes  $P_1(C/\mathbb{F}_q, T)$  in time polynomial in  $D \log q$ .

<sup>&</sup>lt;sup>13</sup>see [Sta18, Tag 03RR]

**Algorithm 4** Verifying  $P_1(T)$  of a variety

- Input: A smooth projective variety  $X_0/\mathbb{F}_q$  of dimension n > 1 and degree D, presented as a system of m homogeneous polynomials  $f_1, \ldots, f_m$  of degree  $\leq d$  in the projective space  $\mathbb{P}^N$ .
- **Pre-processing:** We first move to a field extension  $\mathbb{F}_Q/\mathbb{F}_q$  that affords enough smooth hyperplane sections as in Proposition 4.2 and satisfies the bound of Theorem 4.7. We may reduce to a surface Y by intersecting X with n-2 generic hyperplanes. Next, Y is fibred as a Lefschetz pencil of hyperplane sections following Algorithm 3. Denote by  $U \subset \mathbb{P}^1$ the open subscheme parameterising the smooth fibres, and  $Z := \mathbb{P}^1 \setminus U$  the finitely many singular ones.
- Conditions: Merlin provides a candidate P(T) for  $P_1(X/\mathbb{F}_Q, T)$  and Arthur engages in a protocol with Merlin to determine the veracity of the claim.
- 1: Arthur: Pick randomly distinct  $u_i \in U(\mathbb{F}_Q)$ , for  $1 \le i \le 2$  following Theorem 4.7.
- 2: Merlin: Provide  $P_1(Y_{u_i}/\mathbb{F}_Q, T)$ , for  $1 \le i \le 2$ .
- 3: Arthur: Verify that the  $P_1(Y_{u_i}/\mathbb{F}_Q, T)$  are as claimed by calling Algorithm 2. Compute their greatest common divisor G(T), using e.g., Euclid's algorithm. Accept iff G(T) = P(T).

**Theorem 1.3 (restated).** There exists a quantum algorithm that computes  $P_1(X/\mathbb{F}_q, T)$  in time polynomial in  $D \log q$ , for any  $X_0$  as given above.

Proof. Similarly to Algorithm 4, we begin by reducing to the case of a surface Y (obtained via successive hyperplane sections of X) and fibring as a Lefschetz pencil of hyperplane sections. We then move to a large enough field extension  $\mathbb{F}_Q/\mathbb{F}_q$  as before, and sample  $u_i \in U(\mathbb{F}_Q)$  uniformly randomly for  $i \in \{1, 2\}$ . Now, using Theorem 4.8, we may compute  $P_1(Y_{u_i}/\mathbb{F}_Q, T)$  of the curves  $Y_{u_i}$  and take their gcd. With probability > 2/3, the result is  $P_1(Y/\mathbb{F}_Q, T) = P_1(X/\mathbb{F}_Q, T)$ . We use the technique of Lemma 2.11 to recover the characteristic polynomial  $P_1(X/\mathbb{F}_q, T)$  of the base Frobenius as well.

We now recall the following result to compute the zeta function of a smooth curve of fixed degree.

**Theorem 4.9** (Pila, Huang-Ierardi). Let  $C \subset \mathbb{P}^N$  be a smooth projective curve over  $\mathbb{F}_q$ , of fixed degree D. Then, there exists an algorithm that computes  $P_1(C/\mathbb{F}_q,T)$  in time  $O((\log q)^{\Delta})$ , where  $\Delta$  is independent of q and polynomial in D.

*Proof.* Move to a plane nodal model C' of C via Lemma 2.2 and apply [HI98, Theorem 1.1].

**Theorem 1.4 (restated).** There exists a randomised algorithm that, given  $X_0$  as above of fixed degree D, computes  $P_1(X/\mathbb{F}_q,T)$  in time polynomial in  $\log q$ .

*Proof.* Similar to the proof of Theorem 1.3, use the algorithm of Huang-Ierardi from Theorem 4.9 to compute  $P_1(Y_{u_i}/\mathbb{F}_Q, T)$  and, then take their gcd. Use Lemma 2.11 to recover the characteristic polynomial of the base Frobenius.

#### 5. CONCLUSION

We have presented randomised methods to efficiently compute and certify the characteristic polynomial of the geometric Frobenius on the first  $\ell$ -adic étale cohomology of smooth varieties. The immediate question is for higher cohomologies: to begin with, how do we compute  $P_n(T)$  for a smooth projective variety of dimension n > 1 over  $\mathbb{F}_q$  in time polynomial in  $\log q$ ? In another direction (for variable  $D \log q$ ), one may ask for deterministic verification, i.e., an NP  $\cap$  coNP protocol for  $P_1(T)$  and more generally for  $P_i(T)$ .

#### Acknowledgements

N.S. thanks the funding support from DST-SERB (CRG/2020/45 + JCB/2022/57) and N. Rama Rao Chair. M.V. is supported by a C3iHub research fellowship. M.V. thanks Chris Hall, Donu Arapura and Jeff Achter for email conversations and Hyuk Jun Kweon for pointing out the work [Kwe21], which led to Proposition 3.6. We thank Vasudevan Srinivas, Kiran Kedlaya and Partha Mukhopadhyay for enthusiastic discussions about the related problems. We thank the many researchers who provided valuable feedback by attending our talks and reading the draft.

#### References

- [AB09] Sanjeev Arora and Boaz Barak. Computational Complexity: A Modern Approach. Cambridge University Press, 2009. 3, 4, 9
- [ABCL02] Simon Abelard, Elena Berardini, Alain Couvreur, and Grégoire Lecerf. Computing Riemann–Roch spaces via Puiseux expansions. *Journal of Complexity*, 73, 2002. 7
  - [AH01] Leonard M Adleman and Ming-Deh Huang. Counting rational points on curves and abelian varieties over finite fields. *Journal of Symbolic Computation*, 32(6):171–189, 2001. 3
  - [Bal03] Edoardo Ballico. An effective Bertini theorem over finite fields. Advances in Geometry, 3(4):361–363, 2003. 16
  - [Ber86] Pierre Berthelot. Géométrie rigide et cohomologie des variétés algébriques de caractéristique p. Groupe de travail d'analyse ultramétrique, 9(3):J1–J18, 1986. 3
  - [Ber20] Elena Berardini. Algebraic geometry codes from surfaces over finite fields. PhD thesis, Université d'Aix-Marseille, 2020. 2
  - [BHZ87] Ravi B Boppana, Johan Hastad, and Stathis Zachos. Does co-NP have short interactive proofs? Information Processing Letters, 25(2):127–132, 1987. 4
  - [BPR06] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics). Springer-Verlag, Berlin, Heidelberg, 2006. 2
    - [BS86] Zenon Ivanovich Borevich and Igor Rostislavovich Shafarevich. Number theory. Academic press, 1986. 15
    - [BS10] Peter Bürgisser and Peter Scheiblechner. Counting irreducible components of complex algebraic varieties. *computational complexity*, 19:1–35, 2010. 2
- [CDV06] Wouter Castryck, Jan Denef, and Frederik Vercauteren. Computing zeta functions of nondegenerate curves. International Mathematics Research Papers, 2006:72017, 2006.
- [CFA<sup>+</sup>05] Henri Cohen, Gerhard Frey, Roberto Avanzi, Christophe Doche, Tanja Lange, Kim Nguyen, and Frederik Vercauteren. Handbook of elliptic and hyperelliptic curve cryptography. CRC press, 2005. 2, 3
  - [Cha97] Nick Chavdarov. The generic irreducibility of the numerator of the zeta function in a family of curves with large monodromy. Duke Mathematical Journal, 87(1):151 – 180, 1997. 14, 15, 19
  - [Del74] Pierre Deligne. La conjecture de Weil : I. Publications Mathématiques de l'IHÉS, 43:273–307, 1974. 2, 13, 19
  - [Del80] Pierre Deligne. La conjecture de Weil : II. Publications Mathématiques de l'IHÉS, 52:137–252, 1980. 5, 6, 12, 19
  - [DV06] Jan Denef and Frederik Vercauteren. Counting points on  $C_{ab}$  curves using Monsky-Washnitzer cohomology. Finite Fields and Their Applications, 12(1):78–102, 2006. 3
  - [Dwo60] Bernard Dwork. On the rationality of the zeta function of an algebraic variety. Amer. J. Math, 82:631–648, 1960. 3
  - [FK13] Eberhard Freitag and Reinhardt Kiehl. Étale cohomology and the Weil Conjecture, volume 13. Springer Science & Business Media, 2013. 2, 13

- [FR94] Gerhard Frey and Hans-Georg Rück. A remark concerning m-divisibility and the discrete logarithm in the divisor class group of curves. Mathematics of computation, 62(206):865– 874, 1994. 4
- [Fu11] Lei Fu. Etale cohomology theory, volume 13. World Scientific, 2011. 16
- [G<sup>+</sup>77] Alexander Grothendieck et al. Cohomologie *l*-adique et fonctions L (SGA V). Lecture Notes in Math, 589, 1977. 2, 3
- [Gab83] Ofer Gabber. Sur la torsion dans la cohomologie *l*-adique d'une variété. CR Acad. Sci. Paris Sér. I Math, 297(3):179–182, 1983. 12
- [GLP83] Laurent Gruson, Robert Lazarsfeld, and Christian Peskine. On a theorem of Castelnuovo, and the equations defining space curves. *Inventiones mathematicae*, 72:491–506, 1983. 18
- [Gro57] Alexander Grothendieck. Fondements de la géométrie algébrique. Commentaires. Séminaire Bourbaki, 7:297–298, 1957. 13
- [GS86] Shafi Goldwasser and Michael Sipser. Private coins versus public coins in interactive proof systems. In Proceedings of the eighteenth annual ACM symposium on Theory of computing, pages 59–68, 1986. 5
- [GSS19] Zeyu Guo, Nitin Saxena, and Amit Sinhababu. Algebraic dependencies and PSPACE algorithms in approximative complexity over any field. Theory of Computing, 15(1):1– 30, 2019. 4
- [Hal08] Chris Hall. Big symplectic or orthogonal monodromy modulo ℓ. Duke Mathematical Journal, 141(1):179 – 203, 2008. 6, 13, 19
- [Hal23] Chris Hall. personal communication. 2023. 13
- [Har05] Gilbert Harman. On the greatest prime factor of p-1 with effective constants. Mathematics of Computation, 74(252):2035–2041, 2005. 11
- [Har13] Robin Hartshorne. Algebraic geometry, volume 52. Springer Science & Business Media, 2013. 7
- [Har15] David Harvey. Computing zeta functions of arithmetic schemes. Proceedings of the London Mathematical Society, 111(6):1379–1401, 2015. 4
- [Has36] Helmut Hasse. Zur Theorie der abstrakten elliptischen Funktionenkörper III. Die Struktur des Meromorphismenrings. Die Riemannsche Vermutung. 1936. 8
- [HI94] Ming-Deh Huang and Doug Ierardi. Efficient Algorithms for the Riemann Roch problem and for addition in the Jacobian of a curve. *Journal of Symbolic Computation*, 18(6):519 - 539, 1994. 7, 8
- [HI98] Ming-Deh Huang and Doug Ierardi. Counting points on curves over finite fields. *Journal* of Symbolic Computation, 25(1):1–21, 1998. 3, 5, 21
- [HW79] Godfrey Harold Hardy and Edward Maitland Wright. An introduction to the theory of numbers. Oxford university press, 1979. 14
- [JBW<sup>+</sup>84] Carl Gustav Jakob Jacobi, Carl Wilhelm Borchardt, Karl Weierstrass, Peter Gustav Lejeune Dirichlet, Leopold Kronecker, et al. CGJ Jacobi's Gesammelte werke. Supplementband. 1884. 1
  - [JS12] Uwe Jannsen and Shuji Saito. Bertini theorems and Lefschetz pencils over discrete valuation rings, with applications to higher class field theory. *Journal of Algebraic Geometry*, 21(4):683–705, 2012. 17
  - [Kat73a] Nicholas M Katz. Etude cohomologique des pinceaux de lefschetz. pages 254–327, 1973. 13
  - [Kat73b] Nicholas M Katz. Pinceaux de Lefschetz: théoréme d'existence, expose XVII in Groupe de Monodromy en Geometrie Algebrique [SGA 7 II]. Lecture Notes in Math, 340, 1973. 17, 18
  - [Kat11] Nicholas M Katz. Report on the irreducibility of L-functions. In Number Theory, Analysis and Geometry: In Memory of Serge Lang, pages 321–353. Springer, 2011. 13

- [Ked01] Kiran S Kedlaya. Counting Points on Hyperelliptic Curves using Monsky-Washnitzer Cohomology. J. Ramanujan Math. Soc., 2001. 3
- [Ked06] Kiran S Kedlaya. Quantum computation of zeta functions of curves. computational complexity, 15:1–19, 2006. 3, 4, 5, 8, 10, 20
- [Kle68] Steven L Kleiman. Algebraic cycles and the Weil conjectures. Dix Exposes sur la Cohomologie des Schemas, Masson et Cie, 1968. 12
- [KM74] Nicholas M Katz and William Messing. Some consequences of the Riemann hypothesis for varieties over finite fields. *Inventiones mathematicae*, 23(1):73–77, 1974. 12
- [Koz94] Dexter Kozen. Efficient Resolution of Singularities of Plane Curves. Foundation of Software Technology and Theoretical Computer Science, 141:1 – 11, 1994.
- [KS99] Nicholas M Katz and Peter Sarnak. Random matrices, Frobenius eigenvalues, and monodromy, volume 45. American Mathematical Soc., 1999. 14, 19
- [KS06] Neeraj Kayal and Nitin Saxena. Complexity of ring morphism problems. computational complexity, 15:342–390, 2006. 4
- [Kwe21] Hyuk Jun Kweon. Bounds on the torsion subgroups of Néron–Severi groups. Transactions of the American Mathematical Society, 374(1):351–365, 2021. 6, 14, 22
- [Lau04] Alan G.B. Lauder. Deformation theory and the computation of zeta functions. Proceedings of the London Mathematical Society, 88(3):565–602, 2004. 4
- [Lev22] Christophe Levrat. Calcul effectif de la cohomologie des faisceaux constructibles sur le site étale d'une courbe. arXiv preprint arXiv:2209.10221, 2022. 5
- [Lev23] Christophe Levrat. Computing the cohomology of constructible étale sheaves on curves. arXiv preprint arXiv:2306.03283, 2023. 5
- [LL91] Yagati N Lakshman and Daniel Lazard. On the complexity of zero-dimensional algebraic systems. In *Effective methods in algebraic geometry*, pages 217–225. Springer, 1991. 7, 9
- [LPP03] Hendrik Lenstra, Jonathan Pila, and Carl Pomerance. Future directions in algorithmic number theory. Available on-line at http://www.aimath.org/WWN/primesinp/ articles/html/38a, 2003. 3
- [LW06] Alan G.B. Lauder and Daqing Wan. Counting points on varieties over finite fields of small characteristic. Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography, 2006. 3
- [Mac19] Eoin Mackall. A degree-bound on the genus of a projective curve. Available on author's webpage: https://www.eoinmackall.com/s/deggen2.pdf, 2019. 18
- [Mil80] James S Milne. *Etale cohomology (PMS-33)*. Princeton University Press, 1980. 11, 13, 19
- [Mil98] James S Milne. Lectures on étale cohomology. Available on-line at http://www.jmilne. org/math/CourseNotes/LEC.pdf, 1998. 17, 18
- [MO15] David Madore and Fabrice Orgogozo. Calculabilité de la cohomologie étale modulo ℓ. Algebra & Number Theory, 9(7):1647–1739, 2015. 3, 4
- [Mur67] Jacob P Murre. Lectures on an introduction to Grothendieck's theory of the fundamental group. Tata Institute of Fundamental Research Bombay, 1967. 13
- [NC01] Michael A Nielsen and Isaac L Chuang. Quantum computation and quantum information, volume 2. Cambridge university press Cambridge, 2001. 4
- [OT99] Toshinori Oaku and Nobuki Takayama. An algorithm for de rham cohomology groups of the complement of an affine variety via d-module computation. *Journal of Pure and Applied Algebra*, 139(1-3):201–233, 1999. 2
- [Pie03] James Pierpont. Gauss's collected works. 1903. 1
- [Pil90] Jonathan Pila. Frobenius maps of abelian varieties and finding roots of unity in finite fields. Mathematics of Computation, 55(192):745–763, 1990. 3, 5

- [Poo04] Bjorn Poonen. Bertini theorems over finite fields. Annals of mathematics, pages 1099– 1127, 2004. 16
- [Sat00] Takakazu Satoh. The canonical lift of an ordinary elliptic curve over a finite field and its point counting. *Journal-Ramanujan Mathematical Society*, 15(4):247–270, 2000. 3
- [Sch85] René Schoof. Elliptic curves over finite fields and the computation of square roots mod p. Mathematics of computation, 44(170):483–494, 1985. 3
- [Sch07] Peter Scheiblechner. On the complexity of deciding connectedness and computing Betti numbers of a complex algebraic variety. *Journal of Complexity*, 23(3):359–379, 2007. 2
- [Sch19] Peter Scheiblechner. Effective de Rham cohomology—The general case. Communications in Contemporary Mathematics, 21(05):1850067, 2019. 2
- [Ser88] Jean-Pierre Serre. Algebraic groups and class fields. Springer, 1988. 8
- [Ser16] Jean-Pierre Serre. Lectures on  $N_X(p)$ . CRC Press, 2016. 3
- [Sta18] The Stacks Project Authors. Stacks Project. https://stacks.math.columbia.edu, 2018. 20
- [TV13] Michael Tsfasman and Serge G Vladut. Algebraic-geometric codes, volume 58. Springer Science & Business Media, 2013. 2
- [Wei48a] André Weil. Sur les courbes algébriques et les variétés qui s' en déduisent. Number 1041. Actualités Sci. Ind, 1948. 1, 7, 8
- [Wei48b] André Weil. Variétés abéliennes et courbes algébriques, volume 32. Paris, 1948. 7
- [Wei49] André Weil. Numbers of solutions of equations in finite fields. 1949. 1

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, IIT KANPUR, INDIA Email address: diptajit@cse.iitk.ac.in

Current address: https://www.cse.iitk.ac.in/users/nitin/ Email address: nitin@cse.iitk.ac.in

Email address: madhavan@cse.iitk.ac.in