Ramanujan Graphs and the Random Reducibility of Discrete Log on Isogenous Elliptic Curves

David Jao, Stephen D. Miller*, and Ramarathnam Venkatesan

November 13, 2004

Abstract

Cryptographic applications using an elliptic curve over a finite field filter curves for suitability using their order as the primary criterion: e.g. checking that their order has a large prime divisor before accepting it. It is therefore natural to ask whether the discrete log problem (DLOG) has the same difficulty for all curves with the same order; if so it would justify the above practice. We prove that this is essentially true by showing random reducibility of DLOG among such curves, assuming the Generalized Riemann Hypothesis (GRH). Our reduction proof works for curves with (nearly) the same endomorphism rings, but it is unclear if such a reduction exists in general. This suggests that in addition to the order, the conductor of its endomorphism ring may play a role. The random self-reducibility for DLOG over finite fields is well known; the non-trivial part here is that one must relate non-isomorphic algebraic groups of two isogenous curves. We construct certain expander graphs with elliptic curves as nodes and low degree isogenies as edges, and utilize the rapid mixing of random walks on this graph. We also briefly look at some recommended curves, compare "random" type NIST FIPS 186-2 curves to other special curves from this standpoint, and suggest a parameter to measure how generic a given curve is.

*Partially supported by NSF grant DMS-0301172 and an Alfred P. Sloan Foundation Fellowship.

Keywords: random reducibility, discrete log, elliptic curves, isogenies, modular forms, L-functions, generalized Riemann hypothesis, Ramanujan graphs, expanders, rapid mixing.

1 Introduction

Public key cryptosystems based on the elliptic curve discrete logarithm (DLOG) problem [22, 33] have received considerable attention because they are currently the most widely used systems whose underlying mathematical problem has yet to admit subexponential attacks (see [2, 31, 45]). Hence it is important to formally understand various aspects of DLOG and selection of elliptic curves. This turns out to be more intricate than the corresponding problem of DLOG over finite fields and their selection.

To motivate the questions in this paper, we begin with two observations. First, we note that one typically picks an elliptic curve at random, and examines its group order (e.g. to check if it is smooth) to decide whether to keep it, or discard it and pick another one. It is therefore a natural question whether or not DLOG is of the same difficulty on curves with the same number of points. Indeed, it is a theorem of Tate that curves E_1 and E_2 defined over the same finite field have the same number of points if and only if they are *isogenous*, i.e. there exists a nontrivial algebraic group homomorphism $\phi : E_1 \to E_2$ between them. If this ϕ is efficiently computable and has a smooth kernel, we can solve DLOG on E_1, given a DLOG oracle for E_2.

Secondly, we recall the observation that DLOG on $(\mathbb{Z}/p\mathbb{Z})^*$ has *random self-reducibility*: given any efficient algorithm $A(g^x) = x$ that solves DLOG on a polynomial fraction of inputs, one can solve *any* instance $y = g^x$ by an expected polynomial number of calls to A with *random* inputs of the form $A(g^r y)$. Thus DLOG must be hard for all but a negligible fraction of inputs, if it is hard on any polynomial fraction. This is comforting since for cryptographic use we need the DLOG problems to be hard with overwhelming probability when we pick inputs at random. Such a random self-reduction also holds true for any abelian group, and in particular for DLOG on a *fixed* elliptic curve. The question we consider is the following: given an efficient algorithm to solve DLOG on some μ-fraction of isogenous elliptic curves over $\mathbb{F}_q$, can we efficiently solve DLOG for all curves in the same isogeny class? If so, we would conclude that DLOG must either be easy for almost all curves, or else be hard for almost all curves. That of course would also give some justification for the cryptographic practice of selecting curves at random within an isogeny class. Unlike the above self-reductions which work on a fixed group, we must now relate instances of DLOG on an *arbitrary* E_1 to those on a *randomly distributed* E_2; computing $\phi : E_1 \to E_2$ makes the problem non-trivial.

Acknowledgements: It is a pleasure to thank William Aiello, Michael Ben-Or, Dan Boneh, Henryk Iwaniec, Dimitar Jetchev, Neal Koblitz, Peter Sarnak, and Adi Shamir for their discussions and helpful comments. We are also indebted to Peter Montgomery for his factoring assistance in producing Figure 1.

2 Our results

In this paper, we show that to a large extent one can indeed justify selecting curves (over a fixed finite field $\mathbb{F}_q$) at random, as well as treating isogeny class as the only essential invariant for DLOG. We state our results here; the reader may consult Sections 3 and 4 for definitions and background about elliptic curves and expander graphs, respectively.

It is very convenient to give each set of isogenous elliptic curves the structure of graph, whose nodes are arranged in stratified levels. This idea has arisen before in papers of Mestre [32] and Galbraith [10] (see also [9]). First one defines an *isogeny graph*, whose nodes represent elliptic curves over $\mathbb{F}_q$ and whose edges represent *isogenies* (algebraic group homomorphisms) defined over $\mathbb{F}_q$ having *degree* bounded by polylog(q). Here an isogeny $\phi\colon E_1 \to E_2$ of *degree* ℓ sends a point $P = (x, y) \in E_1$ to a point $Q \in E_2$ whose coordinates are given by a rational function of degree ℓ in $\mathbb{F}_q(x, y)$.

Next one places the nodes in levels: two curves E_1, E_2 are in the same *level* if they have identical endomorphism rings (denoted by $\mathrm{End}(E_i)$). For most "*random*" curves, the isogeny graph turns out to have only one level, so one may ignore the level restriction for now as a technicality. Our main contribution is the following theorem and its simple practical consequence stated in the next corollary (see the end of Section 5 for the proof of the Theorem, and Proposition 4.1 for the definition of "nearly Ramanujan" and the implication to the Corollary).

Theorem 2.1. *(Assuming GRH) The restriction of the isogeny graph to each level is an nearly Ramanujan graph.*

Corollary 2.2. *(Assuming GRH) At each level, the isogeny graph has the* rapid mixing property*: that is, starting from* any *given E_1, a random walk over the graph will reach any other curve E_2 with almost uniform probability (i.e. with exponentially (in $\log q$) small error) using a polynomial (in $\log q$) number of steps.*

In particular these graphs are connected. This can not be guaranteed for the ordinary case unless we take B on the order of $(\log q)^2$, whereas in the supersingular case $B \leq 3$ suffices (see also Remark 2.4). We conclude from the above:

Theorem 2.3. *(Assuming GRH) The* DLOG *problem on elliptic curves is random reducible in the following sense: given any algorithm A that solves* DLOG *on some μ-fraction of curves in a level, one can probabilistically solve* DLOG *on any given curve in the same level with $\frac{1}{\mu}\,\mathrm{polylog}(q)$ expected queries to A with random inputs.*

We call the algorithm A in the last theorem a *balanced attack* if it succeeds on a polynomial fraction of each level. For such an A, one may of course drop the level restriction in the theorem.

Remark 2.4. *(a) By earlier results of Pizer* [37]*, in the case of supersingular curves the GRH assumption and the level restriction can be dropped (see Appendix B). (b) The Ramanujan property was first defined in* [29]*. It characterizes the optimal separation between the two largest eigenvalues of the graph adjacency matrix, and implies the expansion property. Invariably the construction of explicit expanders and proof of their properties is nontrivial. The nomenclature stems from the fact that* [29] *used known cases of the Ramanujan conjectures in their proofs. The isogeny graph in the supersingular case is essentially a Ramanujan graph, and has been well studied in the literature* [18, 32, 36, 37]*. However, both the definition of the isogeny graphs in the ordinary case and the use of GRH as a tool for proving expansion are new. The method in fact gives a (conditional) simple new family of expanders on $(\mathbb{Z}/Q\mathbb{Z})^*$ for any $Q > 0$. The use and dependency of GRH here is akin to that in bounding the least quadratic non-residue mod p. (c) Random walk methods were introduced in* [10, 11] *as a heuristic for constructing isogenies. A contribution of this paper is that we validate the heuristic mixing assumptions used in that work, by providing provable bounds for random walk mixing probabilities (under the assumption of GRH). (d) One can perform a tighter* *analysis using the "set avoidance" results of* [1, 14] *or the Chernoff bounds for random walks* [13]*. (e) For subexponential (instead of polynomial time) reductions, one can replace the assumption of GRH by the weaker Lindelöf hypothesis (see the remarks at the end of Section 6).*

2.1 Using conductors to identify generic curves

Somewhat surprisingly, if the attack is not balanced then in addition to $\#E(\mathbb{F}_q)$ another parameter may have a role to play with respect to the hardness of DLOG. We will refer to this parameter as the *conductor* of E, denoted $c(E) = c(\mathrm{End}(E)) \in \mathbb{Z}_{>0}$, though this should not be confused with the arithmetic conductor (see the remarks after Theorem 3.1 for definitions). We define the *conductor gap* between two elliptic curves to be the largest prime which divides one of their conductors but not the other. For an integer n, let $P(n)$ denote the largest prime which divides n. In Section 7 we explain why $P(c(E))$ is typically quite small for random elliptic curves. In contrast, for curves isogenous to an anomalous binary curve or CM curve [23] empirical data suggests that the distribution of $P(c(E))$ is similar to that of $P(n)$ for random n, and thus is often quite large.

Lemma 2.5. *(a) Given two curves with a conductor gap m, an isogeny between them can be constructed in $O(m^4)$ field operations. (b) The conductor gap $m < 2\sqrt{q}$.*

For part (a) see Kohel [24]. Part (b) follows from the description following Theorem 3.1 that the $c(E)$ is a square factor of the discriminant, which is bounded by $4q$.

Remark 2.6. *(a) Theorem 2.3 states that* DLOG *is random reducible on curves on the same level, i.e. which have the same $c(E)$ and hence* $\mathrm{End}(E)$*. The same applies if the gap is bounded by a polynomial in* $\log q$*, since in this case navigation between levels is feasible* [24]*. (b) When the gap $m \approx \sqrt{q}$, even the construction referred to in Lemma 2.5(a) — which is currently the fastest known — becomes impractically slow. Thus if m is large, it is unclear if even a subexponential reduction for* DLOG *exists. Unless algorithms exist to overcome this large conductor gap, the* DLOG *on generic and special curves may indeed have different hardness; we may also not know which is harder— akin to factorization of random integers, the problem may be easier in the average case. Random walks cannot bridge large gaps between levels (see*

Theorem 3.1, part 5). Moreover, even given a polynomial time algorithm to attack DLOG *on elliptic curves, it is currently unclear whether or not there is a subexponential algorithm to compute an isogeny between curves that have a conductor gap of size q^{ε}, for any fixed $\varepsilon > 0$.*

Table of $P(c(E))$ for recommended curves: We can view the above levelled graph as a pyramid, with the number of curves at each level rapidly increasing as we go down (i.e. as $c(E)$ increases). Most "random" or generic curves belong to pyramids with very few levels, since $c(E)$ is small (approximately 1), the exception being anomalous binary curves or CM curves, which sit on the top level of a large pyramid containing curves with large $P(c(E))$. A fuller discussion is found in Section 7. The structure of the edges between levels is described in Theorem 3.2.

Remark 2.7. *One may consider $c(E)$ as a clear way to measure how generic a given curve is. Sometimes in order to convince others that a curve is not specially chosen, one gives the seed of a secure hash based generator for it. However, the seed may have been picked with a large number of trials or the hash function may have admitted some compromise. For this reason, it may be a good standard practice to reveal $c(E)$, preferably by giving the complete factorization of the discriminant of the characteristic polynomial of Frobenius. In Figure 1 we have indeed computed $c(E)$ for all the curves in the FIPS 186-2 standard; the computations took roughly a week on a cluster of 100 computers and would not be within the reach of most users.*

3 Preliminaries

Let E_1 and E_2 be elliptic curves defined over a finite field $\mathbb{F}_q$ of characteristic $p > 3$. An isogeny $\phi\colon E_1 \to E_2$ defined over $\mathbb{F}_q$ is a non-constant rational map defined over $\mathbb{F}_q$ which is also a group homomorphism from $E_1(\mathbb{F}_q)$ to $E_2(\mathbb{F}_q)$ [41, §III.4]. The degree of an isogeny is its degree as a rational map. Every isogeny of degree greater than 1 can be factored into a composition of isogenies of prime degree [41, §II.2.12 and §III.4.12]. For any elliptic curve $E\colon y^2 = x^3 + ax + b$ defined over $\mathbb{F}_q$, the Frobenius isogeny is the isogeny $\pi\colon E \to E$ given by the equation $\pi(x, y) = (x^q, y^q)$. It satisfies the equation

$$\pi^2 - \mathrm{Trace}(E)\pi + q = 0,$$

Curve	**c (maximal conductor gap amongst isogenous curves)**	**P(c) = largest prime factor of c**
NIST P-192	1	1
NIST P-256	3	3
NIST P-384	1	1
NIST P-521	1	1
NIST K-163	$45641 \cdot 82153 \cdot 56498081 \cdot P(c)$	86110311
NIST K-233	$5610641 \cdot 85310626991 \cdot P(c)$	150532234816721999
NIST K-283	$1697 \cdot 162254089 \cdot P(c)$	1779143207551652584836995286271
NIST K-409	$21262439877311 \cdot 22431439539154506863 \cdot P(c)$	57030553306655053533734286593
NIST K-571	$3952463 \cdot P(c)$	$9021184135396238924389891(contd)$ $9451926768145189936450898(contd)$ 07769277009849103733654828039
NIST B-163	1	1
NIST B-233	1	1
NIST B-283	1	1
NIST B-409	1	1
NIST B-571	1	1
IPSec 3^{rd}OG,$F_{2^{155}}$	1	1
IPSec $4^{th} OG$, $F_{2^{185}}$	1	1

Figure 1: A table of curves recommended as international standards [16, 35]. Note that the maximum possible conductor gap between each standards curve and its isogenous curves is small (at most 3), except for the curves in the NIST K (=Koblitz curve) family.

where $\mathrm{Trace}(E) = q + 1 - \#E(\mathbb{F}_q)$ is the trace of the curve E over $\mathbb{F}_q$. The polynomial $p(X) := X^2 - \mathrm{Trace}(E)X + q$ is called the characteristic polynomial of E.

An endomorphism of an elliptic curve E defined over $\mathbb{F}_q$ is an isogeny $E \to E$ defined over $\mathbb{F}_{q^m}$ for some m. The set of endomorphisms of E together with the zero map forms a ring under the operations of pointwise addition and composition; this ring is called the endomorphism ring of E and denoted $\mathrm{End}(E)$. The ring $\mathrm{End}(E)$ is isomorphic either to an order in a quaternion algebra or to an order in an imaginary quadratic field [41, V.3.1]; in the first case we say E is supersingular and in the second case we say E is ordinary. In the latter situation, the Frobenius isogeny π can be regarded as an algebraic integer which is a root of the characteristic polynomial.

Two elliptic curves E_1 and E_2 defined over $\mathbb{F}_q$ are said to be isogenous over $\mathbb{F}_q$ if there exists an isogeny $\phi \colon E_1 \to E_2$ defined over $\mathbb{F}_q$. A theorem of Tate states that two curves E_1 and E_2 are isogenous over $\mathbb{F}_q$ if and only if $\#E_1(\mathbb{F}_q) = \#E_2(\mathbb{F}_q)$ [42, §3]. Hence the property of being isogenous is an equivalence relation. We define an isogeny class to be an equivalence class of elliptic curves under this relation. Curves in the same isogeny class are either all supersingular or all ordinary. We assume for the remainder of this paper that we are in the ordinary case, which is the more interesting case from the point of view of cryptography anyhow. The supersingular case is discussed further in Appendix B.

The following theorem describes the structure of elliptic curves within an isogeny class from the point of view of their endomorphism rings.

Theorem 3.1. *Let E and E' be ordinary elliptic curves defined over $\mathbb{F}_q$ which are isogenous over $\mathbb{F}_q$. Let K denote the imaginary quadratic field containing* $\mathrm{End}(E)$*, and write $\mathcal{O}_K$ for the maximal order (i.e. ring of integers) of K.*

1. *The order* $\mathrm{End}(E)$ *satisfies the property* $\mathbb{Z}[\pi] \subseteq \mathrm{End}(E) \subseteq \mathcal{O}_K$.
2. *The order* $\mathrm{End}(E')$ *also satisfies* $\mathrm{End}(E') \subset K$ *and* $\mathbb{Z}[\pi] \subseteq \mathrm{End}(E') \subseteq \mathcal{O}_K$.
3. *The following are equivalent:*
   (a) $\mathrm{End}(E) = \mathrm{End}(E')$.
   (b) *There exist two isogenies $\phi\colon E \to E'$ and $\psi\colon E \to E'$ of relatively prime degree.*
   (c) $[\mathcal{O}_K : \mathrm{End}(E)] = [\mathcal{O}_K : \mathrm{End}(E')]$.
   (d) $[\mathrm{End}(E) : \mathbb{Z}[\pi]] = [\mathrm{End}(E') : \mathbb{Z}[\pi]]$.
4. *Let $\phi : E \to E'$ be an isogeny from E to E' of prime degree ℓ. Then either* $\mathrm{End}(E)$ *contains* $\mathrm{End}(E')$ *or* $\mathrm{End}(E')$ *contains* $\mathrm{End}(E)$*, and the index of the smaller in the larger divides ℓ.*
5. *Suppose ℓ is a prime that divides one of* $[\mathcal{O}_K : \mathrm{End}(E)]$ *and* $[\mathcal{O}_K : \mathrm{End}(E')]$*, but not the other. Then every isogeny $\phi\colon E \to E'$ has degree equal to a multiple of ℓ.*

Proof. [24, §4.2]. □

For any order $\mathcal{O} \subseteq \mathcal{O}_K$, the conductor of $\mathcal{O}$ is defined to be the integer $[\mathcal{O}_K : \mathcal{O}]$. The field K is called the CM field of E. We write c_E for the conductor of $\mathrm{End}(E)$ and c_π for the conductor of $\mathbb{Z}[\pi]$. Note that this is not the same thing as the arithmetic conductor of an elliptic curve [41, §C.16], nor is it related to the conductance of an expander graph [21]. It follows from [4, (7.2) and (7.3)] that $\mathrm{End}(E) = \mathbb{Z} + c_E\mathcal{O}_K$ and $D = c_E^2 d_K$, where D (respectively, d_K) is the discriminant of the order $\mathrm{End}(E)$ (respectively, $\mathcal{O}_K$). Applying the same reasoning to $\mathbb{Z}[\pi]$, we find that the characteristic polynomial $p(X)$ has discriminant $\mathrm{disc}(p(X)) = \mathrm{Trace}(E)^2 - 4q = \mathrm{disc}(\mathbb{Z}[\pi]) = c_\pi^2 d_K$, with $c_\pi = c_E \cdot [\mathrm{End}(E) : \mathbb{Z}[\pi]]$.

Following [9] and [10], we say that an isogeny $\phi : E_1 \to E_2$ of prime degree ℓ is "down" if $[\mathrm{End}(E_1) : \mathrm{End}(E_2)] = \ell$, "up" if $[\mathrm{End}(E_2) : \mathrm{End}(E_1)] = \ell$, and "horizontal" if $\mathrm{End}(E_1) = \mathrm{End}(E_2)$. The following theorem classifies the number of degree ℓ isogenies of each type in terms of the Legendre symbol $\left(\frac{D}{\ell}\right)$.

Theorem 3.2. *Let E be an ordinary elliptic curve over $\mathbb{F}_q$, with endomorphism ring $\mathrm{End}(E)$ of discriminant D. Let ℓ be a prime different from the characteristic of $\mathbb{F}_q$.*

- *Assume $\ell \nmid c_E$. Then there are exactly $1 + \left(\frac{D}{\ell}\right)$ horizontal isogenies $\phi : E \to E'$ of degree ℓ.*
  - *If $\ell \nmid c_\pi$, there are no other isogenies $E \to E'$ of degree ℓ.*
  - *If $\ell \mid c_\pi$, there are $\ell - \left(\frac{D}{\ell}\right)$ down isogenies of degree ℓ.*
- *Assume $\ell \mid c_E$. Then there is one up isogeny $E \to E'$ of degree ℓ.*
  - *If $\ell \nmid \frac{c_\pi}{c_E}$, there are no other isogenies $E \to E'$ of degree ℓ.*
  - *If $\ell \mid \frac{c_\pi}{c_E}$, there are ℓ down isogenies of degree ℓ.*

Proof. [9, §2.1] or [10, §11.5]. □

It follows that the maximal conductor gap referred to in Section 2.1 is achieved between a curve at the top level (with $\mathrm{End}(E) = \mathcal{O}_K$) and a curve at the bottom level (with $\mathrm{End}(E) = \mathbb{Z}[\pi]$). The structure of these isogenies is diagrammed in Appendix D.

3.1 Isogeny Graphs

Recall from Section 1 that an isogeny graph is a graph whose nodes consist of all elliptic curves in $\mathbb{F}_q$ belonging to a fixed isogeny class, up to $\mathbb{F}_q$-isomorphism (so that two elliptic curves which are isomorphic over $\mathbb{F}_q$ represent the same node in the graph). We define two curves E_1 and E_2 to have the same level if $\mathrm{End}(E_1) = \mathrm{End}(E_2)$. Note that a horizontal isogeny always goes between two curves of the same level; likewise, an up isogeny enlarges the size of the endomorphism ring and a down isogeny reduces the size. Since there are fewer elliptic curves at higher levels than at lower levels, the isogeny graph under the level interpretation visually resembles a "pyramid" or a "volcano" [9], with up isogenies ascending the structure and down isogenies descending.

The edges of the graph consist of the isogenies between such elliptic curves having prime degree less than the bound $(\log q)^{2+\delta}$ for some fixed $\delta > 0$. The degree bound must be small enough to permit the isogenies to be computed, but large enough to allow the graph to be connected and to have the rapid mixing properties that we want. We will show in Section 5 that a bound of $(\log q)^{2+\delta}$ satisfies all the requirements, provided that we restrict the isogeny graph to a single level.

Accordingly, fix a level of the isogeny graph, and let $\mathrm{End}(E) = \mathcal{O}$ be the common endomorphism ring of all of the elliptic curves in this level. Denote by $\mathcal{G}$ the graph whose vertices are elliptic curves over $\mathbb{F}_q$ with endomorphism ring $\mathcal{O}$, and whose edges are horizontal isogenies defined over $\mathbb{F}_q$ of prime degree $\leq (\log q)^{2+\delta}$. By standard facts from the theory of complex multiplication [4, §10], each ideal $\mathfrak{a} \subset \mathcal{O}$ produces an elliptic curve $\mathbb{C}/\mathfrak{a}$ defined over some number field $L \subset \mathbb{C}$ (called the ring class field of $\mathcal{O}$) [4, §11]. The curve $\mathbb{C}/\mathfrak{a}$ has complex multiplication by $\mathcal{O}$, and two different ideals yield isomorphic curves if and only if they belong to the same ideal class. Likewise, each ideal $\mathfrak{b} \subset \mathcal{O}$ defines an isogeny $\mathbb{C}/\mathfrak{a} \to \mathbb{C}/\mathfrak{a}\mathfrak{b}^{-1}$, and the degree of this isogeny is the norm $N(\mathfrak{b})$ of the ideal $\mathfrak{b}$. Moreover, for any prime ideal $\mathfrak{P}$ in L lying over p, the reductions mod $\mathfrak{P}$ of the above elliptic curves and isogenies are defined over $\mathbb{F}_q$, and every elliptic curve and every horizontal isogeny in $\mathcal{G}$ arises in this way [10, §3]. Therefore, the graph $\mathcal{G}$ is isomorphic to the corresponding graph $\mathcal{H}$ whose nodes are elliptic curves $\mathbb{C}/\mathfrak{a}$ with complex multiplication by $\mathcal{O}$, and whose edges are complex analytic isogenies represented by ideals $\mathfrak{b} \subset \mathcal{O}$ and subject to the same degree bound as before. This isomorphism preserves the degrees of isogenies, in the sense that the degree of any isogeny in $\mathcal{G}$ is equal to the norm of its corresponding ideal $\mathfrak{b}$ in $\mathcal{H}$.

The graph $\mathcal{H}$ has an alternate description as a Cayley graph on the ideal class group $\mathrm{Cl}(\mathcal{O})$ of $\mathcal{O}$. Indeed, each node of $\mathcal{H}$ is an ideal class of $\mathcal{O}$, and two ideal classes $[\mathfrak{a}_1]$ and $[\mathfrak{a}_2]$ are connected by an edge if and only if there exists a prime ideal $\mathfrak{b}$ of norm $\leq (\log q)^{2+\delta}$ such that $[\mathfrak{a}_1\mathfrak{b}] = [\mathfrak{a}_2]$. Therefore, the graph $\mathcal{H}$ (and hence the graph $\mathcal{G}$) is isomorphic to the Cayley graph of the group $\mathrm{Cl}(\mathcal{O})$ with respect to the generators $[\mathfrak{b}] \in \mathrm{Cl}(\mathcal{O})$, as $\mathfrak{b}$ ranges over all prime ideals of $\mathcal{O}$ of norm $\leq (\log q)^{2+\delta}$.

Remark 3.3. *The graph $\mathcal{G}$ consists of objects defined over the finite field $\mathbb{F}_q$, whereas the objects in the graph $\mathcal{H}$ are defined over the number field L. One passes from $\mathcal{H}$ to $\mathcal{G}$ by taking reductions mod $\mathfrak{P}$, and from $\mathcal{G}$ to $\mathcal{H}$ by using Deuring's Lifting Theorem* [7, 10, 25]. *There is no known polynomial* *time or even subexponential time algorithm for computing the isomorphism between $\mathcal{G}$ and $\mathcal{H}$* [10, §3]. *For our purposes, such an explicit algorithm is not necessary, since we only use the complex analytic theory to prove abstract graph-theoretic properties of $\mathcal{G}$.*

Remark 3.4. A priori*, the graph $\mathcal{G}$ is a directed graph, since an isogeny $\phi\colon E_1 \to E_2$ is an asymmetric relation between E_1 and E_2. However, the graph is in fact symmetric, because every isogeny ϕ has a unique dual isogeny $\hat{\phi}\colon E_2 \to E_1$ of the same degree as ϕ in the opposite direction* [41, §III.6]. *From the viewpoint of $\mathcal{H}$, an isogeny represented by an ideal $\mathfrak{b} \subset \mathcal{O}$ has its dual isogeny represented simply by the complex conjugate $\bar{\mathfrak{b}}$. This symmetry allows us to treat $\mathcal{G}$ as undirected and to apply known results about undirected expander graphs (as in the following section) to $\mathcal{G}$.*

4 Expander Graphs

Let $G = (\mathcal{V}, E)$ be a finite graph on h vertices $\mathcal{V}$ with undirected edges $\mathcal{E}$. Suppose G is a regular graph of degree k, i.e. exactly k edges meet at each vertex. Given a labelling of the vertices $\mathcal{V} = \{v_1, \ldots, v_h\}$, the adjacency matrix of G is the symmetric $h \times h$ matrix A whose ij-th entry $A_{i,j} = 1$ if an edge exists between v_i and v_j and 0 otherwise.

It is convenient to identify functions on $\mathcal{V}$ with vectors in $\mathbb{R}^h$ via this labelling, and therefore also think of A as a self-adjoint operator on $L^2(\mathcal{V})$. All of the eigenvalues of A satisfy the bound $|\lambda| \leq k$. Constant vectors are eigenfunctions of A with eigenvalue k, which for obvious reasons is called the trivial eigenvalue λ_{triv}. A family of such graphs G with $h \to \infty$ is said to be a sequence of *expander graphs* if all other eigenvalues of their adjacency matrices are bounded away from $\lambda_{\text{triv}} = k$ by a fixed amount.[1] In particular, no other eigenvalue is equal to k; this implies the graph is connected. A *Ramanujan graph* [29] is a special type of expander which has $|\lambda| \leq 2\sqrt{k-1}$ for any nontrivial eigenvalue which is not equal to $-k$ (this last possibility happens if and only if the graph is bipartite). The supersingular isogeny graphs in Appendix B are sometimes Ramanujan, while the ordinary isogeny graphs in Section 3.1 are often not considered to be, partly because

[1]Expansion is usually phrased in terms of the number of neighbors of subsets of G, but the spectral condition here is equivalent for k-regular graphs and also more useful for our purposes.

their degree is not bounded. Nevertheless, they still share the most important properties of expanders as far as our applications are concerned. In particular their degree k grows slowly (as a polynomial in $\log |\mathcal{V}|$), and they share a qualitatively similar eigenvalue separation: instead the nontrivial eigenvalues λ can be arranged to be $O(k^{1/2+\varepsilon})$ for any desired value of $\varepsilon > 0$. Of course the exponent of k can sometimes be reduced by increasing k (which in fact happens for our isogeny graphs). Obtaining *any* nontrivial exponent is really the key challenge for many applications, and accordingly we shall focus on what we call "nearly Ramanujan" graphs: families of graphs whose nontrivial eigenvalues satisfy the bound $\lambda = O(k^{\beta})$ for some $\beta < 1$.

A fundamental use of expanders is to prove the rapid mixing of the random walk on $\mathcal{V}$ along the edges $\mathcal{E}$. For convenience we present the rapid mixing result below and its proof in Appendix A. For more information, see [5, 28, 39].

Proposition 4.1. *Let G be a regular graph of degree k on h vertices. Suppose that the eigenvalue λ of any nonconstant eigenvector satisfies the bound $|\lambda| \leq c$ for some $c < k$. Let S be any subset of the vertices of G, and x be any vertex in G. Then a random walk of length at least $\frac{\log 2h/|S|^{1/2}}{\log k/c}$ starting from x will land in S with probability at least $\frac{|S|}{2h} = \frac{|S|}{2|G|}$.*

In our application the quantities k, $\frac{k}{c} - 1$, and $\frac{h}{|S|}$ will all be bounded by polynomials in $\log(h)$. Under these hypotheses, the probability is at least $1/2$ that some polylog(h) trials of random walks of polylog(h) length starting from x will reach S at least once. This mixing estimate is the source of our polynomial time random reducibility (Theorem 2.3).

5 Proofs of our results

5.1 Navigating the Isogeny Graph

Let $\mathcal{G}$, $\mathcal{H}$, and $\mathcal{O}$ be as in Section 3.1. The graph $\mathcal{G}$ has exponentially many nodes and thus is too large to be stored. However, given a curve E and a prime ℓ, it is possible to efficiently compute the curves which are connected to E by an isogeny of degree ℓ. These curves E' have j-invariants which can be found by solving the modular polynomial relation $\Phi_\ell(j(E), j(E')) = 0$; the cost of this step is $O(\ell^3)$ field operations [10, 11.6]. In this way, it is possible to navigate the isogeny graph locally without computing the entire graph. We shall see that it suffices to have the degree of the isogenies in the graph be bounded by $(\log q)^{2+\delta}$ to assure the Ramanujan properties required for $\mathcal{G}$ to be an expander.

5.2 θ-Functions and Graph Eigenvalues

The graph $\mathcal{H}$ (and therefore also the graph $\mathcal{G}$) has one node for each ideal class of $\mathcal{O}$. Therefore, the total number of nodes in the graph $\mathcal{G}$ is the ideal class number of the order $\mathcal{O}$, and the vertices $\mathcal{V}$ can be identified with ideal class representatives $\{\alpha_1, \ldots, \alpha_h\}$. Using the isomorphism between $\mathcal{G}$ and $\mathcal{H}$, we see that the generating function $\sum M_{\alpha_i,\alpha_j}(n)q^n$ for degree n isogenies between the vertices α_i and α_j of $\mathcal{G}$ is given by

$$\sum_{n=1}^{\infty} M_{\alpha_i,\alpha_j}(n)\, q^n \quad := \quad \frac{1}{e} \sum_{z \,\in\, \alpha_i^{-1}\alpha_j} q^{N(z)/N(\alpha_i^{-1}\alpha_j)}\,, \tag{5.1}$$

where e is the number of units in $\mathcal{O}$ (which always equals 2 for $\operatorname{disc}(\mathcal{O}) > 4$). The sum on the right hand side depends only on the ideal class of the fractional ideal $\alpha_i^{-1}\alpha_j$; by viewing the latter as a lattice in $\mathbb{C}$, we see that $N(z)/N(\alpha_i^{-1}\alpha_j)$ is a quadratic form of discriminant D where $D := \operatorname{disc}(\mathcal{O})$ [4, p. 142]. That means this sum is a θ-series, accordingly denoted as $\theta_{\alpha_i^{-1}\alpha_j}(q)$. It is a holomorphic modular form of weight 1 for the congruence subgroup $\Gamma_0(|D|)$ of $SL(2,\mathbb{Z})$, transforming according to the character $\left(\frac{D}{\cdot}\right)$ (see [19, Theorem 10.9]).

Before discussing exactly which degrees of isogenies to admit into our isogeny graph $\mathcal{G}$, let us first make some remarks about the simpler graph on $\mathcal{V} = \{\alpha_1, \ldots, \alpha_h\}$ whose edges represent isogenies of degree exactly equal to n. Its adjacency matrix is of course the $h \times h$ matrix $M(n) = \left[M_{\alpha_i,\alpha_j}(n)\right]_{\{1 \le i,j \le h\}}$ defined by series coefficients in (5.1). It can be naturally viewed as an operator which acts on functions on $\mathcal{V} = \{\alpha_1, \ldots, \alpha_h\}$, by identifying them with h-vectors according to this labelling. We will now simultaneously diagonalize all $M(n)$, or what amounts to the same, diagonalize the matrix $A_q = \sum_{n \ge 1} M(n)q^n$ for any value of $q < 1$ (where the sum converges absolutely). The primary reason this is possible is that for each fixed n this graph is an abelian Cayley graph on the ideal class group $\mathrm{Cl}(\mathcal{O})$, with generating set equal to those classes α_i which represent an n-isogeny. The eigenfunctions of the adjacency matrix of an abelian Cayley graph are always given by characters of the group (viewed as functions on the graph), and their respective eigenvalues are sums of these characters over the generating set. This can be seen in our circumstance as follows. The ij-th entry of A_q is $\frac{1}{e}\theta_{\alpha_i^{-1}\alpha_j}(q)$, which we recall depends only on the ideal class of the fractional ideal $\alpha_i^{-1}\alpha_j$. If χ is any character of $\mathrm{Cl}(\mathcal{O})$, viewed as the h-vector whose i-th entry is $\chi(\alpha_i)$, then the i-th entry of the vector $A_q\chi$ may be evaluated through matrix multiplication as

$$(A_q\chi)(\alpha_i) \;=\; \frac{1}{e}\sum_{\alpha_j\,\in\,\mathrm{Cl}(\mathcal{O})} \theta_{\alpha_i^{-1}\alpha_j}(q)\,\chi(\alpha_j) \;=\; \frac{1}{e}\left(\sum_{\alpha_j\,\in\,\mathrm{Cl}(\mathcal{O})}\chi(\alpha_j)\,\theta_{\alpha_j}(q)\right)\chi(\alpha_i)\,, \tag{5.2}$$

where in the last equality we have reindexed $\alpha_j \mapsto \alpha_i\,\alpha_j$ using the group structure of $\mathrm{Cl}(\mathcal{O})$. Therefore χ is in fact an eigenvector of the matrix eA_q, with eigenvalue equal to the sum of θ-functions enclosed in parentheses, known as a *Hecke θ-function* (see [19, §12]). These, which we shall denote $\theta_\chi(q)$, form a more natural basis of modular forms than the ideal class θ-functions θ_{α_j} because they are in fact they are Hecke eigenforms. Using (5.1), the L-functions of these Hecke characters can be written as

$$L(s,\chi) \;=\; L(s,\theta_\chi) \;=\; \sum_{\text{integral ideals }\mathfrak{a}\subset K}\chi(\mathfrak{a})\,(N\mathfrak{a})^{-s} \;=\; \sum_{n=1}^{\infty} a_n(\chi)\,n^{-s}\,, \tag{5.3}$$

where

$$a_n(\chi) \;=\; \sum_{\substack{\text{integral ideals }\mathfrak{a}\subset K\\ N\mathfrak{a}=n}}\chi(\mathfrak{a}) \tag{5.4}$$

is in fact simply the eigenvalue of $e\,M(n)$ for the eigenvector formed from the character χ as above, which can be seen by isolating the coefficient of q^n in the sum on the right hand side of (5.2).

5.3 The Isogeny Graph

Our isogeny graph is a superposition of the previous graphs $M(n)$, when n is a prime bounded by a parameter m (which we recall is $(\log q)^{2+\delta}$ for some fixed $\delta > 0$). This corresponds to a graph on the elliptic curves represented by ideal classes in an order $\mathcal{O}$ of $K = \mathbb{Q}(\sqrt{d})$, whose edges represent isogenies of prime degree $\leq m$. The graphs with adjacency matrices $\{M(p) \mid p \leq m\}$

above share common eigenfunctions (the characters χ of $\mathrm{Cl}(\mathcal{O})$), and so their eigenvalues are of the form

$$\lambda_\chi \quad = \quad \frac{1}{e}\sum_{p\leq m} a_p(\chi) \quad = \quad \frac{1}{e}\sum_{p\leq m}\ \sum_{\substack{\text{integral ideals } \mathfrak{a}\subset K\\ N\mathfrak{a}=p}} \chi(\mathfrak{a})\,. \tag{5.5}$$

When χ is the trivial character, λ_{triv} equals the degree of the regular graph $\mathcal{G}$. Since roughly half of rational primes p split in K, and those which do split into two ideals of norm p, λ_{triv} is roughly $\frac{\pi(m)}{e} \sim \frac{m}{e\log m}$ by the prime number theorem. This eigenvalue is always the largest in absolute value, as can be deduced from (5.5) because $|\chi(\mathfrak{a})|$ always equals 1. For the polynomial mixing of the random walk in Corollary 2.2 we will require a separation between the trivial and nontrivial eigenvalues of size $1/\mathrm{polylog}(m)$. This would be the case, for example, if for each nontrivial character χ there merely exists one ideal $\mathfrak{a}$ of prime norm $\leq m$ with $\mathrm{Re}\,\chi(\mathfrak{a}) \leq 1 - \frac{1}{\mathrm{polylog}(m)}$. This is analogous to the problem of finding a small prime nonresidue modulo say a large prime Q, where one merely needs to find any cancellation at all in the character sum $\sum_{p\leq m}\left(\frac{p}{Q}\right)$. However, the latter requires a strong assumption from analytic number theory, such as the Generalized Riemann Hypothesis (GRH). In the next section we will accordingly derive such bounds for λ_χ, under the assumption of GRH. As a consequence of the more general Lemma 6.3 we will show the following.

Lemma 5.1. *Let $D<0$ and let $\mathcal{O}$ be the quadratic order of discriminant D. If χ is a nontrivial ideal class character of $\mathcal{O}$, then the Generalized Riemann Hypothesis for $L(s,\chi)$ implies that the sum (5.5) is bounded by $O(m^{1/2}\log|D|)$ when $m = \mathrm{polylog}(|D|)$.*

Proof of Theorem 2.1: This follows from the lemma, as the eigenvalues of the adjacency matrix for a given level are given by (5.5). In particular we have chosen $m = (\log q)^{2+\delta}$, and so the nontrivial eigenvalues are bounded by $\lambda_\chi = O(\lambda_{\text{triv}}^{\beta})$ for any $\beta > \frac{1}{2} + \frac{1}{\delta+2}$. $\square$

6 The Prime Number Theorem for Modular Form L-functions

In this section we prove Lemma 5.1, assuming the Generalized Riemann Hypothesis (GRH) for the L-functions (5.3). Our argument is more general, and in fact gives estimates for sums of the form $\sum_{p\leq m} a_p$, where a_p are the prime coefficients of any L-function. This can be thought of as an analog of the Prime Number Theorem because for the simplest L-function, $\zeta(s)$, $a_p = 1$ and this sum is in fact exactly $\pi(x)$. As a compromise between readability and generality, we will restrict the presentation here to the case of modular form L-functions (including (5.3)). Background references for this section include [19, 20, 34]; for information about more general L-functions see also [12, 38].

We shall now consider a classical holomorphic modular form f, with Fourier expansion

$$f(z) \;=\; \sum_{n=0}^{\infty} c_n\, e^{2\pi i n z}\,. \tag{6.1}$$

We will assume that f is a Hecke eigenform, since this condition is met in the situation of Lemma 5.1 (see the comments between (5.2) and (5.3)). It is natural to study the renormalized coefficients $a_n = n^{-(k-1)/2}c_n$, where $k \geq 1$ is the weight of f (in Section 5.2 $k = 1$, so $a_n = c_n$). The L-function of such a modular form can be written as the Dirichlet series

$$L(s,f) \;=\; \sum_{n=1}^{\infty} a_n\, n^{-s} \;=\; \prod_p (1-\alpha_p\, p^{-s})^{-1}\,(1-\beta_p\, p^{-s})^{-1}\,, \tag{6.2}$$

the last equality using the fact that f is a Hecke eigenform. The L-function $L(s,f)$ is entire when f is a cusp form (e.g. $a_0 = 0$). The Ramanujan conjecture (in this case a theorem of Deligne [6]) asserts that $|\alpha_p|, |\beta_p| \leq 1$.

Lemma 5.1 is concerned with estimates for the sums

$$S(m,f) \;:=\; \textstyle\sum_{p\,\leq\, m} a_p\,. \tag{6.3}$$

As with the prime number theorem, it is more convenient to instead analyze the weighted sum

$$\psi(m,f) \;:=\; \textstyle\sum_{p^k} b_{p^k}\, \log p \tag{6.4}$$

over prime powers, where the coefficients b_n are those appearing in the Dirichlet series for $-\frac{L'}{L}(s)$:

$$-\tfrac{L'}{L}(s) \;=\; \textstyle\sum_{n=1}^{\infty} b_n\, \Lambda(n)\, n^{-s} \;=\; \sum_{p,\,k} b_{p^k}\, \log(p)\, p^{-k\,s}\,,$$

i.e. $b_{p^k} \;=\; \alpha_p^k \;+\; \beta_p^k$.

Lemma 6.1. *For a holomorphic modular form f one has* $\psi(m,f) = \sum_{p\le m} a_p \log p \;+\; O(m^{1/2})$.

Proof. The error term represents the contribution of proper prime powers. Since $|b_{p^k}| \le 2$, it is bounded by twice

$$\sum_{\substack{p^k \le m \\ k\ge 2}} \log p \quad\le\quad \sum_{\substack{p\le m^{1/2} \\ 2\le k \le \frac{\log m}{\log p}}} \log p \quad\le\quad \sum_{p\le m^{1/2}} \log p \,\frac{\log m}{\log p} \quad\le\quad \pi(m^{1/2})\log m\,, \tag{6.5}$$

which is $O(m^{1/2})$ by the Prime Number Theorem. □

Lemma 6.2. *(Iwaniec [20, p. 114]) Assume that f is a holomorphic modular cusp form of level[2] N and that $L(s,f)$ satisfies GRH. Then $\psi(m,f) = O(m^{1/2}\log(m)\log(m^2N))$.*

We deduce that

$$S'(m,f) \;:=\; \textstyle\sum_{p\le m} a_p \log p \;=\; O(m^{1/2}\log(m)\log(N)) \quad \text{for} \;\; m \;=\; O(N)\,. \tag{6.6}$$

Finally we shall estimate sums $S(m,f)$ from (6.3) by removing the $\log p$ using a standard partial summation argument. We have included a proof in Appendix C for completeness.

Lemma 6.3. *Suppose that f is a holomorphic modular cusp form of level N and $L(s,f)$ satisfies GRH. Then for $m = O(N)$ one has that $S(m,f) = O(m^{1/2}\log N)$.*

Subexponential Reductions via Lindelöf Hypothesis: In this last lemma we have assumed GRH. It seems very difficult to get a corresponding unconditional bound for $S(m,f)$. However, a slightly weaker statement can be proven by assuming only the Lindelöf hypothesis (which is a consequence of GRH). Namely, one has that

[2] Actually in [20] N equals the conductor of the L-function, which in general may be smaller than the level. The lemma is of course nevertheless valid.

$$\sum_{n \le m} a_n \;=\; O_\varepsilon(m^{1/2+\varepsilon} N^\varepsilon)\ , \quad \text{for any } \ \varepsilon > 0 \tag{6.7}$$

([19, (5.61)]). The fact that this last sum is over all $n \le m$, not just primes, is not of crucial importance for our applications. However, the significant difference here is that the dependence on N is not polynomial in $\log N$, but merely subexponential. This observation can be used to weaken the hypothesis in Theorem 2.3 from GRH to the Lindelöf hypothesis, at the expense of replacing "polynomial" by "subexponential".

7 Distribution of $c(E)$

Recall that the largest prime divisor of $c(\pi)$ for "random" NIST curves in Figure 1 was ≤ 3, while the corresponding values for Koblitz curves were quite large. We now present some elementary arguments to explain these observations.

First we consider the small size of $c(\pi)$ for the random NIST curves. Recall from Section 3 that $c(\pi)$ is the square part of the discriminant D, which is equal to $D = t^2 - 4q$, where $t = \mathrm{Trace}(E)$. Statistically speaking, most integers (a proportion of $\frac{6}{\pi^2} \approx .61$) are square-free, explaining why $c(E)$ would often be 1 or at least fairly small [44].

One can say more about the expected sizes of the largest prime factor of $c(\pi)$, i.e. the largest prime which divides D to order at least 2. By a result of Lenstra [26], the trace t has a fairly uniform distribution — at least as far as the q-aspect is concerned — in $[-2\sqrt{q}, 2\sqrt{q}]$. Namely, for any subset S of integers contained in $H = (q+1-2\sqrt{q}, q+1+2\sqrt{q})$,

$$\tfrac{|S|-2}{|H|}\,\tfrac{1}{(\log q)} \;\ll\; \mathrm{Prob}[\#E(F_q) \in S] \;\ll\; \tfrac{|S|}{|H|}\,(\log q\ \log\log q)^2\,;$$

this probability is taken as one samples over a suitable family of elliptic curves. Of course this is still much cruder than the precise conjectured answers (e.g. Sato-Tate) described for example in [40]. Nevertheless, it serves to provide the useful heuristic that $-D = 4q - t^2$ is typically of size q.

With this in mind, let us now return to the issue of estimating how frequently elliptic curves will have a large conductor gap. We mentioned in Remark 2.6 that our random reducibility result applies between curves whose levels have a conductor gap which is bounded by a polynomial in $\log q$. Thus there is a random reducibility to curves with level $c(E) = 1$ from any level for which $c(E)$'s prime factors are all bounded by $\beta = \text{polylog}(q)$. Given the heuristic that $-D$ is a random number of (the much larger) size q, the probability that $c(E)$ has a prime factor exceeding β can be loosely estimated as $O(1/\beta)$. This is because roughly a fraction of $\rho = \prod_{p>\beta}^{\sqrt{q}}(1 - p^{-2})$ integers of size q have no square prime factor $p > \beta$. It is easy to see that $\log(\rho) = O(\sum_{n>\beta} n^{-2}) = O(1/\beta)$, so that $1 - \rho = O(1/\beta)$ as suggested.

Let us now consider the second issue, which is that the values of $P(c(\pi))$ for Koblitz curves $y^2 + xy = x^3 + 1$ over the binary field $\mathbb{F}_{2^n}$ appear to be large. This is because their discriminants have a specially factored form: since a Koblitz curve has CM by $\mathbb{Q}(\sqrt{-7})$, its discriminant is $D = -7c^2$, where c is the conductor (see the remarks after Theorem 3.1). Since $D = t^2 - 4q$ has approximate size 2^n, we have that $c \approx 2^{n/2}$. Assuming c is a random integer of that size, the distribution of its largest prime factor $P(c)$ is governed by the usual smoothness bounds, and hence $P(c)$ is usually large [44]. It is nevertheless conceivable that there are some curves of Koblitz type which have small $P(c)$.

A Appendix: Proof of Proposition 4.1

Proof. There are k^r random walks of length r starting from x. One would expect in a truly random situation that roughly $\frac{|S|}{h}k^r$ of these land in S. The lemma asserts that for $r \geq \frac{\log 2h/|S|^{1/2}}{\log k/c}$ at least half that number of walks in fact do. Denoting the characteristic functions of S and $\{x\}$ as χ_S and $\chi_{\{x\}}$, respectively, we count that

$$\#\,\{\text{walks starting at } x \text{ and landing in } S\} \;=\; \langle\, \chi_S \,,\, A^r \chi_{\{x\}} \,\rangle\,, \qquad \text{(A.1)}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product of functions in $L^2(\mathcal{V})$. We estimate this as follows. Write the orthogonal decompositions of χ_S and $\chi_{\{x\}}$ as

$$\chi_S \;=\; \frac{|S|}{h}\mathbf{1} \,+\, u \qquad \text{and} \qquad \chi_{\{x\}} \;=\; \frac{1}{h}\mathbf{1} \,+\, w\,, \qquad \text{(A.2)}$$

where $\mathbf{1}$ is the constant vector and $\langle u, \mathbf{1}\rangle = \langle w, \mathbf{1}\rangle = 0$. Then (A.1) equals the expected value of $\frac{|S|}{h}k^r$, plus the additional term $\langle u, A^r w\rangle$, which is bounded by $\|u\|\,\|A^r w\|$. Because $w \perp \mathbf{1}$ and the symmetric matrix A^r has spectrum bounded by c^r on the span of such vectors,

$$\|u\|\,\|A^r w\| \;\leq\; c^r\,\|u\|\,\|w\| \;\leq\; c^r\,\|\chi_S\|\,\|\chi_{\{x\}}\| \;=\; c^r\,|S|^{1/2}\,. \qquad \text{(A.3)}$$

For our values of r this is at most half of $\frac{|S|}{h}k^r$, so indeed at least $\frac{1}{2}\frac{|S|}{h}k^r$ of the paths terminate in S as was required. $\square$

B Supersingular Case

In this section we discuss the isogeny graphs for supersingular elliptic curves and prove Theorem 2.1 in this setting. The isogeny graphs were first considered by Mestre [32], and were shown by Pizer [36,37] to have the Ramanujan property. We have decided to give an account here for completeness, mainly following Pizer's arguments. Actually the isogeny graphs we will present here differ from those in the ordinary case in that they are *directed*. This will cause no serious practical consequences, because one can arrange that only a bounded number of edges in these graphs will be unaccompanied by a reverse edge. Also, the implication about rapid mixing used for Corollary 2.2 carries over as well in the undirected setting with almost no modification. It is instructive to compare the proofs for the ordinary and supersingular cases, in order to see how GRH plays a role analogous to the Ramanujan conjectures.

Every supersingular elliptic curve in characteristic p is defined over either $\mathbb{F}_p$ or $\mathbb{F}_{p^2}$ [41], so it suffices to fix $\mathbb{F}_q = \mathbb{F}_{p^2}$ as the field of definition for this discussion. Thus, in contrast to ordinary curves, there is a finite bound g on the number of curves that can belong to any given isogeny class (this bound is in fact the genus of the modular curve $X_0(p)$, which is roughly $\frac{p+1}{12}$). It turns out that all supersingular curves defined over $\mathbb{F}_{p^2}$ belong to the same isogeny class [32]. Because the number of supersingular curves is so much smaller than the number of ordinary curves, correspondingly fewer of the edges need to be included in order to form a Ramanujan graph. For a fixed prime value of $\ell \neq p$, we define the vertices of the supersingular isogeny graph $\mathcal{G}$ to consist of these g curves, with directed edges indexed by equivalence classes of degree-ℓ isogenies as defined below. In fact, we will prove that $\mathcal{G}$ is a directed $k = \ell + 1$-regular graph satisfying the Ramanujan bound of $|\lambda| \leq 2\sqrt{\ell} = 2\sqrt{k-1}$ for the nontrivial eigenvalues of its adjacency matrix. The degree ℓ in particular may be taken to be as small as 2 or 3.

For the definition of the equivalence classes of isogenies — as well as later for the proofs — we now need to recall the structure of the endomorphism rings of supersingular elliptic curves. In contrast to the ordinary setting (Section 3), the endomorphism ring $\mathrm{End}(E)$ is a maximal order in the quaternion algebra $R = \mathbb{Q}_{p,\infty}$ ramified at p and ∞. Moreover, isomorphism classes of supersingular curves E_i isogenous to E are in 1-1 correspondence with the left ideal classes $I_i := \mathrm{Hom}(E_i, E)$ of R. Call two isogenies $\phi_1, \phi_2 : E_i \to E_j$ equivalent if there exists an automorphism (=invertible endomorphism) α of E_j such that $\phi_2 = \alpha\phi_1$. Under this relation, the set of equivalence classes of isogenies from E_i to E_j is equal to $I_j^{-1}I_i$ modulo the units of I_j. This correspondence is degree preserving, in the sense that the degree of an isogeny equals the reduced norm of the corresponding element in $I_j^{-1}I_i$, normalized by the norm of $I_j^{-1}I_i$ itself. This is the notion of equivalence class of isogenies referred to in the definition of $\mathcal{G}$ in the previous paragraph. Thus, for any integer n, the generating function for the number $M_{ij}(n)$ of equivalence classes of degree n isogenies from E_i to E_j (i.e. the number of edges between vertices representing elliptic curves E_i and E_j) is given by

$$\sum_{n=0}^{\infty} M_{ij}(n)\, q^n \quad := \quad \frac{1}{e_j} \sum_{\alpha \in I_j^{-1}I_i} q^{\,N(\alpha)/N(I_j^{-1}I_i)} \tag{B.1}$$

where e_j is the number of units in I_j (equivalently, the number of automorphisms of E_j). One knows that $e_j \leq 6$, and in fact $e_j = 2$ except for at most two values of j – see the further remarks at the end of this section. Proofs for the statements in this paragraph can be found in [15, 37].

The θ-series on the righthand side of (B.1) is a weight 2 modular form for the congruence subgroup $\Gamma_0(p)$, and the matrices

$$B(n) := \begin{pmatrix} M_{11}(n) & \cdots & M_{1g}(n) \\ \vdots & \ddots & \vdots \\ M_{g1}(n) & \cdots & M_{gg}(n) \end{pmatrix}$$

(called Brandt matrices) are simultaneously both the n-th Fourier coefficients of various modular forms, as well the adjacency matrices for the graph $\mathcal{G}$. A fundamental property of the Brandt matrices $B(n)$, in fact, is that they represent that the action of the n^{th} Hecke operator $T(n)$ on a certain basis of modular forms of weight 2 for $\Gamma_0(p)$ (see [36]). Thus the eigenvalues of $B(n)$ are given by the n^{th} coefficients of the weight-2 Hecke eigenforms for $\Gamma_0(p)$. These eigenforms include a single Eisenstein series, with the rest being cusp forms. Now we suppose that $n = \ell$ is prime (mainly in order to simplify the following statements). The n^{th} Hecke eigenvalue of the Eisenstein series is $n + 1$, while those of the cusp forms are bounded in absolute value by $2\sqrt{n}$

according to the Ramanujan conjectures (in this case a theorem of Eichler [8] and Igusa [17]). Thus the adjacency matrix of $\mathcal{G}$ has trivial eigenvalue equal to $\ell+1$ (the degree k), and its nontrivial eigenvalues indeed satisfy the Ramanujan bound $|\lambda| \leq 2\sqrt{k-1}$.

Finally, we conclude with some comments about the potential asymmetry of the matrix $B(n)$. This is due to the asymmetry in the definition of equivalence classes of isogenies. Indeed, if $\mathrm{Aut}(E_1)$ and $\mathrm{Aut}(E_2)$ are different in size, then two isogenies $E_1 \to E_2$ can sometimes have equivalent dual isogenies even if the original isogenies themselves are not equivalent. This problem arises only if one of the curves E_i has complex multiplication by either $\sqrt{-1}$ or $e^{2\pi i/3}$, since otherwise the only possible automorphisms of E_i are the scalar multiplication maps ± 1 [41, §III.10]. In the supersingular setting, one can avoid curves with such unusually rich automorphism groups by choosing a characteristic p which splits in both $\mathbb{Z}[\sqrt{-1}]$ and $\mathbb{Z}[e^{2\pi i/3}]$, i.e. $p \equiv 1 \bmod 12$ (see [36, Prop. 4.6]). In the case of ordinary curves, however, the quadratic orders $\mathbb{Z}[\sqrt{-1}]$ and $\mathbb{Z}[e^{2\pi i/3}]$ both have class number 1, which then renders the issue moot because the isogeny graphs corresponding to these levels each have only one node.

C Proof of Lemma 6.3

Proof. First define $\tilde{a}_p$ to be a_p, if p is prime, and 0 otherwise. Then

$$\sum_{p \leq m} a_p \quad = \quad \sum_{p \leq m} [\tilde{a}_p \log p] \, \frac{1}{\log p} \quad = \quad \sum_{n \leq m} [\tilde{a}_n \log n] \, \frac{1}{\log n} \, .$$

By partial summation over $2 \leq n \leq m$, we then find

$$\begin{aligned} \sum_{p \leq m} a_p \quad &= \quad \sum_{n < m} S'(n,f) \left(\frac{1}{\log(n)} - \frac{1}{\log(n+1)} \right) \; + \; \frac{S'(m,f)}{\log m} \\ &\ll \quad \sum_{n < m} \left(n^{1/2} \log n \log N \right) \left| \frac{d}{dn} \left((\log n)^{-1} \right) \right| \; + \; m^{1/2} \, \log N \\ &\ll \quad \sum_{n < m} n^{1/2} \log(n) \log(N) \frac{1}{n (\log n)^2} \; + \; m^{1/2} \, \log N \end{aligned}$$

so in fact

$$S(m,f) \quad = \quad \sum_{p \leq m} a_p \quad = \quad O(m^{1/2} \log(N)) \quad \text{for} \;\; m \, = \, O(N) \, . \qquad \text{(C.1)}$$

∎

D Diagram of concepts from Section 3

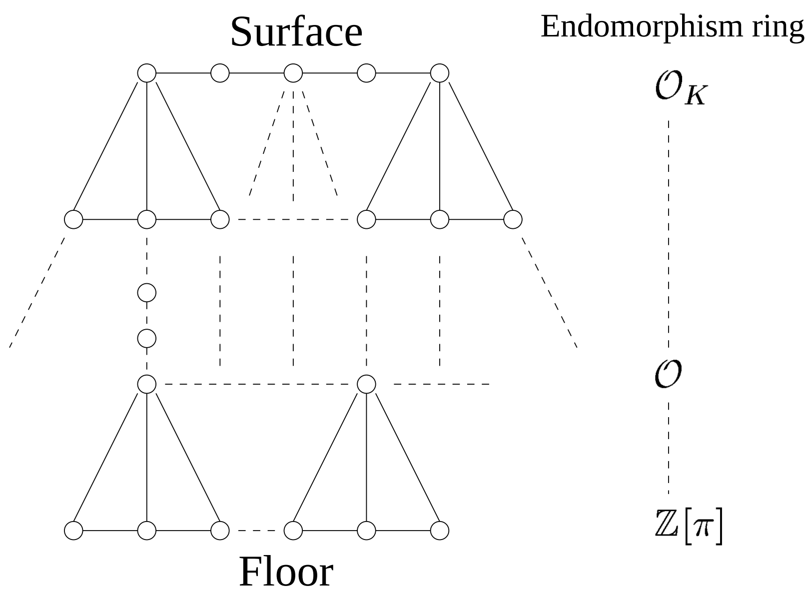

Figure 2: Ordinary elliptic curves over $\mathbb{F}_q$ have complex multiplication by imaginary quadratic number fields $K = \mathbb{Q}(\sqrt{d})$. An isogeny class can be subdivided into collections of curves having the same endomorphism ring, which is always an order in K. The nodes at the top of this figure represent ordinary curves whose endomorphism ring is the maximal order, i.e. the ring of integers of $\mathbb{Q}(\sqrt{d})$. Lower nodes represent curves with smaller endomorphism rings; nodes on different levels are connected by isogenies of potentially high degree ("the conductor gap"), depending on the relative sizes of their associated orders. In particular our results demonstrate that a random walk via low degree isogenies mixes rapidly on a given horizontal level, but does not necessarily bridge the conductor gap when this separation is, say, a large prime.

References

[1] M. Ajtai, J. Komlos, and E. Szemeredi, *Deterministic simulation in LOGSPACE*, Proceedings of the nineteenth annual ACM conference on Theory of computing (1987), 132–140.

[2] I. F. Blake, G. Seroussi, and N. P. Smart, *Elliptic curves in cryptography*, London Mathematical Society Lecture Note Series, vol. 265, Cambridge University Press, Cambridge, 2000.MR 2001i:94048

[3] Denis Charles and Kristin Lauter, *Computing modular polynomials.* preprint. `http://front.math.ucdavis.edu/math.NT/0408051`.

[4] David A. Cox, *Primes of the form $x^2 + ny^2$*, A Wiley-Interscience Publication, John Wiley & Sons Inc., New York, 1989.MR 90m:11016

[5] Giuliana Davidoff, Peter Sarnak, and Alain Valette, *Elementary number theory, group theory, and Ramanujan graphs*, London Mathematical Society Student Texts, vol. 55, Cambridge University Press, Cambridge, 2003.MR 2004f:11001

[6] Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307 (French).MR 49 #5013

[7] Max Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*, Abh. Math. Sem. Hansischen Univ. **14** (1941), 197–272 (German).MR 3,104f

[8] Martin Eichler, *Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion*, Arch. Math. **5** (1954), 355–366 (German).MR0063406 (16,116d)

[9] Mireille Fouquet and François Morain, *Isogeny volcanoes and the SEA algorithm*, Algorithmic number theory (Sydney, 2002), 2002, pp. 276–291.MR2041091

[10] Steven D. Galbraith, *Constructing isogenies between elliptic curves over finite fields*, LMS J. Comput. Math. **2** (1999), 118–138 (electronic).MR 2001k:11113

[11] Steven D. Galbraith, Florian Hess, and Nigel P. Smart, *Extending the GHS Weil descent attack*, Advances in cryptology—EUROCRYPT 2002 (Amsterdam), 2002, pp. 29–44.MR 2004f:94060

[12] Stephen S. Gelbart and Stephen D. Miller, *Riemann's zeta function and beyond*, Bull. Amer. Math. Soc. (N.S.) **41** (2004), no. 1, 59–112 (electronic).2 015 450

[13] David Gillman, *A Chernoff bound for random walks on expander graphs*, IEEE Symposium on Foundations of Computer Science (1993), 680–691.

[14] O. Goldreich, R. Impagliazzo, L.A. Levin, R. Venkatesan, and D. Zuckerman, *Security Preserving Amplification of Hardness* (1990), 318–326.

[15] Benedict H. Gross, *Heights and the special values of L-series*, Number theory (Montreal, Que., 1985), 1987, pp. 115–187.MR 89c:11082

[16] D. Harkins and D. Carrel, *The Internet key exchange (IKE)*, Technical Report IETF RFC 2409, November 1998. `http://www.ietf.org/rfc/rfc2409.txt`.

[17] Jun-ichi Igusa, *Fibre systems of Jacobian varieties. III. Fibre systems of elliptic curves*, Amer. J. Math. **81** (1959), 453–476.MR0104669 (21 #3422)

[18] Yasutaka Ihara, *Discrete subgroups of* PL(2, $k_{\wp}$), Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965), 1966, pp. 272–278.MR0205952 (34 #5777)

[19] Henryk Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17, American Mathematical Society, Providence, RI, 1997.MR 98e:11051

[20] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.2 061 214

[21] M. Jerrum and A. Sinclair, *Conductance and the rapid mixing property for markov chains: the approximation of the permanent resolved*, Symposium on Theory of Computing (May 1988), 235 –243.

[22] Neal Koblitz, *Elliptic curve cryptosystems*, Math. Comp. **48** (1987), no. 177, 203–209.MR 88b:94017

[23] ______, *CM-curves with good cryptographic properties*, Advances in cryptology—CRYPTO '91 (Santa Barbara, CA, 1991), 1992, pp. 279–287.MR1243654 (94e:11134)

[24] David Kohel, *Endomorphism rings of elliptic curves over finite fields*, University of California, Berkeley, 1996, Ph.D thesis.

[25] Serge Lang, *Elliptic functions*, 2nd ed., Graduate Texts in Mathematics, vol. 112, Springer-Verlag, New York, 1987. With an appendix by J. Tate.MR890960 (88c:11028)

[26] H. W. Lenstra Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673.MR916721 (89g:11125)

[27] Wen-Ching Winnie Li, *Elliptic curves, Kloosterman sums and Ramanujan graphs*, Computational perspectives on number theory (Chicago, IL, 1995), 1998, pp. 179–190.MR 98m:11086

[28] Alexander Lubotzky, *Discrete groups, expanding graphs and invariant measures*, Progress in Mathematics, vol. 125, Birkhäuser Verlag, Basel, 1994.MR 96g:22018

[29] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.MR 89m:05099

[30] Alfred J. Menezes, Tatsuaki Okamoto, and Scott A. Vanstone, *Reducing elliptic curve logarithms to logarithms in a finite field*, IEEE Trans. Inform. Theory **39** (1993), no. 5, 1639–1646.MR 95e:94038

[31] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone, *Handbook of applied cryptography*, CRC Press Series on Discrete Mathematics and its Applications, CRC Press, Boca Raton, FL, 1997. With a foreword by Ronald L. Rivest.MR 99g:94015

[32] J.-F. Mestre, *La méthode des graphes. Exemples et applications*, Proceedings of the international conference on class numbers and fundamental units of algebraic number fields (Katata, 1986), 1986, pp. 217–242 (French).MR 88e:11025

[33] Victor S. Miller, *Use of elliptic curves in cryptography*, Advances in cryptology—CRYPTO '85 (Santa Barbara, Calif., 1985), 1986, pp. 417–426.MR 88b:68040

[34] M. Ram Murty, *Problems in analytic number theory*, Graduate Texts in Mathematics, vol. 206, Springer-Verlag, New York, 2001. Readings in Mathematics.MR1803093 (2001k:11002)

[35] National Institute of Standards and Technology, *Digital Signature Standard (DSS)*, Technical Report FIPS PUB 186–2, January 2000. `http://www.csrc.nist.gov/publications/fips/fips186-2/fips186-2-change1.pdf`.

[36] Arnold K. Pizer, *Ramanujan graphs and Hecke operators*, Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 1, 127–137.MR 90m:11063

[37] ______, *Ramanujan graphs*, Computational perspectives on number theory (Chicago, IL, 1995), 1998, pp. 159–178.MR 99b:11046

[38] Zeév Rudnick and Peter Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), no. 2, 269–322. A celebration of John F. Nash, Jr.

[39] Peter Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, vol. 99, Cambridge University Press, Cambridge, 1990.MR 92k:11045

[40] Jean-Pierre Serre, *Abelian l-adic representations and elliptic curves*, McGill University lecture notes written with the collaboration of Willem Kuyk and John Labute, W. A. Benjamin, Inc., New York-Amsterdam, 1968.MR0263823 (41 #8422)

[41] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1994.MR 95m:11054

[42] John Tate, *Endomorphisms of abelian varieties over finite fields*, Invent. Math. **2** (1966), 134–144.MR0206004 (34 #5829)

[43] John T. Tate, *The arithmetic of elliptic curves*, Invent. Math. **23** (1974), 179–206.MR0419359 (54 #7380)

[44] Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995. Translated from the second French edition (1995) by C. B. Thomas.MR1342300 (97e:11005b)

[45] Lawrence C. Washington, *Elliptic curves*, Discrete Mathematics and its Applications (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2003.MR 2004e:11061

David Jao
Microsoft Research
1 Microsoft Way
Redmond, WA 98052
davidjao@microsoft.com

Stephen D. Miller
Department of Mathematics
Hill Center-Busch Campus
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854
miller@math.rutgers.edu

Ramarathnam Venkatesan
Microsoft Research
1 Microsoft Way
Redmond, WA 98052
venkie@microsoft.com