Publications mathématiques de Besançon
Algèbre et théorie des nombres

Mark Watkins, Stephen Donnelly, Noam D. Elkies, Tom Fisher, Andrew Granville, and Nicholas F. Rogers

Ranks of quadratic twists of elliptic curves

2014/2, p. 63-98.

<http://pmb.cedram.org/item?id=PMB_2014____2_63_0>

Publication éditée par le laboratoire de mathématiques
de Besançon, UMR 6623 CNRS/UFC

Publications mathématiques de Besançon – 2014/2, 63-98

RANKS OF QUADRATIC TWISTS OF ELLIPTIC CURVES

by

Mark Watkins, Stephen Donnelly, Noam D. Elkies, Tom Fisher, Andrew Granville
and Nicholas F. Rogers

Abstract. — We report on a large-scale project to investigate the ranks of elliptic curves in a quadratic twist family, focussing on the congruent number curve. Our methods to exclude candidate curves include 2-Selmer, 4-Selmer, and 8-Selmer tests, the use of the Guinand-Weil explicit formula, and even 3-descent in a couple of cases. We find that rank 6 quadratic twists are reasonably common (though still quite difficult to find), while rank 7 twists seem much more rare. We also describe our inability to find a rank 8 twist, and discuss how our results here compare to some predictions of rank growth vis-à-vis conductor. Finally we explicate a heuristic of Granville, which when interpreted judiciously could predict that 7 is indeed the maximal rank in this quadratic twist family.

Résumé. — Nous donnons un compte rendu d'un projet de grande envergure sur les rangs des courbes elliptiques dans une famille de tordues quadratiques en nous focalisant sur les courbes associées aux nombres congruents. Afin d'exclure certaines courbes, nos méthodes incluent des tests sur les 2,4,8-groupes de Selmer, l'utilisation de la formule explicite de Guinand-Weil et également des 3-descentes dans quelques cas. Nous constatons que les tordues quadratiques de rang 6 sont assez répandues (bien que toujours assez difficile à trouver), alors que celles de rang 7 semblent bien plus rares. Nous décrivons aussi notre incapacité à obtenir des tordues quadratiques de rang 8 et expliquons en quoi nos résultats peuvent se comparer à certaines prédictions sur la croissance du rang en fonction du conducteur. Enfin, nous expliquons une heuristique due à Granville, qui, lorsqu'elle est interprétée judicieusement, pourrait prédire que le rang maximal pour cette famille est en effet égal à 7.

1. Introduction

Let $E : y^2 = x^3 - x$ be the congruent number curve, so that d is a congruent number precisely when the dth quadratic twist $E_d : dy^2 = x^3 - x$ has positive rank. We are interested in how the

Mathematical subject classification (2010). — 11D25, 11D45, 11G05, 11G40, 14H52.
Key words and phrases. — elliptic curves, quadratic twists, Selmer groups, explicit formula, Birch–Swinnerton-Dyer conjecture, algebraic rank.
Acknowledgements. — Watkins and Donnelly are members of the Magma computational algebra group, partially supported by the Australian Research Council. Elkies was partially supported by NSF grants DMS-1100511 and DMS-0501029. Granville was partially supported by NSERC.

rank behaves as d varies. More generally, we could fix a different elliptic curve $y^2 = x^3+ax+b$, and enquire about its quadratic twists.

Honda [26, §4, p. 98] conjectures the rank is bounded[1] in any such family, basing this on an analogy [26, Theorems 5-6] between Mordell-Weil groups of abelian varieties and Dirichlet's unit theorem. Schneiders and Zimmer [53] presented some preliminary evidence for Honda's conjecture back in 1989, though it seems that modern thought has preferred to intuit that ranks are unbounded in quadratic twist families, partially relying on function field analogues such as in [55] and [56].[2]

In this paper, we give some experimental data regarding quadratic twists of rank 6-8 for the congruent number curve. In particular, we find "lots" of rank 6 examples (over 1000), while we know of no rank 8 example. A number of rank 7 examples are also given (27 currently). We also explicate a heuristic of Granville that might lead one to suspect that rank 7 is the maximal rank in this family.

1.1. Acknowledgements. — Some of our computations catalogued below were done on `sage.math.washington.edu` (and associated machines) acquired under National Science Foundation (USA) Grant No. DMS-0821725. We thank W. A. Stein for the use of this.

We thank K. Rubin and A. Silverberg for feedback, and M. O. Rubinstein for noting some errors in a previous version. W. B. Hart assisted with a couple of issues in coding, while A. R. Booker and M. O. Rubinstein were instrumental in convincing us that the explicit formula (§7) could be usefully applied to our situation. The anonymous referee also provided us with a thorough reading and helpful comments.

This paper is written from the standpoint of the primary author (Watkins), who might better be termed the "project manager" for some aspects of the research. Indeed, the paper has taken a much different route than might have been guessed following the initial data accumulations in 2008. The primary author apologises for the delay in publication — however, the topic of ranks seems to be of sufficient interest to merit the additional computational time that was needed. Similarly, the somewhat lengthy descriptions of elements which could be considered tangential (such as Mestre-Nagao sums) seems warranted as guidance to future investigators.

2. Outline of experimental methodology

In this section we give a general outline of our experimental methods, and then give details in the following sections. Our goal was typically to find quadratic twists of the congruent number curve whose rank was at least as large as a target rank r. In some cases we also wished to be able to give an upper bound on the rank for some given collection of twists,

[1] The conjecture as printed appears to contain a typographical error, as it asserts an *equality* for the rank rather than an upper bound.

[2] We are indebted to F. Rodriguez Villegas for indicating to us that Néron, in a 1950 footnote in Poincaré's collected works [42, III, p. 495, Footnote 3], stated that it was considered "probable" there was a universal upper bound for ranks of elliptic curves over $\mathbf{Q}$ (not just in quadratic twist families): *On ignore s'il existe pour toutes les cubiques rationnelles, appartenant à un corps donné une borne absolute du rang. L'existence de cette borne est cependant considérée comme probable.*

One can contrast the abstract of [35] (Mestre, 1982): *Au vu de cette méthode, il semble que l'on puisse sérieusement conjecturer que le rang des courbes elliptiques définies sur* $\mathbf{Q}$ *n'est pas borné.* [This method seems to suggest that the rank of elliptic curves defined over $\mathbf{Q}$ is not bounded.]

particularly to say that the ranks were less than r. Here is an enumeration of the methods we used.

- Generate prospective quadratic twists of large rank. We used two principal techniques here. The first (chronologically in our experiments) was to adapt the method used by Rogers in [46] (see §3.1 for more history), corresponding to a rank 1 parametrisation of elliptic curves. This method allows us to find a high-rank twist provided it has a point of small height.

  The second method was a nearly exhaustive search of twists up to a given d-bound. Here we used Monsky matrices (see §4) and variants thereof due to Rogers, so as to efficiently bound the possible rank from 2-Selmer information.

  Finally, in §3.2 we discuss whether a higher rank parametrisation could be of use, and in §3.3 we give a lattice-based method to find twists.

- Bound the 2-Selmer rank via a Monsky matrix computation (using linear algebra over $\mathbf{F}_2$). We also used variants of Monsky's formula (derived by Rogers, see §4) applicable to isogenous curves.

- Apply a Mestre-Nagao heuristic (§5), ignoring curves (for instance) for which the average a_p-value for $p \leq 10^5$ indicated that large rank was not too likely. This step is largely heuristic, and essentially says that curves with "many" local points (on average) are more likely to have large rank.

- Bound the 4-Selmer rank (on all the isogenous curves) via the Cassels-Tate pairing (§6), as implemented in Magma by Donnelly.

- Bound the 8-Selmer rank via higher descent pairings due to Fisher (§6.1), again implemented in Magma. These rely on the rationality of the 2-torsion.

- Apply the (Guinand-Weil) explicit formula to attempt to bound the analytic rank upon assuming a suitable Riemann hypothesis (§7).

Usually we would also search for points at some stage, typically after the number of curves was suitably reduced, for instance after the Cassels-Tate pairing was applied. Due to the experimental nature of our work, we were willing to assume a parity conjecture, that is, we would accept a curve with $(r-1)$ independent points as being rank r if this was implied by parity. We typically only searched on 2-covers (including isogenous curves), but for a few curves Fisher computed 4-covers (extending methods of Bremner and Cassels [6], see [21, §5]), and searched on those (see §9.1).[3]

3. Generating prospective twists

We now give more details about our methods for generating prospective high-rank quadratic twists of the congruent number curve. The first method parametrises curves by a point of "small" height u/v, with $d = uv(u+v)(u-v)$. We then comment on parametrisations

[3] Indeed, for 150 of our 1486 putative rank 6 curves we still only know 5 independent points, and hope to remedy this via a more large-scale usage of 4-covers in the near future. (Added in proof: This is now done.)

by multiple small points, and give a lattice-based method that loops over square divisors of $uv(u+v)(u-v)$ rather than u and v. Finally, in §4 we present a method of Rogers (from his 2004 thesis) that gives a near exhaustive method of finding high-rank quadratic twists with d up to a given bound.

3.1. Parametrisation via a point of small height. — We review the "standard" method to find high rank quadratic twists of a given elliptic curve. This is given in [46], with Rogers noting that Rubin and Silverberg suggested this algorithm following the ideas of [48], who in turn note that one of the principal observations dates back to Gouvêa and Mazur [22]. Letting $E : y^2 = f(x)$ be the given elliptic curve, any rational number $x = u/v$ is the x-coordinate for exactly one of the twists $E_d : dy^2 = f(x)$, namely when $d = f(u/v)$, and here we can reduce the problem by taking the square-free part of d. Alternatively, we can homogenise the equation to note that for $f(x) = x^3+ax+b$ we have $f(u/v) = (1/v^3)(u^3+auv^2+bv^3)$, and are thus interested in the square-free part of $v(u^3 + auv^2 + bv^3)$ for a given u/v. In analogy with more general attempts to find high-rank curves, we might hope that d-values that appear for "many" small-height points u/v will be more likely to have large rank.

The main step in finding the d-values involves computing the square-free part of the specialisation of a binary quartic form. This is done most easily when the quartic splits completely, which is the case of full 2-torsion for the elliptic curve.

3.1.1. *The case of the congruent number curve.* — For the particular case of the congruent number curve given in the introduction, we have $f(x) = x^3 - x$ so that $v^4 f(u/v) = uv(u - v)(u + v)$, and we can note that if u and v are coprime and not both odd, then the four terms on the right side are all pairwise coprime. This implies that we need only compute the square-free part of each term individually, which is most conveniently done by pre-computed table lookup. We also note that for the congruent number curve we can restrict to $u > v > 0$ and to u, v of opposite parity, both of these due to the 2-torsion.[4]

The method of computation then proceeds by looping over $1 \leq v < u \leq L$ up to some limit L, and for each u/v computing the associated d-value as indicated above. It is at this point that the methods that are used tend to vary. Rogers [46] used a hash table of the d that are found (to find d's given by multiple u/v), and also some requirements related to the 2-Selmer group, such as demanding that d have sufficiently many prime factors (possibly of a certain size). The searches in [46] reached $L = 10^5$. Rogers also used alternative methods to find high rank twists (including three of rank 7), such as full 2-descent on all the isogenous curves as a filter for which d to consider (see §4.3).

Dujella, Janfada, and Salami [14] appear to first compute the 2-Selmer rank via a formula of Monsky [25, Appendix], and then use Mestre-Nagao heuristics (see §5) to select which curves to investigate further. They work additionally with the 2-isogenous twist $dy^2 = x^3 + 4x$ corresponding to the quartic form $uv(u^2 + v^2)/2$ from $x = 2u/v$; as the quadratic term here does not factor, this limits their possible range of (u, v) and indeed they take $L = 10^5$.

We chose to use only the congruent number curve in the u/v stage. We first computed the 2-Selmer rank via Monsky's formula (via linear algebra over $\mathbf{F}_2$, which is quite fast), and then filtered via Mestre-Nagao sums. However, we later used code from Rogers that computes the 2-Selmer rank of the isogenous curves again via linear algebra over $\mathbf{F}_2$, using a variant of Monsky's formula and quadratic reciprocity in the Euclidean rings $\mathbf{Z}[\sqrt{-1}]$ and $\mathbf{Z}[\sqrt{2}]$. This

[4] Namely, translation by $(0, 0)$ gives $(u : v) \to (v : -u)$; by $(1, 0)$ yields $(u : v) \to (u + v : u - v)$.

allowed us to eliminate approximately 90% more curves before computing Mestre-Nagao sums. Note these formulæ require us to factor d, which is again feasible via table-lookup in our case. We do not use a hash table of popular d-values as Rogers did; rather, having one (u, v) point is enough for us to pass the resulting d to the Selmer machinery. We searched up to $L = 10^8$ in some cases, with the caveat that we ignored curves for which d was rather large. More information appears below in §8

The 2-Selmer machinery was able to process about 10^5 twists per core-second. In particular, in our $L = 10^7$ experiment we computed about $\frac{2}{\pi^2}10^{14}$ such Selmer groups, taking about 6-7 core-years. With the $L = 10^8$ experiments, many d-values were pruned due to exceeding our size restriction, but a similar number of Selmer group tests seem to have been applied (we do not have an exact accounting).

3.1.2. *Previous computations and records.* — In November 2003, Rogers [47, p.45] found that $d = 797507543735$ yields E_d of rank 7. He found this rank 7 curve via testing all possible d (up to some bound) that allowed rank 7 from the 2-descent information [47, §4.4]. Rogers also found 14 quadratic twists of rank 6, listed in [14]. The work of Dujella, Janfada, and Salami [14] determined about 25 more such rank 6 quadratic twists.

The twist $d = 797507543735$ is first found via the (u, v)-searching method with $(v, u) = (79873, 235280)$, taking less than a core-hour to find.[5] The first rank 6 twist is $d = 6611719866$ and for rank 5 it is $d = 48242239$. The purported growth rate for the first rank r twist is obscure, but one suggestion has been along the lines of 2^{r^2}. The principal rationale here is that (perhaps from L-function considerations) the rank might be roughly as large as $\sqrt{\log d}$, with the speculation then being phrased in a simple form (though see Footnote 32, and compare [20]). However, alternative schools of thought (e.g., descent bounds or function field analogues [56], or rank bounds under a Riemann hypothesis [7, 36]) suggest the rank might be as large as $\frac{\log d}{\log\log d}$. The numerics are conflated by the fact that the smallest d of rank 7 is rather small (perhaps abnormally so), for the second smallest d is more than 2500 times as large. In particular, an estimate of $2^{8^2}/2^{7^2} \approx 30000$ might be quite low for the rank 8 versus rank 7 ratio. In §10 we discuss more fully our thoughts on whether our computer searches should have found a rank 8 twist if one exists.

3.2. Searching in larger rank families. — Another search method could be to restrict to d which are in a parametrised family with larger rank, for instance the rank 3 family $6(u^{12} - 33u^8 - 33u^4 + 1)y^2 = x^3 - x$ given in [49, Theorem 4.5]. However, the degree 12 polynomial has 2 quartic factors, making it difficult to accumulate data.

Instead, as proposed by Elkies, we tested a couple of rank 2 families. The first (I) equates y-values, solving $s^3 - s = t^3 - t$ via $(s, t) = (-\frac{2w+1}{w^2+w+1}, \frac{1-w^2}{w^2+w+1})$, obtaining $(s^3 - s)y^2 = x^3 - x$, or $w(w + 1)(w - 1)(w + 2)(2w + 1)(w^2 + w + 1)y^2 = x^3 - x$. Writing $w = m/n$, via the symmetries of the homogeneous octic polynomial in m and n, we can restrict to m, n that are not congruent modulo 3. The second (II) takes $x = \frac{w^2+2}{w^2-1}$ as one point, and then $x + 1$ also yields a point on $dy^2 = x^3 - x$ where $d = 3(w + 1)(w - 1)(w^2 + 2)(2w^2 + 1)$ here.[6] For both

[5] There are two rank 7 twists with smaller u, found in Table 5 of §9.2.1 below. However, to find the first d this fast, one filters out larger d (say $d > 2^{50}$), greatly reducing the number of Selmer tests.

[6] When $w^2 + 1$ is square, upon writing $w = (u^2 - 1)/2u$ we recover the above rank 3 family, and there is a third (independent) small point with x-coordinate $-(w^2 + 2)/(2w^2 + 1)$.

families, for the purposes of comparison, we considered w up to height 10^4, which meant only minor modifications to our factoring tables.[7]

One goal of this experiment (or "pilot test") was to see how many rank 6 curves are found up to height 10^4 – if the number does not exceed the total from the rank 1 family, it is probably not worth trying to find a rank 8 example in these families.[8] However, there are various problems with a direct comparison. Consider Family I, where the d-values now have 7 factors (six, plus one more for projectivisation) in their polynomial factorisation, compared to 4 previously. This means that the Selmer ranks are likely to be higher, and thus Ш tends to abound. The d-values also now come from a polynomial of degree 8, rather than one of degree 4, and will thus typically be much larger. This implies that searching for points on 2-covers of the surviving curves is much less likely to actually find anything, for the regulators will typically also be larger. Furthermore, the existence of another point of small height means that the other generators will tend to have larger height, again diminishing the expected returns from point searching. Finally, the larger d implies larger conductor, which makes the explicit formula computations (see §7) less valuable. The situation is slightly better with Family II, though similar concerns might apply.

We give the results of the comparisons in Section §12.1 below.

3.3. A variant search method.

— A second search method, suggested and implemented by Elkies[9] (with further implementations by Hart/Watkins), parametrises the square divisors of $uv(u+v)(u-v)$ via

$$d_1^2|u,\ \ d_2^2|v,\ \ d_3^2|(u+v),\ \ d_4^2|(u-v),$$

and then loops over pairwise coprime (d_1, d_2, d_3, d_4), looking for short vectors in the (u, v) lattice. The lattice has determinant $D^2 = \prod_i d_i^2$, which gives a measure of what size of (u, v) to expect. One restricts (say) to $5 \le d_i \le 400$ and $H_1 \le d_1d_2d_3d_4 \le H_2$ for parameters such as $H_1 \approx 10^6, H_2 \approx 10^8$, the lower bounds being imposed to preclude the "trivial" cases from swamping the calculation.[10]

For instance, the 4-tuple $(16, 167, 9, 389)$ yields $d = 797507543735$ (the first rank 7 twist) from the pair $(u, v) = (3764480, 2705233)$. The time taken to find this is comparable to the method of §3.1.1; an exact analysis depends upon various cut-off ranges that one uses. There is a (minor) side issue, in that if (u, v) is sufficiently small then the point was already considered via the previous searches, while otherwise computing the square-reduced part (and factorisation) via table lookup might not be feasible. We give some preliminary data about this method in §12.2 below.

[7] For Family I, instead of factoring $m^2 + mn + n^2$, one could (say) loop over (a, b, c, d) up to some limit T, and generate (m, n) up to L from $(a + b\zeta_3)(c + d\zeta_3) = (m + n\zeta_3)$, thus allowing factorisation by table on each part of the LHS. One could also take T slightly larger than $\sqrt{L}$ here, perhaps $(L, T) = (10^7, 10^4)$ would be useful. Elkies notes that the use of lattice reduction should allow one to similarly loop over suitable factors of the quadratics in Family II.

[8] As Bober remarks, one nice feature of the rank 1 parametrisation is that it contains *all* twists of positive rank, while the corresponding statement for the other families is not true.

[9] The lattice displayed here is also given by Rubin and Silverberg in [50, §9].

[10] Indeed, trivially $(d_1, d_2, d_3, d_4) = (1, 1, 1, 1)$ generates all (u, v), but the "short" vector enumeration for lattices of such small determinant is typically infeasible.

3.3.1. *Searching via restriction to (u, v) with large square divisors.* — In the same vein, considering the parametrisation in §3.1.1, it might be noted that for $L = 10^7$ we would already have $d \approx 10^{28}$ if we did not take the squarefree part. Indeed, in our table of rank 7 twists (Table 5 in §9.2.1) we find the point $(v, u) = (2202624, 98856259)$ with $2281^2|u$ and $96^2|v$. It thus might be feasible to enlarge the L-bound (in a heuristic sense) in conjunction with a demand that $uv/s(u)s(v)$ be of decent size (where here $s(x)$ is the minimal positive integer such that $xs(x)$ is square).

The exact correspondence between this method and that given in the previous subsection has not been fleshed out completely.

4. 2-Selmer ranks and Monsky matrices

As noted above, Rogers was able to derive Monsky-like matrices for the isogenous curves of $dy^2 = x^3 - x$. We briefly recall the construction of Monsky [39], and then give the generalisation of Rogers. We then describe how to use the 2-Selmer information sequentially, considering d as a product of primes and building up the matrices one prime at a time. The effect that appending primes has on the 2-Selmer ranks can be bounded, which then gives a lower bound on how many additional primes must be included if a target rank is to be met.

4.1. Monsky's matrix. — Factor $d = \prod_i p_i$ into primes p_i for d odd, and similarly with $d/2 = \prod_i p_i$ for d even. Monsky defines a matrix A over $\mathbf{F}_2$ as follows. For $i \neq j$ let A_{ij} be 0 or 1 according to whether $\left(\frac{p_j}{p_i}\right) = +1$ or not. Then define A_{ii} so that the row sums are all zero. Define D_u to be the diagonal matrix with entries given by 0 or 1 depending on whether $\left(\frac{u}{p_i}\right) = +1$ or not. Then the Monsky matrix M is given alternately for d odd or even as

$$M = \begin{pmatrix} A + D_2 & D_2 \\ D_2 & A + D_{-2} \end{pmatrix} \text{ and } M = \begin{pmatrix} D_2 & A + D_2 \\ A^T + D_2 & D_{-1} \end{pmatrix}.$$

This is a square matrix of size $2w$, where w is the number of odd prime factors of d. The 2-Selmer rank (modulo torsion) of E_d is the nullity of this matrix.

4.2. The variant of Rogers. — For the isogenous curve $E'_d : y^2 = x^3 + 4dx$, Rogers uses arithmetic over $\mathbf{Z}[\sqrt{-1}]$ to determine a matrix similar to Monsky's. For d even, we factor $d = \prod_j q_j$ into elements q_j corresponding to prime ideals, where $q_1 = 2$ and the other q_j are $a_j + b_j\sqrt{-1}$ with a_j odd (so b_j even), being either part of a conjugate pair, or simply $q_j = a_j > 0$ for an inert prime. For $i, j \geq 2$ define R_{ij} for $i \neq j$ to be 0 if q_i and q_j are either conjugate or squares modulo each other, and 1 otherwise. For the prime above 2, define R_{j1} to be 0 if the norm of q_j is 1 mod 8 and 1 otherwise, and define R_{1j} to be 1 if b_j is ± 2 and 0 otherwise. Finally, define R_{11} to be 0, and the other diagonal R_{ii} entries to make the row sums be 0. The 2-Selmer rank (modulo torsion) of E'_d is then one less than the nullity of R. For odd d, most of the Rogers matrix is the same, while the conditions at 2 need to be modified. There are two cases with the latter, depending on whether d is ± 1 modulo 8 or not; both are sufficiently involved that we omit them here. Letting w be the number of odd prime factors of d over $\mathbf{Z}[\sqrt{-1}]$, the Rogers matrix has size $(w + 1)$ when d is even or ± 3 modulo 8, and $(w + 2)$ when d is ± 1 modulo 8.

The other isogenous curves involve arithmetic over $\mathbf{Z}[\sqrt{2}]$. Again the bulk of the matrix involves quadratic residue symbols, with more complicated computations for the prime $\sqrt{2}$.

The case of conjugate q_i and q_j also differs, as one determines the matrix entry via considering if $\sqrt{2}$ is a square modulo q_i.

Rogers has implemented the above in both GP/PARI and C. We adapted the former into our own C code. Via the use of parity, one finds a quite robust test for correctness, as the computed 2-Selmer ranks (modulo torsion) of the isogenous curves should all have the same parity.

The quadratic reciprocity law over $\mathbf{Z}[\sqrt{-1}]$ is rather simple for odd primes, for q_i is a square modulo q_j only if q_j is a square modulo q_i. Additionally, inert primes are always squares modulo each other, and squareness is preserved upon conjugating all the primes involved. The latter two statements remain true for $\mathbf{Z}[\sqrt{2}]$.

4.3. Sifting for small d via 2-Selmer tests. — Rogers has used his variants of the Monsky matrices as the basis of a method for finding small d which have large 2-Selmer rank for all isogenous curves. Namely, one recurses over primes, noting the 2-Selmer rank (modulo torsion) is determined from the nullity of a matrix, while the possible increase in nullity from appending an additional prime can be computed by considering the choices of diagonal elements, the other entries staying constant.[11]

Given a rank bound r, to reach a d-limit of D should be approximately linear in D, with the constant depending essentially on the probability that a random $\mathbf{F}_2$-matrix of the prescribed type has augmented nullity (that is, the nullity when considering all relevant choices of diagonal elements) of at least $(r-2)+1$.

There is a minor issue about having a small core-product times one large prime, for instance Rogers notes 20162 has augmented nullity 5 while for 723558 it is 6, which then allow[12] ranks 6 and 7 respectively when appending one more prime (albeit with congruence conditions). There are eventually many such core-products, so in our experiments we curtailed the size of this final prime at 10^8. Comparatively, with the (u, v) searching method of §3 one knows that no prime dividing d can exceed $2L$.

Using these methods, Rogers was able to find the first three rank 7 examples, namely 797507543735, 2067037027955295, and 2210857604820494. We catalogue our results from this method in §8.2 below.

5. Mestre-Nagao sums

There are various ways of forming a sum over small primes that heuristically correlates with curves of large(r) rank. The typical underlying idea derives from the original form of the BSD conjecture [2, (A)]. Namely, fixing an elliptic curve E of rank r and writing $N_p = p+1-a_p$ for the number of points of $E/\mathbf{F}_p$ (ignoring bad primes), we should have the asymptotic relation $\prod_{p\le Y} N_p/p \sim C_E(\log Y)^r$ as $Y \to \infty$, for some constant C_E depending on the elliptic curve E. Then by taking logarithms and expanding in a power series we find

$$\sum_{p\le Y} \log(1-a_p/p+1/p) \sim -\sum_{p\le Y}\left(\frac{a_p-1}{p}+\frac{(a_p-1)^2}{2p^2}\right) \sim r\log\log Y,$$

[11] See also [18, §4] which mentions a similar method for 3-descent on $X^3+Y^3=k$.

[12] A prime can raise augmented nullity 5 to nullity 7, then subtract one for the 2-Selmer bound.

and by analytic properties of the symmetric-square L-function we have that a_p^2 is p on average (independent of whether the curve has complex multiplication), so this yields that a_p is $\frac{1}{2} - r$ on average.[13] There are historical reasons why $\frac{L'}{L}(s)$ has been considered as opposed to $L(s)$, and thus more frequently the heuristics have been described with an extra weighting of $\log p$. In our experiments we followed this tradition largely out of inertia, though below (§5.2) we give some consideration as to whether or not this weighting is the most useful for our purposes.

Thus the sums we consider are more analogous to

$$\sum_{p\le Y} \frac{a_p}{p}(\log p) \sim (1/2 - r)\log Y.$$

As a heuristic, these seem more useful when the conductor is "small" relative to the rank (recall $r \lesssim \frac{1}{2}\frac{\log N}{\log\log N}$ under suitable hypotheses [7, §2.11]), as then the effect of a_p being significantly negative (particularly for small p) tends to be more pronounced.

Note that the standard deviation of the above sum should also be of size $\log Y$ (and thus the "$\sim$" symbol is not truly correct), as the variance should resemble[14]

$$\sum_{p\le Y} \frac{\big(a_p - (1/2 - r)\big)^2}{p^2}(\log p)^2 \sim \sum_{p\le Y} \frac{a_p^2(\log p)^2}{p^2} \sim \sum_{p\le Y} \frac{(\log p)^2}{p} \sim \frac{1}{2}(\log Y)^2.$$

But there is a secondary term here, and this very well might not be negligible for the Y we use — for instance, we can recall an equivalent form of a related theorem of Mertens [34], which states that $\sum_{p\le Y}(\log p)/p \sim \log Y - 1.33258\ldots$

As Mestre-Nagao sums are written[15] in terms of $(2 - a_p)$, we thus might expect[16]

$$\sum_{p\le Y} \frac{2 - a_p}{p}(\log p) \approx (r + 3/2 + \sigma/\sqrt{2})\log Y,$$

where σ is a random Gaussian variable with mean 0 and deviation 1.

In particular, both r and σ are multiplied by $\log Y$ in this main term, so taking Y larger does not have much impact. If we wanted to find a given rank 6 curve with probability $1 - \frac{1}{30000}$, thus corresponding to σ about -4, this would say that we might restrict to considering curves with the above sum exceeding $(15/2 - 4/\sqrt{2})\log Y$, though as above, secondary effects might be apparent.

A typical choice of Nagao [40] is to consider

$$\sum_{p\le Y} \frac{2 - a_p(E)}{\#E(\mathbf{F}_p)}(\log p) \quad \text{and/or} \quad -\frac{1}{Y}\sum_{p\le Y} a_p(E)\log p,$$

[13] We are indebted to M. O. Rubinstein who indicated a flaw in our earlier computation. The work of Nagao [41, §3] contains the correct balance with 1/2. See [10] for the general context.

[14] This model for the a_p is probably most useful when Y is sufficiently smaller than $\sqrt{N}$, else effects from modularity of the L-function may play a part. When $Y \to \infty$, analysing the variance via the explicit formula might be preferred, see Rubinstein's method [51] noted in §5.2 below.

[15] The historical reason for this could derive from a possible typo, namely $(p-1)$ appears in the numerator rather than $(p+1)$ in [38], which thus lists the formula $\sum_p\big(\frac{p-1}{\#E(\mathbf{F}_p)} - 1\big)\log p$.

[16] Here we switch to the $\approx$ notation rather than $\sim$. This is to indicate we are using a finite Y in our experiments, but still largely ignoring the error in the asymptotic.

where the denominator in the first can be re-written as $p+1-a_p$. These appear to derive from Mestre [35]. We used

$$\Sigma_5 = \sum_{p\le 10^5} \frac{2-a_p(E)}{p+1}(\log p).$$

It is unclear exactly what effect these minor modifications have.

5.1. Observations from obtained data. — We record here some data about Σ_5 from our experiments. For instance, in the first experiment below, we had almost 38 million curves of even parity for which 2-descent allowed rank 6 (or more) on all isogenous curves. The maximum $\Sigma_5 \approx 70.429$ was for $d = 141486274882017786$ of rank 2, and the first rank 6 curve was 9th in the list at $\Sigma_5 \approx 69.216$, namely $d = 718916589348840586$. The top 1% of the data (after imposing $\Sigma_5 \ge 35$ as we indicate below) went down to $\Sigma_5 \approx 50.1$, and contained about 1175 curves which survived the 4-descent test (there were about 35000 twists overall that survived this test). Of these, after applying a strict $d \le 2^{60}$ bound, there were about 530 for which we found at least 5 independent points. The second percentile reached $\Sigma_5 \approx 48.4$, and had about 600 curves that survived the 4-descent test, of which about 170 yielded at least 5 independent points. The 50% percentile reached $\Sigma_5 \approx 37.7$, with 332 curves surviving the 4-descent test. There were similarly almost 300 survivor curves in the last percentile, so the number of 4-descent survivors does not decrease too much. However, the number of rank 6 curves fell considerably, as the entire bottom half of the data produced only 18 curves on which we found enough independent points (compared to 1243 for the top half). See also the comments at the end of §9.1.3.

5.2. Variations of Mestre-Nagao heuristic. — Elkies notes that removing the weighting by $\log p$ from Σ_5, or indeed returning to the original BSD-weighting, has an advantage asymptotically, as a rank increment is of size $\log\log Y$, while the deviation is only $\sqrt{\log\log Y}$. Furthermore, the constant in the analogue of the Mertens theorem is now positive (approximately 0.265) rather than negative. Both of these aspects help to increase the ratio of the rank increment to the deviation, doing so not only asymptotically, but also for practical values of Y (such as 10^5).

However, for practical Y this change also has the negative aspect that larger r will then induce greater deviation, upon accounting a_p^2 more precisely in our computations as $(a_p-(1/2-r))^2$. Our preference in using a Mestre-Nagao heuristic was to try to miss as few curves of rank 6 or more as possible; while the modified heuristics proposed by Elkies might do better in reducing false positives, the question of whether they miss more borderline curves is comparatively not so clear.[17]

[17] Here is an analysis. Suppose we fix $r=6$ as our target rank. As above, we propose the ratio of rank increment to the deviation as a useful metric. For the congruent number twists, the rank increment seems best modelled by $\sum_p' 2/p$ (or multiplied by $\log p$) where the sum is over primes up to Y that are 1 mod 4. So a rank increment for $Y=10^5$ is approximately 1.87 when considering just $\sum_p a_p/p$, and is about 9.30 when weighting by $\log p$.

The deviation for the unweighted sum is modelled by $\sqrt{\sum_p \bigl(a_p-(1/2-r)\bigr)^2/p^2}$, which is about 1.30 for $r=0,1$, but then starts rising to 1.57, 2.05, 2.61, 3.22, and finally 3.85 for $r=6$, so that a rank increment is actually quite a lot less than a deviation for rank 6 curves. But with the log-weighted sum, the deviation only changes from 7.78 for $r=0$ to 9.14 for $r=6$, still (slightly) less than a rank increment. Finally, the effect of small primes dividing d would presumably also affect the log-less sum more than the log-weighted, which is an aspect we ignored here.

Unfortunately, it seems difficult to construct a relevant data set for comparison between these heuristics without redoing much of our experiment. More specifically, we would typically be looking for "borderline" curves that were eliminated by one heuristic but accepted by the other, and then would need to determine which of these curves had rank 6. One might only expect a handful of relevant curves to be produced (say 10 or 20), and the statistical significance of the end result from such an experiment would probably be rather iffy.[18]

Another alternative method for a rank heuristic is to perform an integrated estimate via the explicit formula. For instance, Rubinstein [51, (1.15)] reports the bias for various curves in terms of such an integrated quantity. As we observe later (§7), the explicit formula can actually be used to produce an upper bound on the (analytic) rank when assuming GRH, but here we are more interested in heuristic aspects.

These methods are dependent on the choice of a test function, but (following [51]) upon taking a suitable sum over primes up to X one essentially has an indicator I that is $(r - 1/2) + \frac{1}{\log X}\sum_t \frac{X^{it}}{it(1+it)}$ where the sum is over noncentral imaginary parts of zeros (assumed to be on the central 1-line). With a prediction of $2\pi/\log N$ for the lowest height zero, the secondary terms are of size $\frac{\log N}{\log X}$ though admittedly with fairly reasonable constants. However, in our experiments, we might have $N = 32d^2 \approx 10^{37}$ and $X \approx 10^5$, and thus (just as with the Mestre-Nagao sums) we would likely have to sort through a lot of examples where the estimation from I was higher than the actual rank. Below we shall indicate a case where we had perhaps 10^9 rank 7 survivors (meaning here that they survived the 2-descent test) to which we wished to apply a rank heuristic, and this obliged us to use one that was relatively fast. As with the Mestre-Nagao heuristic, it seems this method should be more useful when the rank is large compared to the conductor (or d).

A final idea is to produce (say) two independent statistical measures from non-overlapping ranges of primes. For instance, one might take the primes up to 10^3 as a first cut-off, and then the primes from 10^3 to 10^7 as another. However, the second such sum here must be taken quite long if its deviation is to be (say) equal in size to a rank increment; with the above choices, a rank increment is $(7-3)(\log 10) \approx 9.2$, while the deviation is $\sqrt{\frac{(\log 10)^2}{2}(7^2-3^2)} \approx 10.3$. Again there are secondary terms, and in the end we concluded that all this was too much work for a mere heuristic.

6. Use of the Cassels-Tate pairing and beyond

The 2-Selmer test, even when combined with a Mestre-Nagao heuristic, still leaves a large number of curves that could attain the prescribed target rank. For the surviving curves we used the Magma [5] implementation of the `CasselsTatePairing` due to Donnelly [13]. This has the feature that it requires solving a conic only over the base field (whereas [8] demanded nontrivial calculations in the 2-torsion field).

We have a pairing $\mathrm{Sel}^2(E) \times \mathrm{Sel}^2(E) \to \mathbf{Z}/2\mathbf{Z}$ such that a 2-covering $C \in \mathrm{Sel}^2(E)$ trivially pairs with all C' if and only if C is in the image of $\mathrm{Sel}^4(E) \to \mathrm{Sel}^2(E)$; so we obtain precisely the same information as doing 4-descent on E. Thus by taking a basis of the 2-covers (after

[18] Comparing the heuristics on a smaller data set is not unreasonable, but could produce bias as the curves with relatively small d for a given r tend to have more significant Mestre-Nagao sums, while as stated above, we would be more interested in the borderline ones.

removing 2-torsion), the 2-covers that lift to 4-covers are exactly those in the kernel of the pairing matrix on such a basis.

Each call to `CasselsTatePairing` takes typically about 0.2 seconds[19] (though it can depend upon the size of d and the target rank), and in some cases computing the 2-covers themselves is nontrivial. However, in the most common case where our target rank equals the 2-Selmer rank modulo torsion, we need only find one nontrivial CTP result to conclude the rank is smaller than the target. Note also that one can apply the `CasselsTatePairing` to each of the isogenous curves.

6.1. Use of higher-descent pairings. — We are indebted to Fisher for the description of these, and for running the Magma programmes that implement them. Given a 2-isogeny $\phi : E \to E'$, Fisher computes the image of $\mathrm{Sel}^{4\phi}(E) \to \mathrm{Sel}^{\phi}(E)$ and similarly for the dual isogeny. This involves the solution of quadratic forms (over $\mathbf{Q}$) of ranks 3 or 4, then the computation of local points, minimisation and reduction steps, and finally linear algebra over $\mathbf{F}_2$. See [21], which extends the method in [6].

For our congruent-number twists, there are three choices of 2-isogeny, thus giving three upper bounds for the rank from the higher descents. Furthermore, Fisher has a similar pairing corresponding to full 8-descent that is applicable when the curve has (as in our case) full 2-torsion. We do not have precise timings for these higher descents, but a typical example might take 10 seconds for each 4ϕ-computation.

7. Bounding the analytic rank via the explicit formula

The use of the explicit formula to bound the analytic rank seems to be first described by Mestre [36]. Other examples appear in [43] and [3].

Let F be even, continuous, and supported on $[-1, 1]$, and write $F_S(x) = \frac{1}{S}F(x/S)$. Define $\hat{F}(t) = \int e^{ixt}F(x)\,dx$ so that $\hat{F}_S(x) = \hat{F}(Sx)$. The Guinand-Weil explicit formula [24, 59] applied to an elliptic curve L-function (see [36, p.219]) then gives

$$\sum_{\gamma} \hat{F}(S\gamma) = F_S(0)\log\frac{N}{4\pi^2} - 2\int_0^\infty \left(\frac{F_S(x)}{e^x - 1} - F_S(0)\frac{e^{-x}}{x}\right)dx$$
$$- 2\sum_{p^m}(\alpha_p^m + \beta_p^m)F_S(\log p^m)\frac{\log p}{p^m}$$

Here the γ-sum is over nontrivial zeros $1 + i\gamma$ (with multiplicity) of the L-function, N is the conductor, and $\alpha_p + \beta_p = a_p$ with $\alpha_p\beta_p = 0$ if $p|N$ else $\alpha_p\beta_p = p$. When $\hat{F}(0) = 1$, as $S \to \infty$ the left side converges to the analytic rank (from the $\gamma = 0$ terms).

A basic principle of L-functions ([30], [9, §10.3]) is that we can approximate values (via the functional equation) using $O(\sqrt{N})$ terms of the Dirichlet series. Here we hope to do better than this, as we only want an upper bound on the analytic rank. If we choose $\hat{F}$ to be nonnegative on the critical line, then the evaluation of the right side for any S-value will give an upper bound on the analytic rank (assuming GRH). The method does best when the closest noncentral zeros are not too near to the central point, as then (heuristically) a smaller S-value will suffice to make $S\gamma_n$ large enough so that the decay in $\hat{F}$ dominates. One expects

[19] Some of this is implementation-dependent; the discriminant of an auxiliary conic has a factor that is the size of the twisting parameter d, but this conic construction could be modified if needed.

the (low-height) zeros to be spaced at about $2\pi/\log N$, though elevated analytic rank might tend to make the first noncentral zero a bit larger. We often found that taking $e^S \approx N^{1/4}$ or even as small as $N^{1/6}$ or $N^{1/8}$ would suffice to give a suitable rank bound.

The condition that $\hat{F}$ be nonnegative can be recast as saying that it is the square of some entire function, and so F is the self-convolution of some function supported on $[-1/2, 1/2]$.[20] We will normalise so that $\hat{F}(0) = 1$. We want $\hat{F}$ to be concentrated as much as possible around $t = 0$. For instance, we might want to minimise $\int \hat{F}(t)\,dt = 2\pi F(0)$. Note that we do not require that F itself be nonnegative, though this essentially follows from our other conditions.

Mestre uses Odlyzko's function $\hat{F}(t) = \pi^4(\cos t/2)^2/(t^2-\pi^2)^2$, though for us perhaps simply $\hat{F}(t) = (\sin t/2)^2/(t/2)^2$ is superior. Booker [4, §3] introduces a more complicated method,[21] for instance taking

$$\hat{F}(t) = \frac{(\sin t/2M)^2}{(t/2M)^2}\left(\sum_{k=0}^{M-1} c_k \cos\frac{kt}{2M}\right)^2$$

for some (large) integer M, where the c_k are undetermined coefficients.[22] Indeed, one gets a quadratic form in the c_k, which can be minimised subject to the condition that $\hat{F}(0) = \sum_k c_k = 1$.

Booker's method can give a slightly sharper upper bound on the analytic rank, but usually the gain is rather small (say 0.05). The point seems to be that the use of an interpolating sum can quell contributions from zeros that are not too close to $t = 0$, but the uncertainty principle precludes the method from sharply distinguishing zeros at $t = 0$ from those that are close by. In the hardest examples, we expect that there are indeed such zeros that are fairly near to the central point.

We implemented the above method (without Booker's extension) for the function $F(x) = 1 - |x|$ compactly supported on $[-1, 1]$, that is, $\hat{F}(t) = (\sin t/2)^2/(t/2)^2$. Our C-based implementation could compute with $S = 26$ in about an hour. It does not spend much time computing the a_p, as this can be done quite efficiently for the congruent number curve via a pre-computation which enumerates over $a^2 + 4b^2$ in annuli (rather than solving $p = a^2 + 4b^2$ for each p). The Kronecker symbol computations (for a given d) are nonnegligible though not dominant, and similarly with the time spent by the memory-management subsystem in looking up pre-computed a_p values.

Booker's method incurs some overhead regarding the bookkeeping with the quadratic form. However, the additional computations with minimising the quadratic form are typically negligible, and indeed, in a case where the a_p-computation dominates the running time, there would be no reason (other than simplicity of implementation) not to use it.

As an illustrative example, we consider $S = 18$ for $d = 32058375240488794$. The upper bound on the analytic rank from the direct method is 7.1379, while Booker's method with $M = 10$ gives 7.0901 and increasing to $M = 40$ yields 7.0885, so it seems that we have essentially reached the point of diminishing returns. The minimising vector $\vec{c}$ for $M = 10$ is given

[20] The Paley-Wiener theorem [44] reinterprets supp(F) in terms of the exponential type of $\hat{F}$.

[21] Booker and Dwyer showed (see [3, Remark 1.2]) the Elkies curve [16] has analytic rank at most 28 (under GRH) via this, but both Booker and Bober tell us the method of [3] suffices.

[22] He actually takes $(\sin t/4M)^4/(t/4M)^4$ as the multiplier, as the contribution from trivial zeros is a bit simpler (to write) when $F(t)$ is continuously differentiable.

approximately by

$$(0.02354, 0.18034, 0.03976, 0.18067, 0.03592,\ 0.17445, 0.03035, 0.16604, 0.01320, 0.15573).$$

In Figure 1, we plot $\hat{F}(t) = (\sin t/2)^2/(t/2)^2$ versus the optimal $\hat{B}_{10}(t)$ given by Booker's method for $(M, S, d) = (10, 18, 32058375240488794)$, the former being the solid line. For this d, with $S = 27$ the direct method gives a rank bound of 5.99, so that (assuming BSD and GRH) the rank is not 6 (we have 4 independent points on this curve, so the rank is presumably 4).

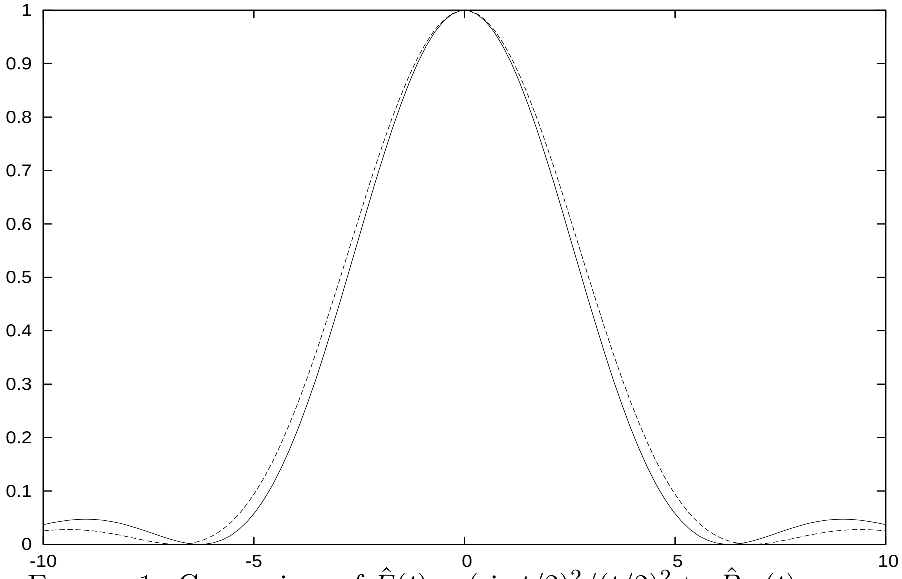

FIGURE 1. Comparison of $\hat{F}(t) = (\sin t/2)^2/(t/2)^2$ to $\hat{B}_{10}(t)$

We have $\log N \approx 78.78$ and $S = 18$ here, so that the re-scaled mean spacing of low-height zeros is $2\pi S/\log N \approx 1.44$. However, the effect at having 4 zeros at the central point must be taken into account in this. Also, noting that the maximum of $\hat{F}(t) - \hat{B}_{10}(t)$ is about 0.0226 near $t \approx 8.45$, while $\hat{B}_{10}(t)$ exceeds $\hat{F}(t)$ by twice this amount around $t \approx 4$, the placing of zeros must be rather delicate to simultaneously have both $\sum_\gamma \hat{F}(S\gamma) \approx 7.1379$ and $\sum_\gamma \hat{B}_{10}(S\gamma) \approx 7.0901$. Looking at Figure 2, one sees that the minor contributions from distant zeros are approximately halved with $\hat{B}_{10}$. For instance, there should be approximately 4 zeros between 4π and 6π, contributing maybe 0.02 more to $\hat{F}$ than $\hat{B}_{10}$, this then being doubled by evenness.

8. Our experiments and results

We performed three experiments with the first method (§3.1.1), which we describe in this section. We keep the notation of §3.1.1, searching for $(v : u)$ pairs with $1 \leq v < u \leq L$ up to some limit L, with d the square-free part of $uv(u + v)(u - v)$. We write $s(x)$ for the

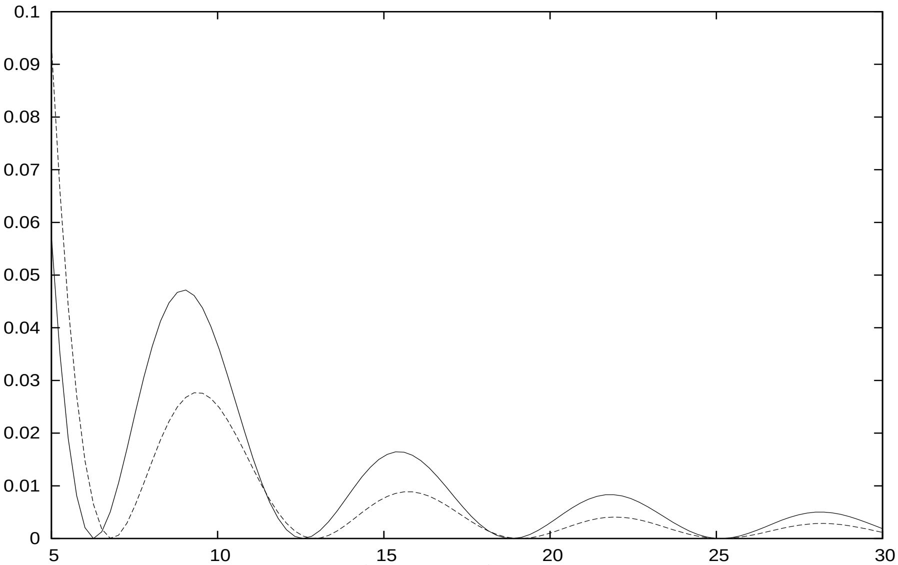

FIGURE 2. Comparison of $\hat{F}(t)$ and $\hat{B}_{10}(t)$ for larger values of t

square-free part of x (which is the minimal positive integer such that $xs(x)$ is square), and $B(u,v) = \lfloor \log_2 s(uv) \rfloor + \lfloor \log_2 s(u^2 - v^2) \rfloor$.

In Experiments P_{6a} and P_{7a} we took $L = 10^8$ and restricted to $B(u,v) \leq 60$. This kept us within 64-bit arithmetic (though latterly we determined this was not a real excuse), and more importantly greatly limited the number of d that got sent to the 2-Selmer tests. One goal of this first data collection was to find as many twists of rank 6 or more as possible in a given d-range.

In Experiment P_{7b} (and P_{8b}), we took $L = 10^7$ but did not restrict d at all. Here our goal was to find as many rank 7 twists as possible, and possibly one of rank 8.

In Experiment P_{8c} we took $L = 5 \cdot 10^7$, and restricted $B(u,v) \leq 80$. Here our goal was simply to try to find a twist of rank 8.

In all cases, we used Monsky's formula (and possibly the extension by Rogers) before turning to our heuristic Mestre-Nagao sum Σ_5. With Experiments P_{6a}, P_{7a}, and P_{8c} we required $\Sigma_5 \geq 35$, while for P_{7b} and P_{8b} we required $\Sigma_5 \geq 40$. A rough estimate (with $L = 10^5$ for $r = 6$) is that about 81.5% of the curves were eliminated by the $\Sigma_5 \geq 35$ condition. From both the analysis in §5 and the data we obtained, we expect that only a few curves of the desired rank were accidentally eliminated by this criterion. Each of the above experiments took roughly 6-7 core-years (for instance, about 4 months on 19 cores).

8.1. Data from these experiments. — We label the first experiment by P_{6a} and P_{7a}, splitting the obtained data into parity classes of the rank. Similarly, the data from the second experiment falls under P_{7b} and P_{8b}, and the third consists of P_{8c}.

In Table 1, for the experiments we list the L-limit, the B-limit (if any), the Mestre-Nagao bound Σ_5, the target rank r; then the number of curves that survived the 2-Selmer tests, then the 4-descent Cassels-Tate pairing, then the number that survived the 4ϕ- and 8-descent pairings of Fisher; and finally the number N_r for which we found at least $(r-1)$ independent points. The others were eliminated (under BSD/GRH) by an explicit formula calculation, with 3-descent being used for 3 curves from P_{6a}.

	L	B	Σ_5	r	2-Selmer	CTP	$F_{4\phi}$	F_8	N_r
P_{6a}	100 million	60	35	6	37873578	21016⋆	3691	2006	1261
P_{7a}	100 million	60	35	7	1912493	71	17	16	13
P_{7b}	10 million	∞	40	7	217704329	4902	67	21	12
P_{8b}	10 million	∞	40	8	8576723	3	0	0	0
P_{8c}	50 million	80	35	8	22516203	9	0	0	0

Table 1. Conditions for experiments (first method), and survivor counts

Note that with Experiments P_{6a} and P_{7a}, we actually applied the Monsky-like formulæ of Rogers (on the isogenous curves) in a separate run, for these experiments were started before we received the code from Rogers. They were thus partially intertwined into the further CTP sifting, and the numeric accounting listed was actually determined after the fact.

Furthermore, there was a minor programming bug in one of the cases in some initial versions of our code, and it thus failed to properly eliminate about 4% of curves in some cases; these were easily detected and eliminated upon applying the `CasselsTatePairing` as in §6. With this caveat, the curves that survive the 2-Selmer tests have a sufficiently large 2-Selmer rank on all isogenous curves.

For comparison with the data from the sifting method (§8.2), we note that 352 of the (presumed) rank 6 curves have $d \leq 2^{50}$. The rank 7 data obtained from Experiments P_{7a} and P_{7b} have large overlap, and in the end we only get 15 twists of rank 7 from combining these.

In the even parity case of the first experiment, we actually had approximately 35000 curves that survived the CTP sifter, but in order not to skew the counts in the rank 6 statistics we enforced a bound of $d < 2^{60}$ rather than $B(u, v) \leq 60$. This is the reason for the asterisk on the P_{6a}/CTP entry (21016) in Table 1.

For the first experiment, processing the 40 million curves took around a core-year. The even parity data for the second experiment, comprising about 5 times fewer curves, took about the same amount of time,[23] and the data for the third experiment was of the same magnitude. The odd parity data for the second experiment was processed in 3 weeks on a cluster of 128 threads, so about 7-8 thread-years.[24]

8.2. Results from 2-Selmer sifting. — Recall that Rogers used his variants of Monsky matrices (§4.3) for a nearly exhaustive search for d up to a given bound for a given target rank (on all isogenous curves). As noted there, we exclude[25] prime factors larger than 10^8. The largest prime divisor in our rank 6 data is 78988561 from $d = 681563383055674$, and four others[26] have a prime factor exceeding 10^7.

In Table 2 we give a list of the smallest d for each rank $r \leq 7$, with the proof of correctness following as in §4.3. The "num" column indicates where the curve stands (in order of d)

[23] The d are larger, making each test slightly harder, while the chance of hitting a nontrivial CTP result is lower when the target rank is 8 rather than 6.

[24] Specialised code could undoubtedly make the CTP test faster, but we have not pursued this.

[25] In the rank 8 data up to 2^{60}, we find 10.2% of the survivors with a prime factor exceeding 10^7, another 21.9% with one exceeding 10^6, while the majority (67.9%) have no prime factor this large.

[26] The most notable are $d = 336476810846858$ with $p = 22425577 > \sqrt{d}$ and 132233570668249 which is $41 \cdot 227089 \cdot 14202401$. The latter and $119222067089 = 3769 \cdot 4729 \cdot 6689$ are currently the only known rank 6 twists that only have 3 prime factors.

among those that survive the 2-descent test, for instance the first rank 5 curve is the 742nd to survive the 2-descent test for rank 5.

r	d	factorisation	num
1	5		1
2	34	$2 \cdot 17$	1
3	1254	$2 \cdot 3 \cdot 11 \cdot 19$	3
4	29274	$2 \cdot 3 \cdot 7 \cdot 17 \cdot 41$	2
5	48272239	$23 \cdot 31 \cdot 79 \cdot 857$	742
6	6611719866	$2 \cdot 3 \cdot 17 \cdot 31 \cdot 449 \cdot 4657$	1346
7	797507543735	$5 \cdot 7 \cdot 17 \cdot 97 \cdot 173 \cdot 79873$	4388

Table 2. Smallest rank r quadratic twists $y^2 = x^3 - d^2x$

8.2.1. *Data and comments.* — One difficulty with this method is the large amount of unwanted 2-Selmer survivors that then fail a 4-Selmer test. Comparatively, the (u, v)-search sidesteps the (substantial) rank 0 subset. Similarly, a point on a rank 1 curve is not likely to be of low height, so the (u, v)-search avoids most rank 1 curves.

We implemented a variant of the C code that Rogers gave us. With the upper limit of 10^8 on prime divisors of d, it takes about 2 minutes to find all 955353 rank 6 survivors up to 2^{40}, about 2.5 minutes to find all 529011 rank 7 survivors up to 2^{45}, and again about 2.5 minutes to find all 36771 rank 8 survivors up to 2^{50} (all timings on one core), with the behaviour close to linear in the D-bound.

In Table 3 we catalogue our experiments: we list the target rank, the D-bound for the computations, the Mestre-Nagao bound we used (if any), the total number of survivors, the number of survivors that exceeded the Mestre-Nagao bound,[27] the number that survived the Cassels-Tate 4-descent pairings, the number that survived Fisher's 4ϕ- and 8-descent pairings, and the number N_r for which we have r independent points. All 8-descent survivors not included in the final column were eliminated via the explicit formula (assuming GRH/BSD).

	r	D	Σ_5	survivors	large Σ_5	CTP	$F_{4\phi}$	F_8	N_r
R_{6a}	6	2^{50}	35.0	2343956262	53082687	3455	849	819	577
R_{7a}	7	2^{55}	35.0	1382722102	31455895	29	9	9	8
R_{7b}	7	2^{60}	45.0	55406567157	26360572	61	28	28	23
R_{8a}	8	2^{60}		193727581	193727581	2	0	0	0
R_{8b}	8	2^{65}	35.0	9668039478	218261949	2	0	0	0
R_{8c}	8	2^{70}	45.0	413434136874	193744327	2	0	0	0

Table 3. Data for the 2-Selmer searches for small d

In particular, a putative rank 8 twist with $d \leq 2^{60}$ must have a prime factor exceeding 10^8, and we have fair confidence there is no rank 8 twist with $d \leq 2^{70}$.

[27] Note the percentages for a given Σ_5 will differ from the earlier experiments (those were biassed toward curves with a point of small height). Also, the speed of the Mestre-Nagao computations becomes non-negligible, even dominant in some ranges, particularly when 64-bit arithmetic is exceeded.

Note that the rank 6 data has 50.8% of its “large Σ_5” data with odd d, while this percentage is 25.3% for rank 7, and 40.8% for rank 8 (similarly for survivor counts).

8.2.2. *Comparison to the first method.* — We find that the parametrisation from a point of small height (§3.1.1) found 352 of the 577 rank 6 twists up to 2^{50}. The smallest Σ_5 in this range is approximately 38.064 (for $d = 562073132513082$); as this is sufficiently greater than our bound of 35.0, it seems reasonable to expect that we did not miss any rank 6 curves due to this. The smallest d that was missed by the other method was $d = 855100330394$, with a point of (naïve) height 7980742225.

9. Data for high ranks

9.1. Rank 6 data. — Upon applying the 4-descent Cassels-Tate pairing (on all isogenous curves) in Experiment P_{6a}, we were left[28] with 21016 rank 6 survivors with $d < 2^{60}$. Fisher then used pairings for higher descents to reduce the count of rank 6 survivors to 2006. We were then able to find at least 5 independent points (via searching to height 10^5 on the 2-covers of all isogenous curves) on 1230 of these.

For the 776 remaining curves, we turned to the explicit formula methodology. Applying this with $S = 26$ (or smaller values), we were able to show that the rank was less than 6 (assuming BSD and GRH) for all but 49 of them. Raising this to $S = 30$ left 34 twists. For 22 of the remaining curves, we were able to find at least 5 independent points via searching to height $3 \cdot 10^6$ on the 2-covers of all isogenous curves. Fisher then provided us with enough points on 9 of the 12 remaining curves, by searching on 4-covers that he computed as in [21, §5]. We probably could have eliminated the final 3 rank 6 survivors by the explicit formula, but instead chose to use the 3-Selmer machinery of Magma. This was originally implemented by Stoll based on [52]. Due to recent class group improvements by Donnelly, each run took only a few days (assuming GRH), and found each 3-Selmer rank to be 4 (as expected).

For Experiment R_{6a} using the method of §4.3, we had about 2.3 billion 2-Selmer survivors for $d \leq 2^{50}$, of which around 53 million had sufficiently high Mestre-Nagao indicator Σ_5. The 4-descent Cassels-Tate pairing reduced this to 3455 survivors, and Fisher's pairings left us with 819 of which 577 have rank 6, while using $S \leq 26$ with the explicit formula eliminated the other 242 upon assuming BSD and GRH.

For 150 of the 1486 rank 6 curves we currently only have 5 independent points, and thus are relying on a Parity Conjecture (we expect that 4-cover computations should resolve most of these in the near future).

9.1.1. *Selmer rank data.* — We can ask how often these (presumed) rank 6 curves have nontrivial even part of Ш. We summarise this data in Table 4. The curve E_d is $y^2 = x^3 - d^2x$, the curve E'_d is $y^2 = x^3 + 4d^2x$, and $E_d^{\pm}$ are $y^2 = x^3 - 11d^2x \pm 14d^3$. The next 3 columns indicate how many curves have reduced 2-Selmer rank (that is, modulo torsion) of the given size, while the fourth column gives the number of curves with reduced 4-Selmer rank of 8. Note that if any isogenous curve has reduced 4-Selmer rank 8, then the reduced 2-Selmer rank of E_d must be at least 8. The information is then replicated (to the right) for the 577 rank 6 twists with $d \leq 2^{50}$.

[28] The total amount with $B(u, v) \leq 60$ was about 35000, but we chose to switch to a hard d-bound.

	6	8	10	8	6	8	10	8
E_d	1309	162	15	1	513	58	6	0
E'_d	1184	300	2	17	516	61	0	4
E^+_d	1369	117	0	0	553	24	0	0
E^-_d	1218	265	3	5	513	64	0	0

Table 4. Selmer rank data for rank 6 twists

There are 12 rank 6 curves that have reduced 2-Selmer rank 8 on all isogenous curves. Of these, two ($d = 457038106894219, 13449178862457819$) have reduced 2-Selmer rank 10, while $d = 86784274056751354$ has reduced 4-Selmer rank 8 on both E_d and E'_d. Three others from these 12 have reduced 4-Selmer rank 8 on one of the isogenous curves.[29]

9.1.2. *The ratios prediction for the rank 6 data.* — As already suggested by the analysis in §5, the d which yield rank 6 are not equidistributed to various moduli. A similar phenomenon was described (for instance) in [58, §3]. A particular prediction from random matrix theory might be that, fixing a prime modulus p, the ratio of the number of rank 6 d-values that are nonzero quadratic residues (QR twists) to those that are nonquadratic residues should be about $(\frac{p+1+a_p}{p+1-a_p})^k$ for some k, possibly $k = 3/2 - r$ so $k = -9/2$ here.[30]

The data agree with this qualitatively quite well. We consider the 577 curves with $d \leq 2^{50}$ (the set of 1486 curves reveal similar data), and primes p up to 10^4. When a_p is not too close to zero, say $a_p^2 \geq p$ (which by sector equidistribution is 2/3 of the primes that are 1 mod 4), the quantity of QR twists almost always exceeds the quantity of non-QR twists precisely when a_p is negative. For $p \leq 10^4$ this fails only for $p \in \{4153, 5573, 8581, 9293\}$.

The quantitative fit is also not bad. Writing $S_p^\pm$ for respectively the number of QR and non-QR twists, the best-fit log-log slope derived from the 609 data points $\big[(\frac{p+1+a_p}{p+1-a_p}), S_p^+/S_p^-\big]$ for p up to 10^4 that are 1 mod 4 is approximately -3.5.

Another minor comment about this congruence data is that certain primes, namely those that are 1 mod 8, tend to divide d with large rank more often, as might be guessed from an analysis of quadratic residue symbols (say) in the Monsky matrix. For instance, 346 of the 1486 d-values are divisible by 13, while 729 are divisible by 17. Similarly, 373 are divisible by 41 compared to 208 divisible by 29.

9.1.3. *Rank 6 distribution.* — One might ask whether we can determine a putative distribution of rank 6 twists from our limited data of 577 curves. A graphical representation is in Figure 3, which is a log-log plot of rank 6 counts versus d, with the x-axis (the d-value) labelled in powers of 2.

If the rank 6 count satisfies a power law, the graph should be close to linear. The best-fitting $e^c D^a$ is $e^{-9.1} D^{0.45}$ for the 577 data points, though it could be imprudent to speculate from such limited data. One can similarly best-fit to $e^c D^a (\log D)^b$, but here the data seem

[29] Except for $d = 301980419090843394$ (where we only have 5 independent points), the isogeny invariance of the BSD-quotient implies that none of the rank 6 curves has any 8-torsion in Ш.

[30] Predictions along these lines (first seen in [11, Conjecture 2]), are sometimes called the "ratios conjecture" though that phrase has now taken upon a different meaning in the field of number theoretical random matrix theory, so we prefer "ratios prediction" instead.

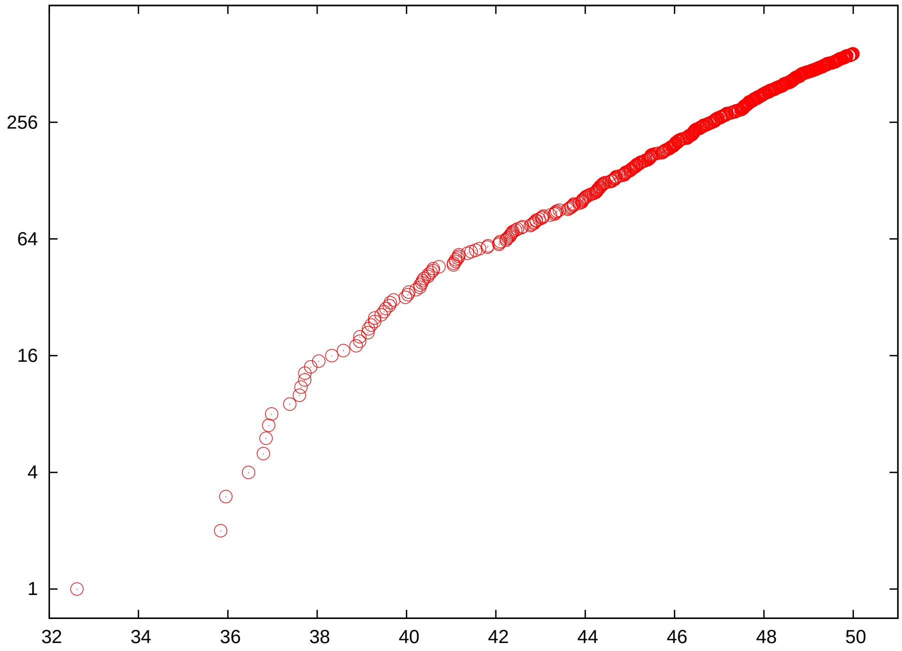

FIGURE 3. Log-log plot of rank 6 counts vs d-value

totally inadequate, yielding the nonsensical $(a, b) = (-0.52, 29.9)$. The best-fit $e^c(\log D)^b$ is approximately $e^{-42.6}(\log D)^{13.8}$.

9.2. Rank 7 data. — In Experiment P_{7a}, searching to $L = 10^8$ with $B(u, v) \leq 2^{60}$, we had almost 2 million rank 7 survivors of the 2-Selmer tests, while an application of the Cassels-Tate pairing left a mere 71 rank 7 survivors. For 13 of these we found 7 independent points via searching (see §9.2.1 and Table 5).

Fisher then eliminated 54 of the survivors via an additional isogeny descent (corresponding to a degree 32 map), and a 55th ($d = 326800477198750566$) via a pairing corresponding to full 8-descent. The remaining 3 curves were eliminated via the explicit formula (see §7 and Table 6).

Given that the lowest observed Σ_5 (at least in this d-range) for a rank 7 twist was 49.5, which is comfortably above our bound of 35, it seems reasonable to suspect that we did not miss any rank 7 twists due to the imposition of this bound.

For Experiment P_{7b} we searched up to $L = 10^7$ with no d-limit, with the purpose being to try to find as many rank 7 twists as possible. The Mestre-Nagao limit used here was 40. This found just 2 new rank 7 twists, though $d = 24951070826189778270$ has a point $(v, u) = (34440, 145343)$ with smaller height than for $d = 797507543735$. Three of the curves found in Experiment P_{7a} were not found here (cf. Table 5), as the height of the point is too large (namely $u \geq 10^7$).

The Cassels-Tate pairing for 4-descent (on all isogenous curves) eliminated all but 4902 of the 217.5 million 2-Selmer survivors. Fisher then reduced this to 67 via a pairing corresponding to an extra isogeny descent, and then further to 21 from the full 8-descent information. Table 5 includes the 12 that have rank 7, while the other 9 were eliminated by the explicit formula (see Table 6).

Experiment R_{7a} considered $d \leq 2^{55}$ and produced 1.38 billion 2-Selmer survivors, of which about 31 million had sufficiently large Σ_5. Of these, only 29 survived the 4-descent CTP test, and Fisher reduced the survivor count to 9. One of these was eliminated by the explicit formula, while the other 8 do indeed have rank 7, and in fact 5 of the 8 were not found by the methods of §3.1.1.
Experiment R_{7b} considered $d \leq 2^{60}$ and produced over 55 billion 2-Selmer survivors, of which about 26 million had sufficiently large Σ_5. Of these, only 61 survived the 4-descent CTP test, and Fisher reduced the survivor count to 28. Five of these were eliminated by the explicit formula, while the other 23 do indeed have rank 7. Here it appears reasonable to assume that the Mestre-Nagao bound $\Sigma_5 \geq 45.0$ did not exclude any rank 7 curves (the smallest Σ_5 in Table 5 is 49.5).

9.2.1. *Twists of rank 7.* — Table 5 lists the 27 rank 7 twists we found. The smallest (canonical) height $\hat{h}$ of a point on the curve is given, and if $u \leq 10^8$ we list $(v : u)$. The Mestre-Nagao sum Σ_5 is listed, while the last four columns record the 2-Selmer rank (modulo torsion) of the isogenous curves as per the notation with Table 4. The E'_d curve for $d = 674252816149274406$ has nontrivial 4-torsion in Ш (thus the asterisk)
The first 8 of these 27 were found by Experiment R_{7a} and the next 15 were additionally found by Experiment R_{7b}. The 13 twists with $u \leq 10^8$ and $d \leq 2^{61}$ (the final two due to the difference between $B(u, v) \leq 60$ and $d \leq 2^{60}$) were found by Experiment P_{7a}, while Experiment P_{7b} found the 12 twists with $u \leq 10^7$.

9.3. Data for rank 8. — Experiment P_{8b} with $L = 10^7$ and $\Sigma_5 \geq 40.0$ produced 8576723 twists of even parity that could have rank at least 8 from the 2-descent information. It took about 1 thread-year to process these with `CasselsTatePairing`.
We found no rank 8 curves, but did note 3 examples where the 4-descent information allows rank 8 on all isogenous curves, namely

$$d = 211348261439238289719306, d = 999813059534639477880290,$$
$$\text{and } d = 143336924388134266044361386.$$

Again Fisher ran his higher degree pairings on these twists, and they showed that the first has rank at most 4, the second at most 2, and the third at most 4.
Recall Experiment P_{8c} was aimed at trying to find a rank 8 twist, searching up to $L = 50 \cdot 10^6$ with $B(u, v) \leq 2^{80}$. We used a Mestre-Nagao cut-off of $\Sigma_5 \geq 35.0$. This yielded 22516203 curves that were possibly of rank 8. This is less than thrice as many as from the previous experiment, even though we increased the L-bound by a factor of 5 (so that one might expect 25x more curves). The reason is that most of the d lie outside our $B(u, v)$-bound. We found no curves of rank 8. There were only 9 curves that survived the `CasselsTatePairing` test, and Fisher could show that each of these has rank no more than 6 (in fact, each has rank 4 or less).
The 4-descent Cassels-Tate pairing was almost completely successful in Experiments R_{8a}, R_{8b}, and R_{8c}, leaving only a handful of curves to be eliminated by Fisher's 8-descent pairings. Again we found no curves of rank 8 (see Table 3).

9.4. Curves unresolved by 8-descent pairings. — As noted above, Fisher's higher degree pairings were able to eliminate a significant number of curves from our consideration.

d	$\hat{h}$	v	u	Σ_5	s_d	s'_d	s_d^+	s_d^-
797507543735	12.507	79873	235280	68.4	7	7	7	7
2067037027955295	18.758			66.8	7	7	7	7
2210857604820494	15.464	1492992	4542367	64.2	7	7	7	7
7616488732945534	23.857			72.1	7	7	7	7
11805305708568790	19.736			67.5	7	7	7	7
17056825108852669	23.999			55.8	7	7	7	7
17260078859287719	13.540	57175	733552	57.3	7	7	7	7
34872135169596005	19.814			50.3	7	7	7	7
46521610872080974	15.835	717241	7213568	49.8	9	7	7	9
55423105368015838	19.879			59.5	7	7	7	7
72909257919534679	22.254			56.8	9	7	7	7
82449281639107110	13.199	376909	383264	55.7	7	7	7	7
88770882541545735	19.146			49.6	7	7	7	7
187756280391835974	15.236	1441834	3511291	61.6	7	7	7	7
204817995109385574	12.369	62936	207689	63.7	7	7	7	7
254563891000186614	27.328			66.8	7	7	7	7
262456590553161245	18.419	2202624	98856259	64.2	7	7	7	7
344926532953988286	22.695			51.6	7	9	7	7
361526994851532510	13.912	70699	1069440	55.1	7	7	7	7
626123180330580614	33.543			52.9	7	7	7	7
667159490914887399	13.910	1577	1098816	49.5	9	7	9	9
674252816149274406	17.284	22664923	22702950	58.5	9	$9^\star$	7	9
763168101947645646	27.164			49.7	7	9	7	7
1500797991496877286	16.670	9221704	13454667	51.9	7	7	7	7
1584837449477135854	16.196	7545824	7677377	57.8	7	7	7	7
2495107082618977827 0	11.987	34440	145343	58.2	7	7	7	7
12301422184906259851584 6	14.621	770953	1901416	53.8	7	9	7	7

Table 5. Twenty-seven known rank 7 quadratic twists

However, some curves were still left, and for these we turned to using the explicit formula methods (§7).

With rank 6, as noted in §9.1 the explicit formula (with $S = 26$ necessary for a couple of curves) eliminated all extraneous 242 survivors of 8-descent for Experiment R_{6a}. Similarly, using $S = 26$ or smaller eliminated 727 of the 8-descent survivors from Experiment P_{6a}, and raising this to $S = 30$ pruned out 15 more.

We still had 3 curves of unknown rank from Experiment P_{6a}, which as noted above, Donnelly was to eliminate by using the 3-descent machinery in Magma. The d-values here were 54638221936676081, 120250527896300074, 529340421036976874. The explicit formula bounds with $S = 26$ were respectively $6.54, 6.48, 6.30$, and with $S = 30$ were $6.08, 6.01, 6.04$, so perhaps Booker's method would have sufficed.

With the rank 7 experiments, Fisher's 8-descent pairings only left 15 survivors for which we could not find sufficiently many independent points, and the explicit formula was able to eliminate all of these. Table 6 gives data[31] concerning these 8-descent survivors. We give the logarithm L_N of the conductor, the cutoff parameter S so that we consider coefficients up to e^S, and the analytic rank bound b (assuming GRH) from our computation. The r-column indicates how many independent points we found.

d	v	u	Σ_5	r	L_N	S	b
9216040803197470	34587520	58049161	50.5	5	76.3	18	6.89
248384796376777526	237402050	241161007	56.3	5	82.9	22	6.95
540517599062334679	648491137	779723300	47.9	5	85.1	20	6.84
570241054482429926	10299572\ 68797169	10303548\ 45705538	52.9	5	84.5	22	6.97
907957034379641662	37945242475	609479331936	48.8	5	85.4	18	6.76
1488977816607274326	18570553	18637634	45.7	5	86.5	21	6.84
1606761724662540886	56910150	63632821	35.3	5	86.6	17	6.91
9665275681497383942	1732721	3279154	46.1	3	94.8	21	6.87
132645399823739432742	1141600	3057073	51.2	3	95.4	20	6.84
493240121331611079055	228769	8313104	53.4	4	98.8	26	6.92
2229007454996999309574	1685977	3440594	41.0	1	101.1	18	6.95
15976519624716905845534	3141233	8517664	43.9	1	105.1	20	6.75
139720610704182979487414	90008	7952921	41.9	1	109.4	21	6.86
6737702895723796083999229	3438756	5072533	40.2	1	117.8	20	6.91
40044726772560104885558214	6739177	6803666	40.2	1	120.7	26	6.82

Table 6. Fifteen rank 7 survivors not eliminated by 8-descent pairings

9.5. Comments about higher descents.

— Fisher also noted about 15 curves (of target rank 7) for which his programme returns a bound of rank 1 for all three choices of 2-torsion point, even though 4-descent allows a rank of 7. The largest example here is $d = 1176708781878011833746605406$. This could be of interest, for we know a point of quite small height, namely $(v, u) = (4748713, 7465946)$, so that either #Ш must be quite large (much more so that the 4^6 from the 2-part), or the L'-value must be quite small (assuming that BSD holds). To be precise, we should have $L'(E_d, 1)/\#Ш_{\mathrm{odd}} \approx 1/388292$.

9.6. Bottlenecks to computing further.

— We briefly state the limiting factors for extending our experiments. With Experiments P_{6a} and P_{7a}, the d-cutoff of 2^{60} means most (u, v) are simply ignored. For Experiments P_{7b} and P_{8b} the time for the 4-descent Cassels-Tate pairings becomes dominant, due to the large number of false positives (indeed, we raised the Mestre bound Σ_5 due to this). In Experiment P_{8c} the d-cutoff (now 2^{80}) again eliminates the great majority of (u, v) pairs.

For the experiments using the method of Rogers (§4.3), the great number of false positives (thus Cassels-Tate pairings) again dominates, though perhaps with a sufficiently large Mestre bound the time for the 2-Selmer tests would be comparable.

[31] The twist $d = 9216040803197470$ has another small point $(v, u) = (156217, 322592)$.

10. Musings about a rank 8 twist

Here we speculate about when one might expect to find a rank 8 twist (if one exists). The conductor N_d will also appear below; for squarefree d it is either $32d^2$ or $16d^2$ depending on whether d is odd or even.

There are competing conjectures for the size of the smallest d of a given rank. The first is $r \sim \frac{1}{2}\frac{\log N_d}{\log\log N_d}$, which is an upper bound under GRH ([7, §2.11]). This is also the upper bound one obtains for the congruent number curve from 2-descent (counting the number of prime factors of d). One impetus for this guess is the function field analogue, though [56, Conjecture 10.5] is not restricted to twist families. The second guess is approximately the square root of this [45, Corollary in Appendix] (only an upper bound is stated), possibly with the $\log\log(N_d)$ in the numerator [20, (5.20)].[32] Again neither of these suggestions was originally restricted to twist families.

The first guess looks contraindicated by the data, as it predicts that by $d \approx 2^{60}$ we should already have seen rank 9. Of course one can speculate a smaller constant than $\frac{1}{2}$, though for an analogous problem [19, §5] it seems that the secondary terms actually seem to *increase* the main term for small conductors. Furthermore, one still must muse on an explanation of the $r = 7$ data. Even if one uses $d \approx 10^{15}$ (close to the second $r = 7$ data point) and takes ratios so as to dismiss the constant, one has[33] $F(2^{60})/F(10^{15}) \approx \frac{8}{7}$ for $F(d) = \frac{\log N_d}{\log\log N_d}$, implying 2^{60} to be the expected size for rank 8 under this model.

The situation is not quite so bad with a guess such as[34] $r \approx c\sqrt{\log N_d}$, at least if one (rather rashly) assumes the $r = 7$ data to be an outlier point, and so substitutes $d \approx 10^{15}$ for it. But even upon this assumption one finds (by taking ratios as above, and solving $G(x)/G(10^{15}) \approx 8/7$ with $G(d) = \sqrt{\log N_d}$) that $r = 8$ should appear around $6.6 \cdot 10^{19}$, bordering the range where we have guarded confidence.[35]

11. A variant of a heuristic of Granville

We now give a heuristic of Granville, which purports to bound the rank of elliptic curves in various families, most specifically quadratic twists. We then give extensions of this heuristic, and some warnings concerning similar problems.

Fix $E : Y^2 = f(X) = X^3 + aX + b$, and consider quadratic twists in projective form as

$$E_d : y^2 z = x^3 + ad^2xz^2 + bd^3z^3.$$

Granville's idea is that we can guess an upper bound on the number of (integral) (d, x, y, z) points on this surface (in some range, considering d as a variable) while *one* twist of sufficiently large rank will produce more points than this upper bound.

[32] One could also use Heath-Brown's result [25] on the density of d with a given 2-Selmer rank to conjecture something similar. Namely, the density of d with (reduced) 2-Selmer rank of r is proportional to $2^r/\prod_{j\le r}(2^j - 1)$, which on a logarithmic scale is $\sim 1/2^{r^2/2}$. If this predicts the size of the smallest d as $2^{r^2/2}$, inversion gives $r \sim \sqrt{2\log d/\log 2} \sim \sqrt{\log N_d/\log 2}$. Perhaps using isogenous curves could sharpen this, or one might consider an analogue for 2^l-Selmer ranks.

[33] Here we use $N_d = 32d^2$ as with odd d; the adjustments for $N_d = 16d^2$ with even d are minor.

[34] If we included a log log in the square root the d-estimate for $r = 8$ then goes down by about 10; it goes up by about 50 (to around 2^{72}) if one inserts the reciprocal of log log.

[35] With $r = 6$ and $d \approx 6 \cdot 10^{10}$ (the 2nd rank 6 twist), we similarly get $r = 8$ around $5.5 \cdot 10^{19}$.

11.1. Heuristic for the number of integral points. — The above formula for E_d has some implications for primitive integral points (x, y, z): first z is cube, say $\tilde{z} = \sqrt[3]{z}$; then $x \equiv 0 \pmod{\tilde{z}}$; and also $\bar{f}(x, dz) \equiv 0 \pmod{y^2}$ with $\bar{f}(X, Z) = X^3 + aXZ^2 + bZ^3$. We next split the variables into dyadic-like intervals,[36] taking $|d| \sim D$, then $|x| \sim T$ and $z \sim U/D$. We also assume that $(x^3 + ad^2xz^2 + bd^3z^3)$ does not generically have much cancellation, so that typically we have $y \sim \sqrt{DV^3/U}$ where $V = \max(T, U)$.
Following Granville's lead, we then proceed to estimate the number $N_D(T, U)$ of (d, x, y, z) points with $|d| \sim D$ and $|x| \sim T$ and $z \sim U/D$ as

$$N_D(T,U) \underset{?}{\ll} \sum_{d\sim D} \sum_{y\sim\sqrt{DV^3/U}} \sum_{\tilde{z}\sim\sqrt[3]{U/D}} \sum_{\substack{x\sim T,\tilde{z}|x \\ \bar{f}(x,d\tilde{z}^3)\equiv 0\,(y^2)}} 1.$$

The y^2-congruence has a density of solutions given by approximately $\sigma_f(y^2)/y^2$, where $\sigma_f(y^2)$ is the number of roots of f modulo y^2.
Granville uses this density to make the heuristic guess that

$$N_D(T,U) \underset{??}{\ll} \sum_{d\sim D} \sum_{y\sim\sqrt{DV^3/U}} \frac{\sigma_f(y^2)}{y^2} \sum_{\tilde{z}\sim\sqrt[3]{U/D}} \frac{T}{\tilde{z}} \ll TD\sqrt{\frac{U}{DV^3}}(\log DV^3/U)^{\eta-1},$$

where $\eta \in \{1, 2, 3\}$ is the average number of roots of f modulo primes.
Summing dyadically over T, U up to a bound G accrues an extra logarithm (from the $T = U$ contributions), and this gives us an overall bound of

$$C_D(G) \underset{??}{\ll} \sqrt{D}(\log G)^{\eta}$$

for the number $C_D(G)$ of points (d, x, y, z) with $|x|, z \leq G$ and $|d| \sim D$.

11.1.1. *Remarks.* —

- Granville notes that something like this should be provable for $G \ll D^{\delta}$ for some $\delta > 0$ via sieve theory, but he applies it for $G \approx e^{D^l}$ with $l > 0$.
  Compare the work of Hooley [27] regarding solutions to Pell equations.

- The original Granville heuristic dealt with $dY^2 = Z(X^3 + aXZ^2 + bZ^3)$, where one seems to lose a logarithm due to the Z-factor on the right.

- At a cruder level, writing $x = \tilde{x}\tilde{z}$, we have $\tilde{z} \sim S$, $\tilde{x} \sim S^2D$, and so $y^2 \sim S^6D^3$. Taking $(\tilde{x}, \tilde{z})$ pairs and (crudely) asking for y^2 to be a square of this size gives the probability $S^3D/\sqrt{S^6D^3}$. Summing over d yields $\sqrt{D}$ integral points per dyadic interval. Counting local roots for f is more precise in obtaining logarithms.

11.2. Relating this to high rank curves. — Next we count the number of points of "small" height on an elliptic curve of rank r and regulator R, where asymptotically the number of points up to (canonical) height H as $H \to \infty$ is $H^{r/2}/\sqrt{R}$. We assume (from ellipsoids) this is a lower bound for $H \gg R^{1/r}$ and that canonical and naïve heights are close. The conjectural BSD formula implies $R_d \ll \sqrt{D} \cdot L^{(r)}(E_d, 1)$ for twists $d \sim D$ (the

[36] In this section we use the notation $a \sim A$ to denote a in a dyadic interval $A \leq a \leq 2A$, or if necessary $A \leq a \leq A + A/F(A)$ where $F(x) \to \infty$ slowly as $x \to \infty$, say $F(x) = \log\log x$.

variation of the real period being dominant), and a Lindelöf-like hypothesis would imply $L^{(r)}(E_d,1) \ll_\epsilon D^\epsilon$ for all $\epsilon > 0$. From this we obtain

$$\frac{H^{r/2}}{D^{1/4+\epsilon}} \ll_\epsilon \frac{H^{r/2}}{\sqrt{R}} \ll \begin{array}{l}\text{\# of pts up to height } H\\ \text{on one rank } r \text{ twist of size } D\end{array} \ll C_D(e^H) \ll \sqrt{D}H^\eta.$$

Finally, we must guess how large we can take $H = D^l$. Plugging into the above, we get $r \leq 2\eta + \frac{3}{2l}$ as $D \to \infty$, so in particular any $l > 0$ gives an upper bound on ranks in twist families. Contrarily, allowing $l > 3/2$ implies $r \leq 2$ for the generic case $\eta = 1$, while the data of [58] suggest otherwise. Granville offers, in relation to the size of solutions to Pell equations, that $l = \frac{1}{2}$ seems reasonable, leading to $r \leq 2\eta + 3$.

For curves with full 2-torsion ($\eta = 3$) there is an additional subtlety, as every such curve is isogenous to one with only one 2-torsion point ($\eta = 2$), and it is unclear whether the bound for the latter should dominate. If so, one obtains an asymptotic bound of $r \leq 7$ for quadratic twists of an elliptic curve with 2-torsion. Obvious additions allow heuristic guesses about densities.

11.3. Adaptation to cubic twists. — One can easily adapt Granville's heuristic to cubic twists of $X^3 + Z^3 = 1$. Here $dY^3 = X^3 + Z^3$ for the twists, which gives the congruence $X^3 + Z^3 \equiv 0 \pmod{Y^3}$. Upon writing $\sigma(Y^3)$ for the number of cube roots of unity modulo Y^3, in each dyadic T-range we get a heuristic bound for the number of (d, X, Y, Z)-points with $d \sim D$ and $X, Z \sim T$ as

$$\sum_{Y \sim T/D^{1/3}} \mathop{\sum\sum}_{\substack{X,Z \sim T\\ X^3+Z^3 \equiv 0\,(Y^3)}} 1 \ll T^2 \sum_{Y \sim T/D^{1/3}} \frac{\sigma(Y^3)}{Y^3} \ll D^{2/3}(\log T)^{\nu-1},$$

where $\nu = 2$ is the average number of cube roots of unity modulo p as $p \to \infty$.

Summing dyadically over T gives a heuristic upper bound of $D^{2/3}(\log G)^2$, which is to be compared to the lower bound of $H^{r/2}/D^{1/6}$ coming from counting points in ellipsoids. With $H = \log G = D^l$, this implies $\frac{rl}{2} - \frac{1}{6} \leq \frac{2}{3} + 2l$, or $r \leq 4 + \frac{5}{3l}$. Granville suggests that $l = \frac{1}{3}$ is appropriate here, yielding $r \leq 9$. Here the finitely many exceptions would include the examples found by Elkies and Rogers [18] that have rank 11.

As noted by Elkies, the above curves are isogenous to $dY^3 = XZ(X+Z)$, where the right-side factors completely. From this, we might speculate that the situation of "arithmetic influence" is analogous to quadratic twists of curves with full 2-torsion.

11.4. Other twists, and the family of all elliptic curves. — It does not seem that Granville's heuristic can be directly adapted to quartic or sextic twists. However, probabilisitic reasoning of a similar sort, which seems not to go past what is already available via conjectures of Lang [29], suggests the following bounds: that $r \leq 21$ except for finitely many elliptic curves; that $r \leq 13$ for all but finitely many (Mordell) curves of the shape $y^2 = x^3 + k$; and that $r \leq 11$ for all but finitely many quartic twists of the congruent number curve, given by $y^2 = x^3 - nx$. Furthermore, one might presume that each of the "borderline" cases should have no faster than log-power growth as the parameter tends to infinity.

However, as with the cubic twist case above, the current records exceed the bounds suggested above, namely rank 14 with

$$y^2 = x^3 + 402599774387690701016910427272483x$$

for quartic twists (listed in [1, Acknowledgements]), rank 15 for the Mordell curve

$$y^2 = x^3 + 46974552981863676115647417$$

(see [17]),[37] and rank 28 for elliptic curves over the rationals (see [16]).[38]

11.5. Further warnings. — In addition to the previous paragraph, there are similar problems where one can notice possible failings of probabilistic models. For instance, Elkies notes (in a letter to Zagier [15] reproduced in the appendix to [57]), one can achieve large integral points on elliptic curves $Y^2 = X^3 + AX + B$ (where all four variables are parameters), namely an infinite family with $\frac{\log X}{\log H} \to 12$ where $H = \max(|A|^{1/2}, |B|^{1/3})$, whereas the probabilistic heuristic suggests a limiting upper bound of 10 for this quotient.[39]
Similarly, for integral points of small height on curves $y^2 = x^3 + Ax + B$, say $|x| \leq H^2$, $|y| \leq H^3$ with $|A| \leq H^4$ and $|B| \leq H^6$, upon looping over (x, y, A) and solving for B, one finds that there are H^9 such (x, y, A, B) with an integral point of small height. The paper [19] extends this observation to *pairs* of integral points, essentially indicating a bound like $H^8(\log H)^\bullet$ (for some unspecified power of $\log H$). One can pass from "integral points" of small height to "rational points" of small height via increasing the power of logarithm, and a similar logarithmic effect should come about from enlarging the notion of "small height" to any polynomial bound in H. However, the rank 11 (or 12) families of Mestre [37, 38] have 11 (or 12) independent points, with all of these of small height (indeed, for the points to be written down easily, they must of necessity be of polynomial height). So the natural extrapolation of this heuristic about k-tuples of points of small height appears to break down.

11.6. Data about Granville's heuristic. — There are several possible avenues of trying to collect data about Granville's heuristic, particularly the first consideration (in §11.1) regarding the bound on integral points on the twist surface. For instance, again sticking with the congruent number curve, one could take the twisting parameter d in a dyadic-like range around 10^4, so that the regulators might typically be of size 100. This is sufficiently small that one could expect to be able to find Mordell-Weil generators on all the twists, either by Heegner points for twists of rank 1, or by 4-descent and point searches for twists of higher rank. In fact, taking $d \sim 10^5$ or even $d \sim 10^6$ may be feasible, but one still must consider whether the asymptotic behaviour is beginning to be seen.[40]

[37] Elkies has "many" such $r = 13$ curves, but it is unclear if the growth rate exceeds a log-power.

[38] It so happens that this curve of Elkies has $\frac{\log N}{2 \log\log N} \approx 28.16$, and [19, §5] again gives some evidence (admittedly a bit tenuous) for such growth. Meanwhile, the rank 24 curve [32] actually has $\frac{\log N}{2 \log\log N} \approx 20.39$ (and Elkies has a rank 24 curve with 19.86 here) indicating that the secondary terms should play some rôle here. It has also been suggested that 28 is so large that one should not expect it to be a natural barrier. However, given the discussion in the next subsection, namely that parametrisations often allow one to beat the strong Lang surmises by a small amount, one might alternatively propose that 28 does not really exceed 21 to an irrefragable extent.

[39] Elkies considers $Q(t)Y(t)^2 = X(t)^3 + A(t)X(t) + B(t)$ where $\deg(X, Y, A, B) = (4, 5, 0, 1)$ and Q is quadratic; via a parameter count there should be a nondegenerate 0-dimensional solution variety, which happens to yield a rational point here. Upon scaling appropriately, the sparse (Pellian) set of t-values for which $Q(t)$ is a square then give large integral points via specialisation.

[40] One of Granville's concerns was whether one should really expect the bound $\sqrt{D}(\log G)^\eta$ as opposed to $\sqrt{D}(\log D)^\eta$, particularly in ranges where $\log G$ is much larger than $\log D$.

Another idea, especially as Granville uses solutions of Pell equations to intuit that $l = 1/2$, is to investigate solution sizes of conic twists, namely $dy^2 = f(x)$ with f quadratic. Some calculations in this regard have been carried out by Lacasse [28].

12. Data from affiliated experiments

Here we give the data for the experiments described in §3.2 and §3.3.

12.1. Initial data from rank 2 families. — We recall the rank 2 families introduced in §3.2. The first (I) has $w(w+1)(w-1)(w+2)(2w+1)(w^2+w+1)y^2 = x^3 - x$. Writing $w = m/n$, via the symmetries of the homogeneous octic polynomial in m and n, we can restrict to m, n that are not congruent modulo 3. The second (II) family has $3(w+1)(w-1)(w^2+2)(2w^2+1)y^2 = x^3 - x$. For both families, for the purposes of comparison, we considered w up to height 10^4, which meant only minor modifications to our factoring tables. One goal of this experiment was to see how many rank 6 curves are found up to height 10^4 – unless the count exceeds that from the rank 1 family, it is probably not worth trying to find a rank 8 example in I or II.

Table 7 lists the data we obtained from these experiments. It lists the number of m/n up to height 10^4 that survived the 2-Selmer test (note that a few d appear twice), the number of d that survived the 4-descent Cassels-Tate pairing,[41] then the number that survived Fisher's higher pairings, the number of curves on which we found at least $(r-1)$ independent points, followed by the number of unknowns (after applying the explicit formula machinery), and finally the smallest example (if any) for the target rank. The first two lines correspond to the full (u, v) experiment, and the latter ones to the indicated families. In these comparisons we omitted the Mestre-Nagao filtration step.[42] For some curves, Fisher's implementation took too long to compute a bound; the number of such failures is noted by a plus sign in the tabulation. For instance, 11 of the 6135 Family I rank 6 survivors took too long at the 4ϕ-step, and 342 of the 1437 remaining hit time constraints with the 8-descent pairing.

Fam	r	num	CTP	$F_{4\phi}$	F_8	N_r	?	first
(u, v)	6	16692	36	21	18	17	0	779/134
	7	740	0	0	0	0	0	
I	6	447030	6135	1437+11	122+342+11	1	449	227/210
	7	56197	38	3	0	0	0	
II	6	546208	7839	1416	380+6	18	133	103/41
	7	80275	76	5	4	0	0	

Table 7. Comparison with rank 2 families up to height 10^4

Family II contains a number of examples of notable rank. For instance, Elkies notes $m/n = 18$ gives the rank 5 twist $d = 205015206$ found by Rogers.[43] Similarly $m/n = 103/41$ gives

[41] There was also a missing step in the Magma integer factorisation code (not detecting powers before entering ECM in all cases), which caused `CasselsTatePairing` to take hours occasionally.

[42] There was one d in the (u, v) family which survived the descent tests but had rank only 4 (using the explicit formula); this $d = 344333282586$ has $\Sigma_5 \approx 34.566$, less than our cutoff of 35.

[43] This is cited in [50, Table 2] as appearing in [46], though [46] lists 4132814070 for rank 5.

the smallest rank 6 example, while both $m/n = 229/146$ and $m/n = 248/203$ yield $d = 61471349610$. We found 18 curves of rank 6 via searching to 10^5 on 2-covers (increasing to 10^6 yielded no additional ones), and currently have 133 remaining curves, having pruned the original list of 386 via the explicit formula with $S = 26$. As a general conclusion, this Family produces a comparable amount of high-rank curves to the (u, v) family, but also produces a lot more false positives.

Our results are largely inconclusive for Family I, due to computational difficulties.[44] Note that Family II had 4 rank 7 survivors after the 8-descent pairing was applied (all were eliminated by the explicit formula with $S = 22$ or less), while Family I had none. Also, the largest Σ_5 for Family I is only 54.569 (for $m/n = 7382/1531$), much less than 65.873 for Family II.

12.2. Data concerning the lattice search method. — We recall here the lattice-based method given in §3.3. We parametrise the square divisors of $uv(u + v)(u - v)$ via

$$d_1^2|u,\ d_2^2|v,\ d_3^2|(u+v),\ d_4^2|(u-v),$$

and then loop over pairwise coprime (d_1, d_2, d_3, d_4), looking for short vectors in the (u, v)-lattice. The lattice determinant $D^2 = \prod_i d_i^2$ gives an expected (u, v) size.

In Table 8 we list points on known rank 7 twists that have $\max_i(d_i) \le 10^3$ (where d_i^2 is the maximal square dividing the relevant expression). The measure u/D indicates how much enumeration of lattice points would be necessary to find such a point, while $T = \max(1, u/D)^2 \cdot \max_i(d_i)^4$ quantifies the total work needed to find it.

One can see that the (u, v) that are "easily" obtained are already known from the other experiments; indeed there is quite a large correlation between points of "small" height ($u \le 10^8$) and those found by this method. We estimate that searching the range $d_i \le 2500$ would take maybe a core-year with optimised code, somewhat ignoring various issues with taking squarefree parts and exclusion of small d_i (either individually or product-wise).

13. Complementary ideas

13.1. Short vector distribution. — Rubin and Silverberg give equivalent conditions for the unboundedness of ranks in quadratic twist families in [48]. This involves short vectors in the lattices $\mathcal{L}_{\alpha,d,d'} = \{(u, v) \in \mathbf{Z}^2 : d^2|(u-\alpha v), d'^2|v\}$ where α satisfies $d^2|f(\alpha)$ with $y^2 = f(x)$ defining the elliptic curve E.

Furthermore, their Remark 5.2 discusses that if an elliptic curve has at least one quadratic twist with rank exceeding 8, then there is a lack of uniformity in the distribution of these vectors.[45] This becomes somewhat problematic when one *starts* with a curve of rank 9 or more, and then twists it. Granville's heuristic attempts to bypass this issue by requiring that the twisting parameter d be in a dyadic-like interval. Thus for small d the implied constants can presumably be so large that larger ranks are allowed.

[44] Admittedly, we did not make so much effort, only applying the explicit formula with $S = 24$ (eliminating 25 of 475 curves), and searching on 2-covers up to 10^5.

[45] The cutoff of 8 here appears to be related to their use of a specific statistic to measure non-uniformity, and thus might be lowered by a sharpened heuristic.

d	v	u	u/D	T	$[d_1, d_2, d_3, d_4]$
797507543735	79873	235280	6.93	2.36E10	$[4, 1, 57, 149]$
	351232	367625	8.42	1.78E11	$[5, 224, 3, 13]$
	2705233	3764480	0.40	2.29E10	$[16, 167, 9, 389]$
2210857604820494	1492992	4542367	19.1	2.04E14	$[1, 864, 11, 25]$
17260078859287719	57175	733552	643	1.38E13	$[76, 5, 1, 3]$
	162377	5552384	138	3.51E14	$[368, 1, 109, 1]$
46521610872080974	717241	7213568	95.3	4.31E11	$[16, 1, 57, 83]$
55423105368015838	291489400	308455417	2.40	5.74E12	$[677, 10, 19, 999]$
82449281639107110	376909	383264	4166	4.86E12	$[4, 1, 23, 1]$
	37841	1068704	1429	1.70E11	$[4, 1, 11, 17]$
	6320405	6991648	101	7.42E13	$[292, 1, 237, 1]$
	1421797795	6888910758	0.09	7.71E11	$[199, 473, 833, 937]$
88770882541545735	19060673	197870608	4.87	4.96E12	$[676, 373, 7, 23]$
187756280391835974	1441834	3511291	211	1.09E13	$[7, 19, 125, 1]$
204817995109385574	62936	207689	4154	6.74E12	$[1, 2, 25, 1]$
	10030425	30779392	27.2	6.59E12	$[32, 5, 307, 23]$
361526994851532510	70699	1069440	2191	6.65E13	$[8, 61, 1, 1]$
667159490914887399	1577	1098816	19622	1.58E12	$[8, 1, 7, 1]$
	1595984	55443691	86.9	6.68E11	$[47, 4, 35, 97]$
674252816149274406	22664923	22702950	626	4.74E12	$[15, 41, 59, 1]$
1500797991496877286	9221704	13454667	151	3.80E13	$[9, 202, 1, 49]$
1584837449477135854	7545824	7677377	897	7.44E11	$[23, 4, 31, 3]$
2495107082618977827 0	34440	145343	72672	8.45E10	$[1, 2, 1, 1]$
1230142218490625985 15846	770953	1901416	316903	8.13E12	$[2, 1, 1, 3]$

TABLE 8. Values of (d_1, d_2, d_3, d_4) for some points on rank 7 twists.

13.2. Using visibility to bound rank.

— It was suggested to us by N. Bruin that we might use visibility (see [12] for instance) to bound the rank for some of the curves for which we still have not found enough independent points. The idea is to find l such that E_{dl} is known to have positive rank r_l, while one can suitably bound the 2-Selmer rank (modulo torsion) of E_d over $\mathbf{Q}(\sqrt{l})$ as s_l. Then one bounds the rank E_d as $\leq s_l - r_l$, hoping this is less than the desired target.[46] The quadratic field extension should trivialise an "unknown" 2-cover (or a pair of independent such covers), that is, one for which it is not known whether it has points over $\mathbf{Q}$.

However, in our case where we have already applied higher Selmer tests, we would have to take more than one quadratic extension – that is, the Ш[4] will become Ш[2] at each visibility step, and similarly for Ш[8] becoming Ш[4]. Recalling that Fisher has performed an 8-descent pairing on our curves of interest, we would need to make Ш[8] visible, and thus need to work (minimally) with a triquadratic extension.

Also, as B. Creutz pointed out to us (see [33, Proposition 4.3]), when there is full 2-torsion the dimension of the Selmer group over $\mathbf{Q}(\sqrt{l})$ increases significantly depending on the number

[46] In effect, one is making the Ш[2] of E_d visible upon passing to the quadratic field extension.

of primes dividing l. The idea is that various homogeneous spaces become everywhere locally soluble upon making the field extension.

13.3. Quadratic twists of other curves. — The first part of the above method generalises naturally to other curves, particularly those that have full 2-torsion. In general, we can use SQUFOF [54, 23] to factor $(u^3 + auv^2 + bv^3)$ when it is less than 2^{60}, and this is somewhat efficient. However, it is not clear to us how to compute an upper bound on the 2-Selmer rank as simply as with Monsky's formula. Another idea is to investigate curves with $\mathbf{Z}/2 \times \mathbf{Z}/6$ torsion, where one could additionally try to exploit the 3-isogeny.

Rubinstein notes that one could also change the problem slightly: consider (say) all curves in Cremona's database (the maximum rank is 4) – can you find a quadratic twist of rank 8 (or rank 6 if there is no 2-torsion) of *any* of these? This would already say something about Granville's heuristic.

14. Concluding Remarks

We briefly review what we consider to be the main findings of our work.

- It seems relatively easy (though certainly nontrivial) to find rank 6 quadratic twists of the congruent number curve. We found 577 in a nearly exhaustive search up to 2^{50}, and an additional 909 more up to 2^{60} (each of the latter has at least one point of small height).
  - Only 352 of the 577 curves up to 2^{50} had a point of "small" height.
  - Statistics regarding rank 6 curves are still unclear as to a prediction of a growth rate (§9.1.3).
  - The ratios prediction regarding popular congruence classes of high rank twists is qualitatively corroborated (§9.1.2).

- It seems rather difficult to find rank 7 quadratic twists. We found 23 up to 2^{60} (we expect this to be exhaustive, but our methods are statistical), and 4 more from points of small height.
  - Again only 15 of these 27 curves had a point of "small" height.

- We did not find a rank 8 quadratic twist of the congruent number curve.
  - Our searches for a rank 8 twist are sufficiently broad (up to 2^{70}, using a Mestre-Nagao heuristic as a filter) for this to cast doubt on some guesses about growth rates of ranks (§10).
  - However, as noted previously, the data point for $r = 7$ appears to be abnormally small, which might throw off any mundane curve-fitting analysis. In particular, a (slower) growth rate like $r \sim \sqrt{\log d / \log\log d}$ is not completely rejected by the experimental data.
  - As another measure, we have searched hundreds of millions of 2-Selmer survivors for rank 8, while for rank 7 the first curve appeared with the 4388th such survivor (see Table 2).
  - Granville gives a heuristic which, when suitably interpreted, could predict that 7 is the largest rank in the family (§11).

- There are other methods that could be used to try to find high rank twists.
  - Foremost of these would seem to be using the rank 2 parametrisation given by Family II in §3.2. Such a parametrisation could be a way to confound Granville's heuristic, in parallel with the examples in §11.5. However, even if it beats the "probabilistic" estimate on the rank, it might only do so by 1 or 2 at most.
  - The use of different search methods (§3.3) is another path to explore.

15. Electronic availability

The 27 twists of rank 7 and 1486 of (presumed) rank 6 are available for download from `http://magma.maths.usyd.edu.au/~watkins/PTS.r6r7`, the format being a Magma file that takes about 10 seconds to load, and this gives a set of (known) independent points for each twist in the corresponding arrays `RANK6` and `RANK7`.
Some of the code we used in this project can be downloaded from
`http://magma.maths.usyd.edu.au/~watkins/CONGCODE.tar`

References

[1] J. Aguirre, F. Castañeda, J. C. Peral, *High rank elliptic curves with torsion group Z/(2Z).* Math. Comp. **73** (2004), 323–331 `http://dx.doi.org/10.1090/S0025-5718-03-01547-3`

[2] B. J. Birch, H. P. F. Swinnerton-Dyer, *Notes on elliptic curves. II.* J. reine angew. Math. **218** (1965), 79–108. `http://resolver.sub.uni-goettingen.de/purl?GDZPPN002181169`

[3] J. Bober, *Conditionally bounding analytic ranks of elliptic curves.* In *ANTS X (2013)*, Proceedings of the Tenth Algorithmic Number Theory Symposium, edited by E. W. Howe and K. S. Kedlaya, The Open Book Series, Mathematical Sciences Publishers, 135–144.
`http://dx.doi.org/10.2140/obs.2013.1.135`

[4] A. R. Booker, *Artin's Conjecture, Turing's Method, and the Riemann Hypothesis.* Experiment. Math **15** (2006), 385–408. `http://projecteuclid.org/euclid.em/1175789775`

[5] W. Bosma, J. Cannon, C. Playoust, *The Magma algebra system. I. The user language.* In *Computational algebra and number theory*, Proceedings of the 1st MAGMA Conference held at Queen Mary and Westfield College, London, August 23–27, 1993. Edited by J. Cannon and D. Holt, published by Elsevier Science B.V., Amsterdam (1997), 235–265. Cross-referenced as J. Symbolic Comput. **24** (1997), no. 3-4, 235–265. `http://magma.maths.usyd.edu.au`

[6] A. Bremner, J. W. S. Cassels, *On the equation* $Y^2 = X(X^2+p)$. Math. Comp. **42** (1984), no. 165, 257–264. `http://dx.doi.org/10.1090/S0025-5718-1984-0726003-4`

[7] A. Brumer, *The average rank of elliptic curves I.* Inventiones math. **109** (1992), 445–472.
`http://dx.doi.org/10.1007/BF01232033`

[8] J. W. S. Cassels, *Second descents for elliptic curves.* J. reine angew. Math **494** (1998), 101–127. See also *Arithmetic on curves of genus 1, IV. Proof of the Hauptvermutung.* J. reine angew. Math **211** (1962), 95–112.
`http://resolver.sub.uni-goettingen.de/purl?GDZPPN002179873`

[9] H. Cohen, *Advanced Topics in Computational Number Theory.* Graduate Texts in Mathematics **193** (2000), Springer.

[10] K. Conrad, *Partial Euler products on the critical line.* Canad. J. Math. **57** (2005), 267–297. http://dx.doi.org/10.4153/CJM-2005-012-6

[11] J. B. Conrey, J. P. Keating, M. O. Rubinstein, N. C. Snaith, *On the frequency of vanishing of quadratic twists of modular L-functions.* In *Number theory for the millennium, I* (Urbana, IL, 2000), Papers from the conference held at the University of Illinois at Urbana–Champaign, Urbana, IL, May 21–26, 2000. Edited by M. A. Bennett, B. C. Berndt, N. Boston, H. G. Diamond, A. J. Hildebrand and W. Philipp, published by A K Peters, Ltd., Natick, MA (2002), 301–315. Preprint at http://www.hpl.hp.com/techreports/2000/HPL-BRIMS-2000-23.pdf

[12] J. E. Cremona, B. Mazur, *Visualizing elements in the Shafarevich-Tate group.* Experiment. Math. **9**, no. 1 (2000), 13–28. http://projecteuclid.org/euclid.em/1046889588

[13] S. Donnelly, *Computing the Cassels-Tate pairing on* Ш$(E)[2]$ *in Magma.* Unpublished notes (2008).

[14] A. Dujella, A. S. Janfada, S. Salami, *A search for high rank congruent number elliptic curves.* J. Integer Seq. **12** (2009), 09.5.8. http://www.cs.uwaterloo.ca/journals/JIS/VOL12/Janfada/janfada3.ps

[15] N. D. Elkies, *Letter to Prof D. B. Zagier* (1988). Second part of the Appendix to [57].

[16] N. D. Elkies, Z^{28} *in E(Q), etc.* NMBRTHRY list posting, May 2006. http://listserv.nodak.edu/cgi-bin/wa.exe?A2=ind0605&L=NMBRTHRY&P=50

[17] N. D. Elkies, *j=0, rank 15; also 3-rank 6 and 7 in real and imaginary quadratic fields*, NMBRTHRY list posting, Dec 2009. http://listserv.nodak.edu/cgi-bin/wa.exe?A2=ind0912&L=NMBRTHRY&P=14012

[18] N. D. Elkies, N. F. Rogers, *Elliptic Curves* $x^3 + y^3 = k$ *of High Rank.* In *Algorithmic number theory,* ANTS-VI (Burlington 2004), ed. by D. Buell, Springer LNCS **3076** (2004), 184–193. http://dx.doi.org/10.1007/978-3-540-24847-7_13

[19] N. D. Elkies, M. Watkins, *Elliptic curves of large rank and small conductor.* In *Algorithmic number theory,* ANTS-VI (Burlington 2004), ed. by D. Buell, Springer LNCS **3076** (2004), 42–56. http://dx.doi.org/10.1007/978-3-540-24847-7_3

[20] D. W. Farmer, S. M. Gonek, C. P. Hughes, *The maximum size of L-functions.* J. reine Angew. Math. **609** (2007), 215–236. http://dx.doi.org/10.1515/CRELLE.2007.064

[21] T. A. Fisher, *Higher descents on an elliptic curve with a rational 2-torsion point*, preprint, http://www.dpmms.cam.ac.uk/~taf1000/papers/higherdesc.html

[22] F. Gouvêa, B. Mazur, *The square-free sieve and the rank of elliptic curves.* J. Amer. Math. Soc. **4** (1991), 1–23. http://www.jstor.org/stable/2939253

[23] J. E. Gower, S. S. Wagstaff Jr., *Square form factorization.* Math. Comp. **77** (2008), 551–588. http://dx.doi.org/10.1090/S0025-5718-07-02010-8

[24] A. P. Guinand, *Summation Formulae and Self-reciprocal Functions (III).* Quarterly J. Math., **13** (1942), 30–39. http://dx.doi.org/10.1093/qmath/os-13.1.30

[25] D. R. Heath-Brown, *The size of Selmer groups for the congruent number problem, II.* Inv. Math. **118**, no. 1 (1994), 331–370. http://dx.doi.org/10.1007/BF01231536

[26] T. Honda, *Isogenies, rational points and section points of group varieties.* Japan. J. Math., **30** (1960), 84–101.

[27] C. Hooley, *On the Pellian equation and the class number of indefinite binary quadratic forms.* J. reine Angew. Math., **353** (1984), 98–131. http://eudml.org/doc/152668

[28] M.-A. Lacasse, *Sur la répartition des unités dans les corps quadratiques réels.* Master's thesis, Université de Montréal, December 2011. `http://hdl.handle.net/1866/6881`

[29] S. Lang, *Hyperbolic and Diophantine analysis*, Bull. AMS, **14** (1986), 159–205. `http://dx.doi.org/10.1090/S0273-0979-1986-15426-1`

[30] A. F. Lavrik, *On functional equations of Dirichlet functions.* Mat. USSR Izv. **1**, no. 2 (1967), 421–432. `http://dx.doi.org/10.1070/IM1967v001n02ABEH000565`

[31] S. McMath, F. Crabbe, D. Joyner, *Continued Fractions and Parallel SQUFOF*, Int. J. Pure Appl. Math. **34**, no. 1 (2007), 19–38. `http://ijpam.eu/contents/2007-34-1/2/2.pdf`

[32] R. Martin, W. McMillen, *An elliptic curve over Q with rank at least 24.* NMBRTHRY Listserver, May 2000. `http://listserv.nodak.edu/cgi-bin/wa.exe?A2=ind0005&L=NMBRTHRY&P=2105`

[33] K. Matsuno, *Elliptic curves with large Tate-Shafarevich groups over a number field.* Math. Res. Lett. **16** (2009), 449–461. `http://tinyurl.com/muoz5zl`

[34] F. Mertens, *Ein Beitrag zur analytischen Zahlentheorie.* (German) [A contribution to analytic number theory]. J. reine angew. Math. **78** (1874), 46–62. `http://resolver.sub.uni-goettingen.de/purl?GDZPPN002155656`

[35] J.-F. Mestre, *Construction d'une courbe elliptique de rang* $\geq$ 12. (French) [Construction of an elliptic curve with rank at least 12]. C. R. Acad. Sci. Paris, **295**, série I (1982), 643–644.

[36] J.-F. Mestre, *Formules explicities et minorations de conducteurs de variétés algébriques.* (French) [Explicit formulas and lower bounds on conductors of algebraic varieties]. Compositio Math. **58** (1986), 209–232. `http://eudml.org/doc/89769`

[37] J.-F. Mestre, *Courbes elliptiques de rang* $\geq$ 11 *sur Q(t)* (French). [Elliptic curves of rank at least 11 over $\mathbf{Q}(t)$]. C. R. Acad. Sci. Paris, **313**, série I (1991), 139–142. `http://gallica.bnf.fr/ark:/12148/bpt6k57325582/f143`

[38] J.-F. Mestre, *Courbes elliptiques de rang* $\geq$ 12 *sur Q(t)* (French). [Elliptic curves of rank at least 12 over $\mathbf{Q}(t)$]. C. R. Acad. Sci. Paris, **313**, série I (1991), 171–174. `http://gallica.bnf.fr/ark:/12148/bpt6k57325582/f175`

[39] P. Monsky, appendix to [25].

[40] K.-i. Nagao, *Examples of elliptic curves over Q with rank* $\geq$ 17. Proc. Japan Acad. Ser. A Math. Sci. **68**, no. 9 (1992), 287–9. `http://dx.doi.org/10.3792/pjaa.68.287`

[41] K.-i. Nagao, *Construction of high-rank elliptic curves.* Kobe J. Math. **11** (1994), 211–219.

[42] A. Néron, Footnote in *Œuvres de Henri Poincaré.* Tome V, 1950, Gabay. `http://www.univ-nancy2.fr/poincare/bhp/pdf/hp1950oea.pdf`

[43] S. Omar, *Non-vanishing of Dirichlet L-functions at the central point.* In *Algorithmic Number Theory*, ANTS-VIII (Calgary 2008), edited by A. J. van der Poorten and A. Stein. Springer LNCS **5011** (2008), 443–453. `http://dx.doi.org/10.1007/978-3-540-79456-1_30`

[44] R. Paley, N. Wiener, *Fourier transforms in the Complex Domain.* AMS Coll. Publ. **19**, 1934. Twelfth printing 2000, `http://www.ams.org/bookstore-getitem/item=COLL-19`

[45] C. S. Rajan, *On the size of the Shafarevich-Tate group of elliptic curves over function fields.* Compositio Math. **105** (1997), no. 1, 29–41. `http://dx.doi.org/10.1023/A:1000105104709`

[46] N. F. Rogers, *Rank computations for the congruent number elliptic curves.* Experiment. Math. **9**, no. 4 (2000), 591–4. http://projecteuclid.org/euclid.em/1045759524

[47] N. F. Rogers, *Elliptic curves $x^3 + y^3 = k$ with high rank.* Doctoral thesis, Harvard University (Cambridge, MA, USA), April 2004.

[48] K. Rubin, A. Silverberg, *Ranks of elliptic curves in families of quadratic twists.* Experiment. Math. **9** no. 4 (2000), 583–590. http://projecteuclid.org/euclid.em/1045759523

[49] K. Rubin, A. Silverberg, *Rank frequencies for quadratic twists of elliptic curves.* Experiment. Math. **10**, no. 4 (2001), 559–570. http://projecteuclid.org/euclid.em/1069855256

[50] K. Rubin, A. Silverberg, *Ranks of Elliptic Curves.* Bull. AMS, **39** no. 4 (2002), 455–474. http://dx.doi.org/10.1090/S0273-0979-02-00952-7

[51] M. O. Rubinstein, *Elliptic curves of high rank and the Riemann zeta function.* Experiment. Math **22**. no. 4 (2013), 465–480. http://dx.doi.org/10.1080/10586458.2013.840870

[52] E. F. Schaefer, M. Stoll, *How to do a p-descent on an elliptic curve.* Trans. Amer. Math. Soc. **356** (2004), 1209–1231. http://dx.doi.org/10.1090/S0002-9947-03-03366-X

[53] U. Schneiders, H. G. Zimmer, *The rank of elliptic curves upon quadratic extension.* In *Computational number theory* (Debrecen 1989), ed. A. Pethő, M. E. Pohst, H. C. Williams, and H. G. Zimmer, published by Walter de Gruyter & Co., Berlin (1991), 239–260.

[54] D. Shanks, *On Gauss and composition, II.* In *Number Theory and Applications* (1989), edited by R. A. Mollin. Proceedings of the NATO Advanced Study Institute (Banff 1988), NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, **265**. Kluwer Academic Publishers Group, Dordrecht, 1989, 179–204. See also [31].

[55] D. T. Tèǐt [J. T. Tate], I. R. Šafarevič, *The rank of elliptic curves.* Soviet Math. Dokl. **8** (1967), 917–920. [trans. J. W. S. Cassels from *Dokl. Akad. Nauk SSSR* **175** (1967), 770–773.]

[56] D. Ulmer, *Elliptic curves with large rank over function fields.* Ann. of Math. (2) **155** (2002), no. 1, 295–315. http://www.jstor.org/stable/3062158

[57] M. Watkins, *A note on integral points on elliptic curves.* J. Théor. Nombres Bordeaux, **18** (2006), no. 3, 707–719. http://eudml.org/doc/249658

[58] M. Watkins, *On elliptic curves and random matrix theory.* J. Théor. Nombres Bordeaux, **20** (2008), no. 3, 829–845. http://dx.doi.org/10.5802/jtnb.653

[59] A. Weil, *Sur les "formules explicites" de la théorie des nombres premiers.* (French) [On the "explicit formulæ" in the theory of prime numbers]. In *Festkrift tillaegnad Marcel Riesz*, edited by C. W. K. Gleerup. Supplementary tome of Meddelanden Frän Lunds Univ. mat. Sem. [Comm. Sém. Math. Univ. Lund], 1952, 252–265. Also in Weil's *Œuvres Scientifiques* [Collected Works], Vol. **2** (1952b), 48–61.

April 7, 2014

Mark Watkins, School of Mathematics and Statistics, Carslaw Building (F07), University of Sydney, NSW 2006, AUSTRALIA • *E-mail :* `watkins@maths.usyd.edu.au`

Stephen Donnelly, School of Mathematics and Statistics, Carslaw Building (F07), University of Sydney, NSW 2006, AUSTRALIA • *E-mail :* `donnelly@maths.usyd.edu.au`

Noam D. Elkies, Department of Mathematics, Harvard University, One Oxford Street, Cambridge MA 02138, UNITED STATES OF AMERICA • *E-mail :* `elkies@math.harvard.edu`

Tom Fisher, Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WB, UNITED KINGDOM *E-mail :* `T.A.Fisher@dpmms.cam.ac.uk`

Andrew Granville, Départment de mathématiques et de statistique, Pavilion André-Aisenstadt, 2920, chemin de la Tour, Montréal (Québec) H3T 1J4, CANADA • *E-mail :* `andrew@dms.umontreal.ca`

Nicholas F. Rogers, UR Mathematics, 915 Hylan Building, University of Rochester, RC Box 270138, Rochester, NY 14627, UNITED STATES OF AMERICA • *E-mail :* `rogers@math.rochester.edu`