NOTES ON 4-SELMER AND 8-CLASS RANKS

Version of 24 Oct 2020.

MARK WATKINS

Abstract. We continue our attempts to understand the recent work of Smith.

He has two significant ideas that are used in going beyond the 2-Selmer case. One of these is co-homological and involves writing the sum of higher Selmer pairings over a hypercube of vertices in terms of an Artin symbol.

Secondly, assuming there is suitable equi-distribution in the Artin symbols from these hypercube sums, one then wishes to show that the Selmer pairings themselves are equi-distributed, so in particular the distribution of the higher Selmer ranks therein can be determined. This step, at least when the co-homological input is suitably aligned, is mostly combinatorial in nature.

Indeed, there are notable technical difficulties with glueing together these two steps, though in the case of the 4-Selmer group (where there is still a governing field) one sees various simplifications. It is this case that we choose to highlight for now, essentially to try to digest Smith's combinatorial methodology without yet bringing in the co-homological machinery.

The analogue for class groups is the 8-rank, and we actually mostly follow the preprint of Chan, Koymans, Milovic, and Pagano, who consider the situation for Gaussian discriminants.

1. Introduction

In our previous work [21] we described some recent (2017) methods of Smith [15] to determine the 2-Selmer rank distribution of quadratic twists of a given elliptic curve $E/\mathbf{Q}$ with full 2-torsion and no 4-torsion, and similarly for the narrow 4-rank class group distribution for quadratic fields, including the special case of Gaussian discriminants (positive fundamental discriminants that have no prime divisors that are 3 mod 4).

Smith's methods are (much) more powerful than what we discussed in [21], and here we aim to move up to the next cases: namely 4-Selmer ranks of suitable elliptic curves, and 8-ranks of narrow class groups of quadratic fields. This still hides most of the co-homological machinery that he uses to go to higher 2^k-ranks, but we still have a number of novel combinatorial ideas that can be presented without getting too weighed down in Galois co-homology.

The preprint of Chan, Koymans, Milovic, and Pagano [4] discusses the case of narrow 8-ranks for Gaussian discriminants, and it is indeed this work that we shall largely follow.

Another useful preprint is one from 2016 by Smith [14], where he describes the 4-Selmer and 8-class distributions under GRH.

Finally, there is a preprint from 2018 of Koymans and Pagano [9], which generalizes Smith's 2017 work to the l^∞-part of the class group of degree l fields with cyclic Galois group. This goes in a much different direction,[1] particularly being heavy on the Galois co-homology, though some of the combinatorial ideas that we highlight here are still present in a somewhat alternative presentation to Smith's setup. They also assume GRH, as analytic questions are not their main focus.

1.1. We now describe the contents of this exposition.

We let $\mathcal{G}$ be the set of Gaussian discriminants, which are positive fundamental discriminants that have no prime factor that is 3 mod 4. We write $e_4(d)$ for the

[1] As the last paragraph of their Introduction indicates, the situation is closer to real quadratic fields, while Smith's work considers only imaginary quadratic.

narrow 4-rank of the class group of $\mathbf{Q}(\sqrt{d})$, and $\mathbf{R}^8(d)$ for a pairing matrix on the narrow 4-class group that will be defined later (§4.3). For now, it is enough to know that this matrix is defined over $\mathbf{F}_2$ and has e_4 rows and (e_4+1) columns, and the dimension of the left kernel of this matrix is the narrow 8-rank for $\mathbf{Q}(\sqrt{d})$.

1.1.1. Recalling our notation (from [21, §10.2.1]) for the proposition of Gaussian discriminants with 4-rank equal to e as

$$\gamma_{\mathrm{G}}(e)=\frac{1}{2^{e(e+1)/2}}\prod_{u=1}^{\infty}(1+1/2^u)^{-1}\prod_{j=1}^{e}(1-1/2^j)^{-1},$$

the main result we show is the following. (As per [21], our results here are effective).

Theorem 1.1.2. *For any $M\in\mathrm{Mat}(e,e+1,\mathbf{F}_2)$ we have*

$$\frac{\#\{d\le X: d\in\mathcal{G}\mid e_4(d)=e,\mathbf{R}^8(d)=M\}}{\#\{d\le X: d\in\mathcal{G}\}}=\frac{\gamma_{\mathrm{G}}(e)}{2^{e(e+1)}}+O\Big(\frac{1}{(\log\log X)^{1/4}}\Big).$$

This was shown in [4] with a weaker error term (as we discuss below).

Of course, such a statement really only makes sense if we have some sort of fixed basis for the 8-rank pairing matrices, which indeed we discuss in §5.3.1. On the other hand this somewhat loose phrasing already allows some useful corollaries. We write $e_4^{\mathrm{o}}(d)$ for the ordinary 4-rank of the class group of $\mathbf{Q}(\sqrt{d})$, and recall Fouvry and Klüners [8, Theorem 2] showed that a relative proportion $1/2^e$ of the time the narrow and ordinary 4-ranks are equal; the methods here recover this result.

Corollary 1.1.3. *We have*

$$\frac{\#\{d\le X: d\in\mathcal{G}\mid e_4(d)=e,e_4^{\mathrm{o}}(d)=e\}}{\#\{d\le X: d\in\mathcal{G}\}}=\frac{\gamma_{\mathrm{G}}(e)}{2^{e}}+O\Big(\frac{1}{(\log\log X)^{1/4}}\Big).$$

This follows upon identifying the instances of $e_4(d)=e_4^{\mathrm{o}}(d)$ with $\mathbf{R}^8(d)$ having a rightmost column of zeros (§4.6.1). One naturally has the complementary result: the proportion with the ordinary and narrow 4-ranks differing is $\gamma_{\mathrm{G}}(e)(1-1/2^e)$.

Next we can accumulate instances of M having a left kernel of dimension e_8.

Corollary 1.1.4. *We have*

$$\frac{\#\{d\le X: d\in\mathcal{G}\mid e_4(d)=e_4,e_8(d)=e_8\}}{\#\{d\le X: d\in\mathcal{G}\}}=\gamma_{\mathrm{G}}(e_4)\mathbf{P}(e_4,e_8)+O\Big(\frac{1}{(\log\log X)^{1/4}}\Big)$$

where $\mathbf{P}(e_4,e_8)$ is the proportion of matrices in $\mathrm{Mat}(e_4,e_4+1,\mathbf{F}_2)$ that have a left kernel of dimension e_8.

Moreover, we can consider the subcase where the narrow and ordinary 4-ranks are equal, with this signified by the notation $e_4^{\mathrm{n,o}}(d)=e_4$.

Corollary 1.1.5. *We have*

$$\frac{\#\{d\le X: d\in\mathcal{G}\mid e_4^{\mathrm{n,o}}(d)=e_4,e_8(d)=e_8\}}{\#\{d\le X: d\in\mathcal{G}\}}=\gamma_{\mathrm{G}}(e_4)\mathbf{P}_{\mathrm{o}}(e_4,e_8)+O\Big(\frac{1}{(\log\log X)^{1/4}}\Big)$$

where $\mathbf{P}_{\mathrm{o}}(e_4,e_8)$ is the proportion of matrices in $\mathrm{Mat}(e_4,e_4+1,\mathbf{F}_2)$ that have a rightmost column of zeros and a left kernel of dimension e_8.

Of course, $\mathbf{P}_{\mathrm{o}}(e_4,e_8)$ is also $1/2^{e_4}$ of the proportion of matrices in $\mathrm{Mat}(e_4,e_4,\mathbf{F}_2)$ that have a kernel of dimension e_8.

1.1.6. One can write down more elaborate expressions for $\mathbf{P}(e_4, e_8)$ if desired.

We recall the accounting of the number $N_q(m \times n, r)$ of matrices of size m-by-n over $\mathbf{F}_q$ of rank r (cf. [16, Proposition 2.3] for instance). We have that

$$N_q(m \times n, r) = \begin{bmatrix} n \\ r \end{bmatrix}_q \cdot N_q^0(m, r)$$

where the Gaussian coefficient $\left[\begin{smallmatrix} n \\ r \end{smallmatrix}\right]_q$ is the number of r-dimensional subspaces of a vector space of dimension n over $\mathbf{F}_q$ and $N_q^0(m, r)$ is the number of surjective maps from an m-dimensional vector space to an r-dimensional space over $\mathbf{F}_q$ (in other words, the number of m-by-r matrices of full rank). One can write the Gaussian coefficient variously as

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \prod_{l=0}^{r-1} \frac{q^{n-l} - 1}{q^{r-l} - 1} = \prod_{l=1}^{n} (q^l - 1) \Big/ \prod_{j=1}^{r} (q^j - 1) \prod_{k=1}^{n-r} (q^k - 1).$$

Meanwhile, a result anciently due to Landsberg [11] implies that

$$N_q^0(m, r) = q^{mr} \prod_{j=0}^{r-1} \bigl(1 - 1/q^{m-j}\bigr) = q^{r(r-1)/2} \prod_{j=0}^{r-1} (q^{m-j} - 1).$$

It is my preference to write things in terms of $F_q(u) = \prod_{i=1}^{u} (1 - 1/q^i)$, so that

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \frac{q^{\binom{n}{2}} F_q(n)}{q^{\binom{r}{2}} F_q(r) \cdot q^{\binom{n-r}{2}} F_q(n-r)} \quad \text{and} \quad N_q^0(m, r) = q^{mr} F_q(m)/F_q(m-r).$$

One can also note that

$$\gamma_{\mathrm{G}}(e) = \frac{1}{2^{e(e+1)/2}} \prod_{u=1}^{\infty} (1 + 1/2^u)^{-1} \prod_{j=1}^{e} (1 - 1/2^j)^{-1} = \frac{1/F_2(e)}{2^{e(e+1)/2}} \prod_{u=1}^{\infty} (1 + 1/2^u)^{-1}.$$

I'm not sure how useful it is to write down more expressions for $\mathbf{P}(e_4, e_8)$ in general. In the case of $\mathbf{P}_{\mathrm{o}}(e_4, e_8)$ we are interested in $(m, n, r, q) = (e_4, e_4, e_4 - e_8, 2)$ and have a factor of $1/2^{e_4}$ from the rightmost column so that

$$\mathbf{P}_{\mathrm{o}}(e_4, e_8) = \frac{1}{2^{e_4}} N_2(e_4 \times e_4, e_4 - e_8)/2^{e_4^2},$$

where $N_2(e_4 \times e_4, e_4 - e_8)$ is

$$\frac{2^{\binom{e_4}{2}} F_2(e_4)}{2^{\binom{e_8}{2}} F_2(e_8) \cdot 2^{\binom{e_4 - e_8}{2}} F_2(e_4 - e_8)} \cdot 2^{e_4(e_8 - e_4)} \frac{F_2(e_4)}{F_2(e_8)} = 2^{(e_4 - e_8)^2} \frac{F_2(e_4)^2}{F_2(e_8)^2 F_2(e_4 - e_8)}$$

and so

$$\mathbf{P}_{\mathrm{o}}(e_4, e_8) = \frac{2^{(e_4 - e_8)^2}}{2^{e_4} 2^{e_4^2}} \frac{F_2(e_4)^2}{F_2(e_8)^2 F_2(e_4 - e_8)}.$$

Perhaps we can highlight the $e_8 = 0$ case (of interest since it immediately implies the negative Pell equation is solvable), where

$$\gamma_{\mathrm{G}}(e_4)\mathbf{P}_{\mathrm{o}}(e_4, 0) = \frac{1/F_2(e_4)}{2^{e_4(e_4+1)/2}} \prod_{u=1}^{\infty} (1 + 1/2^u)^{-1} \frac{F_2(e_4)}{2^{e_4}} = \frac{1}{2^{e_4(e_4+3)/2}} \prod_{u=1}^{\infty} (1 + 1/2^u)^{-1},$$

recovering the phrasing of [4, Theorem 1.1].

1.1.7. As noted above, our main reference text is the preprint of Chan, Koymans, Milovic, and Pagano [4]. Most of their §5 is replaced by our previous §§5-6 in reworking Smith's estimates for equi-distribution of $(p_i|p_j)$. It is their §2 and §6 that we emulate most; here §2 is rather specific to the case of 8-ranks (and indeed Gaussian discriminants at some points), and they/we follow Stevenhagen's rendition [18] of Rédei symbols therein. Meanwhile, their §6 is largely a suitable versioning of Smith's combinatorial methods, with some distinctions in the notion of genericity that are necessitated due to the sparsity of the set of Gaussian discriminants.

The main technical advance we achieve is the removal[2] of "extravagant spacing" as introduced by Smith. Note that extravagant spacing (see [15, Theorem 5.4 (3)] or [4, Theorem 4.1(iii)]) actually has quite a poor relative error term (saving only a power of $\log\log\log X$) and it is principally for this reason that we improve upon the error in [4].

1.2. We give a rather brief outline of the proof technique (which indeed originates from Smith), and then list the contents herein.

1.2.1. Recall that we represent fundamental discriminants d from boxes of Cartesian products $\prod_j T_j$ where the T_j are the primes from dyadic-like intervals (all being 1 mod 4 in the Gaussian discriminant case), and then subject said primes to various mutual Legendre symbol conditions to determine the narrow 4-rank. This then gives a basis of the narrow 4-class group, and we can write a pairing matrix on it that will in turn give (as its kernel) the narrow 8-class group.

What Smith notes is that these pairing matrices, or more particularly the values of characters for their ambient matrix space, can be related amongst various d; he has sums of these over vertices of a hypercube (in our case in no more than 3 dimensions), with the result then being related to a Frobenius element in a certain field. (One can replicate these sums over 2^3 vertices many times over via different choices of which 2 primes to use from each T_j, so that equi-distribution of the summatory Frobenius elements will give equi-distribution of the character values).

He then takes a selection of the indices of the Cartesian product $\prod_j T_j$, designed so that one of these indices is large (greater than 5/8 of the total number of indices) and the others are small (between 1/10 and 1/8 of the total), with the 0-1 pattern of the 4-class basis vectors at these indices specified so as to induce a desirable result (in particular, to show cancellation) with the above hypercube sums. Since we typically have many indices (on the scale of $\log\log X$, for the number of prime divisors of d) but only a few basis vectors (more than $99\sqrt{\log\log\log X}$ will be non-generic) finding such indices (which we call hypercube co-ordinates) is feasible.

Smith then fixes a prime from every T_j not in the hypercube co-ordinates (so that only three of the components are varying for us), and considers a sum over the upper hypercube co-ordinate. In other words, we are (roughly) working on[3]

$$t^{\downarrow} \times \prod_{u\in\tilde{\mathcal{V}}_\psi} T_u \times t^{\uparrow} \times T_{\mathrm{s}_\psi}$$

[2] I am told that Smith also did this, and indeed for our case it is mostly a matter of arranging averaging techniques properly. One might also try to apply a large sieve for Artin representations, as recently developed by Thorner and Zaman [20], which is a somewhat more streamlined and direct version of the averaged Chebotarev results of Pierce, Turnage-Butterbaugh, and Wood [12].

[3] Here $t^{\downarrow}$ is split from $t^{\uparrow}$ somewhat arbitrarily at half the number of prime divisors of d; the grids we describe below will depend on $t^{\downarrow}$ but not $t^{\uparrow}$.

where $t^{\downarrow}$ and $t^{\uparrow}$ are fixed tuples of primes with one from each T_j for j that is not a hypercube co-ordinate, while $\prod_u T_u$ is the product over the lower hypercube co-ordinates, and T_{s_ψ} is the set of primes correponding to the upper hypercube co-ordinate. (Here ψ is a multiplicative character on the matrix space). By the Chebotarev theorem, the primes in T_{s_ψ} should be Frobenius equi-distributed in a field defined by the primes from the lower hypercube co-ordinates. However, there are too many members in the sets T_u corresponding to the latter (leading to a large field degree in the Chebotarev usage), and Smith has to show that we can suitably cover these $\prod_u T_u$ from the lower hypercube co-ordinates by (much) smaller grids in a nearly uniformly manner. This part of the argument is purely combinatorial in our case. Once the field parameters are adequately bounded, the Chebotarev usage then gives us the desired Frobenius equi-distribution when summing over the upper hypercube co-ordinate, which in turn gives us cancellation of the pairing matrix character sums, and thus equi-distribution of the pairing matrices themselves.

Of course, in practice much of this is true only "on average" over the fixed selections (here $t^{\downarrow}$ and $t^{\uparrow}$) away from the hypercube co-ordinates, and indeed Smith simplifies his argument somewhat by using "extravagant spacing" to skip such averaging in some cases.

For the higher 2^k-ranks, Smith has $(k-1)$ lower hypercube co-ordinates and there is much more delicacy in arranging the Galois co-homology arguments to ensure that cancellation can be detected for the higher 2^k-pairings, and moreover to show that there are applicable choices of grids for $\prod_u T_u$. We do not delve into this here, other than to make some comments at the very end in §13.6.3.

1.2.2. Let us list the contents of each section below.

In §2 we catalogue some notation we use, though the reader should also see our setup review in §5, and §6.6 for notation with box-restrictions.

We then give some brief general background facts in §3. In particular, we recall the bilinear estimate for $\sum_i \sum_j \alpha_i \beta_j (p_i|p_j)$ and list the version of the Chebotarev density theorem that we will use. We also give a short combinatorial Lemma (arising from Smith in this context) that allows to pass from known cancellation for a "derivative" function (from $H \times H$ to $\mathbf{F}_2$) to then show cancellation for the anti-derivatives (from H to $\mathbf{F}_2$).

In §4 we give background on the 8-ranks of narrow quadratic class groups, largely following Stevenhagen's description from Rédei symbols.

In §5 we review our previous setup of splitting up discriminants d into boxes, and then in §6 we introduce hypercube co-ordinates, and show that suitable ones exist in generic circumstances.

We then apply the results about Rédei symbols in §7 to determine hypercube sums in terms of Frobenius elements, and then in §8 we apply the Chebotarev density theorem to show said Frobenius elements are equi-distributed.

This equi-distribution is dependent upon a suitable selection of "grids" that form a nearly uniform cover of the sets from the lower hypercube co-ordinates, and in §9 we give Smith's combinatorial arguments that allow us to do this. Then in §10 we show that indeed this grid-based cover of the lower hypercube co-ordinates induces a suitable uniform cover of the entire Cartesian product (with respect to the Legendre symbol conditions, which we largely elided in our description above). In §11 we put everything together and show the main results for Gaussian discriminants.

In §12 we discuss the modifications necessary to handle the general quadratic case, and in §13 consider the 4-Selmer case for quadratic twists of elliptic curves.

2. Notation

We have $\mathbf{Q}(\sqrt{D})$ as a quadratic field with D a fundamental discriminant, and $\mathcal{C}_D$ its narrow class group. The Rédei matrix $\mathbf{R}^4$ determines the 4-rank (being one less than the dimension of its kernel), and similarly with $\mathbf{R}^8$ for the 8-rank. We write e for the 4-rank, or sometimes e_4 when it is necessary to clarify with the 8-rank e_8. We write $[A,B,C]$ for the Rédei symbol.

The number of prime divisors of D is r. We write $\tilde{d}$ for the odd part of $|D|$, and $\tilde{r}$ for the number of its prime divisors. We then have $(\mathcal{K},\mathcal{L})$ as a residue and Legendre symbol specification for $\tilde{d}$.

We write $E_2^L(u,X)=\exp\exp(u\log\log X)$ and note $\log E_2^L(u,X)=(\log X)^u$.

In §6 we have the multiplicative character ψ on the matrix space $\mathrm{Mat}(e,e{+}1,\mathbf{F}_2)$, and its exponent array c in terms of basic characters. We then have h_ψ^{z} and h_ψ^{s} as distinguished basis indices. We have $\mathcal{V}_\psi$ as a set of hypercube co-ordinates, with s_ψ and z_ψ and possibly z'_ψ being its elements, and $\tilde{\mathcal{V}}_\psi$ referring to $\mathcal{V}_\psi$ with s_ψ removed.

We let $r_{\mathrm{g}}=\lfloor(\alpha_{\mathcal{P}}/2)\log\log X\rfloor$ be a demarcation point for the indices of the prime divisors of $\tilde{d}$, and call the hypercube co-ordinates well-gapped if $\mathrm{s}_\psi\ge(5/4)r_{\mathrm{g}}$ and $r_{\mathrm{g}}/5\le i\le r_{\mathrm{g}}/4$ for $i\in\tilde{\mathcal{V}}_\psi$. We then let $\mathcal{I}_\downarrow$ be the set of indices $\le r_{\mathrm{g}}$ and $\mathcal{I}_\uparrow$ those exceeding this, and $\mathcal{I}_\downarrow^\star$ and $\mathcal{I}_\uparrow^\star$ the sets of these upon removing the hypercube co-ordinates in $\tilde{\mathcal{V}}_\psi$. We will then have $r_{\mathrm{g}}^\uparrow$ as the size of $\mathcal{I}_\uparrow^\star$ and $r_{\mathrm{g}}^\Uparrow$ the size of $\mathcal{I}_\uparrow$, with $r_{\mathrm{g}}^\downarrow$ the size of $\mathcal{I}_\downarrow^\star$.

We then have various notation for box products with these demarcations. Recall that a box $\bar{T}$ represents squarefree integers via $T=\prod_l T_l$. We have $T^\downarrow(\mathcal{K})$ and $T^\uparrow(\mathcal{K})$, the Cartesian products of $T_l(\mathcal{K})$ for l respectively in $\mathcal{I}_\downarrow^\star$ and $\mathcal{I}_\uparrow^\star$, with then $t^\downarrow$ and $t^\uparrow$ elements of these, perhaps better thought of as a product of singleton sets in some instances. We will have $T_{\mathrm{s}_\psi}(\mathcal{K})$ corresponding to the upper hypercube co-ordinate, and will reserve the subscript u herein as $T_u(\mathcal{K})$ for $u\in\tilde{\mathcal{V}}_\psi$.

We then have grids $\check{Z}\subset\prod_{u\in\tilde{\mathcal{V}}_\psi}T_u(\mathcal{K})$, which have components $\check{Z}^u\in T_u(\mathcal{K})$ of a fixed size $\#\check{Z}^u=B$. Here $B=\lfloor\sqrt{\log\log X}/999\rfloor$, and we write $n_\psi=\#\tilde{\mathcal{V}}_\psi$ for the number of lower hypercube co-ordinates. These grids will approximately cover the $\mathcal{L}$-compatible subset of $\prod_u T_u(\mathcal{K})$ to multiplicity roughly $\bar{R}$, where here we have that $\bar{R}=\lfloor E_2^L(0.05,X)\rfloor=\lfloor\exp\exp(0.05\log\log X)\rfloor$.

We then have various restrictions of box components, such as $T_j(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]$, which is the subset of $T_j(\mathcal{K})$ whose members meet the Legendre conditions from $\mathcal{L}$ for $t^\downarrow$ and $\check{Z}$. There is also a version $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ of this, where here the the double-brackets refer to members of $T^\downarrow(\mathcal{K})$ themselves being $\mathcal{L}$-compatible. Finally we have $T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle$, where this is the subset of $T(\mathcal{K},\mathcal{L})$ with the components for the indices in $\mathcal{I}_\downarrow^\star$ being $t^\downarrow$.

In §7 we have fields $L_{\check{Z}}$ and $K_{\check{Z}}$ that depend on a grid $\check{Z}$, and we write $\psi^\star$ for the additive version of the character ψ. We first demonstrate cancellation over a product $\mathcal{A}(t^\downarrow,\check{Z},t^\uparrow)=t^\downarrow\times\check{Z}\times t^\uparrow\times T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]$, and then (in §8) progress to

$$\mathcal{B}^\star(t^\downarrow\times\check{Z})=\bigl(t^\downarrow\times\check{Z}\times\prod_{j\in\mathcal{I}_\uparrow}T_j(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]\bigr)\cap T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle.$$

The final intersection ensures that everything is $\mathcal{L}$-compatible, and we also have

$$\mathcal{B}(t^{\downarrow} \times \check{Z}) = t^{\downarrow} \times \check{Z} \times \prod_{j \in \mathcal{I}_{\uparrow}} T_j(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}].$$

We then collate grids $\check{Z}$ into $\mathcal{Z}(t^{\downarrow})$ in §9, and write $Y_{t^{\downarrow}}$ for the $\mathcal{L}$-compatible subset of $\prod_u T_u(\mathcal{K})$, with this being covered to multiplicity $\bar{R}$ by the grids except for an exceptional set $\tilde{Y}_{t^{\downarrow}}(\mathcal{Z})$. We write $R_{\mathcal{Z}}(y)$, or $R_{t^{\downarrow}}(y)$ once $\mathcal{Z}(t^{\downarrow})$ is fixed, for the number of grids in which $y \in Y_{t^{\downarrow}}$ appears. We then (§10) have the counting function $\Lambda_{t^{\downarrow}}(\vec{t})$ for $\vec{t} \in T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow}\rangle$, which counts how many grids have $\vec{t}$ as an element in $\mathcal{B}(t^{\downarrow} \times \check{Z})$.

We will sometimes write $\tilde{c}$ for an unspecified constant, for instance in error term of the prime number theorem.

3. Various background material

3.1. We recall the bilinear estimate (originating from Heilbronn) for sums over primes linked by a Legendre symbol.

Lemma 3.1.1. *Let $\{\alpha_p\}_p$ be complex numbers bounded by 1 and supported on odd primes p with $P \leq p \leq 2P$, and similarly for $\{\beta_q\}_q$. Then*

$$\sum_{p \sim P} \sum_{q \sim Q} \alpha_p \beta_q (p|q) \ll \frac{PQ}{\min(P,Q)^{1/9}}. \tag{1}$$

This is Proposition 13.3.5 in [21] (where a substantial history is also given).

3.1.2. We will also need to use the Chebotarev density theorem, which contains primes in arithmetic progressions as a special case.

Suppose that $M/\mathbf{Q}$ is a Galois extension of degree m, and ρ a nontrivial irreducible Artin representation for M of degree v and conductor f whose L-function is entire. Then

$$\sum_{p \leq U} \chi_\rho(p) \log p \ll U^\beta + U \exp\Bigl(-\frac{(\log U)/(mv)^4}{\sqrt{\log U} + 3\log f}\Bigr)(mr \log mfU)$$

where χ_ρ is the character of ρ, and β is a prospective exceptional zero (by our usage of pleasant boxes this will never be a hindrance). By summing this over all irreducible representations we get an equi-distribution result for Frobenius classes.

We will use this for fields of degree $\leq 2 \cdot 4^{(B-1)^{n_\psi}}$ where B will be the size of our grid components and n_ψ the number of lower hypercube co-ordinates. In our case the latter is ≤ 2, and we take $B = \lfloor \sqrt{\log\log X}/999 \rfloor$ to ensure the degree is suitably small as $m \ll (\log X)^{1/10^5}$. The conductor will be $\leq (8K)^v$ where K is the product of $n_\psi B$ primes each (say) of size $\ll \exp\exp(0.51 \log\log X)$, while the primes in question will be $\gg \exp\exp(0.62 \log\log X)$. Thus $(\log U)/(mv)^4 \gg (\log X)^{0.61}$ while $\log f \ll (\log X)^{0.52}$, so we obtain a suitable savings.

3.2. We next provide a way of showing that if a differential map has cancellation in its value-distribution, then any "anti-derivative" of it must have cancellation too.

Let H be a Cartesian product $\prod_j H_j$ of h finite nonempty sets, and write $\tilde{\mathbf{F}}_2^H$ for the set of functions from H to $\mathbf{F}_2$, and $\tilde{\mathbf{F}}_2^{H \times H}$ for the set of functions from $H \times H$ to $\mathbf{F}_2$. From this we then define the linear map $\partial_H : \tilde{\mathbf{F}}_2^H \to \tilde{\mathbf{F}}_2^{H \times H}$ from the rule given by $\bigl(\partial_H F\bigr)(\vec{u}^1, \vec{u}^2) = \sum_{\vec{c} \in \mathbf{F}_2^H} F\bigl((u_1^{c_1}, ..., u_j^{c_j})\bigr)$ for a given map $F \in \tilde{\mathbf{F}}_2^H$.

We can note the kernel of ∂_H consists of factorable functions that are constant on a factorable component, that is, we can write $F(\vec{x}) = F_i(x_i) + \check{F}_i(\check{x}_i)$ where F_i is a constant function (either all zeros or all ones) on H_i, and $\check{F}_i$ is some function on the product of the H_j's with H_i removed.

In particular, the image of ∂_H has dimension $\prod_j(\#H_j - 1)$.

3.2.1. We will ultimately be able to show cancellation for functions $g \in \mathrm{im}(\partial_H)$ over such $\mathbf{F}_2$-hypercubes, and it turns out we may choose such g rather arbitrarily; naturally, we want to transfer this cancellation (or equi-distribution) back to functions on H itself. In particular, we want to take g so that any F with $g = \partial_H F$ has a reasonable amount of cancellation.[4]

Taking the contrapositive of the terminology given in [4, §3], we define an ϵ-good function F on H to be one that has $|\#F^{-1}(0) - \#H/2| \le \epsilon\#H$ (where $F^{-1}(0)$ is the inverse image of 0), and $g \in \mathrm{im}(\partial_H)$ to be ϵ-good if every F with $\partial_H F = g$ is ϵ-good. With the above kernel dimension computation, it is an exercise in tails of the binomial distribution[5] to show that when the sets H_j are large compared to their number h, there is a suitable g that is ϵ-good. Namely, we have the following.

Lemma 3.2.2. [4, Theorem 3.3] *(see also* [15, Proposition 4.3]*). We have*

$$\frac{\#\{g \in \mathrm{im}(\partial_H) : g \textit{ is } \epsilon\textit{-bad}\}}{\#\{g \in \mathrm{im}(\partial_H)\}} \le 2 \cdot 2^{\#H - \prod_j(\#H_j - 1)} \exp(-2\epsilon^2\#H).$$

In practice, we will apply this with $h = 2$ or $h = 3$, with each H_j of size B. We can then take $\epsilon \sim \sqrt{h \log \sqrt{2}}/\sqrt{B}$ and ensure the existence of a map $g \in \mathrm{im}(\partial_H)$ that is ϵ-good. This essentially means that the error in our cancellation detection will be of relative size $1/\sqrt{B}$, and thus we want to take B as large as possible. Of course, in the previous subsection we saw a barrier therein, namely in terms of the degree of various fields when using the Chebotarev density theorem. This then indicates that our ultimate relative error estimate will be no better than roughly $\ll 1/(\log\log X)^{1/2n_\psi} \le 1/(\log\log X)^{1/4}$, somewhat worse than the result for the 4-rank. (Moreover, for higher 2^k-ranks we would need $n_\psi = k - 1$).

4. Background on 8-ranks of narrow quadratic class groups

We give some background on narrow quadratic class groups, and in particular how to compute their 2-rank, 4-rank, and 8-rank. The latter we will then relate to Rédei symbols. Largely we follow Stevenhagen's presentation [18], with some indications from [4, §2] when we say more about the case of Gaussian discriminants.

4.1. Let $K = \mathbf{Q}(\sqrt{D})$ be a quadratic field, where $D = p_1 \cdots p_r$ is a fundamental discriminant with r prime factors. We let $D = \prod_i \hat{p}_i$ be its factorization into discriminantal divisors, adjusting at 2 as necessary (so $\hat{2} \in \{-4, -8, 8\}$).

The 2-part of the narrow class group $\mathcal{C}_D$ has rank $(r - 1)$, and is generated by ambiguous ideal classes $[\mathfrak{p}_i]$ where each $\mathfrak{p}_i$ is a ramified prime dividing p_i; these

[4] One might note that this is easy when $h = 1$ and $\#H_1$ is even; we take F to be any function that is split equally between the values 0 and 1, and as the kernel is generated by the constant function A equal to 1, both F and $F + A$ have an equal split of function values.

[5] By this phrase we mean the following bound: for bit-strings of length l, the number that have more than $(l/2 + r)$ or less than $(l/2 - r)$ bits equal to 0 is $\ll \sum_{j=0}^{l/2-r} \binom{l}{j} \ll 2^l 2^{-r^2/l}$.

generators are subject to a unique relation.[6] The dual $\hat{\mathcal{C}}_D[2]$ is generated by the quadratic characters $\chi_{\hat{p}_i}$ corresponding to $\mathbf{Q}(\sqrt{\hat{p}_i})$, and here the relation identifies $\chi_{\hat{p}_i}$ with $\chi_{D/\hat{p}_i}$, essentially that $K(\sqrt{\hat{p}_i}) = K(\sqrt{D/\hat{p}_i})$.

We write H_K for the narrow Hilbert class field of K, and can note that

$$\mathrm{Gal}(H_K/\mathbf{Q}) \cong \mathrm{Gal}(H_K/K) \rtimes \mathrm{Gal}(K/\mathbf{Q}) \cong \mathcal{C}_D \rtimes \langle\tau\rangle$$

where $\mathrm{Gal}(K/\mathbf{Q}) = \langle\tau\rangle$ and τ acts by inversion on $\mathcal{C}_D$. Letting H_K^2 be the genus subfield, this then is the maximal subfield of H_K that is abelian over $\mathbf{Q}$, so we have $\mathrm{Gal}(H_K^2/\mathbf{Q}) = \mathrm{Gal}(H_K/\mathbf{Q})^{\mathrm{ab}} \cong \mathcal{C}_D/2\mathcal{C}_D \times \langle\tau\rangle$, with this isomorphic to $\mathbf{F}_2^r$.

4.2. There is a natural map from $\mathcal{C}_D[2]$ to $\mathcal{C}_D/2\mathcal{C}_D$, and the dimension of the kernel of this map is the 4-rank of the narrow class group. Indeed, this kernel is $\mathcal{C}_{\mathcal{D}}[2] \cap 2\mathcal{C}_D$, and we have a sequence of maps

$$\mathbf{F}_2^r \to \mathcal{C}_D[2] \to \mathcal{C}_D/2\mathcal{C}_D \to \mathrm{Gal}(H_K^2/\mathbf{Q}) \cong \mathbf{F}_2^r.$$

We write $\chi_{\hat{p}_i}^\star$ for the additive version of $\chi_{\hat{p}_i}$, taking values in $\mathbf{F}_2$.

The Rédei matrix $\mathbf{R}^4$ is then a representation of the above composition of maps, with the entries $\mathbf{R}_{ij}^4$ given by $\chi_{\hat{p}_i}^\star(p_j)$ for $i \neq j$, and then $\mathbf{R}_{jj}^4 = \sum_{i\neq j} \mathbf{R}_{ij}^4$ to ensure the column[7] sums are zero. In any case, we have that the 4-rank of the narrow class group of $\mathbf{Q}(\sqrt{D})$ is one less than the dimension of the kernel of $\mathbf{R}^4$.

Due to the column-sums being zero, the left kernel of $\mathbf{R}^4$ has an obvious vector $(1,\ldots,1)$, corresponding to the character relation that χ_D is trivial in the dual class group.

4.3. We can then continue and consider the 8-rank. There is a natural map from $\mathcal{C}_D[2] \cap 2\mathcal{C}_D$ to $2\mathcal{C}_D/4\mathcal{C}_D$. If we restrict the map from $\mathbf{F}_2^r$ to $\mathcal{C}_D[2]$ to the right kernel of $\mathbf{R}^4$ and write H_K^4 for the narrow 4-Hilbert class field we have a chain of maps

$$\ker(\mathbf{R}^4) \to \mathcal{C}_D[2] \cap 2\mathcal{C}_D \to 2\mathcal{C}_D/4\mathcal{C}_D \cong \mathrm{Gal}(H_K^4/H_K^2) \cong \mathbf{F}_2^{e_4}$$

where e_4 is the narrow 4-rank. The 8-rank of the narrow class group of $\mathbf{Q}(\sqrt{D})$ will then be one less than the dimension of the kernel of this map.

We wish to write down explicit generators for for $\hat{\mathcal{C}}_D[4]/\hat{\mathcal{C}}_D[2]$. For this, it is profitable to first describe which characters in $\hat{\mathcal{C}}_D[2]$ are 2-divisible. This was done by Rédei and Reichardt [13] in terms of cyclic quartic extensions – although this might seem unnecessarily complicated compared to the other criteria we give, it is this that allows us to write a natural pairing for the entries of the $\mathbf{R}^8$-matrix.[8]

[6] This relation is that the ideal class $[(\sqrt{D})]$ is trivial when $D < 0$, and similarly for $D > 0$ when the fundamental unit has norm -1. Otherwise it is more complicated.

[7] Conventions differ as to whether the column or row sums are taken to be zero, and indeed some authors confuse themselves on this point.

Also, in [16, Proposition 2.2] Stevenhagen defines the matrix of having entries $(\hat{p}_i|\hat{p}_j)^\star$ (and sloppily has "d_j is even" instead of d_j having odd 2-valuation for the relevance of $(x|d_j)$), but I don't think this is as useful.

[8] Our pairing matrix on characters and ideals will naturally be in $\mathrm{Mat}(e_4, e_4+1, \mathbf{F}_2)$, and thus it seems natural to write the $\mathbf{R}^8$-matrix on the left, so that as a map from an $\mathbf{F}_2$-space of dimension (e_4+1) to a space of dimension e_4 it will indeed have e_4 rows and (e_4+1) columns.

4.3.1. Indeed, as Stevenhagen notes [18, Lemma 4.2] some equivalent conditions for $\chi_a \in \hat{\mathcal{C}}_D[2]$ to be 2-divisible are:

- an unramified cyclic quartic extension M/K containing $\mathbf{Q}(\sqrt{a}, \sqrt{D/a})$ exists;
- all ramified primes of K split completely in $\mathbf{Q}(\sqrt{a}, \sqrt{D/a})$;
- for primes $p|a$ we have $\bigl((D/a)|p\bigr) = +1$, and for $p|(D/a)$ we have $(a|p) = +1$.

The final condition can perhaps be rewritten more succinctly in terms of Hilbert symbols. We write $(a,b)_{\mathbf{Q}}^{\Pi}$ for the product of places v of $\mathbf{Q}$ such that $(a,b)_v = -1$, and thus $(a,b)_{\mathbf{Q}}^{\Pi} = 1$ is equivalent to saying that $(a,b)_v = +1$ for all v. The last condition is then equivalent to $(a,-D)_{\mathbf{Q}}^{\Pi} = 1$ for discriminantal divisors a of D.

It is also useful to write down generators for $2\mathcal{C}_D[4]$. Given $b|D$ with $b > 0$ we write $\mathcal{J}_D(b)$ for the unique integral ideal of norm b; and for $b|D$ with $b < 0$ we let it be $\mathcal{J}_D(|b|)$ multiplied by the ideal $(\sqrt{D})$. One can then show that $\mathcal{J}_D(b) \in 2\mathcal{C}_D$ exactly when $(b,D)_{\mathbf{Q}}^{\Pi} = 1$.

4.3.2. Let us give a couple of examples.

Take $D = 3 \cdot 7 \cdot 19 \cdot 103$ so $(309|7) = (309|19) = (133|3) = (133|103) = +1$, and thus χ_{133} is 2-divisible in $\mathcal{C}_D$, and this is equivalent to $(133,-D)_{\mathbf{Q}}^{\Pi} = 1$. Meanwhile we have $(7,D)_{\mathbf{Q}}^{\Pi} = 1$, and indeed $[\mathfrak{p}_7]$ is twice the class of an ideal of norm 17. (Here of course we also have $(-D/7,D)_{\mathbf{Q}}^{\Pi} = (7,D)_{\mathbf{Q}}^{\Pi}(-D,D)_{\mathbf{Q}}^{\Pi} = (7,D)_{\mathbf{Q}}^{\Pi} = 1$, leading to $\mathfrak{p}_3\mathfrak{p}_{19}\mathfrak{p}_{103} \cdot \mathfrak{p}_3\mathfrak{p}_7\mathfrak{p}_{19}\mathfrak{p}_{103}$, whose class is again the same as $[\mathfrak{p}_7]$).

As another example, when $D = -8 \cdot 7 \cdot 11 \cdot 109$ we have $(22,D)_{\mathbf{Q}}^{\Pi} = 1$ and thus $[\mathfrak{p}_2\mathfrak{p}_{11}]$ is 2-divisible (being twice the class of the fifth power of an ideal of norm 13), while $(2 \cdot 11 \cdot 109, -D)_{\mathbf{Q}}^{\Pi} = (-7,-D)_{\mathbf{Q}}^{\Pi} = +1$ so that χ_{-7} is 2-divisible (and indeed $\bigl((D/-7)|7\bigr) = (-7|2) = (-7|11) = (-7|109) = +1$).

On the other hand, in the special case for the Gaussian discriminants we have that $(a,D)_v = (a,-D)_v$ for all v and positive $a|D$, and thus the same divisors of D generate both $2\mathcal{C}_D[4]$ and $2\hat{\mathcal{C}}_D[4]$, leading to a somewhat simpler situation.

4.4. We have a basis of $\mathbf{F}_2$-vectors for the right kernel of $\mathbf{R}^4$ of dimension (e_4+1). We notate these as given by the ideal classes $[\mathfrak{m}_j]$, which satisfy a unique relation (corresponding to the class relation amongst ramified primes); moreover, they all have trivial Artin symbol on the genus field H_K^2.

We can then take quartic characters θ_i for $1 \le i \le e_4$ that span $\hat{\mathcal{C}}_D[4]/\hat{\mathcal{C}}_D[2]$. These quartic characters are only defined up to a quadratic character, as indeed the unramified extension M/K in the first equivalent condition is not unique (what is unique is the quadratic extension of H_K^2 that it generates). In any event, the quadratic characters $2\theta_i \in \hat{\mathcal{C}}_D[2]$ correspond to vectors $\vec{v}_i \in \mathbf{F}_2^r$ in the left kernel of $\mathbf{R}^4$, and indeed together with the obvious vector of all 1's these form a basis for the kernel. The quadratic extensions M_iH_K/H_K^2 then span H_K^4, and we find that $\mathbf{R}^8$ is represented by the matrix in $\mathrm{Mat}(e_4, e_4{+}1, \mathbf{F}_2)$ whose entries are $\theta_i^\star([\mathfrak{m}_j])$. The rank e_8 of the narrow class group of $\mathbf{Q}(\sqrt{D})$ is then equal to the dimension of the left kernel of $\mathbf{R}^8$, and is one less than the dimension of the right kernel.

4.4.1. We can then re-interpret these character values in terms of the Artin symbols of $\mathfrak{m}_j$ for the extension $M_iH_K^2/K$, and since the $\mathfrak{m}_j$ have trivial Artin symbol on the genus field, we find that these take values in $\mathrm{Gal}(M_iH_K^2/H_K^2)$, which we can identify with $\mathbf{F}_2$.

We can also refer to these values in pairing notation as $\langle \chi_a, b\rangle_D$, where $\chi_a \in 2\hat{\mathcal{C}}_D$ and $\mathcal{J}_D(b) \in 2\mathcal{C}_D$. This also was classically the way that Rédei symbols $[a, D/a, b]$ were defined, though we can approach the problem via somewhat different means, and ultimately show the desired equivalence.

4.5. The most natural setting for Rédei symbol inputs[9] will be to take them as elements of $\mathbf{Q}^\star/(\mathbf{Q}^\star)^2$. We let A, B be two such elements such that $(A, B)^\Pi_{\mathbf{Q}} = 1$. It will be convenient to exclude the cases where either A or B is trivial.

4.5.1. The Hilbert symbol relation $(A, B)^\Pi_{\mathbf{Q}} = 1$ is equivalent ([18, Lemma 5.1]) to there being a Galois extension $L/\mathbf{Q}$ that contains $\mathbf{Q}(\sqrt{A}, \sqrt{B})$ with $L/\mathbf{Q}(\sqrt{AB})$ cyclic of order 4. When $A = B$ this $L/\mathbf{Q}$ is itself just a cyclic quartic extension, and otherwise it is an octic dihedral extension.

As discussed by Stevenhagen [18, §7], one wants to select a minimally ramified L as above, and this is unique up to twisting by quadratic characters. We let $\mathcal{F}^{\mathrm{mr}}_{A,B}$ be the set of such fields.[10] With suitable care at 2 and ∞, in particular writing Δ_A for the associated fundamental discriminant associated to A, etc., this then leads us to the definition of a Rédei symbol ([18, §7.8]).

4.5.2. Suppose $A, B, C \in \mathbf{Q}^\star/(\mathbf{Q}^\star)^2$ are nontrivial and $\gcd(\Delta_A, \Delta_B, \Delta_C) = 1$ with $(A, B)^\Pi_{\mathbf{Q}} = (A, C)^\Pi_{\mathbf{Q}} = (B, C)^\Pi_{\mathbf{Q}} = 1$. Take $L \in \mathcal{F}^{\mathrm{mr}}_{A,B}$ and let $\mathfrak{c}$ be an integral ideal of norm $|C|$ for the integer ring of $\mathbf{Q}(\sqrt{AB})$. When $C > 0$ the Rédei symbol $[A, B, C]$ is then defined as the Artin symbol of $\mathfrak{c}$ in $L/\mathbf{Q}(\sqrt{AB})$, and indeed this is in $\mathrm{Gal}\big(L/\mathbf{Q}(\sqrt{A}, \sqrt{B})\big) = \mathbf{F}_2$. When $C < 0$ the Rédei symbol is analogously the Artin symbol of $\mathfrak{c}\infty$. Finally, if any of A, B, or C is trivial we have $[A, B, C] = 0$.

In terms of the above Artin pairing $\langle \chi_a, b\rangle_D$ we indeed have $\langle \chi_a, b\rangle_D = [a, D/a, b]$. Here we can also note that $\gcd(\Delta_a, \Delta_{D/a}) = 1$ for discriminantal divisors a of D.

4.5.3. The Rédei symbol is palpably linear in third input and symmetric in the first two. A reciprocity law [18, §8] (again somewhat tricky at 2 and ∞) implies it is also symmetric in the third, thus trilinear. That is, $[A, B, C] = [B, A, C] = [A, C, B]$ and $[A, B, C] + [A', B, C] = [AA', B, C]$.

We also have the relation $[A, B, -AB] = 0$ and thus $[A, B, C] = [A, B, -ABC]$, and as Stevenhagen notes (Proposition 7.10), this follows essentially from defining the symbol correctly at ∞.

Chan, Koymans, Milovic and Pagano [4] proceed to give various results about sums of Rédei symbols in their Theorems 2.9 to 2.12, applying merely linearity and reciprocity for 2.9 and 2.11, with $[A, B, C] = [A, B, -ABC]$ additionally employed for 2.10 and 2.12. Therein they note that there is no analogue of these latter two in Smith's work.

We simply derive such results as needed in §7 below.

4.5.4. The concept of Rédei reciprocity has been generalized by Koymans and Pagano [10, Theorem 3.3]. In particular, they are able to use their results to detect when the ordinary and narrow 8-ranks are equal.

[9] We will use this term to try to avoid referring to everything as variables.

[10] It appears we can always take L to be unramified in our situations: we have A as a product of two primes that can be required to be in the same residue class mod 8; thus A is 1 mod 8, so the fact that we have $B = -1$ in some cases is irrelevant vis-à-vis Stevenhagen's Definition 7.6.

4.6. Finally we make some notes applicable to the case of Gaussian discriminants that we consider in detail.

Here the Rédei matrix $\mathbf{R}^4$ is symmetric, and thus we can take the "same" kernel basis for both the characters and ideals, that is, we have $\chi_a \in 2\hat{\mathcal{C}}_D$ precisely when $\mathcal{J}_D(a) \in 2\mathcal{C}_D$ (note all discriminantal a dividing D are positive). This can also be seen from $(a, D)^{\Pi}_{\mathbf{Q}}$ being equivalent to $(a, -D)^{\Pi}_{\mathbf{Q}}$ in this case.

4.6.1. Moreover, the obvious vector $(1, \ldots, 1)$ is also in the right kernel of $\mathbf{R}^4$, and we take it as the basis vector with $[\mathfrak{m}_{e+1}]$, thus corresponding to the class $[(\sqrt{D})]$. Recall that the ordinary class group and narrow class group are the same precisely when this $[(\sqrt{D})]$ is trivial. In particular, if the rightmost column of $\mathbf{R}^8$ has a nonzero entry, then it is necessarily nontrivial. Conversely, one can show that if the rightmost column of $\mathbf{R}^8$ is zero, then the ordinary and narrow 4-class groups are the same. The equi-distribution of $\mathbf{R}^8 \in \mathrm{Mat}(e, e+1, \mathbf{F}_2)$ will thus imply that the proportion of Gaussian discriminants with narrow 4-rank e that have ordinary 4-rank e is $1/2^e$, thereby recovering the result of Fouvry and Klüners [8, Theorem 2].

5. Review of our setup

We recall the setup we introduced in our previous work [21, §4].

5.1. We have a set of primes $\mathcal{P}$ that consists of all the primes in various residue classes $\mathcal{R}_{\mathcal{P}}$ to the fixed modulus $M_{\mathcal{P}}$, which we assume is divisible by 4. The most pertinent constant associated to $\mathcal{P}$ is the number $\xi_{\mathcal{P}}$ of coprime residue classes that it contains, and we write $\alpha_{\mathcal{P}} = \xi_{\mathcal{P}}/\phi(M_{\mathcal{P}})$, and assume this is nonzero.

In the case of Gaussian discriminants we will have $M_{\mathcal{P}} = 8$ and $\mathcal{R}_{\mathcal{P}} = \{1, 5\}$, so $\alpha_{\mathcal{P}} = 1/2$. For the general case of fundamental discriminants we have $M_{\mathcal{P}} = 8$ and $\alpha_{\mathcal{P}} = 1$. For quadratic twist families of elliptic curves with full 2-torsion we can write the set of bad primes (including 2) of E as Ω, and then $M_{\mathcal{P}}$ is 4 times the product of the primes in Ω, and again $\alpha_{\mathcal{P}} = 1$.

5.1.1. We then consider positive squarefree $\tilde{d} \leq X$ all of whose prime factors come from $\mathcal{P}$, writing $S^{\mathcal{P}}(X)$ for this set. Almost all such $\tilde{d}$ have roughly $\alpha_{\mathcal{P}} \log\log X$ prime factors, and we place the prime divisors into basic intervals (of a dyadic nature), which then form boxes when taken as a Cartesian product. The main intention is that almost all $\tilde{d}$ then be represented by "pleasant" boxes, which satisfy various technical properties depending on parameters (η_0, η_1, η_s).

First, for a box, we take a given η_0 with $0 < \eta_0 < 1$, and an $(X, \eta_0, \mathcal{P})$-box $\bar{T}$ is a Cartesian product $\prod_u \{\mathbf{p}_u\} \times \prod_t \bar{T}_t$ where the primes $\mathbf{p}_u$ are distinct and in $\mathcal{P}$, with the $\mathbf{p}_u$ ordered increasingly with each $\leq P_0 = \exp\bigl((\log\log X)^{\eta_0}\bigr)$; while the $\bar{T}_t$ are a strictly increasing sequence of basic (dyadic-like) intervals $(A_t, B_t]$ with $\prod_u \mathbf{p}_u \prod_t B_t \leq X$ (and $A_t > P_0$).

The squarefree integers with prime divisors from $\mathcal{P}$ that are represented by a box come from the set $T = \prod_u \{\mathbf{p}_u\} \times \prod_t T_t$ where T_t consists of the primes in $\bar{T}_t$ that are in $\mathcal{P}$. Indeed, there is a natural injective map from this set to $S^{\mathcal{P}}_{\tilde{r}}(X)$, recalling the latter is the set of (positive) integers up to X with exactly $\tilde{r}$ prime factors, all of which are in $\mathcal{P}$. We say that $\tilde{d}$ is represented by a box if it is in the image $\hat{T}$ of this map. Having fixed a selection of non-overlapping basic intervals, every $\tilde{d}$ is represented by at most one such box.

5.1.2. For a box $\bar{T}$ we let k_0 be the number of primes $\mathbf{p}_u \leq P_0 = \exp\bigl((\log\log X)^{\eta_0}\bigr)$ corresponding to singleton sets, and let $\tilde{r}$ be k_0 plus the number of basic intervals.

We introduce another parameter η_1 and write k_1 for the number of singletons plus basic intervals that are less than Q_1, which is $\exp\exp\bigl((\log\log X)^{\eta_1}\bigr)$.

We expect that k_0 is roughly $\alpha_{\mathcal{P}}\log\log P_0 = \alpha_{\mathcal{P}}\eta_0\log\log\log X$. We will allow k_0 to be as large as $\kappa_0\alpha_{\mathcal{P}}\log\log P_0$ for some parameter $\kappa_0 > 3$. The error from excluding $\tilde{d}$ with larger k_0 will (essentially) be of relative size $1/(\log P_0)^{\alpha_{\mathcal{P}}\kappa_0(\log\kappa_0 - 1)}$ which is the same as $(1/\log\log X)^{\alpha_{\mathcal{P}}\eta_0[\kappa_0(\log\kappa_0-1)]}$.

We expect that k_1 is roughly $\alpha_{\mathcal{P}}\log\log Q_1 = \alpha_{\mathcal{P}}(\log\log X)^{\eta_1}$. We allow k_1 to be as large as $\kappa_1\alpha_{\mathcal{P}}\log\log Q_1$ for some parameter $\kappa_1 = 3$, and the error from ignoring $\tilde{d}$ with larger k_1 will be of relative size $1/(\log Q_1)^{\alpha_{\mathcal{P}}\kappa_1(\log\kappa_1 - 1)}$, which saves more than any power of $(\log\log X)$ asymptotically.

Finally, we introduce $P_{\mathrm{s}} = \exp\bigl((\log\log X)^{\eta_{\mathrm{s}}}\bigr)$, dependent on the parameter η_{s}. This is the size of allowable moduli in our usage of the prime number theorem in arithmetic progressions for which we do not care if there is an exceptional zero or not, as the gains are large enough anyway. We will thus need to exclude conductors $\geq P_{\mathrm{s}}$ associated to exceptional zeros.

This then gives us the definition of a $(\kappa_0, \eta_1, \eta_{\mathrm{s}})$-pleasant $(X, \eta_0, \mathcal{P})$-box. Namely it is a box with k_0 and k_1 restricted as above (with respect to κ_0 and η_1), and such that there is no $\tilde{d}$ represented by the box that is a multiple of the coprime-to-$\mathcal{P}$ part of some element of the sequence $\{\mathcal{M}_i\}$ of exceptional (Siegel) conductors that is $\geq P_{\mathrm{s}} = \exp\bigl((\log\log X)^{\eta_{\mathrm{s}}}\bigr)$.

5.1.3. In particular ([21, §4.5]), with $\eta_0/2 > \eta_1 > \eta_{\mathrm{s}} > 0$ and $\kappa_0\eta_0 = 1/\sqrt{\log 1/\eta_0}$, upon taking η_0 sufficiently small we find that the exceptional set of $\tilde{d} \in S^{\mathcal{P}}(X)$ that are not represented by such pleasant boxes is of relative size $\ll 1/(\log\log X)^{99}$.

Thus we can consider such pleasant boxes instead of $\tilde{d}$.

(In fact, in §6.2.2 below we shall append more regularity conditions to boxes, and the resulting very pleasant boxes will still cover $S^{\mathcal{P}}(X)$ with acceptable error).

5.2. We then consider [21, §5] restricting boxes $\bar{T}$ in various ways. The first to specify a residue class modulo $M_{\mathcal{P}}$ for each prime divisor of $\tilde{d}$.

We thus define the $\mathcal{K}$-trimming of a box. For each l with $1 \leq l \leq \tilde{r}$ we let $\mathcal{K}_l$ be a residue class modulo $M_{\mathcal{P}}$. Recall that the squarefree integers represented by a box $\bar{T}$ naturally lie in a Cartesian product $\prod_l T_l$, where each T_l is a singleton set or the set of primes in a basic interval $(A_l, B_l]$ that are in $\mathcal{P}$. Assuming that each singleton set meets its requisite $\mathcal{K}$-condition (otherwise we just take $T(\mathcal{K})$ as empty), we define $T(\mathcal{K}) = \prod_l T_l(\mathcal{K})$ where $T_l(\mathcal{K})$ is the set of primes in the basic interval $(A_l, B_l]$ that are in the residue class specified by $\mathcal{K}_l$ (in other words, it is the subset of T_l that meets said $\mathcal{K}_l$-condition).

This procedure of $\mathcal{K}$-trimming does not lose much in our estimates because we are simply taking progressions to a fixed modulus $M_{\mathcal{P}}$.

Note also that $\mathcal{K}$ might contain some "global" information. For instance, in the case of quadratic fields we have $D = 2^f\varepsilon\tilde{d}$ for some $(f, \varepsilon) \in \{0, 2, 3\} \times \{+, -\}$, which we prefer to be subject to conditions that ensure D is a fundamental discriminant, such as $\tilde{d} \equiv 3\,(4)$ when $(f, \varepsilon) = (0, -)$. This latter condition can then be encoded into $\mathcal{K}$ as saying the number of primes dividing $\tilde{d}$ that are 3 mod 4 is odd.

In the Selmer case we also more naturally have $\tilde{\mathcal{K}}$, which does not specify a congruence class but rather a class in $\mathbf{Q}_l^\star/(\mathbf{Q}_l^\star)^2$ for all $l \in \tilde{\Omega}$. Since $\mathcal{K}$ makes a stronger restriction, if we prove theorems for it, they also naturally hold for $\tilde{\mathcal{K}}$.

5.2.1. We let $\mathcal{L}$ be a set of Legendre symbol specifications, so that for $1 \le i < j \le \tilde{r}$ we take $\mathcal{L}_{ij} \in \{\pm 1\}$. We then define the $(\mathcal{K}, \mathcal{L})$-restriction of a box. This is the set of $\tilde{d} \in \hat{T}(\mathcal{K})$ with $\tilde{d} = p_1 \cdots p_{\tilde{r}}$ such that $(p_i|p_j) = \mathcal{L}_{ij}$ for all $1 \le i < j \le \tilde{r}$. This is more severe than the $\mathcal{K}$-trimming, as in general it will no longer be a Cartesian product. (If desired, we can define $\mathcal{L}_{ji}$ from $\mathcal{L}_{ij}$ by quadratic reciprocity since $\mathcal{K}$ specifies each prime modulo 4).

We write $\mathcal{D}(\tilde{r}, \mathcal{P})$ for the set of all applicable $(\mathcal{K}, \mathcal{L})$; for Gaussian discriminants this is all of them, while for general discriminants it only has those $\mathcal{K}$ that yield a fundamental discriminant from the (f, ε)-selection; and for the Selmer case it is again all of them (we twist by squarefree integers, so there is no bookkeeping issue).

5.3. The basic algebraic fact that we will use in the quadratic field case is that the 4-tuple $(f, \varepsilon, \tilde{\mathcal{K}}, \mathcal{L})$ determines the 4-rank of the narrow class group of $\mathbf{Q}(\sqrt{D})$. This can be made explicit by the Rédei matrix above .

Similarly, in the case where E is an elliptic curve with full 2-torsion, we know that the 4-tuple $(E, \varepsilon, \tilde{\mathcal{K}}, \mathcal{L})$ determines the 2-Selmer rank of E_d for $d = \varepsilon\tilde{d}$; however, the explicit matrix is more ponderous here [21, §§7-8], having two rows/columns for each prime divisor of d, and indeed similarly for each place in $\tilde{\Omega}$.

5.3.1. As the Rédei matrix $\mathbf{R}^4$ only depends on Legendre symbols for the primes that divide D, we can rewrite it in terms of "formal symbols" similarly to as we did for the 2-Selmer pairing matrix in [21, §8]. Indeed, this will allow us to fix a choice of $\mathbf{R}^4$-basis for all D that meet a given $(f, \varepsilon, \mathcal{K}, \mathcal{L})$ specification, so that comparison between such D at the 8-rank level is then possible.

Given u, we define the set $\mathbf{P}_u$ of u formal symbols $\dot{p}_i$ for $1 \le i \le u$, and the formal symbol $\dot{d}$, which is $2^f\varepsilon$ times the product of these u symbols; we consider the symbols to commute under concatenation/multiplication, and whether one has $\dot{2}$ (or even $\dot{\varepsilon}$) as a formal symbol or not is largely a matter of taste/expediency.

Given $(\tilde{\mathcal{K}}, \mathcal{L})$ we then define Legendre/Hilbert symbols for the formal symbols as $(q|\dot{p}_j)^\star = (q, \dot{p}_j)_{\dot{p}_j} = \tilde{\mathcal{K}}^\star_{qj}$ for $q \in \tilde{\Omega}$ (with here $q = \infty$ as $q = -1$) and $1 \le j \le \tilde{r}$, and $(\dot{p}_i|\dot{p}_j)^\star = (\dot{p}_i, \dot{p}_j)_{\dot{p}_j} = \mathcal{L}^\star_{ij}$ for $1 \le i \ne j \le \tilde{r}$, and $(\dot{p}_j, \dot{p}_j)_{\dot{p}_j} = (-1, \dot{p}_j)_{\dot{p}_j} = \tilde{\mathcal{K}}^\star_{\infty j}$. The Rédei matrix $\mathbf{R}^4$ for $(f, \varepsilon, \mathcal{K}, \mathcal{L})$ can then be described in terms of these (note in particular that $\mathcal{K}$ gives conditions modulo 8, so $(2|\dot{p}_j)$ is again determinable).

6. Hypercube co-ordinates

For a fixed $(f, \varepsilon) \in \{0, 2, 3\} \times \{+, -\}$, we consider a residue and Legendre specification $(\mathcal{K}, \mathcal{L}) \in \mathcal{D}(\tilde{r}, \mathcal{P})$, with e its 4-rank and $\{\vec{w}^j\}_j$ a basis for the left kernel of the pairing matrix $\mathbf{R}^4(f, \varepsilon, \mathcal{K}, \mathcal{L})$, with this matrix being r-by-r (so the vectors are in $\mathbf{F}_2^r$). We assume $\vec{w}^{e+1} = (1, 1, \ldots, 1)$, which we call the obvious vector (and/or basis element).

From now until §12 we will only consider the case of Gaussian discriminants. In particular, this implies that $\mathbf{R}^4$ is symmetric, so the bases for the left/right kernels can be taken to be the same. However, other than the enumeration of various types of characters in §6.1.1 (which in general will need some modifications), much of what we say in this section will be applicable in more generality.

6.1. We let ψ be a multiplicative character on the space of matrices over $\mathbf{F}_2$ with e rows and $(e+1)$ columns, and write $\psi^\star$ for its additive version (taking values in $\mathbf{F}_2$). We can write ψ as a product $\prod_i \prod_j \psi_{i,j}^{c(i,j)}$ for some exponent-array c of zeros and ones where $\psi_{i,j}^\star \in \widehat{\mathrm{Mat}}(e, e+1, \mathbf{F}_2)$ is the basic character that is dual to the matrix whose (i,j)-entry is 1 with the rest 0.

Depending on the particularities of the c-array, we then divide into three cases.

6.1.1. The first case (Type I) is when for some $1 \le i, j \le e$ there are asymmetric entries $c(i,j) \neq c(j,i)$ in the c-array. Here we will take h_ψ^{z} and h_ψ^{s} to be basis indices with $c(h_\psi^{\mathrm{s}}, h_\psi^{\mathrm{z}}) = 1 \neq c(h_\psi^{\mathrm{z}}, h_\psi^{\mathrm{s}})$. The set $\mathcal{V}_\psi$ will have three hypercube co-ordinates:[11] first s_ψ, which is taken to be in the set of co-ordinates such that nonobvious basis elements are zero except the (h_ψ^{s})th, which is nonzero at said co-ordinate; a second z_ψ for which all nonobvious basis elements are zero except the (h_ψ^{z})th, which is nonzero at said co-ordinate; and a third co-ordinate z'_ψ such that all nonobvious basis elements are zero at it.[12]

The second case (Type II) is when the exponent array $c(i,j)$ is symmetric in its left e-by-e block, and has a nonzero rightmost column or a nonzero diagonal entry (or indeed, possibly both). In this case the obvious vector $\vec{w}^{e+1} = (1, \ldots, 1)$ plays a rôle. We let h_ψ^{z} be a basis index at which either $c(h_\psi^{\mathrm{z}}, e+1) = 1$ or $c(h_\psi^{\mathrm{z}}, h_\psi^{\mathrm{z}}) = 1$, and call the corresponding basis vector $\vec{w}^{h_\psi^{\mathrm{z}}}$ the *z*-distinguished basis vector. We then take the set $\mathcal{V}_\psi$ of hypercube co-ordinates to have two elements:[13] first s_ψ, which is taken to be in the set of co-ordinates at which all nonobvious basis elements are zero; and a second co-ordinate z_ψ for which all nonobvious basis elements are zero except the (h_ψ^{z})th, which is nonzero at said co-ordinate. (One can note that when $e = 1$ we are always in Type II).

The third case (Type III) is when $c(i,j)$ has a rightmost column and diagonal of zero, and the remaining entries are symmetric with $c(i,j) = c(j,i)$. Here we take h_ψ^{z} and h_ψ^{s} to be basis indices with $c(h_\psi^{\mathrm{z}}, h_\psi^{\mathrm{s}}) = 1$, and indeed there is some choice of these since the $c(i,j)$ are not all zero. As previously, we call the associated basis vectors the *z*-distinguished and *s*-distinguished basis vectors. Here we again take $\mathcal{V}_\psi$ to have two elements: first s_ψ, which should again be a co-ordinate for which all nonobvious basis elements are zero except the (h_ψ^{s})th, which is nonzero at said co-ordinate; and another co-ordinate z_ψ, for which the same holds but for (h_ψ^{z}) instead of (h_ψ^{s}).

We give a pictorial representation of this, with the left side corresponding to Type I or III (the latter needs no z'_ψ), and the right to Type II.

$$\begin{array}{r@{\ }*{12}{c@{}}}
 & & & 0 & & & & 0 & & & & 0 & \\
 & & & 0 & & & & 0 & & & & 0 & \\
(h_\psi^{\mathrm{z}})\text{th basis vector} & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
 & & & 0 & & & & 0 & & & & 0 & \\
(h_\psi^{\mathrm{s}})\text{th basis vector} & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
 & & & 0 & & & & 0 & & & & 0 & \\
 & & & \mathrm{z}_\psi & & & & \mathrm{z}'_\psi & & & & \mathrm{s}_\psi &
\end{array}
\qquad
\begin{array}{r@{\ }*{12}{c@{}}}
 & & & 0 & & & & & & & & 0 & \\
 & & & 0 & & & & & & & & 0 & \\
(h_\psi^{\mathrm{z}})\text{th basis vector} & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
 & & & 0 & & & & & & & & 0 & \\
\text{obvious vector} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 & & & 0 & & & & & & & & 0 & \\
 & & & \mathrm{z}_\psi & & & & & & & & \mathrm{s}_\psi &
\end{array}$$

[11] Smith calls these "variable indices", but there are so many indices and variables floating about, I think it is better to denote them more distinctly.

[12] This is simpler than what is proposed in [4, Proposition 6.8], namely that one only need that the third index not be 1 for both the distinguished basis vectors.

[13] Smith always has three elements in his variable indices in his analogue (8-rank of imaginary quadratic fields), but the arrangement is different, so a comparison is not that easy. We follow [4].

6.2. When $(\mathcal{K},\mathcal{L})$ is sufficiently generic we can readily show that suitable hypercube co-ordinates exist, as indeed approximately $r/2^e$ co-ordinates should be zero at all nonobvious basis elements, etc. In practice, we shall also require s_ψ to be large and the other hypercube co-ordinates to be small.

The reader might note the discrepancy between r, which is the number of prime factors of D and thus the size of the $\mathbf{R}^4$ matrix, and $\tilde{r}$, which is one less when $f\neq 0$ (so $2|D$). It is convenient to index the co-ordinates so that $\tilde{r}$ is always the largest one, thus allowing the 0th to correspond to 2 (if extant). In the Selmer case, the matrix is of size $2(\tilde{r}+\tilde{\Omega})$, and similarly one prefers to index the co-ordinates with $\tilde{r}$ the largest, with 2 bits for each.

6.2.1. We define $r_{\mathrm{g}}=\lfloor(\alpha_{\mathcal{P}}/2)(\log\log X)\rfloor$, which is as good of a demarcation point as any (one could make it in terms of $\tilde{r}$ instead of $\log\log X$ if preferred).

We write $\tilde{\mathcal{V}}_\psi$ for $\mathcal{V}_\psi$ with s_ψ removed, and will call a selection of hypercube co-ordinates *well-gapped* if $\mathrm{s}_\psi\geq(5/4)r_{\mathrm{g}}$ and $r_{\mathrm{g}}/5\leq i\leq r_{\mathrm{g}}/4$ for $i\in\tilde{\mathcal{V}}_\psi$. (The numerology here is irrelevant; one only needs the upper co-ordinate to be a constant factor above r_{g}, and the lower co-ordinates a constant factor below it).

6.2.2. Let us expand on this a bit.

The upper candidate co-ordinates are the integers in $[(5/4)r_{\mathrm{g}},\tilde{r}]$, and the lower candidate co-ordinates are the integers in $[r_{\mathrm{g}}/5,r_{\mathrm{g}}/4]$. We write ν_{z} and ν_{s} for the numbers of such upper/lower candidate co-ordinates; the former is $\approx(3/4)r_{\mathrm{g}}$ and the latter is $\approx(1/20)r_{\mathrm{g}}$, with both of these $\gg\log\log X$.

It is then convenient to require regularity on the prime divisors of $\tilde{d}$ (in other words, regularity in the box-sizing) at various demarcation points. We introduce the notation $E_2^L(u,X)=\exp\exp(u\log\log X)$, noting that $\log E_2^L(u,X)=(\log X)^u$.

Recall we expect $\alpha_{\mathcal{P}}\log\log p_j\sim j$. For a box $\bar{T}$ we will require that all the primes in T_l for $l\geq(5/8)\alpha_{\mathcal{P}}\log\log X$ are $\geq\exp\exp\bigl(0.62(\log\log X)\bigr)=E_2^L(0.62,X)$. This will ensure the upper hypercube co-ordinate is large. We then require all the primes in T_l for $l\leq r_{\mathrm{g}}\sim(\alpha_{\mathcal{P}}/2)\log\log X$ to be small-ish, namely $\leq E_2^L(0.51,X)$. Similarly, we require all the primes in T_l for $l>r_{\mathrm{g}}$ to be $\geq E_2^L(0.49,X)$. Finally all the primes in T_l for $l\geq r_{\mathrm{g}}/5\sim(1/10)\alpha_{\mathcal{P}}(\log\log X)$ should be $\geq E_2^L(0.09,X)$, and this ensures the lower hypercube co-ordinates are of some size; while similarly for $l\leq r_{\mathrm{g}}/4\sim(1/8)\alpha_{\mathcal{P}}(\log\log X)$ all the primes in T_l should be $\leq E_2^L(0.13,X)$, to ensure that they are not too large.

(This is somewhat of a minimal set of regularity conditions, and one could append a few more to ease/sharpen later bounds if desired).

We enhance the definition of pleasant to include these, referring to such boxes as very pleasant. We will typically drop reference to the parameters, with the implicit understanding that we have $\eta_0/2>\eta_1>\eta_{\mathrm{s}}>0$ and $\kappa_0\eta_0=1/\sqrt{\log 1/\eta_0}$, and consider the limit $\eta_0\to 0$. As with [21, §4.3] the exceptional sets are sparse.

Lemma 6.2.3. *The number of $\tilde{d}$ that are represented by a pleasant box that is not very pleasant is $\ll\Phi^{\mathcal{P}}(X)/(\log X)^{\tilde{c}}$.*

The proof (left to the reader) is easier than in [21, Lemma 4.3.2] as here the demarcations are proportional to $\log\log X$.

6.3. Given a large subset of the r co-ordinates, it is not hard to show (via tails of the binomial distribution) that having a nonobvious nontrivial kernel vector with an abnormal number of 0's or 1's at said co-ordinates is rare. However, this statement only regards the vectors one at a time, and we want such a result when considering a "joint distribution" on a set of them – that is, if we take b basis vectors, then approximately $1/2^b$ of co-ordinate selections should match any binary pattern therein.

The following Lemma allows us to pass from single vectors to sets of them.

Lemma 6.3.1. [4, Lemma 6.9]. *Suppose every nontrivial element of a subspace G of vectors on $\mathbf{F}_2^l$ has $l/2+\Theta(\lambda E)$ co-ordinates that are zero where $\lambda = (3+\sqrt{17}/4)$.*

Then for every linearly independent subset $\mathcal{S}$ of b vectors from G and every pattern P in $\mathbf{F}_2^b$ the number of co-ordinates matching P on $\mathcal{S}$ is $l/2^b + \Theta(\lambda^b E)$.

Here we have used the Θ-notation, which is the big-Oh notation but with an implicit constant of 1.

Proof. We induct on b, with $b = 0$ being trivial and $b = 1$ the supposition.

We let $u(\mathcal{S}, P)$ be the number of co-ordinates that match the pattern P for the vectors in $\mathcal{S}$. We write $\mathcal{S} = \{\vec{v}_1, \vec{v}_2, \ldots,\}$ and define P_{10} to be the pattern P with the first entry (corresponding to $\vec{v}_1$) flipped, with similarly P_{01} flipping the second entry, and P_{11} flipping both of these.

We can then note that sums over such perturbed patterns reduce the problem, as $u(\mathcal{S}, P) + u(\mathcal{S}, P_{10}) = u(\mathcal{S}\backslash\{\vec{v}_1\}, P_1')$ and $u(\mathcal{S}, P) + u(\mathcal{S}, P_{01}) = u(\mathcal{S}\backslash\{\vec{v}_2\}, P_2')$, where P_i' removes the ith item from P. A slightly more subtler computation is that $u(\mathcal{S}, P) + u(\mathcal{S}, P_{11}) = u(\mathcal{S}\backslash\{\vec{v}_1, \vec{v}_2\} \cup \{\vec{v}_1 + \vec{v}_2\}, \tilde{P}')$, where $\tilde{P}'$ is the relevant pattern for the new basis. We can thus estimate these sums by induction.

Of course, the reason why we chose *two* special co-ordinates is that we also have $u(\mathcal{S}, P) + u(\mathcal{S}, P_{10}) + u(\mathcal{S}, P_{01}) + u(\mathcal{S}, P_{11}) = u(\mathcal{S}\backslash\{\vec{v}_1, \vec{v}_2\}, P_{12}'')$, which we can also estimate by induction. We then interpose a beneficial linear combination as

$$\begin{aligned} 2\big(u(\mathcal{S}, P) - l/2^b\big) = \big(&3u(\mathcal{S}, P) + u(\mathcal{S}, P_{01}) + u(\mathcal{S}, P_{10}) + u(\mathcal{S}, P_{11}) - 6l/2^b\big) \\ &- \big(u(\mathcal{S}, P) + u(\mathcal{S}, P_{01}) + u(\mathcal{S}, P_{10}) + u(\mathcal{S}, P_{11}) - 4l/2^b\big) \end{aligned}$$

which we see is equal to

$$\begin{aligned} &= \big([u(\mathcal{S}, P) + u(\mathcal{S}, P_{01})] + [u(\mathcal{S}, P) + u(\mathcal{S}, P_{10})] + [u(\mathcal{S}, P) + u(\mathcal{S}, P_{11})] - 6l/2^b]\big) \\ &\qquad - \big([u(\mathcal{S}, P) + u(\mathcal{S}, P_{01}) + u(\mathcal{S}, P_{10}) + u(\mathcal{S}, P_{11})] - 4l/2^b\big) \\ &= \big(3l/2^{b-1} + \Theta(3\cdot\lambda^{b-1}E) - 6l/2^b\big) - \big(l/2^{b-2} + \Theta(\lambda^{b-2}E) - 4l/2^b\big) \\ &= \Theta\big((3/\lambda + 1/\lambda^2)\lambda^b E\big), \end{aligned}$$

and indeed $\lambda = (3+\sqrt{17})/4$ implies $2 = 3/\lambda + 1/\lambda^2$ to complete the proof. □

6.4. One could take the definition of generic $(\mathcal{K}, \mathcal{L})$ simply to be that there is a suitable choice of basis for the kernel of $\mathbf{R}^4$ such that well-gapped hypercube co-ordinates exist (for all characters ψ). It is rather more convenient to give a sufficient condition for this to be true.

We define a *rotten* vector in $\mathbf{F}_2^r$ to be one that is nonobvious and nontrivial and fails either to have $\nu_z/2+\Theta(\lambda(\log\log X)^{3/4})$ zeros/ones on the ν_z lower candidate co-ordinates, or to have $\nu_s/2+\Theta(\lambda(\log\log X)^{3/4})$ zeros/ones on the ν_s upper candidate co-ordinates (where $\lambda = (3+\sqrt{17})/4 \approx 1.781$).

A specification $(\mathcal{K}, \mathcal{L})$ is *pattern-generic* if there are no rotten vectors in the kernel of $\mathbf{R}^4(\mathcal{K}, \mathcal{L})$. A specification is *size-generic* if $e \leq 99\sqrt{\log\log\log X}$. It is *generic* if it is both pattern- and size-generic.

Lemma 6.4.1. *If $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde r,\mathcal{P})$ is generic, then for sufficiently large X there is a selection of well-gapped hypercube co-ordinates for every $\psi^\star\in\widehat{\mathrm{Mat}}(e,e+1,\mathbf{F}_2)$.*

Proof. Given ψ we then have its c-array, and take h_ψ^{z} based upon its shape (§6.1.1), and similarly with h_ψ^{s} for Types I or III. These indices correspond to basis vectors of $\ker(\mathbf{R}^4)$. We then wish to select well-gapped hypercube co-ordinates based upon these distinguished indices.

Since $(\mathcal{K},\mathcal{L})$ is pattern-generic, every nontrivial nonobvious vector in the kernel has nearly the expected number of zeros/ones on the upper and lower candidate co-ordinates.

We want there to be an upper candidate co-ordinate with the pattern from the nonobvious basis vectors being all zeros for Type II, or all zeros except for h_ψ^{s} for Type I or III. By the preceding Lemma, the number of upper candidate co-ordinates that match such a pattern is $\nu_{\mathrm{s}}/2^e+\Theta\bigl(1.781^e(\log\log X)^{3/4}\bigr)$, and since we have $e\le 99\sqrt{\log\log\log X}$ and $\nu_{\mathrm{s}}\gg\log\log X$, this is positive for sufficiently large X.

Similarly, we want there to be a lower candidate co-ordinate whose pattern is all zeros except for h_ψ^{z}, and the number of lower candidate co-ordinates with such a pattern is $\nu_{\mathrm{z}}/2^e+\Theta\bigl(1.781^e(\log\log X)^{3/4}\bigr)$, again positive for sufficiently large X.

Finally, for Type I we want there to be a lower candidate co-ordinate whose pattern on the nonobvious basis vectors is all zeros, and the same computation as above shows one exists. □

6.4.2. Next we show that the effect of the non-generic $(\mathcal{K},\mathcal{L})$ can be placed in the error term. We recall that $k_0\le\kappa_0\eta_0\alpha_{\mathcal{P}}\log\log\log X$ and $k_1\le 3(\log\log X)^{\eta_1}$ in terms of our parameters.

Lemma 6.4.3. *The proportion of specifications $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde r,\mathcal{P})$ that are not size-generic is $\ll 1/(\log\log X)^{99}$.*

For a pleasant box $\bar T$, the sum of $\#T(\mathcal{K},\mathcal{L})$ over $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde r,\mathcal{P})$ that are not size-generic is $\ll\#T\cdot 2^{k_0/2}k_1/\sqrt{\log\log X}\ll\#T/(\log\log X)^{49/99}$ (as $\eta_0\to 0$).

Proof. The first statement follows immediately since the proportion of $(\mathcal{K},\mathcal{L})$ with a given e-value is $\ll 1/2^{e(e+1)/2}$.

The latter claim follows in the manner of [21, Proposition 6.3.1, §6.4], as we have

$$\tilde r!\sum_{\substack{(\mathcal{K},\mathcal{L})\\ e_f^{\varepsilon}(\mathcal{K},\mathcal{L})=e}}\#T(\mathcal{K},\mathcal{L})=\sum_{\substack{(\mathcal{K},\mathcal{L})\\ e_f^{\varepsilon}(\mathcal{K},\mathcal{L})=e}}\sum_{\sigma\in\mathrm{Sym}_{\tilde r}}\#T(\mathcal{K}^\sigma,\mathcal{L}^\sigma)=\tilde r!\sum_{\substack{(\mathcal{K},\mathcal{L})\\ e_f^{\varepsilon}(\mathcal{K},\mathcal{L})=e}}\frac{\#T}{2^{\binom{\tilde r}{2}}\xi_{\mathcal{P}}^{\tilde r}}+O\Bigl(\frac{\tilde r!\cdot\#T2^{k_0/2}k_1}{\sqrt{\log\log X}}\Bigr),$$

the final step from Proposition 6.3.1 therein, and the result follows. □

6.4.4. Next we make an accounting of how often each vector (in particular, the rotten ones) appears in kernels. The version here is specific to Gaussian discriminants, and we will employ a more complicated analysis in the general case (Lemma 12.3.2).

Lemma 6.4.5. *Every nonobvious nonzero vector in $\mathbf{F}_2^r$ is in the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ for a proportion $2/2^r$ of the $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde r,\mathcal{P})$.*

Proof. The Rédei matrix is symmetric and r-by-r, with row sums equal to 0. We remove (say) the rightmost column and bottom row, and the vectors $\vec v$ of the kernel of the remaining matrix M will give vectors in the kernel of $\mathbf{R}^4$ as $(\vec v,0)$ and $(\vec v+\vec 1,1)$.

Thus we are interested in how often a vector $\vec{v}$ is in the kernel of a symmetric $\mathbf{F}_2$-matrix M of size $(r-1)$. For any nonzero vector $\vec{v}$, it is in exactly $1/2^{r-1}$ of the M; for instance $(1,0,0,\cdots,0)$ is in the left kernel precisely when the left column is zero, and by symmetry the same argument works for any nonzero vector.

When D is odd we have $r=\tilde{r}$ and there is a one-to-one correspondence between the $2^{\binom{\tilde{r}}{2}}$ choices of $\mathcal{L}$ and such M (any $\mathcal{K}$-conditions do not matter here), and the result follows. On the other hand, when D is even we collate any $\mathcal{K}$-conditions into $\tilde{\mathcal{K}}$-conditions modulo 8 (thus two possibilities for each prime, either 1 mod 8 or 5 mod 8) and then the number of $(\tilde{\mathcal{K}},\mathcal{L})$ is $2^{\binom{r}{2}}$, again in one-to-one correspondence with the M of above. □

Lemma 6.4.6. *The proportion of* $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})$ *that are not pattern-generic is* $\ll \exp(-\tilde{c}\sqrt{\log\log X})$.

For a pleasant box $\bar{T}$ *the sum of* $\#T(\mathcal{K},\mathcal{L})$ *over* $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})$ *that are not pattern-generic is* $\ll \#T/(\log\log X)^{99}$.

Proof. By Stirling's approximation and tails of the binomial distribution, the number of rotten vectors is $\ll 2^r\exp(-\tilde{c}_{\mathcal{P}}\sqrt{\log\log X})$. On the other hand every nontrivial nonobvious vector is in a proportion $\ll 1/2^r$ of the kernels. So the proportion of nongeneric $(\mathcal{K},\mathcal{L})$ is as stated.

Then by [21, Lemma 5.5.1] we lose only a factor of $\ll 2^{k_0k_1}$ when passing to T-sizing, and as the exponent here is $k_0k_1\le\kappa_0\eta_0(\log\log\log X)\cdot 3(\log\log X)^{\eta_1}$, we see that this is dominated by $\exp(-\tilde{c}\sqrt{\log\log X})$. □

6.4.7. By these last two Lemmata and the above Lemma 6.2.3 concerning the sparsity of $\tilde{d}$ that are represented by not very pleasant boxes, we can restrict our future considerations to very pleasant boxes and generic $(\mathcal{K},\mathcal{L})$-specifications, and thus will always have well-gapped hypercube co-ordinates.

6.5. The question remains what to do with these hypercube co-ordinates.

Given a box $\bar{T}$ with $T=\prod_l T_l$, we will specialize every component T_l to a single prime, except those corresponding to a hypercube co-ordinate. It will be natural to do so in two steps, first for the part $t^{\downarrow}$ up to $r_{\mathrm{g}}\sim(\alpha_{\mathcal{P}}/2)\log\log X$, and then for the part $t^{\uparrow}$ beyond this.

We thus let $T^{\downarrow}$ be the product of the T_l with $l\le r_{\mathrm{g}}$ and $l\notin\tilde{\mathcal{V}}_\psi$, and $T^{\uparrow}$ be the product of the T_l with $l>r_{\mathrm{g}}$ and $l\ne \mathrm{s}_\psi$. Allowing a suitable re-ordering of the components of the Cartesian product, we then have

$$T(\mathcal{K})=T^{\downarrow}(\mathcal{K})\times\prod_{u\in\tilde{\mathcal{V}}_\psi}T_u(\mathcal{K})\times T^{\uparrow}(\mathcal{K})\times T_{\mathrm{s}_\psi}(\mathcal{K}).$$

We shall let the variable in $T_{\mathrm{s}_\psi}(\mathcal{K})$ range freely, and show cancellation over it; however, this will depend on Frobenius-equidistribution over a field defined from elements of the $T_u(\mathcal{K})$ from the lower hypercube co-ordinates. It will thus be profitable to break up $\prod_u T_u(\mathcal{K})$ into smaller parts, which we call grids.[14]

[14] Smith uses this term totally differently (as a union of product of real intervals) on page 51, but later (see the supergrid Z^l on page 81, or the grids Z'_{AR} on page 83) appears to use it to refer to a notion more similar to ours.

6.5.1. We will take grids $\check{Z} = \prod_u \check{Z}^u$ where $\check{Z}^u \subset T_u(\mathcal{K})$ for $u \in \tilde{\mathcal{V}}_\psi$, and each $\check{Z}^u$ is of size B. We write n_ψ for the number of $u \in \tilde{\mathcal{V}}_\psi$, that is, the number of lower hypercube co-ordinates.

It can get notationally messy if we had to refer to elements $\check{z}_i^u$ of $\check{Z}^u$, with $\check{Z}$ itself being in a sequence, etc. Fortunately, in our case we have $n_\psi \leq 2$, and thus can write more simply either $\check{Z} = Z$ or $\check{Z} = Z \times Z'$.

(Of course, we will eventually ensure that all these specializations with $t^\downarrow$, $t^\uparrow$, and $\check{Z}$ are compatible with $\mathcal{L}$).

6.5.2. We will take a collection $\{\check{Z}_i\}_i$ for which every intersection is of size at most 1, with the multiset $\bigcup_i \check{Z}_i$ covering the $\mathcal{L}$-compatible part of $\prod_u T_u(\mathcal{K})$ fairly uniformly, for instance, having all but a small proportion of elements appearing a fixed number $\bar{R}$ of times in the multi-set union.

If we had $n_\psi = 1$ this would perhaps be a familiar problem: for instance, given a set of size 10^{999} take a collection of subsets each of size 10^9 which intersect in at most one element and cover the original set nearly uniformly to multiplicity 10^{99}.

However, I don't know of any precursor result already with $n_\psi > 1$, and we will also need to ensure $\mathcal{L}$-conditions on $\check{Z}$, etc. Smith instead gives a completely combinatorial Lemma that we replicate in §9 below. This will allow us to reduce the situation to detection of ψ-cancellation on Cartesian products such as

$$t^\downarrow \times \check{Z} \times t^\uparrow \times T_{\mathrm{s}_\psi}(\mathcal{K})$$

where $t^\uparrow \in T^\uparrow$ and $t^\downarrow \in T^\downarrow$ perhaps would be more properly written as products of singleton sets (rather than as tuples). For the next two sections we will work with this ansatz; first, we develop suitable notation to refer to various box components and restrictions.

6.6. We assume that we have a very pleasant box $\bar{T}$, a residue and Legendre specification $(\mathcal{K}, \mathcal{L})$, a basis of size $(e+1)$ for the kernel of $\mathbf{R}^4(\mathcal{K}, \mathcal{L})$ that contains the obvious vector, a nontrivial multiplicative character ψ on $\mathrm{Mat}(e, e+1, \mathbf{F}_2)$, and a set of well-gapped hypercube co-ordinates $\mathcal{V}_\psi$.

To remind ourselves that we are in the case of Gaussian discriminants, we will refer to $\bar{T}$ as a $\mathcal{P}_4$-box.

Recall that $r_\mathrm{g} = \lfloor (\alpha_\mathcal{P}/2) \log\log X \rfloor$. We let $\mathcal{I}_\downarrow$ be the set of indices[15] $l \leq r_\mathrm{g}$ and $\mathcal{I}_\downarrow^\star$ when excluding $l \in \mathcal{V}_\psi$, and $T^\downarrow$ be the Cartesian product of the T_l over such $l \in \mathcal{I}_\downarrow^\star$. Similarly, we let $\mathcal{I}_\uparrow$ be the set of indices $l > r_\mathrm{g}$, with $\mathcal{I}_\uparrow^\star$ removing $l = \mathrm{s}_\psi$, and $T^\uparrow$ the Cartesian product of the T_l over such l.

Upon abusing notation by allowing permuting of the components, we then have

$$T(\mathcal{K}) = T^\downarrow(\mathcal{K}) \times \prod_{u \in \mathcal{V}_\psi^\star} T_u(\mathcal{K}) \times T^\uparrow(\mathcal{K}) \times T_{\mathrm{s}_\psi}(\mathcal{K}).$$

(We perhaps could drop the notational $\mathcal{K}$-references here, but particularly with the Selmer case in mind I prefer to retain them).

We will also write $r_\mathrm{g}^\uparrow$ for the size $\mathcal{I}_\uparrow^\star$, and similarly $r_\mathrm{g}^\downarrow$ for the size of $\mathcal{I}_\downarrow^\star$. We refer to the indices $l \leq r_\mathrm{g}$ as the lower co-ordinates, and $l > r_\mathrm{g}$ as the upper co-ordinates.

[15] Here we speak of indices (as opposed to co-ordinates), since the T-components correspond to divisors of $\tilde{d}$. Any distinction here is mostly irrelevant except for bookkeeping.

6.6.1. We write $t^{\downarrow}$ for an element of $T^{\downarrow}(\mathcal{K})$, though notationally it may also represent a product of singleton sets therein. Note that many $t^{\downarrow}$ will not even meet the Legendre specifications for the indices in $\mathcal{I}_{\downarrow}^{\star}$. We say that $t^{\downarrow}$ is $\mathcal{L}$-compatible if it does meet these, and write $T^{\downarrow}(\mathcal{K})[\![\mathcal{L}]\!]$ for the set of such $t^{\downarrow}$.

We write $T_j(\mathcal{K})[\mathcal{L}|t^{\downarrow}]$ for the subset of $T_j(\mathcal{K})$ that meets the $\mathcal{L}$-specifications corresponding to $t^{\downarrow}$. This will typically be about $1/2^{r_{\mathrm{g}}^{\downarrow}}$ smaller than $T_j(\mathcal{K})$. Note here that we used the single-bracket notation when referring to a single component that is being Legendre-restricted from other components, while the double-bracket notation of above refers to the $\mathcal{L}$-compatibility of elements of $T^{\downarrow}(\mathcal{K})$ themselves.

Similarly, given a set $\check{Z} = \prod_u \check{Z}^u \subset \prod_{u \in \mathcal{V}_{\psi}^{\star}} T_u(\mathcal{K})$, which we will call a grid (at least when it meets further $\mathcal{L}$-conditions), we write $T_j(\mathcal{K})[\mathcal{L}|\check{Z}]$ for the subset of $T_j(\mathcal{K})$ that meets the $\mathcal{L}$-specifications for all elements of $\check{Z}$. The grid components $\check{Z}^u$ will all be of size B and there will be n_{ψ} of them, so this gives $n_{\psi}B$ Legendre conditions, and the resulting restriction will typically be about $1/2^{n_{\psi}B}$ as large. We will also allow this notation to accumulate, with $T_j(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}]$ restricting $T_j(\mathcal{K})$ by $\mathcal{L}$-specifications from both $t^{\downarrow}$ and $\check{Z}$.

In particular, we will have $T^{\uparrow}(\mathcal{K})[\![\mathcal{L}|t^{\downarrow}, \check{Z}]\!]$ – this is the members of $T^{\uparrow}(\mathcal{K})$ that are $\mathcal{L}$-compatible amongst themselves (in the upper co-ordinates), and furthermore are $\mathcal{L}$-compatible with respect to $t^{\downarrow}$ and all members of $\check{Z}$.

6.6.2. Finally we introduce $T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow}\rangle$, which is the subset of $T(\mathcal{K}, \mathcal{L})$ such that the restriction to the $T^{\downarrow}$-component is $t^{\downarrow}$. Note that the elements of this set meet all the $\mathcal{L}$-conditions, not just those associated to $t^{\downarrow}$. We also have $T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow} \times y\rangle$, where y is an element of $\prod_u T_u(\mathcal{K})$ associated to the lower hypercube co-ordinates.

7. Applying results from Rédei symbols

As above with §6.6, we assume we have a very pleasant $\mathcal{P}_4$-box $\bar{T}$, a residue and Legendre specification $(\mathcal{K}, \mathcal{L}) \in \mathcal{D}(\tilde{r}, \mathcal{P})$ that is generic, a basis of size $(e+1)$ for the kernel of $\mathbf{R}^4(\mathcal{K}, \mathcal{L})$ that contains the obvious vector, a nontrivial multiplicative character ψ on $\mathrm{Mat}(e, e{+}1, \mathbf{F}_2)$, and a set of well-gapped hypercube co-ordinates $\mathcal{V}_{\psi}$.

7.1. Corresponding to the selection of n_{ψ} lower hypercube co-ordinates we will have a grid $\check{Z} = \prod_u \check{Z}^u$ with each of the n_{ψ} components $\check{Z}^u$ being of size B, with $n_{\psi} \leq 2$ and $B = \lfloor \sqrt{\log\log X}/999 \rfloor$. Here we have $\check{Z}^u \subset T_u(\mathcal{K})$ for all $u \in \tilde{\mathcal{V}}_{\psi}$. Since we will have $n_{\psi} \leq 2$, we can more simply write either $\check{Z} = Z$ or $\check{Z} = Z \times Z'$ rather than $\check{Z} = \prod_u \check{Z}^u$.

With evident notation as in §6.6, we will be considering

$$\mathcal{A}(t^{\downarrow}, \check{Z}, t^{\uparrow}) = t^{\downarrow} \times \check{Z} \times t^{\uparrow} \times T_{\mathrm{s}_{\psi}}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}, t^{\uparrow}]$$

where $t^{\downarrow}$ is $\mathcal{L}$-compatible, and every element in $\check{Z}$ is $\mathcal{L}$-compatible with itself and $t^{\downarrow}$, and $t^{\uparrow}$ is $\mathcal{L}$-compatible with: itself, $t^{\downarrow}$, and every element of $\check{Z}$. Thus, this is by design a subset of $T(\mathcal{K}, \mathcal{L})$, and indeed is a subset of $T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow}\rangle$.

Our goal will be to detect some sort of cancellation (for a ψ-sum) over the final component $T_{\mathrm{s}_{\psi}}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}, t^{\uparrow}]$, and we will achieve this by showing that for almost all choices of $t^{\uparrow}$ this set of primes is Frobenius-equidistributed in a field associated to $\check{Z}$. We will then sum (or take the union) over $t^{\uparrow}$, leaving us to further discuss the intricacies of $\check{Z}$ in §9ff.

7.2. We define a field $L_{\check{Z}}$ from $\check{Z}$, depending on the type of ψ as enumerated in §6.1.1. For now we will consider the case where ψ is of Type III, where we take $L_{\check{Z}}$ to be the compositum of number fields

$$L_{\check{Z}} = \prod_{k=2}^{B} \phi[p_1 p_k, -1]$$

where the p_k (including p_1) are the primes in $\check{Z}$, and $\phi[a,b]$ is any[16] field in $\mathcal{F}^{\mathrm{mr}}_{a,b}$, with the latter (as in §4.5.1) being the set of minimally ramified Galois extensions of $\mathbf{Q}(\sqrt{a},\sqrt{b})$ that are cyclic of degree 4 over $\mathbf{Q}(\sqrt{ab})$.

We let $K_{\check{Z}}$ be the largest multiquadratic extension of $\mathbf{Q}$ inside $L_{\check{Z}}$, which will be generated by $\sqrt{-1}$ and the $\sqrt{p_1 p_k}$ for $p_k \in \check{Z}$, and thus be of degree $2 \cdot 2^{B-1}$. We can note that the primes q in $T_l(\mathcal{K})[\mathcal{L}|\check{Z}]$ (for any applicable l, in particular $l = \mathrm{s}_\psi$) split completely in $K_{\check{Z}}$: first $(-1|q) = +1$, while $(p_i p_j|q) = (p_i|q)(p_j|q) = \mathcal{L}^2_{ul} = +1$ where u is the lower hypercube co-ordinate associated to $\check{Z}$.

7.2.1. We have a natural isomorphism from $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ to the image of $\partial_{\check{Z}}$, where the latter is defined in §3.2 for a set H as the linear map $\partial_H : \tilde{\mathbf{F}}_2^H \to \tilde{\mathbf{F}}_2^{H\times H}$ from the rule given by $\big(\partial_H F\big)(\vec{u}^1, \vec{u}^2) = \sum_{\vec{c}\in\mathbf{F}_2^H} F\big((u_1^{c_1}, ..., u_h^{c_h})\big)$ for a given map $F \in \tilde{\mathbf{F}}_2^H$. This isomorphism sends an element τ to the map that sends (p_i, p_j) to the element in $\mathbf{F}_2$ that signifies whether τ is trivial in $\mathrm{Gal}\big(\phi[p_i p_j, -1]/\mathbf{Q}(\sqrt{p_i p_j}, \sqrt{-1})\big)$. This isomorphism implies that $L_{\check{Z}}/K_{\check{Z}}$ is of degree 2^{B-1}, so $L_{\check{Z}}/\mathbf{Q}$ is of degree $2 \cdot 4^{B-1}$. (We say a bit more about this isomorphism in §7.8.3 below).

7.3. We want to describe how to relate a map $\tilde{g}$ that is in the image of $\partial_{\check{Z}\times[\![B]\!]}$ to a map g from $[\![B]\!] \times [\![B]\!]$ to $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$, and vice-versa. (Here $[\![B]\!]$ is the set of integers from 1 to B).

Suppose we have a map $g : [\![B]\!] \times [\![B]\!] \to \mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ that satisfies $g(i,j) = \tau_i\tau_j$ for some sequence $\{\tau_l\}_l$ (of length B) of elements in $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$. From the above isomorphism $\iota : \mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}}) \to \mathrm{im}(\partial_{\check{Z}})$ we get that $\iota g(i,j) = \partial_{\check{Z}} F_i + \partial_{\check{Z}} F_j$ for some maps $F_m : \check{Z} \to \mathbf{F}_2$. This then gives us a map

$$\tilde{g}(\check{z}_1, \check{z}_2, i, j) = \big(\partial_{\check{Z}} F_i\big)(\check{z}_1, \check{z}_2) + \big(\partial_{\check{Z}} F_j\big)(\check{z}_1, \check{z}_2) = F_i(\check{z}_1) + F_i(\check{z}_2) + F_j(\check{z}_1) + F_j(\check{z}_2),$$

and the latter shows that $\tilde{g}$ is in the image of $\partial_{\check{Z}\times[\![B]\!]}$.

Similarly, if we start from $\tilde{g}$ in this image, for some $\tilde{F} : \check{Z} \times [\![B]\!] \to \mathbf{F}_2$ we then have that $\tilde{g}(\check{z}_1, \check{z}_2, i, j) = \tilde{F}(i, \check{z}_1) + \tilde{F}(i, \check{z}_2) + \tilde{F}(j, \check{z}_1) + \tilde{F}(j, \check{z}_2)$, which then induces some $F_m : \check{Z} \to \mathbf{F}_2$ with $\tilde{g}(\check{z}_1, \check{z}_2, i, j) = \big(\partial_{\check{Z}} F_i\big)(\check{z}_1, \check{z}_2) + \big(\partial_{\check{Z}} F_j\big)(\check{z}_1, \check{z}_2)$, and thereupon a sequence of $\tau \in \mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ with $g(i,j) = \tau_i\tau_j$.

Note that a sequence of $\tau \in \mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ corresponding to g is specified uniquely by an arbitrary choice of one element (say τ_1), with $g(1,j) = \tau_1\tau_j$ then determining the others. We will subsequently work with a specific $\tilde{g}_0$ on $\check{Z} \times [\![B]\!]$ chosen so that it is ϵ-good as defined in §3.2.

[16] The original preprint version of [4] seems to allow any choice (page 44) of $\phi[p_k p_l, -1]$ for all $1 \le k, l \le B$, though correspondence with Koymans clarified that (for instance) in general $\phi[p_1 p_2, -1]$ and $\phi[p_1 p_3, -1]$ should determine what $\phi[p_2 p_3, -1]$ must be (as the twisting factor of Stevenhagen [18, (43)] is then prescribed).

7.4. With g in $\operatorname{im}(\partial_{\check{Z}\times[\![B]\!]})$ as above, suppose that $S = \{s_i\}_i$ is a sequence of B primes taken from $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}, t^{\uparrow}]$ that satisfies the Frobenius relation given by $g(1, j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1) \cdot \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$. We can identify S with $[\![B]\!]$ of above.

Recall we are working on

$$\mathcal{A}(t^{\downarrow}, \check{Z}, t^{\uparrow}) = t^{\downarrow} \times \check{Z} \times t^{\uparrow} \times T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}, t^{\uparrow}]$$

and have a nontrivial multiplicative character ψ that maps $\mathbf{R}^8(\vec{t})$ to $\{\pm 1\}$, with the associated additive character $\psi^\star$ to $\mathbf{F}_2$. As such, we have a map $\bar{\psi}^\star$ that sends $\check{Z}\times S$ to $\mathbf{F}_2$, given by $\bar{\psi}^\star(\check{z}, s) = \psi^\star\big(\mathbf{R}^8(t^{\downarrow} \times \check{z} \times t^{\uparrow} \times s)\big)$.

Given an ϵ-good $\tilde{g}$, we will select the sequence S in such a way that $\tilde{g} = \partial_{\check{Z}\times S}\bar{\psi}^\star$, as this will imply $\sum_{c\in\check{Z}\times S}\bar{\psi}(c)$ has some cancellation, and in particular its absolute value is $\le 2\epsilon(\#\check{Z}\#S)$. Upon showing below (§8) that $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}, t^{\uparrow}]$ is Frobenius-equidistributed (in $L_{\check{Z}}/K_{\check{Z}}$) for almost all choices of $t^{\uparrow}$, we will then have suitable cancellation on such $\mathcal{A}(t^{\downarrow} \times \check{Z} \times t^{\uparrow})$.

7.5. We now use the results on Rédei symbols to show we can ensure $\tilde{g} = \partial\bar{\psi}^\star$ on $\check{Z} \times S$ when ψ is of Type III.

Lemma 7.5.1. *Suppose ψ is of Type III (§6.1.1) and $S \subset T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|\check{Z}]$ is a sequence of primes, and $\tilde{g}$ a map in $\operatorname{im}(\partial_{\check{Z}\times S})$ with $g(1, j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1) \cdot \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$ for $1 \le j \le B = \#S$. Then we have $\tilde{g} = \partial\bar{\psi}^\star$ on $\check{Z} \times S$.*

As $\check{Z} = Z$ in this case, we shall suppress ornamentation on the variables.

Proof. First we compute what $\tilde{g}$ is. Identifying $[\![B]\!]$ with S we see $g(s_1, s_2) = \tau_{s_1}\tau_{s_2}$, so the value of $\tilde{g}\big((z_1, s_1), (z_2, s_2)\big)$ then exactly says whether the image of $\tau_{s_1}\tau_{s_2}$ in $\mathrm{Gal}\big(\phi[z_1z_2, -1]/\mathbf{Q}(\sqrt{z_1z_2}, \sqrt{-1})\big) \cong \mathbf{F}_2$ is trivial or not (in other words, whether the images of τ_{s_1} and τ_{s_2} are the same); meanwhile, the Rédei symbol $[z_1z_2, -1, s_1]$ says whether the image of τ_{s_1} is trivial in this extension, and similarly for s_2, so we have $\tilde{g}\big((z_1, s_1), (z_2, s_2)\big) = [z_1z_2, -1, s_1] + [z_1z_2, -1, s_2] = [z_1z_2, -1, s_1s_2]$.

7.5.2. Then we compute out what $\partial\bar{\psi}^\star$ is, and by definition of ∂ it is

$$(\partial\bar{\psi}^\star)\big((z_1, s_1), (z_2, s_2)\big) = \bar{\psi}^\star(z_1, s_1) + \bar{\psi}^\star(z_1, s_2) + \bar{\psi}^\star(z_2, s_1) + \bar{\psi}^\star(z_2, s_2).$$

We recall that $\psi^\star = \sum_i\sum_j c(i, j)\psi^\star_{i,j}$ is a linear combination of basic characters, and we will compute the contribution from each (i, j) separately. In this case (Type III) of symmetry we have $c(i, j) = c(j, i)$, and in particular the distinguished basis indices h^{s}_ψ and h^{z}_ψ have $c(h^{\mathrm{s}}_\psi, h^{\mathrm{z}}_\psi) = c(h^{\mathrm{z}}_\psi, h^{\mathrm{s}}_\psi) = 1$. Meanwhile the diagonal and right-column entries of c are zero, so $c(j, j) = c(j, e + 1) = 0$ for all j. Also, the (h^{s}_ψ)th basis element is the unique nonobvious one which is nonzero at s_ψ (corresponding to S), and the (h^{z}_ψ)th basis element is the unique nonobvious one which is nonzero at the hypercube co-ordinate z_ψ corresponding to $\check{Z}$.

We write $b_k = \prod_{l\notin\mathcal{V}_\psi} t_l^{v_l}$ where $\vec{v}$ is the kth basis vector and (v_l) its components, and u for the product here where all $v_l = 1$. When we are considering even D (so $f = 3$) we multiply b_k by 8 when the 0th component of v_k is 1, and similarly multiply u by 8.

7.5.3. Suppose first that we have a pair (i,j) where neither i nor j is equal to h_ψ^{s} or h_ψ^{z}. Then we have that the relevant co-ordinates of the ith and jth basis vectors $\vec{w}^i$ and $\vec{w}^j$ are zero, namely $\bigl((w^i_{\mathrm{z}_\psi}, w^i_{\mathrm{z}_\psi}), (w^j_{\mathrm{z}_\psi}, w^j_{\mathrm{s}_\psi})\bigr) = \bigl((0,0),(0,0)\bigr)$. Indeed, we work in the generality of this latter 4-tuple being specified. In terms of the Artin pairing we then simply have $\bar\psi^\star_{i,j}(z,s) = \langle\chi_{b_i}, b_j\rangle_{zsu}$, so that $\partial\bar\psi^\star_{i,j}\bigl((z_1,s_1),(z_2,s_2)\bigr)$ is

$$\langle\chi_{b_i}, b_j\rangle_{z_1s_1u} + \langle\chi_{b_i}, b_j\rangle_{z_2s_1u} + \langle\chi_{b_i}, b_j\rangle_{z_2s_2u} + \langle\chi_{b_i}, b_j\rangle_{z_1s_2u},$$

and writing this in terms of Rédei symbols gives

$$[b_i, z_1s_1b_iu, b_j] + [b_i, z_2s_1b_iu, b_j] + [b_i, z_2s_2b_iu, b_j] + [b_i, z_1s_2b_iu, b_j] = [b_i, 1, b_j] = 0,$$

upon using linearity in the second input.

Next we consider when a pair (i,j) has exactly one entry in common with the distinguished basis indices. Again we work in terms of 4-tuples as above. So suppose that $\bigl((w^i_{\mathrm{z}_\psi}, w^i_{\mathrm{s}_\psi}), (w^j_{\mathrm{z}_\psi}, w^j_{\mathrm{s}_\psi})\bigr) = \bigl((0,0),(0,1)\bigr)$, which occurs when $j = h_\psi^{\mathrm{s}}$ and i is neither h_ψ^{s} nor h_ψ^{z}. We then have that the Artin pairing is $\bar\psi^\star_{i,j}(z,s) = \langle\chi_{b_i}, sb_j\rangle_{zsu}$, so that $\partial\bar\psi^\star_{i,j}\bigl((z_1,s_1),(z_2,s_2)\bigr)$ is

$$\langle\chi_{b_i}, s_1b_j\rangle_{z_1s_1u} + \langle\chi_{b_i}, s_1b_j\rangle_{z_2s_1u} + \langle\chi_{b_i}, s_2b_j\rangle_{z_2s_2u} + \langle\chi_{b_i}, s_2b_j\rangle_{z_1s_2u},$$

and in terms of Rédei symbols this is

$$\begin{aligned}[] [b_i, z_1s_1b_iu, s_1b_j] + [b_i, z_2s_1b_iu, s_1b_j] + [b_i, z_2s_2b_iu, s_2b_j] + [b_i, z_1s_2b_iu, s_2b_j]& \\ = [b_i, z_1z_2, s_1b_j] + [b_i, z_1z_2, s_2b_j] = [b_i, z_1z_2, s_1s_2]&,\end{aligned}$$

where we used linearity in the second input, then linearity in the third. Moving the nonzero entry of the 4-tuple gives similar results; for instance, when $j = h_\psi^{\mathrm{s}}$ we get $[b_i, s_1s_2, z_1z_2]$ which by reciprocity is the same, while if i is h_ψ^{z} or h_ψ^{s} we get $[z_1z_2, s_1s_2, b_j]$ or the equivalent $[s_1s_2, z_1z_2, b_j]$. This is then sufficiently useful in our situation of symmetry: for instance if $c(l, h_\psi^{\mathrm{z}}) = 1$ (with thus $l \neq h_\psi^{\mathrm{z}}$) we obtain a contribution therein of $[b_l, s_1s_2, z_1z_2]$; and then we have $c(h_\psi^{\mathrm{z}}, l) = 1$ also, for which a similar computation yields $[b_l, z_1z_2, s_1s_2]$. Thus their sum is zero.

Finally we get to the cases where both i and j are distinguished basis vector indices (though of course $i \neq j$, since the diagonal entries $c(l,l)$ are 0). Suppose then that $\bigl((w^i_{\mathrm{z}_\psi}, w^i_{\mathrm{s}_\psi}), (w^j_{\mathrm{z}_\psi}, w^j_{\mathrm{s}_\psi})\bigr) = \bigl((1,0),(0,1)\bigr)$. We have $\bar\psi^\star_{i,j}(z,s) = \langle\chi_{zb_i}, sb_j\rangle_{zsu}$, so that $\partial\bar\psi^\star_{i,j}\bigl((z_1,s_1),(z_2,s_2)\bigr)$ is

$$\langle\chi_{z_1b_i}, s_1b_j\rangle_{z_1s_1u} + \langle\chi_{z_2b_i}, s_1b_j\rangle_{z_2s_1u} + \langle\chi_{z_2b_i}, s_2b_j\rangle_{z_2s_2u} + \langle\chi_{z_1b_i}, s_2b_j\rangle_{z_1s_2u},$$

and in terms of Rédei symbols this is

$$\begin{aligned}&[z_1b_i, s_1b_iu, s_1b_j] + [z_2b_i, s_1b_iu, s_1b_j] + [z_2b_i, s_2b_iu, s_2b_j] + [z_1b_i, s_2b_iu, s_2b_j] \\ &\quad = [z_1z_2, s_1b_iu, s_1b_j] + [z_1z_2, s_2b_iu, s_2b_j] \\ &\quad = [z_1z_2, s_1b_iu, -z_1z_2b_ib_ju] + [z_1z_2, s_2b_iu, -z_1z_2b_ib_ju] = [z_1z_2, s_1s_2, -z_1z_2b_ib_ju],\end{aligned}$$

where we used linearity in the first input, then $[A,B,C] = [A,B,-ABC]$, and then linearity in the second input. Again this calculation is not too useful by itself, but the computation with the $\bigl((0,1),(1,0)\bigr)$ 4-tuple gives $[s_1s_2, z_1z_2, -s_1s_2b_ib_ju]$; in our situation of symmetry we will always be summing these, and applying reciprocity and linearity gives their sum as $[z_1z_2, s_1s_2, z_1z_2s_1s_2] = [z_1z_2, s_1s_2, -1]$.

7.5.4. Summing up the above computations, and recalling that the diagonal and rightmost column of c are 0, we find that $\partial\bar{\psi}^\star\big((z_1,s_1),(z_2,s_2)\big)$ is

$$\begin{aligned}\sum_i\sum_j c_{i,j}\partial\bar{\psi}^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big) &= \partial\big[\bar{\psi}^\star_{h^{\mathrm{z}}_\psi,h^{\mathrm{s}}_\psi}+\bar{\psi}^\star_{h^{\mathrm{s}}_\psi,h^{\mathrm{z}}_\psi}\big]\big((z_1,s_1),(z_2,s_2)\big)\\ &= [z_1z_2,s_1s_2,-z_1z_2b_{h^{\mathrm{z}}_\psi}b_{h^{\mathrm{s}}_\psi}u]+[s_1s_2,z_1z_2,-s_1s_2b_{h^{\mathrm{s}}_\psi}b_{h^{\mathrm{z}}_\psi}u]=[z_1z_2,s_1s_2,-1].\end{aligned}$$

By Rédei reciprocity, this is thus equal to our computation of $\tilde{g}$ on $\check{Z}\times S$. □

7.6. We now do the analogous computations for Type II (and then Type I in the next subsection). It is probably easier to first compute the Rédei symbol relations (and thus $\partial\bar{\psi}^\star$), and then define the field $L_{\check{Z}}$ in terms of what we obtain.

Lemma 7.6.1. *Suppose ψ is of Type II, so c is symmetric in its leftmost block and at least one of $c(h^{\mathrm{z}}_\psi,h^{\mathrm{z}}_\psi)$ and $c(h^{\mathrm{z}}_\psi,e+1)$ is nonzero. Then*

$$\partial_{\check{Z}\times S}\bar{\psi}^\star\big((z_1,s_1),(z_2,s_2)\big)=c(h^{\mathrm{z}}_\psi,h^{\mathrm{z}}_\psi)[z_1z_2,s_1s_2,z_1z_2]+c(h^{\mathrm{z}}_\psi,e+1)[z_1z_2,s_1s_2,-1].$$

Proof. We recall that for Type II we have $c(i,j)=c(j,i)$ for $1\le i,j\le e$, and that the distinguished basis index h^{z}_ψ has either $c(h^{\mathrm{z}}_\psi,h^{\mathrm{z}}_\psi)$ or $c(h^{\mathrm{z}}_\psi,e+1)$ nonzero (or possibly both). All nonobvious basis vectors are zero at the hypercube co-ordinate s_ψ, while the (h^{z}_ψ)th basis vector is the only one that is nonzero at the hypercube co-ordinate z_ψ.

We again compute $\partial\bar{\psi}^\star$ as a sum of components $\sum_i\sum_j c(i,j)\partial\bar{\psi}^\star_{i,j}$.

7.6.2. When neither i nor j is equal to h^{z}_ψ or $(e+1)$, we are in the simple situation where $\big((w^i_{\mathrm{z}_\psi},w^i_{\mathrm{s}_\psi}),(w^j_{\mathrm{z}_\psi},w^j_{\mathrm{s}_\psi})\big)=\big((0,0),(0,0)\big)$, and the first computation in §7.5.3 gives $\partial\bar{\psi}^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big)=0$.

Next we consider when i or j is h^{z}_ψ, with the other not $(e+1)$, and moreover $i\neq j$. When $j=h^{\mathrm{z}}_\psi$ with the 4-tuple $\big((w^i_{\mathrm{z}_\psi},w^i_{\mathrm{s}_\psi}),(w^j_{\mathrm{z}_\psi},w^j_{\mathrm{s}_\psi})\big)=\big((0,0),(1,0)\big)$, by the second computation in §7.5.3 we get $[b_i,z_1z_2,s_1s_2]$. Again we can use c-symmetry to pair nonzero $c(l,h^{\mathrm{z}}_\psi)$ with $c(h^{\mathrm{z}}_\psi,l)$; the above computation gives $[b_l,z_1z_2,s_1s_2]$ for $(i,j)=(l,h^{\mathrm{z}}_\psi)$, while the analogous computation for $(i,j)=(h^{\mathrm{z}}_\psi,l)$ with the 4-tuple $\big((1,0),(0,0)\big)$ yields $[s_1s_2,z_1z_2,b_l]$, with the sum ergo being 0.

7.6.3. Next suppose that $i=j=h^{\mathrm{z}}_\psi$, considered in the generality of the 4-tuple $\big((w^i_{\mathrm{z}_\psi},w^i_{\mathrm{s}_\psi}),(w^j_{\mathrm{z}_\psi},w^j_{\mathrm{s}_\psi})\big)=\big((1,0),(1,0)\big)$. Here we have $\bar{\psi}^\star_{i,j}(z,s)=\langle\chi_{zb_i},zb_j\rangle_{zsu}$, so that $\partial\bar{\psi}^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big)$ is

$$\langle\chi_{z_1b_i},z_1b_j\rangle_{z_1s_1u}+\langle\chi_{z_2b_i},z_2b_j\rangle_{z_2s_1u}+\langle\chi_{z_2b_i},z_2b_j\rangle_{z_2s_2u}+\langle\chi_{z_1b_i},z_1b_j\rangle_{z_1s_2u}$$

and in terms of Rédei symbols this is

$$\begin{aligned}&[z_1b_i,s_1b_iu,z_1b_j]+[z_2b_i,s_1b_iu,z_2b_j]+[z_2b_i,s_2b_iu,z_2b_j]+[z_1b_i,s_2b_iu,z_1b_j]\\ &\qquad=[z_1b_i,s_1b_iu,-s_1b_ju]+[z_2b_i,s_1b_iu,-s_1b_ju]+\\ &\qquad\qquad\qquad+[z_1b_i,s_2b_iu,-s_2b_ju]+[z_2b_i,s_2b_iu,-s_2b_ju]\\ &\qquad=[z_1z_2,s_1b_iu,-s_1b_ju]+[z_1z_2,s_2b_iu,-s_2b_ju]\\ &\qquad=[z_1z_2,s_1b_iu,z_1z_2b_ib_j]+[z_1z_2,s_2b_iu,z_1z_2b_ib_j]=[z_1z_2,s_1s_2,z_1z_2b_ib_j],\end{aligned}$$

where we applied $[A,B,C]=[A,B,-ABC]$, then linearity in the first input, then $[A,B,C]=[A,B,-ABC]$ again, and linearity in the second input. Since $i=j$ we have $b_i=b_j$, and this is $[z_1z_2,s_1s_2,z_1z_2]$.

Suppose next $j = e+1$ and $i \neq h^{\mathrm{z}}_{\psi}$, or again we can work in the 4-tuple generality of $\big((w^i_{\mathrm{z}_\psi}, w^i_{\mathrm{s}_\psi}), (w^j_{\mathrm{z}_\psi}, w^j_{\mathrm{s}_\psi})\big) = \big((0,0),(1,1)\big)$. We have $\bar\psi^\star_{i,j}(z,s) = \langle \chi_{b_i}, zsb_j\rangle_{zsu}$, so that $\partial\bar\psi^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big)$ is

$$\langle \chi_{b_i}, z_1s_1b_j\rangle_{z_1s_1u} + \langle \chi_{b_i}, z_2s_1b_j\rangle_{z_2s_1u} + \langle \chi_{b_i}, z_2s_2b_j\rangle_{z_2s_2u} + \langle \chi_{b_i}, z_1s_2b_j\rangle_{z_1s_2u}.$$

which in terms of Rédei symbols is

$$\begin{aligned}&[b_i, z_1s_1b_iu, z_1s_1b_j] + [b_i, z_2s_1b_iu, z_2s_1b_j] + [b_i, z_2s_2b_iu, z_2s_2b_j] + [b_i, z_1s_2b_iu, z_1s_2b_j]\\ &\quad = [b_i, z_1s_1b_iu, -b_ju] + [b_i, z_2s_1b_iu, -b_ju] + [b_i, z_2s_2b_iu, -b_ju] + [b_i, z_1s_2b_iu, -b_ju]\end{aligned}$$

by $[A,B,C] = [A,B,-ABC]$, and this is 0 by linearity in the second input.

Finally, suppose $(i,j) = (h^{\mathrm{z}}_{\psi}, e+1)$, considered in the generality of the 4-tuple $\big((w^i_{\mathrm{z}_\psi}, w^i_{\mathrm{s}_\psi}), (w^j_{\mathrm{z}_\psi}, w^j_{\mathrm{s}_\psi})\big) = \big((1,0),(1,1)\big)$. We have $\bar\psi^\star_{i,j}(z,s) = \langle \chi_{zb_i}, zsb_j\rangle_{zsu}$, so that $\partial\bar\psi^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big)$ is

$$\langle \chi_{z_1b_i}, z_1s_1b_j\rangle_{z_1s_1u} + \langle \chi_{z_1b_i}, z_1s_2b_j\rangle_{z_1s_2u} + \langle \chi_{z_2b_i}, z_2s_2b_j\rangle_{z_2s_2u} + \langle \chi_{z_2b_i}, z_2s_1b_j\rangle_{z_2s_1u}.$$

and in terms of Rédei symbols this is

$$\begin{aligned}&[z_1b_i, s_1b_iu, z_1s_1b_j] + [z_1b_i, s_2b_iu, z_1s_2b_j] + [z_2b_i, s_2b_iu, z_2s_2b_j] + [z_2b_i, s_1b_iu, z_2s_1b_j]\\ &\quad = [z_1b_i, s_1b_iu, -b_ju] + [z_1b_i, s_2b_iu, -b_ju] + [z_2b_i, s_2b_iu, -b_ju] + [z_2b_i, s_1b_iu, -b_ju]\\ &\quad = [z_1z_2, s_1b_iu, -b_ju] + [z_1z_2, s_2b_iu, -b_ju] = [z_1z_2, s_1s_2, -b_ju]\end{aligned}$$

upon using $[A,B,C] = [A,B,-ABC]$ and linearity in the first and second inputs. For the obvious vector we have $b_{e+1} = u$, so this is $[z_1z_2, s_1s_2, -1]$.

7.6.4. Putting these all together, we find in Type II that $\partial\bar\psi^\star\big((z_1,s_1),(z_2,s_2)\big)$ is

$$\begin{aligned}\sum_i\sum_j &c(i,j)\partial\bar\psi^\star_{i,j}\big((z_1,s_1),(z_2,s_2)\big)\\ &= \partial\big[c(h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi})\bar\psi^\star_{h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi}} + c(h^{\mathrm{z}}_{\psi}, e+1)\bar\psi^\star_{h^{\mathrm{z}}_{\psi}, e+1}\big]\big((z_1,s_1),(z_2,s_2)\big)\\ &= c(h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi})[z_1z_2, s_1s_2, z_1z_2] + c(h^{\mathrm{z}}_{\psi}, e+1)[z_1z_2, s_1s_2, -1],\end{aligned}$$

and this gives the statement of the Lemma. □

7.6.5. The above then informs us how to define $L_{\check Z}$ in this case. We let

$$L_{\check Z} = \prod_{k=2}^{B} \phi[p_1p_k, (-1)^{c(h^{\mathrm{z}}_{\psi}, e+1)}(p_1p_k)^{c(h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi})}],$$

where again $\phi[a,b]$ is a field from $\mathcal{F}^{\mathrm{mr}}_{a,b}$ as in §4.5.1 or §7.2 (here $\phi[p_1p_k, p_1p_k]$ will be cyclic quartic rather than octic dihedral, but this does not really matter).

We let $K_{\check Z}$ be the largest multiquadratic extension of $\mathbf{Q}$ inside $L_{\check Z}$, which will be generated by the $\sqrt{p_1p_k}$ for $p_k \in \check Z$ when $c(h^{\mathrm{z}}_{\psi}, e+1) = 0$, and in general will also have $\sqrt{-1}$. We again note that any prime in $T_l(\mathcal{K})[\mathcal{L}|\check Z]$ splits completely in $K_{\check Z}$.

We have a natural isomorphism from $\mathrm{Gal}(L_{\check Z}/K_{\check Z})$ to the image of $\partial_{\check Z}$ that sends an element τ to the map that sends (p_i, p_j) to the element in $\mathbf{F}_2$ to signify whether τ is trivial in $\mathrm{Gal}\big(\phi[p_ip_j, x]/\mathbf{Q}(\sqrt{p_ip_j}, \sqrt{x})\big)$, where $x = (-1)^{c(h^{\mathrm{z}}_{\psi}, e+1)}(p_ip_j)^{c(h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi})}$.

We take a map $\tilde g : (\check Z \times [\![B]\!]) \times (\check Z \times [\![B]\!]) \to \mathbf{F}_2$ in $\mathrm{im}(\partial_{\check Z \times [\![B]\!]})$ as in §7.4, and suppose that $S = \{s_i\}_i$ is a sequence of B primes taken from $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check Z, t^{\uparrow}]$ that satisfies the Frobenius relation given by $g(1,j) = \mathrm{Frob}_{L_{\check Z}/K_{\check Z}}(s_1) \cdot \mathrm{Frob}_{L_{\check Z}/K_{\check Z}}(s_j)$. We have the following Lemma, proven in the same way as Lemma 7.5.1.

Lemma 7.6.6. *Suppose ψ is of Type II (§6.1.1) and $S \subset T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|\check{Z}]$ is a sequence of primes, and $\tilde{g}$ a map in $\mathrm{im}(\partial_{\check{Z}\times S})$ with $g(1,j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1)\cdot\mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$ for $1 \le j \le B = \#S$. Then we have $\tilde{g} = \partial\bar{\psi}^\star$ on $\check{Z}\times S$.*

7.7. The case of Type I is somewhat more notationally burdensome due to the fact that $\check{Z}$ is now a Cartesian product $Z \times Z'$. On the other hand, the occurrence of the third hypercube co-ordinate makes some of the calculations simpler.

We recall that we have an s-distinguished basis vector (indexed by h^{s}_ψ) and a z-distinguished basis vector (indexed by h^{z}_ψ); the former is 1 at the hypercube co-ordinate s_ψ and 0 at z_ψ, while the latter switches the behavior. Also, all the other non-obvious basis vectors are 0 at these hypercube co-ordinates. In terms of the c-coefficients, we have $c(h^{\mathrm{s}}_\psi, h^{\mathrm{z}}_\psi) = 1$ and $c(h^{\mathrm{z}}_\psi, h^{\mathrm{s}}_\psi) = 0$.

We also have a third hypercube co-ordinate z'_ψ at which all the non-obvious basis vectors are zero.

Lemma 7.7.1. *Suppose ψ is of Type I. Then*

$$\partial_{\check{Z}\times S}\bar{\psi}^\star\big((z_1,z_1',s_1),(z_2,z_2',s_2)\big) = [s_1s_2, z_1'z_2', z_1z_2].$$

Proof. First consider (i,j) with $i,j \notin \{e+1, h^{\mathrm{z}}_\psi, h^{\mathrm{s}}_\psi\}$. All 6 relevant co-ordinate values are 0, so $\bar{\psi}^\star_{i,j}(z,z',s) = \langle\chi_{b_i}, b_j\rangle_{zz'su}$, and $\partial\bar{\psi}^\star_{i,j}\big((z_1,z_1',s_1),(z_2,z_2',s_2)\big)$ is

$$\langle\chi_{b_i},b_j\rangle_{z_1z_1's_1u} + \langle\chi_{b_i},b_j\rangle_{z_1z_2's_1u} + \langle\chi_{b_i},b_j\rangle_{z_2z_1's_1u} + \langle\chi_{b_i},b_j\rangle_{z_2z_2's_1u} +$$
$$+ \langle\chi_{b_i},b_j\rangle_{z_1z_1's_2u} + \langle\chi_{b_i},b_j\rangle_{z_1z_2's_2u} + \langle\chi_{b_i},b_j\rangle_{z_2z_1's_2u} + \langle\chi_{b_i},b_j\rangle_{z_2z_2's_2u}.$$

In terms of Rédei symbols this is the 8-fold sum $\sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{s}}[b_i, \tilde{z}\tilde{z}'\tilde{s}b_iu, b_j]$, where each of the tildes indicates a co-ordinate to be summed over. By linearity in the second input this is 0.

Next we have (i,j) with $i = h^{\mathrm{s}}_\psi$ with $j \notin \{i, e+1, h^{\mathrm{z}}_\psi\}$. The only nonzero co-ordinate value is the (s_ψ)th for the ith basis vector, so $\bar{\psi}^\star_{i,j}(z,z',s) = \langle\chi_{sb_i}, b_j\rangle_{zz'su}$, and the 8-fold sum becomes

$$\sum_{\tilde{z}}\sum_{\tilde{z}'}[s_1b_i, \tilde{z}\tilde{z}'b_iu, b_j] + \sum_{\tilde{z}}\sum_{\tilde{z}'}[s_2b_i, \tilde{z}\tilde{z}'b_iu, b_j]$$

in terms of Rédei symbols, with again this being 0 by linearity in the second input. The same holds true if we switch the identity of i and j, or of h^{s}_ψ and h^{z}_ψ.

For the diagonal terms (i,i) with $i = h^{\mathrm{s}}_\psi$ we have $\bar{\psi}^\star_{i,i}(z,z',s) = \langle\chi_{sb_i}, sb_i\rangle_{zz'su}$, and the 8-fold sum is

$$\sum_{\tilde{z}}\sum_{\tilde{z}'}[s_1b_i, \tilde{z}\tilde{z}'b_iu, s_1b_i] + \sum_{\tilde{z}}\sum_{\tilde{z}'}[s_2b_i, \tilde{z}\tilde{z}'b_iu, s_2b_i],$$

which again is 0. The analogous computation holds for (i,i) with $i = h^{\mathrm{z}}_\psi$.

7.7.2. We then will get a main term from $(i,j) = (h^{\mathrm{s}}_\psi, h^{\mathrm{z}}_\psi)$. Here two of the six co-ordinate values are nonzero, and we have $\bar{\psi}^\star_{i,j}(z,z',s) = \langle\chi_{sb_i}, zb_j\rangle_{zz'su}$, and we find the 8-fold sum is

$$[s_1b_i, z_1z_1'b_iu, z_1b_j] + [s_1b_i, z_1z_2'b_iu, z_1b_j] + [s_1b_i, z_2z_1'b_iu, z_2b_j] + [s_1b_i, z_2z_2'b_iu, z_2b_j] +$$
$$+ [s_2b_i, z_1z_1'b_iu, z_1b_j] + [s_2b_i, z_1z_2'b_iu, z_1b_j] + [s_2b_i, z_2z_1'b_iu, z_2b_j] + [s_2b_i, z_2z_2'b_iu, z_2b_j],$$

which by linearity in the second input is

$$[s_1b_i, z_1'z_2', z_1b_j] + [s_1b_i, z_1'z_2', z_2b_j] + [s_2b_i, z_1'z_2', z_1b_j] + [s_2b_i, z_1'z_2', z_2b_j],$$

and then by linearity in the third and first input is

$$[s_1b_i, z_1'z_2', z_1z_2] + [s_2b_i, z_1'z_2', z_1z_2] = [s_1s_2, z_1'z_2', z_1, z_2].$$

Note that $(i,j) = (h_\psi^{\mathrm{z}}, h_\psi^{\mathrm{s}})$ has $c(i,j) = 0$ by the asymmetry assumption, and so does not contribute.

We still need to consider cases where $j = e+1$ (with the obvious basis vector). When $i \notin \{h_\psi^{\mathrm{s}}, h_\psi^{\mathrm{z}}\}$ we have $\bar{\psi}_{i,j}^\star(z, z', s) = \langle \chi_{b_i}, zz'sb_j\rangle_{zz'su}$, with then the 8-fold sum as $\sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{s}}[b_i, \tilde{z}\tilde{z}'\tilde{s}b_iu, \tilde{z}\tilde{z}'\tilde{s}b_j] = \sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{s}}[b_i, \tilde{z}\tilde{z}'\tilde{s}b_iu, -b_ju] = 0$.

Finally we have $(i,j) = (h_\psi^{\mathrm{s}}, e+1)$, where $\bar{\psi}_{i,j}^\star(z, z', s) = \langle \chi_{sb_i}, zz'sb_j\rangle_{zz'su}$ and the 8-fold sum is

$$\begin{aligned}\sum_{\tilde{z}}\sum_{\tilde{z}'}&[s_1b_i, \tilde{z}\tilde{z}'b_iu, s_1\tilde{z}\tilde{z}'b_j] + \sum_{\tilde{z}}\sum_{\tilde{z}'}[s_2b_i, \tilde{z}\tilde{z}'b_iu, s_2\tilde{z}\tilde{z}'b_j]\\ &= \sum_{\tilde{z}}\sum_{\tilde{z}'}[s_1b_i, \tilde{z}\tilde{z}'b_iu, -b_ju] + \sum_{\tilde{z}}\sum_{\tilde{z}'}[s_2b_i, \tilde{z}\tilde{z}'b_iu, -b_ju] = 0,\end{aligned}$$

with an analogous computation for $(i,j) = (h_\psi^{\mathrm{z}}, e+1)$. Putting these all together, we get the Lemma. □

7.7.3. The above then informs us how to define $L_{\check{Z}}$ in this case. We let

$$L_{\check{Z}} = \prod_{k=2}^{B}\prod_{l=2}^{B}\phi[p_1p_k, p_1'p_l']$$

where the p_k come from Z and the p_l' from Z', and $\phi[a,b]$ is a field from $\mathcal{F}_{a,b}^{\mathrm{mr}}$ as in §4.5.1 or §7.2.

We let $K_{\check{Z}}$ be the largest multiquadratic extension of $\mathbf{Q}$ inside $L_{\check{Z}}$, and note that any prime in $T_l(\mathcal{K})[\mathcal{L}|\check{Z}]$ splits completely in $K_{\check{Z}}$.

We have a natural isomorphism from $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ to the image of $\partial_{\check{Z}}$. This sends an element τ to the map that sends $\big((p_i, p_i'), (p_j, p_j')\big)$ to the element in $\mathbf{F}_2$ signifying whether τ is trivial in $\mathrm{Gal}\big(\phi[p_ip_j, p_i'p_j']/\mathbf{Q}(\sqrt{p_ip_j}, \sqrt{p_i'p_j'})\big)$.

We take a map $\tilde{g} : (\check{Z} \times [\![B]\!]) \times (\check{Z} \times [\![B]\!]) \to \mathbf{F}_2$ in $\mathrm{im}(\partial_{\check{Z}\times[\![B]\!]})$ as in §7.4, and suppose that $S = \{s_i\}_i$ is a sequence of B primes taken from $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow, \check{Z}, t^\uparrow]$ that satisfies the Frobenius relation given by $g(1,j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1)\cdot\mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$. We have the following Lemma, proven in the same way as Lemma 7.5.1.

Lemma 7.7.4. *Suppose ψ is of Type I (§6.1.1) and $S \subset T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|\check{Z}]$ is a sequence of primes, and $\tilde{g}$ a map in $\mathrm{im}(\partial_{\check{Z}\times S})$ with $g(1,j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1)\cdot\mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$ for $1 \le j \le B = \#S$. Then we have $\tilde{g} = \partial\bar{\psi}^\star$ on $\check{Z}\times S$.*

7.8. We sum up the calculations of the various types, and phrase the results in terms of cancellation for some $\tilde{g}$ that is ϵ-good (see §3.2.1).

Proposition 7.8.1. *Suppose that ψ is a nontrivial multiplicative character for the space $\mathrm{Mat}(e, e+1, \mathbf{F}_2)$, and let S be a sequence of primes from $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow, \check{Z}, t^\uparrow]$ of length B, and take $\tilde{g} \in \mathrm{im}(\partial_{\check{Z}\times S})$ with $g(1,j) = \mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_1)\cdot\mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(s_j)$ for $1 \le j \le B = \#S$, with $\tilde{g}$ being ϵ-good. Then*

$$\Bigg|\sum_{(\check{z},s)\in\check{Z}\times S}\bar{\psi}\big((\check{z},s)\big)\Bigg| \le 2\epsilon(\#\check{Z}\#S).$$

Proof. From the Lemmata for the various types we have $\partial\bar{\psi}^\star = \tilde{g}$. Since $\tilde{g}$ is ϵ-good the same is true for $\bar{\psi}^\star$, so $|\#(\bar{\psi}^\star)^{-1}(0) - (\#\check{Z}\#S)/2| \leq \epsilon(\#\check{Z}\#S)$. Transforming $\bar{\psi}^\star$ to $\bar{\psi}$ then gives the result. □

7.8.2. Let us record some facts about the fields $L_{\check{Z}}$ and $K_{\check{Z}}$ in the three cases.

For Type III we have

$$L_{\check{Z}} = \prod_{k=2}^{B} \phi[p_1 p_k, -1]$$

where the p_k are from $\check{Z}$, and $\phi[a,b]$ is any field in $\mathcal{F}^{\mathrm{mr}}_{a,b}$, with the latter (as in §4.5.1) being the set of minimally ramified Galois extensions of $\mathbf{Q}(\sqrt{a}, \sqrt{b})$ that are cyclic of degree 4 over $\mathbf{Q}(\sqrt{ab})$. The largest multiquadratic extension $K_{\check{Z}}$ of $\mathbf{Q}$ inside $L_{\check{Z}}$ is generated by $\sqrt{-1}$ and the $\sqrt{p_1 p_k}$ for $p_k \in \check{Z}$, and is thus of degree $2 \cdot 2^{B-1}$. We also have that $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ is isomorphic to $\mathrm{im}(\partial_{\check{Z}})$, and thus $L_{\check{Z}}/K_{\check{Z}}$ has degree 2^{B-1}, with the degree of $L_{\check{Z}}/\mathbf{Q}$ then being $2 \cdot 4^{B-1}$.

Similarly, for Type II we have

$$L_{\check{Z}} = \prod_{k=2}^{B} \phi[p_1 p_k, (-1)^{c(h^{\mathrm{z}}_{\psi}, e+1)} (p_1 p_k)^{c(h^{\mathrm{z}}_{\psi}, h^{\mathrm{z}}_{\psi})}],$$

and when $c(h^{\mathrm{z}}_{\psi}, e+1)$ is nonzero we reach the same conclusions as with Type III. Moreover, things are even easier in the alternative case, as then the $\phi[p_1 p_k, p_1 p_k]$ are all cyclic of degree 4 over $\mathbf{Q}$.

Finally, for Type I we have

$$L_{\check{Z}} = \prod_{k=2}^{B} \prod_{l=2}^{B} \phi[p_1 p_k, p'_1 p'_l],$$

where here there are two components in $\check{Z} = Z \times Z'$. The largest multiquadratic extension $K_{\check{Z}}$ of $\mathbf{Q}$ inside $L_{\check{Z}}$ is generated by $\sqrt{p_1 p_k}$ and $\sqrt{p'_1 p'_l}$ for $p_k \in Z$ and $p'_l \in Z'$, so is of degree $2^{2(B-1)} = 2^{n_\psi (B-1)}$. Meanwhile, the isomorphism from $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ to $\mathrm{im}(\partial_{\check{Z}})$ implies $L_{\check{Z}}/K_{\check{Z}}$ has degree $2^{(B-1)^2} = 2^{(B-1)^{n_\psi}}$.

In particular, each possibility for $L_{\check{Z}}$ is Galois over $\mathbf{Q}$ and its Galois group has order a power-of-two, with this group then being nilpotent (as all finite p-groups are), hence monomial, so that the Artin conjecture is known to hold [1], namely the L-functions of the irreducible representations are entire, and we can apply the Chebotarev density theorem as in §3.1.2.

The degrees v of said irreducible representations will be less than the square root of the size m of the Galois group. Meanwhile, the ramified primes will be those in the union $\bigcup_u \check{Z}^u$ and possibly the prime 2, each of the former being tamely ramified. This gives the conductor of each irreducible representation as $\leq (8K)^v$, where K is the product of the primes in $\bigcup_u \check{Z}^u$, and is thus $\leq E_2(0.13, X)^{n_\psi B} \ll E_2(0.14, X)$ by regularity (see §6.2.2).

7.8.3. (Exercise). Let us verify ι is an isomorphism from $\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ to $\mathrm{im}(\partial_{\check{Z}})$, for Type III. This sends an element τ to the map that sends (p_1, p_2) to the element in $\mathbf{F}_2$ signifying whether τ is trivial in $\mathrm{Gal}\big(\phi[p_1 p_2, -1]/\mathbf{Q}(\sqrt{p_1 p_2}, \sqrt{-1})\big)$. An alternative phrasing is whether the image of τ in $\mathrm{Gal}\big(\phi[p_1 p_2, -1]K_{\check{Z}}/K_{\check{Z}}\big)$ is trivial.

First we note that ι is a homomorphism. It sends τ to $g_\tau(i,j)$ that says whether the image of τ in $\mathrm{Gal}\big(\phi[p_i p_j, -1]K_{\check{Z}}/K_{\check{Z}}\big)$ is trivial, and $g_{\tau_1 \tau_2}$ is then just the sum of the maps g_{τ_1} and g_{τ_2}. Moreover, the kernel of ι is trivial, as any nonidentity τ will be nontrivial in some $\mathrm{Gal}\big(\phi[p_ip_j,-1]K_{\check{Z}}/K_{\check{Z}}\big)$.

Next we check that the image of ι is in $\mathrm{im}(\partial_{\check{Z}})$. Indeed, we let $F_\tau(j)$ determine whether the image of τ in $\mathrm{Gal}\big(\phi[p_1p_j,-1]K_{\check{Z}}/K_{\check{Z}}\big)$ is trivial, and then we have $g_\tau(i,j)=F_\tau(i)+F_\tau(j)$. Similarly, writing this in reverse shows that the image of ι is all of $\mathrm{im}(\partial_{\check{Z}})$; that is, with $g(i,j)=F(i)+F(j)$ there is some τ that is trivial in $\mathrm{Gal}\big(\phi[p_1p_j,-1]K_{\check{Z}}/K_{\check{Z}}\big)$ precisely when $F(j)=0$.

8. Frobenius equi-distribution with the upper hypercube co-ordinate

Almost all of what we do in the sequel below will be applicable more generally than just to Gaussian discriminants, but we still refer to a $\mathcal{P}_4$-box to remind us of this constraint.

As with §6.6, we again assume we have a very pleasant $\mathcal{P}_4$-box $\bar{T}$, a residue and Legendre specification $(\mathcal{K},\mathcal{L})$ that is generic, a basis of size $(e+1)$ for the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ that contains the obvious vector, a nontrivial multiplicative character ψ on $\mathrm{Mat}(e,e+1,\mathbf{F}_2)$, and a set of well-gapped hypercube co-ordinates $\mathcal{V}_\psi$.

With notation as in §6.6, we will be considering

$$\mathcal{A}(t^\downarrow,\check{Z},t^\uparrow)=t^\downarrow\times\check{Z}\times t^\uparrow\times T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]$$

where $t^\downarrow$ and $t^\uparrow$ are respectively products of singleton subsets (and/or elements therein) from $T^\downarrow(\mathcal{K})$ and $T^\uparrow(\mathcal{K})$, and $\check{Z}$ is a suitable subset of $\prod_{u\in\check{V}_\psi}T_u(\mathcal{K})$ associated to the lower hypercube co-ordinates.

Given $t^\downarrow$ and $\check{Z}$, we will now show that almost all choices of $\mathcal{L}$-compatible $t^\uparrow$ have $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]$ equi-distributed over Frobenius classes for $L_{\check{Z}}/K_{\check{Z}}$, and the results of the previous section will then imply that $\psi\big(\mathbf{R}^8(\vec{t})\big)$ has cancellation when summed over $\vec{t}\in\mathcal{A}(t^\downarrow,\check{Z},t^\uparrow)$.

For notation, we define T_l^τ to be the subset of $p\in T_l$ such that the Frobenius image $\mathrm{Frob}_{L_{\check{Z}}/K_{\check{Z}}}(p)$ is τ, and the various restrictions of this in the expected manner. We shall only need this for $l=\mathrm{s}_\psi$.

8.1. The equi-distribution follows from two things: namely $T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]$ is of the expected size for every τ (by the Chebotarev theorem and primes in arithmetic progressions to small moduli); and the number of $t^\uparrow$ with an unexpected size of $T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]$ is small (which will follow from a bilinear estimate).

We recall our notation $E_2^L(u,X)=\exp\exp(u\log\log X)$ which in turn implies that $\log E_2^L(u,X)=(\log X)^u$. Also, $r_{\mathrm{g}}^\downarrow$ is the number of indices in $\mathcal{I}_\downarrow^\star$, and is thus the number of components of the corresponding Cartesian product $T^\downarrow$, while $\check{Z}$ has n_ψ components each of size $B=\lfloor\sqrt{\log\log\log X}/999\rfloor$.

Lemma 8.1.1. *Suppose that $\bar{T}$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K},\mathcal{L})$ is a generic residue and Legendre specification. Let $t^\downarrow$ be an element of $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$, and $\check{Z}$ be a grid that is $\mathcal{L}$-compatible (with respect to $t^\downarrow$), with components of size B for the n_ψ well-gapped lower hypercube co-ordinates, with s_ψ the upper hypercube co-ordinate. Then for every $\tau\in\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ we have (with $E_2^L(u,X)=\exp\exp(u\log\log X)$)*

$$\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]=\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{(B-1)^{n_\psi}}}\biggl[1+O\Bigl(\frac{1}{E_2^L(0.09,X)}\Bigr)\biggr].$$

Proof. There are two aspects here: firstly to use primes in arithmetic progressions (or more particularly: primes with a fixed Legendre symbol) for the moduli with $t^\downarrow$

and $\check{Z}$; and secondly to use the Chebotarev theorem for the equi-distribution of τ. Of course, the former is merely the case of degree 1 Artin representations in the Chebotarev terminology.

Since we have $\mathrm{s}_\psi \geq (5/8)(\alpha_{\mathcal{P}} \log\log X)$ and $\bar{T}$ is very pleasant (see §6.2.2), the primes in question will be $\gg \exp\exp(0.62 \log\log X) = E_2^L(0.62, X)$; meanwhile since $r_{\mathrm{g}} \leq (\alpha_{\mathcal{P}}/2)\log\log X$, again by very pleasantness the moduli will be bounded as $\ll E_2^L(0.51, X)^{r_{\mathrm{g}}^{\downarrow}} E_2^L(0.13, X)^{n_\psi B} \ll E_2^L(0.52, X)$

Applying the prime number theorem for arithmetic progressions, this gives us

$$\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}] = \frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})}{2^{r_{\mathrm{g}}^{\downarrow} + n_\psi B}} + O\big(\tilde{E} \cdot \#T_{\mathrm{s}_\psi}(\mathcal{K})(\log X)^{0.62}\big)$$

where

$$\tilde{E} = \exp\biggl(-\tilde{c}\frac{\log E_2^L(0.62, X)}{\sqrt{\log E_2^L(0.62, X)} + \log E_2^L(0.52, X)}\biggr) + E_2^L(0.62, X)^{-1/\sqrt{P_{\mathrm{s}}}}.$$

Here $P_{\mathrm{s}} = \exp\big((\log\log X)^{\eta_{\mathrm{s}}}\big)$ corresponds to the allowable limit of exceptional zeros for pleasant boxes. We thus get $-\log\log \tilde{E} \gtrsim (0.62 - 0.52)\log\log X$ which implies that $\tilde{E} \ll 1/E_2(0.095, X)$ for sufficiently large X.

8.1.2. We then want to estimate $\#T^{\tau}_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}]$ by the Chebotarev density theorem. Referring to §7.8.2 the degree m of $L_{\check{Z}}/\mathbf{Q}$ is $\leq 2 \cdot 4^{B^{n_\psi}}$, and every irreducible Artin representation has degree bounded by $\sqrt{m}$. The conductors are all $\leq (8K)^{\sqrt{m}}$ where K is the product of all the primes in $\bigcup_u \check{Z}^u$, with $K \leq E_2^L(0.13, X)^{n_\psi B}$ by very pleasantness.

We also append the Legendre specifications from $t^{\downarrow}$ and $\check{Z}$; this amounts to twisting by the various degree 1 Artin represenations associated to the quadratic fields they define, and as above this is harmless in the conductor aspect. Indeed, the total conductor is

$$\ll E_2^L(0.51, X)^{r_{\mathrm{g}}^{\downarrow}} \cdot E_2^L(0.13, X)^{n_\psi B} \cdot 8^{\sqrt{m}} \cdot E_2^L(0.13, X)^{\sqrt{m} n_\psi B} \ll E_2^L(0.52, X),$$

where here we noted that

$$\begin{aligned}\log\log E_2^L(0.13, X)^{\sqrt{m} n_\psi B} &= \log(\sqrt{m} n_\psi B) + \log\log E_2^L(0.13, X)\\ &= \log\sqrt{m} + 0.13 \log\log X + O(\log\log\log X)\\ &\leq B^{n_\psi} \log 2 + 0.13 \log\log X + O(\log\log\log X) \lesssim 0.131 \log\log X\end{aligned}$$

since $B^{n_\psi} \leq (\log\log X)/999^2$.

We then apply the Chebotarev density theorem, giving a bound for the character sums corresponding to the irreducible representations and get a bound of

$$\ll U/U^{1/\sqrt{P_{\mathrm{s}}}} + U\exp\Bigl(-\frac{(\log U)/m^6}{\sqrt{\log U} + 3\log E_2^L(0.52, X)}\Bigr) \cdot \log\big(UmE_2^L(0.52, X)\big) \cdot m^{3/2}$$

where U is the size of primes in $T_{\mathrm{s}_\psi}(\mathcal{K})$. Since $m^6 \leq 64 \cdot 4^{6B^2} \ll (\log X)^{1/10^5}$ and $U \gg E_2^L(0.62, X)$, we thus obtain equi-distribution of the Frobenius classes for $L_{\check{Z}}/\mathbf{Q}$, which then naturally gives the result for $L_{\check{Z}}/K_{\check{Z}}$. Specifically, we obtain

$$\#T^{\tau}_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^{\downarrow}, \check{Z}] = \frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})}{2^{r_{\mathrm{g}}^{\downarrow} + n_\psi B} \cdot 2^{(B-1)^{n_\psi}}}\biggl[1 + O\Bigl(\frac{1}{E_2^L(0.09, X)}\Bigr)\biggr].$$

Combining this with the previous estimate then gives the Lemma. □

8.1.3. We next turn to results for equi-distribution with respect to $t^\uparrow$, which will only hold on average for almost all selections therein.

We recall that $r^\uparrow_{\mathrm{g}}$ is the number of indices in $\mathcal{I}^\star_\uparrow$, and is thus the number of components of the Cartesian product $T^\uparrow$. Meanwhile $T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]$ is as in §6.6.1: the $\mathcal{L}$-compatible subset of $T^\uparrow(\mathcal{K})$ with respect to itself, $t^\downarrow$, and all members of $\check{Z}$.

Lemma 8.1.4. *Suppose that $\bar{T}$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K},\mathcal{L})$ is a generic residue and Legendre specification. Let $t^\downarrow$ be an element of $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$, and $\check{Z}$ be a grid that is $\mathcal{L}$-compatible (with respect to $t^\downarrow$), with components of size B for the n_ψ well-gapped lower hypercube co-ordinates, with s_ψ the upper hypercube co-ordinate. Then for every $\tau\in\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ we have (with $E^L_2(u,X)=\exp\exp(u\log\log X)$ and $r^\uparrow_{\mathrm{g}}$ the number of indices in $\mathcal{I}^\star_\uparrow$)*

$$\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\left|\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]-\frac{\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{r^\uparrow_{\mathrm{g}}}}\right|\ll\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E^L_2(0.47,X)}.$$

Proof. As with various other proofs, we will show this by computing the average and mean value, and applying Cauchy's inequality. Since $\bar{T}$ is very pleasant, by regularity we have that the primes forming $t^\uparrow$ are $\geq E^L_2(0.49,X)$; ergo we can exclusively employ the bilinear estimate (1) for sums with $(t^\uparrow_j|t_{\mathrm{s}_\psi})$. We have

$$\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]=\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\;\sum_{t_{\mathrm{s}_\psi}\in T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}\mathcal{W}(t^\uparrow,t_{\mathrm{s}_\psi},\mathcal{L})$$

where

$$\mathcal{W}(t^\uparrow,t_{\mathrm{s}_\psi},\mathcal{L})=\frac{1}{2^{r^\uparrow_{\mathrm{g}}}}\prod_{j\in\mathcal{I}^\star_\uparrow}\left[1+\mathcal{L}_{j,\mathrm{s}_\psi}(t^\uparrow_j|t_{\mathrm{s}_\psi})\right]$$

with $t^\uparrow_j$ the components of $t^\uparrow$ and $r^\uparrow_{\mathrm{g}}$ the number of indices in $\mathcal{I}^\star_\uparrow$. Multiplying out the product gives a main term of

$$\#T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]\cdot\frac{\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{r^\uparrow_{\mathrm{g}}}}.$$

There are $(2^{r^\uparrow_{\mathrm{g}}}-1)$ other terms; each has a term $(t_a|t_{\mathrm{s}_\psi})$ for some $a\in\mathcal{I}^\star_\uparrow$, and we can bound each by the bilinear estimate (1) (using suitable α,β) to get an error of

$$\begin{aligned}&\ll\frac{2^{r^\uparrow_{\mathrm{g}}}-1}{2^{r_{\mathrm{g}}}}\prod_{j\neq a}\sum_{t_j\in T_j(\mathcal{K})}\left|\sum_{t_a\in T_a(\mathcal{K})}\;\sum_{t_{\mathrm{s}_\psi}\in T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}\alpha_{t_a}\beta_{t_{\mathrm{s}_\psi}}(t_a|t_{\mathrm{s}_\psi})\right|\\&\ll\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})\frac{(C\log X)^2}{\exp\big(\exp(0.49\log\log X)/9\big)}\end{aligned}$$

where $C\sim(\log\log X)^{99}$ is the compression factor of our basic intervals and we bounded $\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]\leq\#T_{\mathrm{s}_\psi}(\mathcal{K})$. This gives us

$$\begin{aligned}&\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K},\mathcal{L})[t^\downarrow,Z,t^\uparrow]\\&\qquad=\#T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]\cdot\frac{\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K},\mathcal{L})[t^\downarrow,\check{Z}]}{2^{r^\uparrow_{\mathrm{g}}}}+O\Big(\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E^L_2(0.48,X)}\Big),\end{aligned}$$

8.1.5. We then do the same for the mean-square value, which is

$$\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\ \sum\sum_{t^{(1)}_{\mathrm{s}_\psi},t^{(2)}_{\mathrm{s}_\psi}\in T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]} \mathcal{W}'(t^\uparrow,t^{(1)}_{\mathrm{s}_\psi},t^{(2)}_{\mathrm{s}_\psi},\mathcal{L})$$

where

$$\mathcal{W}'(t^\uparrow,t^{(1)}_{\mathrm{s}_\psi},t^{(2)}_{\mathrm{s}_\psi},\mathcal{L})=\frac{1}{(2^{r^\uparrow_{\mathrm{g}}})^2}\prod_{j\in\mathcal{I}^\star_\uparrow}\bigl[1+\mathcal{L}_{j,\mathrm{s}_\psi}(t^\uparrow_j|t^{(1)}_{\mathrm{s}_\psi})\bigr]\bigl[1+\mathcal{L}_{j,\mathrm{s}_\psi}(t^\uparrow_j|t^{(2)}_{\mathrm{s}_\psi})\bigr].$$

In the same manner as above we get

$$\begin{aligned}&\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K},\mathcal{L})[t^\downarrow,Z,t^\uparrow]^2\\&\qquad=\#T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]\cdot\frac{\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K},\mathcal{L})[t^\downarrow,\check{Z}]^2}{4^{r^\uparrow_{\mathrm{g}}}}+O\Bigl(\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})^2}{E^L_2(0.48,X)}\Bigr),\end{aligned}$$

and applying Cauchy's inequality then gives the Lemma. □

Corollary 8.1.6. *For every $\tau\in\mathrm{Gal}(L_{\check{Z}}/K_{\check{Z}})$ we have*

$$\sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\biggl|\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]-\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]}{2^{(B-1)^{n_\psi}}}\biggr|\ll\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E^L_2(0.09,X)}.$$

Proof. This follows from the previous two Lemmata. The triangle inequality implies the sum in question is bounded by

$$\begin{aligned}&\sum_{t^\uparrow}\biggl|\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]-\frac{\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{r^\uparrow_{\mathrm{g}}}}\biggr|\\&\qquad+\frac{1}{2^{r^\uparrow_{\mathrm{g}}}}\sum_{t^\uparrow}\biggl|\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]-\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{(B-1)^{n_\psi}}}\biggr|\\&\qquad\qquad+\frac{1}{2^{(B-1)^{n_\psi}}}\sum_{t^\uparrow}\biggl|\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]}{2^{r^\uparrow_g}}-\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]\biggr|,\end{aligned}$$

and Lemma 8.1.4 bounds the first sum, with then Lemma 8.1.1 used for the second, while upon summing Lemma 8.1.4 over τ (or arguing directly as in its proof) we find it bounds the third sum. □

8.2. We now accumulate the results of the previous sections. We assume we are given $t^\downarrow$ and $\check{Z}$ that are $\mathcal{L}$-compatible, and will show cancellation for $\psi\bigl(\mathbf{R}^8(\vec{t})\bigr)$ when summing over $\vec{t}$ in

$$\mathcal{B}^\star(t^\downarrow\times\check{Z})=\bigl(t^\downarrow\times\check{Z}\times\prod_{j\in\mathcal{I}_\uparrow}T_j(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}]\bigr)\cap T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle,$$

where the final intersection ensures that everything is $\mathcal{L}$-compatible. In particular, we can write a sum over $\mathcal{B}^\star(t^\downarrow\times\check{Z})$ as

$$\sum_{\vec{t}\in\mathcal{B}^\star(t^\downarrow\times\check{Z})}=\sum_{\check{z}\in\check{Z}}\ \sum_{t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check{Z}]\!]}\ \sum_{t_{\mathrm{s}_\psi}\in T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z},t^\uparrow]}$$

with the condition that everything in $\check{Z}$ be $\mathcal{L}$-compatible with $t^\downarrow$ and itself, and similarly $t^\downarrow\in T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ to ensure its $\mathcal{L}$-compatibility in the first place.

Proposition 8.2.1. *Suppose that $\bar T$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K},\mathcal{L})$ is a generic residue and Legendre specification. Let $t^\downarrow$ be an element of $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$, and $\check Z$ be a grid that is $\mathcal{L}$-compatible (with respect to $t^\downarrow$), with n_ψ components of size $B=\lfloor\sqrt{\log\log X}/999\rfloor$ for the well-gapped lower hypercube co-ordinates, with s_ψ the upper hypercube co-ordinate. Writing $\mathbf{R}^8(\vec t)$ for the 8-rank pairing matrix for $\vec t$ in terms of a fixed basis for the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$, we have*

$$\sum_{\vec t\in\mathcal{B}^\star(t^\downarrow\times\check Z)}\psi\bigl(\mathbf{R}^8(\vec t)\bigr)\ll\frac{\#\mathcal{B}^\star(t^\downarrow\times\check Z)}{\sqrt B}+\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E_2^L(0.07,X)}.$$

Proof. We are considering the sum

$$\sum_{\vec t\in\mathcal{B}^\star(t^\downarrow\times\check Z)}\psi\bigl(\mathbf{R}^8(\vec t)\bigr)=\sum_{\check z\in\check Z}\sum_{t^\uparrow}\sum_{t_{\mathrm{s}_\psi}\in T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]}\psi\bigl(\mathbf{R}^8(t^\downarrow,\check z,t^\uparrow,t_{\mathrm{s}_\psi})\bigr).$$

We first remove the $t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check Z]\!]$ for which there is $\tau\in\mathrm{Gal}(L_{\check Z}/K_{\check Z})$ with

$$\left|\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]-\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]}{2^{(B-1)^{n_\psi}}}\right|\geq\frac{\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E_2^L(0.08,X)}.$$

By the above Corollary 8.1.6, these bad $t^\uparrow\in T^\uparrow_{\mathrm{b}}$ contribute no more than

$$\#\check Z\sum_{t^\uparrow\in T^\uparrow_{\mathrm{b}}}\#T_{\mathrm{s}_\psi}(\mathcal{K})\leq\#\check Z\cdot E_2^L(0.08,X)\sum_{t^\uparrow}\Bigl|(\cdots)\Bigr|\ll\frac{\#T^\uparrow(\mathcal{K})\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E_2^L(0.085,X)},$$

where the ellipsis indicates the same T-difference as in the previous display.

For each remaining $t^\uparrow\in T^\uparrow(\mathcal{K})[\![\mathcal{L}|t^\downarrow,\check Z]\!]$ we split up $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]$ into the τ-orbits for $\mathrm{Gal}(L_{\check Z}/K_{\check Z})$. We let $a=\min_\tau\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]$, and by the above removal of bad $t^\uparrow$ we see that each such τ-set is approximately this size. We choose subsets $\tilde T^\tau\subset T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]$ each of size a, and then re-partition $\bigcup_\tau\tilde T^\tau$ into sets A with one element from each $\tilde T^\tau$. (Thus, there will be a such sets A, each of size $\#\mathrm{Gal}(L_{\check Z}/K_{\check Z})$).

We then choose an ϵ-good map $\tilde g_0\in\mathrm{im}(\partial_{\check Z\times S})$, and the calculation in §3.2.1 shows that this is possible (for sufficiently large X) for $\epsilon=99/\sqrt B$.

8.2.2. For each A as above and for each $\tau\in\mathrm{Gal}(L_{\check Z}/K_{\check Z})$), we form the unique sequence $S_\tau(A)$ with $g_0(1,j)=\tau\cdot\mathrm{Frob}_{L_{\check Z}/K_{\check Z}}(s_j)$ for $s_j\in A$ for all $1\leq j\leq B$. Note that $S_\tau(A)$ may contain repeats, but in any case when we take the multiset-union over τ we get an uniform covering of A of multiplicity B. For good $t^\uparrow$ we then have

$$\sum_{\check z\in\check Z}\sum_{t_{\mathrm{s}_\psi}\in\bigcup_\tau\tilde T^\tau}\psi\bigl(\mathbf{R}^8(t^\downarrow,\check z,t^\uparrow,t_{\mathrm{s}_\psi})\bigr)=\sum_{\check z\in\check Z}\sum_A\sum_\tau\frac{1}{B}\sum_{s\in S_\tau(A)}\bar\psi\bigl((\check z,s)\bigr).$$

Now the point of making this sub-division is that we can apply Lemma 7.8.1 to the sum over $(\check z,s)$ here, and thus for good $t^\uparrow$ we find

$$\left|\sum_{\check z\in\check Z}\sum_{t_{\mathrm{s}_\psi}}\psi\bigl(\mathbf{R}^8(t^\downarrow,\check z,t^\uparrow,t_{\mathrm{s}_\psi})\bigr)\right|\ll\epsilon\cdot\#\check Z\#T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]+O\Bigl(\frac{\#\check Z\cdot\#T_{\mathrm{s}_\psi}(\mathcal{K})}{E_2^L(0.079,X)}\Bigr),$$

where the t_{s_ψ}-sum is over $T_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]$, and we accounted the error with the τ-equidistribution via $\#T^\tau_{\mathrm{s}_\psi}(\mathcal{K})[\mathcal{L}|t^\downarrow,\check Z,t^\uparrow]-\#\tilde T^\tau\leq 2\#T_{\mathrm{s}_\psi}(\mathcal{K})/E_2^L(0.08,X)$ for good $t^\uparrow$ (with summing over τ giving another factor $\ll 2^{(B-1)^{n_\psi}}\leq\log X$).

Upon summing over good $t^\uparrow \in T^\uparrow(\mathcal{K})\llbracket\mathcal{L}|t^\downarrow, \check{Z}\rrbracket$ and including the error induced from bad $t^\uparrow$, we conclude that for all $\mathcal{L}$-compatible $(t^\downarrow, \check{Z})$ we have

$$\sum_{\vec{t}\in\mathcal{B}^\star(t^\downarrow\times\check{Z})} \psi\big(\mathbf{R}^8(\vec{t})\big) \ll \epsilon \cdot \#\mathcal{B}^\star(t^\downarrow \times \check{Z}) + \frac{\#T^\uparrow(\mathcal{K}) \cdot \#T_{\mathrm{s}_\psi}(\mathcal{K})}{E_2^L(0.07, X)}.$$

With $\epsilon = 99/\sqrt{B}$ this then gives the Proposition. □

It is perhaps worth emphasizing that this is the place where the crucial cancellation comes from.

8.2.3. It will be convenient to refer to consider the Cartesian product in $B^\star(t^\downarrow \times \check{Z})$ when the intersection with $T(\mathcal{K}, \mathcal{L})\langle t^\downarrow\rangle$ is not present, so we define

$$\mathcal{B}(t^\downarrow \times \check{Z}) = t^\downarrow \times \check{Z} \times \prod_{j\in\mathcal{I}_\uparrow} T_j(\mathcal{K})[\mathcal{L}|t^\downarrow, \check{Z}].$$

9. The selection of the grids

We will now collate grids into a fairly uniform cover $\mathscr{Z}(t^\downarrow)$ of $\mathcal{L}$-compatible selections from $\big(\prod_{u\in\tilde{\mathcal{V}}_\psi} T_u(\mathcal{K})\big)\llbracket\mathcal{L}|t^\downarrow\rrbracket$ over the lower hypercube co-ordinates. We shall fix $t^\downarrow \in T^\downarrow(\mathcal{K})\llbracket\mathcal{L}\rrbracket$ and consider $\vec{t} \in T(\mathcal{K}, \mathcal{L})\langle t^\downarrow\rangle$. What we wish to show, and it will take us two sections to complete, is that with a beneficial selection $\mathscr{Z}(t^\downarrow)$ of grids the counting function

$$\begin{aligned}\Lambda_{t^\downarrow}(\vec{t}) &= \#\{\check{Z} : \check{Z} \in \mathscr{Z}(t^\downarrow) \mid \vec{t} \in \mathcal{B}(t^\downarrow \times \check{Z})\}\\ &= \#\{\check{Z} : \check{Z} \in \mathscr{Z}(t^\downarrow) \mid \vec{t} \in t^\downarrow \times \check{Z} \times \prod_{j\in\mathcal{I}_\uparrow} T_j(\mathcal{K})[\mathcal{L}|t^\downarrow, \check{Z}]\}\end{aligned}$$

will be fairly uniform in size, at least outside of a negligible set of $\vec{t} \in T(\mathcal{K}, \mathcal{L})\langle t^\downarrow\rangle$.

With a result like $\Lambda_{t^\downarrow}(\vec{t}) \approx U$ in hand, we will then have

$$\begin{aligned}\sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})} \psi\big(\mathbf{R}^8(\vec{t})\big) &= \sum_{t^\downarrow\in T^\downarrow(\mathcal{K})\llbracket\mathcal{L}\rrbracket} \ \sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \psi\big(\mathbf{R}^8(\vec{t})\big)\\ &\approx \frac{1}{U}\sum_{t^\downarrow} \ \sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t})\psi\big(\mathbf{R}^8(\vec{t})\big) = \frac{1}{U}\sum_{t^\downarrow} \sum_{\check{Z}\in\mathscr{Z}(t^\downarrow)} \ \sum_{\vec{t}\in\mathcal{B}^\star(t^\downarrow\times\check{Z})} \psi\big(\mathbf{R}^8(\vec{t})\big),\end{aligned}$$

with $\mathcal{B}^\star(t^\downarrow \times \check{Z})$ as in §8.2, and indeed the previous section showed adequate cancellation over the inner sum. (Note here we omitted the set of summation for $t^\downarrow$ on the second line, and will continue to use such shorthand below).

There are two aspects to this. First we will select the $\check{Z}$ so that they cover the $\mathcal{L}$-compatible subset of the Cartesian product over lower hypercube co-ordinates nearly uniformly; this is essentially a combinatorial argument that we give in this section. Secondly, we will show that such a selection indeed induces approximate Λ-equi-sizing when considering $\mathcal{B}^\star(t^\downarrow \times \check{Z})$. This will use the bilinear estimate (1) and primes in arithmetic progressions.

9.1. Consider a Cartesian product $H = \prod_j H_j$ of sets. What we wish to show is that if $W \subset H$ has enough density, then there is some product $\prod_j G_j \subset W$ with each $G_j \subset H_j$ of reasonable size. This is clear in the case when the number h of components is 1, and our application will only additionally need the case of $h = 2$; however, an inductive argument readily handles the general case, so we opt for a phrasing in such generality.[17] We follow Smith's presentation [15, Proposition 4.1].

Lemma 9.1.1. *Suppose that we have v sets V_j each of size n, and $W \subset \prod_j V_j$ has density δ. When*

$$l \le n(\delta/2^{v+1})^{l^{v-1}+l^{v-2}+\cdots+1}$$

there are subsets $G_j \subset V_j$ each of size l such that $\prod_j G_j \subset W$.

The appearance of l^{v-1} in the exponent here is the dominating factor, so we will essentially need $l^{v-1} \lesssim (\log n)/\log(2^{v+1}/\delta)$ for this assumption to be met.

Proof. We write $I(l,\delta,v) = (\delta/2^{v+1})^{l^v+l^{v-1}+\cdots+l}$. The result is trivial when $l = 0$. Similarly, when $l = 1$ there is such a subset when $\delta \ge 1/n^v$, and otherwise we have $1 > n(\delta/2^{v+1})^v$ so that the size-supposition is not met.

Our induction hypothesis (on v) will be that: for every Cartesian product $\prod_j V_j$ of v sets with $\#V_j = n$ and $W \subset \prod_j V_j$ of density $\ge \delta$, when $l \le nI(l,\delta,v)^{1/l}$ there are at least $I(l,\delta,v)\binom{n}{l}^v$ tuples of subsets $G_j \subset V_j$ with $\#G_j = l$ and $\prod_j G_j \subset W$. (The positivity of $I(l,\delta,v)$ then implies there is at least one such set). We let $N_l(W)$ be the number of tuples of such subsets, and aim to show $N_l(W) \ge I(l,\delta,v)\binom{n}{l}^v$.

The $v = 1$ case says that when $l \le n(\delta/4)$ there are at least $(\delta/4)^l\binom{n}{l}$ subsets $G_1 \subset W$ with $\#G_1 = l$. Indeed, since $\#W \ge \lceil \delta n \rceil$ the number of such subsets is $\binom{\#W}{l} \ge \binom{\lceil \delta n \rceil}{l} \ge \frac{(\delta n - l)^l}{l!} \ge \frac{(3\delta n/4)^l}{l!} \ge (3\delta/4)^l\binom{n}{l}$.

9.1.2. We let P be the subset of $p \in V_v$ such that $W \cap (V_1 \times \cdots \times V_{v-1} \times \{p\})$ has density at least $\delta/2$ in $V_1 \times \cdots \times V_{v-1}$.

The union of $W \cap (V_1 \times \cdots \times V_{v-1} \times \{p\})$ over all $p \notin P$ contains $\le n^{v-1}(\delta/2) \cdot n$ elements. Since W has at least density δ and thus $\ge \delta n^v$ elements, we thus see that there are $\ge \big[\delta n^v - (\delta/2)n^v\big]/n^{v-1} = (\delta/2)n$ elements in P.

Meanwhile, the number of subsets $A_j \subset V_j$ with $\#A_j = l$ for $1 \le j < v$ is $\binom{n}{l}^{v-1}$. For a given tuple a of such subsets, we denote by u_a the number of $p \in P$ such that $A_1 \times \cdots \times A_{v-1} \times \{p\} \subset W$, so that $N_l(W) \ge \sum_a \binom{u_a}{l}$.

To use this, for $p \in P$ we write $W' \times \{p\} = W \cap (V_1 \times \cdots \times V_{v-1} \times \{p\})$ so that W' has density at least $\delta/2$, and then we can apply the induction hypothesis to $V_1 \times \cdots \times V_{v-1}$ and W', with the same l and n, and (δ, v) replaced by $(\delta/2, v-1)$. We readily verify that

$$I(l,\delta,v) = (\delta/2^v)^{l^v+\cdots+l} \le (\delta/2^v)^{l^{v-1}+\cdots l} = I(l,\delta/2,v-1),$$

so that our assumption $l \le nI(l,\delta,v)^{1/l}$ implies $l \le nI(l,\delta/2,v-1)^{1/l}$. Thus for each $p \in P$ the induction hypothesis tells us there are at least $I(l,\delta/2,v-1)\binom{n}{l}^{v-1}$ tuples a with $A_1 \times \cdots \times A_{v-1} \subset W'$, with each $A_1 \times \cdots \times A_{v-1} \times \{p\}$ then being

[17] Such a result is perhaps somewhere in the combinatorics literature, but again it seems simpler to show it directly.

contained in W, so that

$$\sum_a u_a \geq \#P \cdot I(l,\delta/2,v-1)\binom{n}{l}^{v-1} \geq (\delta/2)n \cdot I(l,\delta/2,v-1)\binom{n}{l}^{v-1}.$$

Comparing this to $N_l(W) \geq \sum_a \binom{u_a}{l}$, we see that (in particular) when n is significantly larger than l there must be some u_a that exceed l, which then contribute nontrivially to the sum. More explicitly, Hölder's inequality says

$$\Bigl(\sum_a u_a\Bigr)^l \leq \Bigl(\sum_a 1^l\Bigr)^{l-1}\Bigl(\sum_a u_a^l\Bigr)$$

so that

$$\sum_a u_a^l \geq \Bigl((\delta/2)n \cdot I(l,\delta/2,v-1)\Bigr)^l \binom{n}{l}^{l(v-1)} \Big/ \binom{n}{l}^{(l-1)(v-1)}.$$

Now when $u_a \geq 2l$ we have $\binom{u_a}{l} \geq (u_a-l)^l/l! \geq (u_a/2)^l/l!$, so in any case we have $\binom{u_a}{l} \geq (u_a/2)^l/l! - l^l/l!$, implying

$$\sum_a \binom{u_a}{l} \geq \sum_a \Bigl(\frac{(u_a/2)^l}{l!} - \frac{l^l}{l!}\Bigr) \geq \frac{1}{l!}\Bigl[\bigl((\delta/4)n \cdot I(l,\delta/2,v-1)\bigr)^l - l^l\Bigr]\binom{n}{l}^{v-1}.$$

Since

$$\begin{aligned} I(l,\delta/2,v-1)^l(\delta/4)^l &= (\delta/2^{v+1})^{l^v+\cdots+l^2}(\delta/4)^l \\ &= (\delta/2^{v+1})^{l^v+\cdots+l^2+l}(2^{v+1}/4)^l = I(l,\delta,v)(2^v/2)^l \end{aligned}$$

we then get

$$N_l(W) \geq \sum_a \binom{u_a}{l} \geq \frac{1}{l!}\bigl[n^l(2^v/2)^l I(l,\delta,v) - l^l\bigr]\binom{n}{l}^{v-1},$$

and use the assumption $n^l I(l,\delta,v) \geq l^l$ and $v \geq 2$ and $l \geq 1$ to bound $N_l(W)$ as

$$\geq \frac{1}{l!}\bigl[n^l I(l,\delta,v)\bigr]\bigl[(2^v/2)^l - 1\bigr]\binom{n}{l}^{v-1} \geq \frac{n^l}{l!} I(l,\delta,v)\binom{n}{l}^{v-1} \geq I(l,\delta,v)\binom{n}{l}^{v}.$$

This verifies the induction hypothesis, and shows the Lemma. □

We then generalize this in a slightly more flexible version (allowing the sets V_j to be different sizes), and phrase it contrapositively (the form in which it is used).

Corollary 9.1.3. *Suppose that we have v sets V_j the smallest of which has cardinality n, and $W \subset \prod_j V_j$ contains no product $\prod_j G_j$ with $\#G_j = l$ for all $1 \leq j \leq v$. Then the density δ of W satisfies $\delta \leq 2^{v+1}/n^{1/l^v}$.*

Proof. First we note that the enlargement of the sets makes no difference, as we can find subsets V_j' all of size n such that the intersection $W \cap \prod_j V_j'$ has density at least δ in $\prod_j V_j'$. Also, when $l = 1$ we would need W to be empty, whence $\delta = 0$.

Since W contains no $\prod_j G_j$, by the Proposition we have $l(2^{v+1}/\delta)^{l^{v-1}+\cdots+1} \geq n$, which implies (the worst case for the latter inequality is when $\delta = 1$ and $v = 1$)

$$\log n \leq \log l + \bigl(l^{v-1} + \cdots + 1\bigr)\log(2^{v+1}/\delta) \leq l^v(\log 2^{v+1}/\delta).$$

The estimate on δ follows. □

In our case we will have $n \geq E_2^L(0.08, X)$ and $l^v \leq \log\log X$ and $v \leq 2$, so that

$$\delta \leq 8/\exp\exp(0.08\log\log X)^{1/\log\log X} \ll 1/E_2^L(0.07, X).$$

9.2. We will now take a selection of grids that will cover the $\mathcal{L}$-compatible subset of $\prod_{u\in\tilde{V}_\psi} T_u(\mathcal{K})$ fairly uniformly, and whose intersections are all at most of size 1.

We let $\bar{R} = \lfloor E_2^L(0.05, X)\rfloor = \lfloor \exp\exp(0.05\log\log X)\rfloor$, and this will be the multiplicity to which we (nearly) cover the $\mathcal{L}$-compatible subset of $\prod_u T_u(\mathcal{K})$. In particular, since $\bar{T}$ is very pleasant the $T_u(\mathcal{K})$ for the lower hypercube co-ordinates u have cardinality $\geq E_2^L(0.08, X)$, which ensures this is much larger than $\bar{R}$. Meanwhile, all of the grids will be of size $B^{n_\psi} \leq B^2 \leq (\log\log X)/999^2$.

9.2.1. We take $t^\downarrow \in T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ and define $Y_{t^\downarrow}$ as $\bigl(\prod_{u\in\tilde{\mathcal{V}}_\psi} T_u(\mathcal{K})\bigr)[\![\mathcal{L}|t^\downarrow]\!]$, which is thus the $\mathcal{L}$-compatible subset of $\prod_u T_u(\mathcal{K})$ therein[18] with respect to $t^\downarrow$. Recall that a grid (or perhaps a B-grid if we want to emphasize the size) is a subset $\check{Z} \subset Y_{t^\downarrow}$ such that $\check{Z} = \prod_u \check{Z}^u$ with each $\#\check{Z}^u = B$.

We make a selection $\mathcal{Z}(t^\downarrow)$ of grids so that:[19] every $y \in Y_{t^\downarrow}$ is in at most $\bar{R}$ of the grids; and $\#(\check{Z}_i \cap \check{Z}_j) \leq 1$ for $i \neq j$. We also want $\mathcal{Z}(t^\downarrow)$ to be maximal with these properties, in that there is no grid that can be appended which preserves them.[20]

9.2.2. The point here is that maximality of $\mathcal{Z}(t^\downarrow)$ implies that almost all $y \in Y_{t^\downarrow}$ appear in exactly $\bar{R}$ grids. Indeed, let us define $R_{\mathcal{Z}}(y)$ to be the number of grids in which $y \in Y_{t^\downarrow}$ appears. (Recall that $E_2(u, X) = \exp\exp(u\log\log X)$).

Proposition 9.2.3. *Suppose that $\mathcal{Z}(t^\downarrow)$ is a maximal $\bar{R}$-selection of grids. Then the set of $y \in Y_{t^\downarrow}$ such that $R_{\mathcal{Z}}(y) \neq \bar{R}$ has density $\ll 1/E_2(0.07, X)$ in $\prod_{u\in\tilde{\mathcal{V}}_\psi} T_u(\mathcal{K})$.*

Proof. Given an $\bar{R}$-selection $\mathcal{Z}(t^\downarrow)$ of grids we define

$$\tilde{Y}_{t^\downarrow}(\mathcal{Z}) = \{y \in Y_{t^\downarrow} \mid y \text{ is in less than } \bar{R} \text{ of the grids in } \mathcal{Z}(t^\downarrow)\},$$

and let λ be the density of $\tilde{Y}_{t^\downarrow}(\mathcal{Z})$ in $\prod_u T_u(\mathcal{K})$.

We then take a maximal $W \subseteq \tilde{Y}_{t^\downarrow}(\mathcal{Z})$ with $\#(W \cap \check{Z}) \leq 1$ for all grids $\check{Z} \in \mathcal{Z}(t^\downarrow)$. The maximality of W implies that is has reasonable density. Indeed, the union U of grids that contain some element of W has cardinality $\leq \#W \cdot \bar{R}B^{n_\psi}$, and if this is smaller than $\#\tilde{Y}_{t^\uparrow}(\mathcal{Z}) = \lambda\#\prod_u T_u(\mathcal{K})$ then W is not maximal (any element of $\tilde{Y}_{t^\uparrow}(\mathcal{Z})\backslash U$ is appendable). Thus the density of W in $\prod_u T_u(\mathcal{K})$ is $\geq \lambda/\bar{R}B^{n_\psi}$.

Furthermore, if there are sets $Z^u \subset T_u(\mathcal{K})$ for $u \in \tilde{\mathcal{V}}_\psi$ such that all $\#Z^u = B$ and $\prod_u Z^u \subset W$, then $\mathcal{Z}(t^\downarrow)$ is not maximal, as we can append $\prod_u Z^u$ to it (since $W \subseteq \tilde{Y}_{t^\downarrow}(\mathcal{Z})$ each $y \in Y_{t^\downarrow}$ will still be in at most $\bar{R}$ grids, while the intersection property will be retained since $\#(W \cap \check{Z}) \leq 1$ for all grids $\check{Z}$).

Thus we can apply the above Corollary 9.1.3 to $W \subseteq \prod_u T_u(\mathcal{K})$ and obtain

$$\lambda/\bar{R}B^{n_\psi} \leq 2^{n_\psi+1}/E_2^L(0.08, X)^{1/\log\log X},$$

so that $\lambda \ll 1/E_2^L(0.07, X)$. □

[18] Both [14, page 80] and [4, page 41] work in the Cartesian product $\prod_u T_u(\mathcal{K})[\mathcal{L}|t^\downarrow]$ rather than our $\prod_u T_u(\mathcal{K})$. I don't particularly see the reason for their choice, as one has to contend (when $n_\psi > 1$) with $\mathcal{L}$-information amongst lower hypercube co-ordinates anyway. Moreover, they then have to verify that each $T_u(\mathcal{K})[\mathcal{L}|t^\downarrow]$ is large, which can only be done on average over $t^\downarrow$.

[19] It is also convenient word-wise to ensure that each $y \in Y_{t^\downarrow}$ appears in at least one grid in $\mathcal{Z}(t^\downarrow)$, so that the "union of the grids" is indeed $Y_{t^\downarrow}$.

[20] The notation $\mathcal{Z}(t^\downarrow)$ could contain $\mathcal{L}$ to note $\mathcal{L}$-compatibility, but I chose to suppress this.

10. Applying grids

Given $t^\downarrow \in T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ with $\mathcal{Z}(t^\downarrow)$ chosen as a maximal $\bar{R}$-selection of grids as in the previous section, we will now show that

$$\Lambda_{t^\downarrow}(\vec{t}) = \#\{\check{Z} : \check{Z} \in \mathcal{Z}(t^\downarrow) \mid \vec{t} \in \mathcal{B}(t^\downarrow \times \check{Z})\}$$

is near its expected size, at least outside of a negligible set of $\vec{t}$. Again the notation might have this be $\Lambda_{\mathcal{Z}}(\vec{t})$, but I've chosen to emphasize the dependence on $t^\downarrow$. Similarly, as we have now fixed a maximal $\bar{R}$-selection $\mathcal{Z}(t^\downarrow)$ of grids, we will refer to $R_{t^\downarrow}(y)$ instead of $R_{\mathcal{Z}}(y)$ for the number of grids that y appears in.

As with the proof of various other Lemmata, the main idea is to compute the average and mean square, and the sums in question can be estimated by either the bilinear estimate or primes in arithmetic progressions (we have no further need of the Chebotarev estimate, as it was only used for the upper hypercube co-ordinate).

Our arrangement of the arguments seems able to remove the need for the notion of "extravagant spacing" introduced by Smith.

10.1. I find it more convenient to first give a Lemma that shows that the size of $T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle$ is approximately reduced by the correct power-of-2 when the T_j for $j > r_{\mathrm{g}}^\uparrow$ are restricted by the full $\check{Z}$ information (rather than just for one y-value therein), as per the definition of $\mathcal{B}(t^\downarrow \times \check{Z})$.

Note that the analogue of this argument for Smith (top page 81) works on each upper co-ordinate individually, comparing the sizes of the T_j-restrictions. This is then emulated by Chan, Koymans, Milovic, and Pagano [4, (6.9ff)]. I cannot say whether this is a real technical advantage for us or not.

In any case, these authors also do not take an average over $t^\downarrow$ and $\check{Z}$, and doing so is what allows us to avoid extravagant spacing.

10.1.1. We recall $E_2^L(u,X) = \exp\exp(u \log\log X)$ and $\bar{R} = \lfloor E_2(0.05,X)\rfloor$, while

$$\mathcal{B}(t^\downarrow \times \check{Z}) = t^\downarrow \times \check{Z} \times \prod_{j\in\mathcal{I}_\uparrow} T_j(\mathcal{K})[\mathcal{L}|t^\downarrow,\check{Z}],$$

and that (as a shorthand) sums over $t^\downarrow$ range over $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$.

Lemma 10.1.2. *Suppose that $\bar{T}$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K},\mathcal{L})$ a generic residue and Legendre specification, and for each $t^\downarrow \in T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ we have a maximal $\bar{R}$-selection $\mathcal{Z}(t^\downarrow)$ of grids that each have n_ψ components of size B, with $Y_{t^\downarrow}$ the union of such grids, so that $Y_{t^\downarrow} = \big(\prod_{u\in\tilde{\mathcal{V}}_\psi} T_u(\mathcal{K})\big)[\![\mathcal{L}|t^\downarrow]\!]$. Then we have*

$$\sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum \left[\#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle \cap \mathcal{B}(t^\downarrow\times\check{Z})\big) - \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}}\right] \ll \frac{\#T(\mathcal{K})}{E_2^L(0.21,X)}$$

where $r_{\mathrm{g}}^{\Uparrow} = \#\mathcal{I}_\uparrow$ is the number of upper indices (including s_ψ). Furthermore,

$$\sum_{t^\downarrow}\sum_{\substack{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow)\\ \check{Z}_1\neq\check{Z}_2\\ y\in Y_{t^\downarrow}\cap(\check{Z}_1\cap\check{Z}_2)}}\sum\sum \left[\#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle\cap\mathcal{B}(t^\downarrow\times\check{Z}_1)\cap\mathcal{B}(t^\downarrow\times\check{Z}_2)\big) - \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle}{2^{2n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}}\right]$$

is similarly $\ll \#T(\mathcal{K})/E_2^L(0.21,X)$. (Note that these estimates do not have absolute values on the summands).

One might prefer to put a multiplier such as $\bar{R}/2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}$ in the error, but since $E_2^L(0.21, X)$ is so much larger than this, it does not matter. We can also note that [4, (6.9ff)] only saves $1/(\log X)^{\tilde{c}}$, as their usage of primes in arithmetic progressions can have a rather large modulus compared to the size of the primes (with these separated by only the "extravagant spacing" gap).

Proof. We will show that

$$\sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum \#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle \approx \frac{1}{2^v}\sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum\prod_{j\in\mathcal{I}_\uparrow}\#T_j(\mathcal{K}) \tag{2}$$

where $v = \binom{r_{\mathrm{g}}^{\Uparrow}}{2} + (r_{\mathrm{g}}^{\downarrow} + n_\psi)r_{\mathrm{g}}^{\Uparrow}$ with $r_{\mathrm{g}}^{\downarrow}$ the number of indices in $\mathcal{I}_\downarrow^\star$, while

$$\sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum \#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle\cap\mathcal{B}(t^\downarrow\times\check{Z})\big) \approx \frac{1}{2^w}\sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum\prod_{j\in\mathcal{I}_\uparrow}\#T_j(\mathcal{K})$$

where $w = \binom{r_{\mathrm{g}}^{\Uparrow}}{2} + (r_{\mathrm{g}}^{\downarrow} + n_\psi B)r_{\mathrm{g}}^{\Uparrow}$, and this will give the first result.

Furthermore, we will show

$$\sum_{t^\downarrow}\sum_{\substack{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap(\check{Z}_1\cap\check{Z}_2)}}\sum\sum \#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle\cap\mathcal{B}(t^\downarrow\times\check{Z}_1)\cap\mathcal{B}(t^\downarrow\times\check{Z}_2)\big)$$

$$\approx \frac{1}{2^b}\sum_{t^\downarrow}\sum_{\substack{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap(\check{Z}_1\cap\check{Z}_2)}}\sum\sum\prod_{j\in\mathcal{I}_\uparrow}\#T_j(\mathcal{K}),$$

where $b = \binom{r_{\mathrm{g}}^{\Uparrow}}{2} + \big(r_{\mathrm{g}}^{\downarrow} + n_\psi(2B-1)\big)r_{\mathrm{g}}^{\Uparrow}$, and this will then give the second result.

10.1.3. We begin by noting that $(t^\downarrow\times y)$ fixes all the lower indices, and that

$$\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle = \frac{1}{2^v}\prod_{j\in\mathcal{I}_\uparrow}\sum_{t_j\in T_j(\mathcal{K})}\mathcal{W}_y(\vec{t})$$

where $v = \binom{r_{\mathrm{g}}^{\Uparrow}}{2} + (r_{\mathrm{g}}^{\downarrow} + n_\psi)r_{\mathrm{g}}^{\Uparrow}$ and $\vec{t}$ is a shorthand for $t^\downarrow$, y, and the t_j, with

$$\mathcal{W}_y(\vec{t}) = P(\mathcal{I}_\uparrow,\mathcal{L},\vec{t})\cdot\prod_{j\in\mathcal{I}_\uparrow}\bigg(\prod_{u\in\check{\mathcal{V}}_\psi}\big[1+\mathcal{L}_{uj}(y_u|t_j)\big]\prod_{i\in\mathcal{I}_\downarrow^\star}\big[1+\mathcal{L}_{ij}(t_i|t_j)\big]\bigg)$$

where y_u denotes a component of y, and

$$P(\mathcal{I}_\uparrow,\mathcal{L},\vec{t}) = \prod_{j_1\in\mathcal{I}_\uparrow}\prod_{\substack{j_2\in\mathcal{I}_\uparrow\\ j_2>j_1}}\big[1+\mathcal{L}_{j_1j_2}(t_{j_1}|t_{j_2})\big]$$

is the Legendre product over the upper indices $\mathcal{I}_\uparrow$. Indeed, this latter product ensures the primes from the upper indices are $\mathcal{L}$-compatible amongst themselves, and the rest of the product ensures they are compatible with $t^\downarrow$ and y.

Thus our target sum is

$$A = \sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum \#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\times y\rangle = \sum_{t^\downarrow}\sum_{\substack{\check{Z}\in\mathcal{Z}(t^\downarrow)\\ y\in Y_{t^\downarrow}\cap\check{Z}}}\sum\frac{1}{2^v}\prod_{j\in\mathcal{I}_\uparrow}\sum_{t_j\in T_j(\mathcal{K})}\mathcal{W}_y(\vec{t}).$$

We then ease the sum over $\check{Z}$ by instead summing over a B-fold copy of the Cartesian product $\prod_u T_u(\mathcal{K})$ over the lower hypercube co-ordinates. This gives us

$$A = \frac{1}{2^v} \sum_{t^\downarrow} \prod_{u \in \tilde{\mathcal{V}}_\psi} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \mathcal{H}_{t^\downarrow}(\textstyle\prod_u t_u^B) \displaystyle\sum_{y \in \prod_u t_u^B} \prod_{j \in \mathcal{I}_\uparrow} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{W}_y(\vec{t})$$

where $T_u(\mathcal{K})^B_<$ is the set of strictly increasingly ordered B-tuples from $T_u(\mathcal{K})$, with t_u^B thus such a B-tuple, and $\prod_u t_u^B$ forms a concatenation of this over $u \in \tilde{\mathcal{V}}_\psi$. Meanwhile $\mathcal{H}_{t^\downarrow}(\prod_u t_u^B)$ counts the number of times (either 0 or 1) that $\prod_u t_u^B$ appears as a grid in $\mathcal{Z}(t^\downarrow)$. In essence, we have re-written the grid $\check{Z}$ as $\prod_u t_u^B$; but as many of the latter do not correspond to grids,[21] we then detect this via $\mathcal{H}$. The notation in the sum over $y \in \prod_u t_u^B$ refers to selecting one element of t_u^B for each u, so that y is an n_ψ-tuple, indeed in $\prod_u T_u(\mathcal{K})$.

10.1.4. We then expand the products in $\mathcal{W}_y$. Every multiplicand therein corresponds to one of:[22] a pair (m,n) with $r_{\mathrm{g}} < m < n$; a pair (l,n) with $l \le r_{\mathrm{g}} < n$ and $l \notin \tilde{\mathcal{V}}_\psi$; or a pair (u,n) with $u \in \tilde{V}_\psi$ and $n > r_{\mathrm{g}}$, with then y_u appearing in $\mathcal{W}_y$ from this. With $\mathcal{S}(y)$ as the set of such pairs (the y-dependence here is tangential, occurring in y_u), we are then summing over nonempty subsets of this for the error estimate, while the main term is indeed given as in (2). More specifically, we have

$$\begin{aligned} A &= \frac{1}{2^v} \sum_{t^\downarrow} \prod_{u \in \tilde{\mathcal{V}}_\psi} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \mathcal{H}_{t^\downarrow}(\textstyle\prod_u t_u^B) \displaystyle\sum_{y \in \prod_u t_u^B} \sum_{S \subseteq \mathcal{S}(y)} \prod_{j \in \mathcal{I}_\uparrow} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{U}_S(\vec{t}) \\ &= \frac{1}{2^v} \prod_{u \in \tilde{\mathcal{V}}_\psi} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \sum_{y \in \prod_u t_u^B} \sum_{S \subseteq \mathcal{S}(y)} \Biggl[\sum_{t^\downarrow} \mathcal{H}_{t^\downarrow}(\textstyle\prod_u t_u^B) \displaystyle\prod_{j \in \mathcal{I}_\uparrow} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{U}_S(\vec{t}) \Biggr] \end{aligned}$$

where (with m, l, u denoting types of elements as above)

$$\mathcal{U}_S(\vec{t}) = \prod_{(m,n) \in S} \mathcal{L}_{mn}(t_m|t_n) \prod_{(l,n) \in S} \mathcal{L}_{ln}(t_l|t_n) \prod_{(u,n) \in S} \mathcal{L}_{un}(y_u|t_n).$$

Suppose first that there is some $(m,n) \in S$ with $m > r_{\mathrm{g}}$. We then isolate these two variables, with the contribution $\mathcal{E}(S)$ bracketed above being

$$\sum_{t^\downarrow} \mathcal{H}_{t^\downarrow}(\textstyle\prod_u t_u^B) \displaystyle\prod_{\substack{j \in \mathcal{I}_\uparrow \\ j \ne m,n}} \sum_{t_j \in T_j(\mathcal{K})} \sum_{\substack{t_m \in T_m(\mathcal{K}) \\ t_n \in T_n(\mathcal{K})}}\sum \alpha_{t_m} \beta_{t_n} (t_m|t_n),$$

where α_{t_m} can contain various terms such as $(t_i|t_m)$, $(t_{j_1}|t_m)$, $(t_m|t_{j_2})$, and $(y_u|t_m)$; and similarly for β_{t_n}. We then use the bilinear estimate (1) on the inner double sum, and since $m > r_{\mathrm{g}}$, by regularity we have $\#T_m(\mathcal{K}) \gg E_2^L(0.49, X)$. Thus the inner double sum is bounded as $\ll (C \log X)^2 \#T_m(\mathcal{K}) \#T_n(\mathcal{K}) / E_2^L(0.49, X)^{1/9}$,

[21] It may seem wasteful to create a B-fold sum from the lower hypercube co-ordinates, but they are so much smaller than what we gain elsewhere that this does not matter.

[22] Although the notation is more complicated here than with the proof of [21, Proposition 5.4.1], the estimates are superior in the end, as every contributor to the error will have a term $(\cdot|t_j)$ with t_j large. Moreover, the immense gains easily outswamp the number of subsets, as recall that previously we were considering t's as small as $P_1 = \exp\exp\bigl((\log\log X)^{\eta_1}\bigr)$ or even $\exp\bigl((\log\log X)^{999}\bigr)$, while here the smallest is $\exp\exp(0.25 \log\log X)$.

where $C \sim (\log\log X)^{99}$ is the compression factor of our basic intervals. Summing over such S, this gives a contribution to A bounded as

$$\ll 2^{r^2} \prod_{u \in \tilde{\mathcal{V}}_\psi} \sum_{t_u^B \in T_u(\mathcal{K})_<^B} \sum_{y \in \prod_u t_u^B} \sum_{t^\downarrow} \frac{\prod_{j \in \mathcal{I}_\uparrow} \#T_j(\mathcal{K})}{E_2^L(0.48, X)}$$

$$\ll \frac{\#T(\mathcal{K})}{E_2^L(0.47, X)} \prod_{u \in \tilde{\mathcal{V}}_\psi} \#T_u(\mathcal{K})^B \ll \frac{\#T(\mathcal{K})}{E_2^L(0.46, X)},$$

where we used $\#T_u(\mathcal{K}) \ll E_2^L(0.13, X)$ for the lower hypercube co-ordinates. (One can estimate the t_u^B-sum with more sharply via $\mathcal{H}$, but there is no reason to do so, and in the next paragraph we do not so clearly have such an option).

Next, if there is some $(l, n) \in S$ such that the primes in T_l exceed $E_2^L(0.25, X)$ we can again use the bilinear estimate. It is a bit more complicated to reduce to an acceptable form, but here we have $\mathcal{E}(S)$ as

$$\sum_{t_{\hat{l}}^\downarrow \in T_{\hat{l}}^\downarrow(\mathcal{K})} \prod_{\substack{j \in \mathcal{I}_\uparrow \\ j \neq n}} \sum_{t_j \in T_j(\mathcal{K})} \sum_{\substack{t_l \in T_l(\mathcal{K}) \\ t_n \in T_n(\mathcal{K})}} \sum \mathcal{H}_{t^\downarrow}\bigl(\textstyle\prod_u t_u^B\bigr)\, \alpha_{t_l} \beta_{t_n} (t_l | t_n),$$

where $t_{\hat{l}}^\downarrow$ omits the lth co-ordinate. The point is that in the inner double sum $\mathcal{H}$ only depends on t_l and not t_n, so we can again use the bilinear estimate. Here also we have summed $t^\downarrow$ over $T^\downarrow(\mathcal{K})$ instead of $T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$, but the Legendre conditions can just be included in the coefficients (for instance, if $t^\downarrow$ is not $\mathcal{L}$-compatible, then we take $\mathcal{H}_{t^\downarrow}$ to be zero). The error contribution here is $\ll \#T(\mathcal{K})/E_2^L(0.24, X)$.

Finally, in the remaining case we can use primes in arithmetic progressions. There is some element (l, n) or (u, n) in S for some n, and we write $\mathcal{E}(S)$ as

$$\sum_{t^\downarrow} \mathcal{H}_{t^\downarrow}\bigl(\textstyle\prod_u t_u^B\bigr) \prod_{\substack{j \in \mathcal{I}_\uparrow \\ j \neq n}} \sum_{t_j \in T_j(\mathcal{K})} \sum_{t_n \in T_n(\mathcal{K})} (M_S | t_n),$$

where M_S is the product of t_l and y_u for (l, n) and (u, n) in S. We have that all t_l in M_S are $\ll E_2^L(0.26, X)$ (else the bilinear estimate would have been used), and $y_u \ll E_2^L(0.13, X)$ by regularity of the lower hypercube co-ordinates. On the other hand, the primes in T_n are $\gg E_2^L(0.49, X)$ by regularity at r_{g}. Thus with $U \gg E_2^L(0.49, X)$ being size of the primes in T_n, the inner sum is bounded by

$$U\left[\exp\left(\frac{-\log U}{\exp\bigl((\log\log X)^{\eta_s}\bigr)}\right) + \exp\left(\frac{-\tilde{c}\log U}{\sqrt{\log U} + 3\log M_S}\right)\right] \ll \frac{U}{E_2^L(0.22, X)},$$

where for the first term we used the pleasantness of $\bar{T}$ to bound the effect of exceptional zeros. This then gives an overall bound of $\ll \#T(\mathcal{K})/E_2^L(0.215, X)$.

Thus (2) holds with an acceptable error for the first statement of the Lemma.

10.1.5. We next note that for $y \in \check{Z}$ (else the intersection is empty) we have

$$\#T(\mathcal{K}, \mathcal{L})\langle t^\downarrow \times y\rangle \cap \mathcal{B}(t^\downarrow \times \check{Z}) = \frac{1}{2^w} \prod_{j \in \mathcal{I}_\uparrow} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{W}'_{\check{Z}}(\vec{t})$$

where $w = \binom{r_{\mathrm{g}}^\Uparrow}{2} + (r_{\mathrm{g}}^\downarrow + n_\psi B) r_{\mathrm{g}}^\Uparrow$ and

$$\mathcal{W}'_{\check{Z}}(\vec{t}) = P(\mathcal{I}_\uparrow, \mathcal{L}, \vec{t}) \cdot \prod_{j \in \mathcal{I}_\uparrow} \Biggl(\prod_{u \in \tilde{\mathcal{V}}_\psi} \prod_{\check{z}^u \in \check{Z}^u} \bigl[1 + \mathcal{L}_{uj}(\check{z}^u | t_j)\bigr] \cdot \prod_{i \in \mathcal{I}_\downarrow^\star} \bigl[1 + \mathcal{L}_{ij}(t_i | t_j)\bigr] \Biggr).$$

This includes the Legendre conditions for members of $\check{Z}$ besides just y.

Thus our target sum is

$$A' = \sum_{t^{\downarrow}} \sum_{\substack{\check{Z} \in \mathcal{Z}(t^{\downarrow}) \\ y \in Y_{t^{\downarrow}} \cap \check{Z}}} \#T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow} \times y \rangle = \sum_{t^{\downarrow}} \sum_{\substack{\check{Z} \in \mathcal{Z}(t^{\downarrow}) \\ y \in Y_{t^{\downarrow}} \cap \check{Z}}} \frac{1}{2^w} \prod_{j \in \mathcal{I}_{\uparrow}} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{W}'_{\check{Z}}(\vec{t}).$$

We again ease the sum over $\check{Z}$ by instead summing over a B-fold copy of the Cartesian product $\prod_u T_u(\mathcal{K})$ over the lower hypercube co-ordinates. This gives us

$$A' = \frac{1}{2^w} \sum_{t^{\downarrow}} \prod_{u \in \tilde{\mathcal{V}}_{\psi}} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \mathcal{H}_{t^{\downarrow}}(\textstyle\prod_u t_u^B) \displaystyle\sum_{y \in \prod_u t_u^B} \prod_{j \in \mathcal{I}_{\uparrow}} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{W}'_{\prod_u t_u^B}(\vec{t})$$

where

$$\mathcal{W}'_{\prod_u t_u^B}(\vec{t}) = P(\mathcal{I}_{\uparrow}, \mathcal{L}, \vec{t}) \cdot \prod_{j \in \mathcal{I}_{\uparrow}} \Bigg(\prod_{u \in \tilde{\mathcal{V}}_{\psi}} \prod_{w_u \in t_u^B} [1 + \mathcal{L}_{uj}(w_u | t_j)] \cdot \prod_{i \in \mathcal{I}_{\downarrow}^{\star}} [1 + \mathcal{L}_{ij}(t_i | t_j)] \Bigg).$$

We then expand the products in $\mathcal{W}'$. Every multiplicand therein corresponds to one of: a pair (m, n) with $r_{\mathrm{g}} < m < n$; a pair (l, n) with $l \le r_{\mathrm{g}} < n$ and $l \notin \tilde{\mathcal{V}}_{\psi}$; or a triple (w_u, u, n) with $u \in \tilde{V}_{\psi}$, $w_u \in t_u^B$, and $n > r_{\mathrm{g}}$. Letting $\mathcal{S}'(\prod_u t_u^B)$ be the set of such pairs and triples, we are then summing over nonempty subsets of this for the error estimate, while the main term is given above. More specifically, we have

$$\begin{aligned} A' &= \frac{1}{2^w} \sum_{t^{\downarrow}} \prod_{u \in \tilde{\mathcal{V}}_{\psi}} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \mathcal{H}_{t^{\downarrow}}(\textstyle\prod_u t_u^B) \displaystyle\sum_{y \in \prod_u t_u^B} \sum_{S \subseteq \mathcal{S}'(\prod_u t_u^B)} \prod_{j \in \mathcal{I}_{\uparrow}} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{U}'_S(\vec{t}) \\ &= \frac{1}{2^w} \prod_{u \in \tilde{\mathcal{V}}_{\psi}} \sum_{t_u^B \in T_u(\mathcal{K})^B_<} \sum_{y \in \prod_u t_u^B} \sum_{S \subseteq \mathcal{S}'(\prod_u t_u^B)} \Bigg[\sum_{t^{\downarrow}} \mathcal{H}_{t^{\downarrow}}(\textstyle\prod_u t_u^B) \displaystyle\prod_{j \in \mathcal{I}_{\uparrow}} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{U}'_S(\vec{t}) \Bigg] \end{aligned}$$

where

$$\mathcal{U}'_S(\vec{t}) = \prod_{(m,n) \in S} \mathcal{L}_{mn}(t_m | t_n) \prod_{(l,n) \in S} \mathcal{L}_{ln}(t_l | t_n) \prod_{(w_u, u, n) \in S} \mathcal{L}_{un}(w_u | t_n).$$

Again if there is $(m, n) \in S$ with $m > r_{\mathrm{g}}$ we can use the bilinear estimate on $(t_m | t_n)$. Similarly, if there is $(l, n) \in S$ where the primes in T_l are $\ge E_2^L(0.25, X)$ we can use the bilinear estimate on $(t_l | t_n)$. Finally, if there is no such pair as above, we can use primes in arithmetic progressions. Here the modulus might have more factors from the lower hypercube co-ordinates with $w_u \in t_u^B$, but as their number is bounded by B, this is harmless. We thus conclude the first part of the Lemma.

10.1.6. Finally, with $b = \binom{r_{\mathrm{g}}^{\Uparrow}}{2} + \big(r_{\mathrm{g}}^{\downarrow} + n_{\psi}(2B-1)\big) r_{\mathrm{g}}^{\Uparrow}$, for $y \in \check{Z}_1 \cap \check{Z}_2$ we have

$$\#T(\mathcal{K}, \mathcal{L})\langle t^{\downarrow} \times y \rangle \cap \mathcal{B}(t^{\downarrow} \times \check{Z}_1) \cap \mathcal{B}(t^{\downarrow} \times \check{Z}_2) = \frac{1}{2^b} \prod_{j \in \mathcal{I}_{\uparrow}} \sum_{t_j \in T_j(\mathcal{K})} \mathcal{W}''_{\check{Z}_1, \check{Z}_2}(\vec{t})$$

where $\mathcal{W}''_{\check{Z}_1, \check{Z}_2}(\vec{t})$ is $P(\mathcal{I}_{\uparrow}, \mathcal{L}, \vec{t})$ times

$$\prod_{j \in \mathcal{I}_{\uparrow}} \Bigg(\prod_{u \in \tilde{\mathcal{V}}_{\psi}} \prod_{\check{z}_1^u \in \check{Z}_1^u} [1 + \mathcal{L}_{uj}(\check{z}_1^u | t_j)] \prod_{\check{z}_2^u \in \check{Z}_2^u} [1 + \mathcal{L}_{uj}(\check{z}_2^u | t_j)] \cdot \prod_{i \in \mathcal{I}_{\downarrow}^{\star}} [1 + \mathcal{L}_{ij}(t_i | t_j)] \Bigg).$$

This includes Legendre conditions for members of $\check{Z}_1 \cup \check{Z}_2$ besides just y. Note that $\{y\} = \check{Z}_1 \cap \check{Z}_2$, and thus exactly one of the conditions is redundant for each j.

The proof then goes in the same manner as before, and we get the result. □

10.2. Next we show $\Lambda_{t^\uparrow}(\vec{t})$ is typically near the expected size. Its definition is

$$\begin{aligned}\Lambda_{t^\downarrow}(\vec{t}) &= \#\{\check{Z} : \check{Z} \in \mathcal{Z}(t^\downarrow) \mid \vec{t} \in \mathcal{B}(t^\downarrow \times \check{Z})\} \\ &= \#\big\{\check{Z} : \check{Z} \in \mathcal{Z}(t^\downarrow) \mid \vec{t} \in t^\downarrow \times \check{Z} \times \prod_{j \in \mathcal{I}_\uparrow} T_j(\mathcal{K})[\mathcal{L}|t^\downarrow, \check{Z}]\big\}.\end{aligned}$$

Lemma 10.2.1. *Suppose that $\bar{T}$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K}, \mathcal{L})$ is a generic residue and Legendre specification, while ψ is a nontrivial multiplicative character of $\mathrm{Mat}(e, e+1, \mathbf{F}_2)$ and $\mathcal{V}_\psi$ a choice of well-gapped hypercube co-ordinates. Suppose that with $\bar{R} = \lfloor E_2^L(0.05, X) \rfloor$ for each $t^\downarrow \in T^\downarrow(\mathcal{K})[\![\mathcal{L}]\!]$ we have a maximal $\bar{R}$-selection $\mathcal{Z}(t^\downarrow)$ of ($\mathcal{L}$-compatible) grids in $\prod_{u \in \tilde{\mathcal{V}}_\psi} T_u(\mathcal{K})$ with each grid having n_ψ components of size $B = \lfloor \sqrt{\log\log X}/999 \rfloor$. Then*

$$\sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow \rangle} \left| \Lambda_{t^\downarrow}(\vec{t}) - \frac{\bar{R}}{2^{n_\psi (B-1) r_{\mathrm{g}}^{\Uparrow}}} \right| \ll \sqrt{\bar{R}} \cdot \#T(\mathcal{K}),$$

with $r_{\mathrm{g}}^{\Uparrow} = \#\mathcal{I}_\uparrow$ the number of upper indices and $E_2^L(u, X) = \exp\exp(u \log\log X)$.

Let us note that $\bar{R}/2^{n_\psi (B-1) r_{\mathrm{g}}^{\Uparrow}}$ is indeed the expected size of $\Lambda_{t^\downarrow}(\vec{t})$. We have that $\vec{t} \in T(\mathcal{K}, \mathcal{L})\langle t^\downarrow \rangle$, and it has some components y from the lower hypercube co-ordinates. In particular, $\vec{t}$ satisfies the Legendre conditions with respect to y, and so is in $t^\downarrow \times y \times \prod_{j \in \mathcal{I}_\uparrow} T_j(\mathcal{K})[\mathcal{L}|t^\downarrow, y]$. Note also that y is in approximately $\bar{R}$ of the grids $\check{Z}$. The effect of including Legendre conditions from $\check{Z}$ (in the definition of Λ) should be to reduce this $\bar{R}$ by a factor of 2 for each j and each prime in $\check{Z}$ that is not in y. The number of such primes is $n_\psi(B-1)$, and the number of j is $r_{\mathrm{g}}^{\Uparrow}$.

Proof. Again we will compute the average and mean square.

10.2.2. We fix the lower hypercube co-ordinate components as y, signify this condition with $\pi_{\tilde{\mathcal{V}}_\psi}(\vec{t}) = y$, and thereby find the sum of $\Lambda_{t^\downarrow}(\vec{t})$ over $\vec{t}$ is

$$\begin{aligned}\sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow \rangle} \Lambda_{t^\downarrow}(\vec{t}) &= \sum_{y \in Y_{t^\downarrow}} \sum_{\substack{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow \rangle \\ \pi_{\tilde{\mathcal{V}}_\psi}(\vec{t}) = y}} \Lambda_{t^\downarrow}(\vec{t}) = \sum_{y \in Y_{t^\downarrow}} \sum_{\substack{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow \rangle \\ \pi_{\tilde{\mathcal{V}}_\psi}(\vec{t}) = y}} \sum_{\substack{\check{Z} \in \mathcal{Z}(t^\downarrow) \\ \vec{t} \in \mathcal{B}(t^\downarrow \times \check{Z})}} 1 \\ &= \sum_{y \in Y_{t^\downarrow}} \sum_{\check{Z} \in \mathcal{Z}(t^\downarrow)} \#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y \rangle \cap \mathcal{B}(t^\downarrow \times \check{Z})\big).\end{aligned}$$

Taking the sum of this over $t^\downarrow$, the previous Lemma 10.1.2 then tells us that

$$\sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow \rangle} \Lambda_{t^\downarrow}(\vec{t}) = \sum_{t^\downarrow} \sum_{\check{Z} \in \mathcal{Z}(t^\downarrow)} \sum_{y \in Y_{t^\downarrow} \cap \check{Z}} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y \rangle}{2^{n_\psi (B-1) r_{\mathrm{g}}^{\Uparrow}}} + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.21, X)}\Big).$$

Switching the order of y and $\check{Z}$, the definition of $R_{t^\downarrow}(y)$ says the main term is

$$\sum_{t^\downarrow} \sum_{y \in Y_{t^\downarrow}} R_{t^\downarrow}(y) \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y \rangle}{2^{n_\psi (B-1) r_{\mathrm{g}}^{\Uparrow}}},$$

and the error of replacing $R_{t^\downarrow}(y)$ by $\bar{R}$ here is estimated by Proposition 9.2.3 as

$$\ll \sum_{t^\downarrow} \bar{R} \sum_{y\in\tilde{Y}_{t^\downarrow}} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} \ll \bar{R} \sum_{t^\downarrow} \frac{\#T(\mathcal{K})\langle t^\downarrow\rangle}{E_2^L(0.07,X)} \ll \frac{\#T(\mathcal{K})}{E_2^L(0.06,X)}.$$

For the main term with $R_{t^\downarrow}(y)$ replaced by $\bar{R}$, executing the sum over $y \in Y_{t^\downarrow}$ removes it from the angle brackets, and so we conclude that

$$\sum_{t^\downarrow} \sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t}) = \bar{R} \sum_{t^\downarrow} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.21,X)}\Big) + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.06,X)}\Big).$$

10.2.3. For the mean square we have

$$\begin{aligned}\sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t})^2 &= \sum_{y\in Y_{t^\downarrow}} \sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle \\ \pi_{\tilde{\mathcal{V}}_\psi}(\vec{t})=y}} \sum\sum_{\substack{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow) \\ \vec{t}\in\mathcal{B}(t^\downarrow\times\check{Z}_1)\cap\mathcal{B}(t^\downarrow\times\check{Z}_2)}} 1 \\ &= \sum_{y\in Y_{t^\downarrow}} \sum\sum_{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow)} \#\big(T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle \cap \mathcal{B}(t^\downarrow\times\check{Z}_1)\cap\mathcal{B}(t^\downarrow\times\check{Z}_2)\big).\end{aligned}$$

Here we note that y must be in both $\check{Z}_1$ and $\check{Z}_2$ for the intersection to be nontrivial, and recall that $\#(\check{Z}_1\cap\check{Z}_2) \le 1$ for $\check{Z}_1 \ne \check{Z}_2$. Thus we split off the diagonal $\check{Z}_1 = \check{Z}_2$ contribution,[23] and upon including the sum over $t^\downarrow$ we estimate it as above as

$$\ll \sum_{t^\downarrow} \sum_{y\in Y_{t^\downarrow}} \bar{R} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.06,X)}\Big) \ll \bar{R}\cdot\#T(\mathcal{K}).$$

For $\check{Z}_1 \ne \check{Z}_2$ we then apply the second part of the previous Lemma 10.1.2 to get a contribution to the mean square of

$$\sum_{t^\downarrow} \sum\sum_{\substack{\check{Z}_1,\check{Z}_2\in\mathcal{Z}(t^\downarrow) \\ \check{Z}_1\ne\check{Z}_2}} \sum_{\substack{y\in Y_{t^\downarrow} \\ y\in\check{Z}_1\cap\check{Z}_2}} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle}{2^{2n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.21,X)}\Big).$$

For a given $y \in Y_{t^\downarrow}$ the number of $(\check{Z}_1,\check{Z}_2)$-pairs that have an intersection of $\{y\}$ is $R_{t^\downarrow}(y)^2 - R_{t^\downarrow}(y)$, so upon using $R_{t^\downarrow}(y) = \bar{R}$ away from $\tilde{Y}_{t^\downarrow}$ the main term here is

$$\begin{aligned}&\sum_{t^\downarrow} \sum_{y\in Y_{t^\downarrow}} \big[R_{t^\downarrow}(y)^2 - R_{t^\downarrow}(y)\big] \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle}{2^{2n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} \\ &\qquad = \bar{R}^2 \sum_{t^\downarrow} \sum_{y\in Y_{t^\downarrow}} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow \times y\rangle}{2^{2n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\Big(\frac{\bar{R}^2\cdot\#T(\mathcal{K})}{E_2^L(0.21,X)}\Big) + O\Big(\frac{\bar{R}^2\cdot\#T(\mathcal{K})}{E_2^L(0.07,X)}\Big).\end{aligned}$$

Putting these together and executing the y-sum then gives

$$\sum_{t^\downarrow} \sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t})^2 = \bar{R}^2 \sum_{t^\downarrow} \frac{\#T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle}{2^{2n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\big(\bar{R}\cdot\#T(\mathcal{K})\big).$$

[23] One reason to take $\bar{R}$ reasonably large is to ensure this diagonal contribution is negligible.

10.2.4. Combining the above results for the average and mean square, since we have $\bar{R} = E_2^L(0.05, X)$ while $n_\psi B r_{\mathrm{g}}^{\Uparrow} \ll (\log\log X)^{3/2}$, we obtain

$$\sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \left| \Lambda_{t^\downarrow}(\vec{t}) - \frac{\bar{R}}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} \right|^2 \ll \bar{R} \cdot \#T(\mathcal{K}),$$

and applying Cauchy's inequality gives the result. □

10.3. Now we fill in the details of the sketch at the start of §9, to conclude that we have ψ-cancellation on $T(\mathcal{K},\mathcal{L})$.

Proposition 10.3.1. *Suppose $\bar{T}$ is a very pleasant $\mathcal{P}_4$-box and $(\mathcal{K},\mathcal{L})$ a generic residue and Legendre specification. We write $\mathbf{R}^8(\vec{t})$ for the 8-rank pairing matrix for $\vec{t}$ in terms of a fixed basis for the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$, and ψ for a nontrivial multiplicative character of* $\mathrm{Mat}(e, e+1, \mathbf{F}_2)$. *For $B = \lfloor \sqrt{\log\log X}/999 \rfloor$ we have*

$$\sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})} \psi\big(\mathbf{R}^8(\vec{t})\big) \ll \frac{\#T(\mathcal{K},\mathcal{L})}{\sqrt{B}} + O\Big(\frac{\#T(\mathcal{K})}{E_2^L(0.03, X)}\Big)$$

Proof. We start by noting that we have

$$\frac{\bar{R}}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})} \psi\big(\mathbf{R}^8(\vec{t})\big) = \frac{\bar{R}}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} \sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \psi\big(\mathbf{R}^8(\vec{t})\big)$$

from specifying $t^\downarrow$, and then break this up as

$$\sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t})\psi\big(\mathbf{R}^8(\vec{t})\big) + \sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \left[\frac{\bar{R}}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} - \Lambda_{t^\downarrow}(\vec{t})\right] \psi\big(\mathbf{R}^8(\vec{t})\big)$$

The second term here is $\ll \sqrt{\bar{R}} \cdot \#T(\mathcal{K})$ by the preceding Lemma 10.2.1, while the first can be re-arranged and estimated by Proposition 8.2.1 as

$$\sum_{t^\downarrow} \sum_{\check{Z} \in \mathcal{Z}(t^\downarrow)} \sum_{\vec{t} \in \mathcal{B}^\star(t^\downarrow \times \check{Z})} \psi\big(\mathbf{R}^8(\vec{t})\big) \ll \frac{1}{\sqrt{B}} \sum_{t^\downarrow} \sum_{\check{Z} \in \mathcal{Z}(t^\downarrow)} \#\mathcal{B}^\star(t^\downarrow \times \check{Z}) + O\Big(\frac{\bar{R} \cdot \#T(\mathcal{K})}{E_2^L(0.07, X)}\Big).$$

Undoing this re-arrangement then gives a bound for the original expression as

$$\ll \frac{1}{\sqrt{B}} \sum_{t^\downarrow} \sum_{\vec{t} \in T(\mathcal{K},\mathcal{L})\langle t^\downarrow\rangle} \Lambda_{t^\downarrow}(\vec{t}) \ll \frac{1}{\sqrt{B}} \frac{\bar{R} \cdot \#T(\mathcal{K},\mathcal{L})}{2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}} + O\big(\sqrt{\bar{R}} \cdot \#T(\mathcal{K})\big)$$

where we estimated the double sum again by Lemma 10.2.1.

Dividing out by $\bar{R}/2^{n_\psi(B-1)r_{\mathrm{g}}^{\Uparrow}}$ then gives the result. □

11. Putting everything together

What we have shown is that, in favorable circumstances, every nontrivial multiplicative character ψ on $\mathrm{Mat}(e, e+1, \mathbf{F}_2)$ has cancellation for $\psi\big(\mathbf{R}^8(\vec{t})\big)$ when summed over $\vec{t} \in T(\mathcal{K},\mathcal{L})$. Let us review what this means, and then derive some consequences below.

11.1. As with §6.6, we have a very pleasant $\mathcal{P}_4$-box $\bar{T}$, a generic residue and Legendre specification $(\mathcal{K},\mathcal{L})$, and a basis of size $(e+1)$ for the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ that contains the obvious vector. There is then an 8-rank pairing matrix $\mathbf{R}^8(\vec{t})$ for $\vec{t}\in T(\mathcal{K},\mathcal{L})$, and we have shown that every nontrivial multiplicative character ψ on $\mathrm{Mat}(e,e+1,\mathbf{F}_2)$ has cancellation for the sum of $\psi\bigl(\mathbf{R}^8(\vec{t})\bigr)$ over $\vec{t}$.

By character duality this gives us equi-distribution of $\mathbf{R}^8(\vec{t})$ in its matrix space, and summing over $(\mathcal{K},\mathcal{L})$ then yields a result for boxes, which can then be summed over to finally get the result for Gaussian discriminants. We now carry this out.

11.1.1. First we sum over residue and Legendre specifications $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})$ with a given narrow 4-rank e, accounting for those that are non-generic (see §6.4).

Lemma 11.1.2. *Let $\bar{T}$ be a very pleasant $\mathcal{P}_4$-box, with $(f,\varepsilon)\in\{(0,+),(3,+)\}$ and $e\geq 0$ given. Then for any matrix $M\in\mathrm{Mat}(e,e+1,\mathbf{F}_2)$ we have*[24]

$$\sum_{\substack{(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})\\ e_\varepsilon^f(\mathcal{K},\mathcal{L})=e}}\ \sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\\ \mathbf{R}^8(\vec{t})=M}} 1=\frac{\gamma_{\mathrm{G}}(e)}{2^{e(e+1)}}\cdot\#T+O\Bigl(\frac{\#T}{(\log\log X)^{1/4}}\Bigr),$$

with $e_\varepsilon^f(\mathcal{K},\mathcal{L})$ the narrow 4-rank for the residue and Legendre specification $(\mathcal{K},\mathcal{L})$, and $\mathbf{R}^8$ the 8-rank pairing matrix for a fixed basis of the kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$, and

$$\gamma_{\mathrm{G}}(e)=\frac{1}{2^{e(e+1)/2}}\prod_{u=1}^{\infty}(1+1/2^u)^{-1}\prod_{j=1}^{e}(1-1/2^j)^{-1}.$$

Proof. We want to estimate

$$A=\sum_{\substack{(\mathcal{K},\mathcal{L})\\ e_\varepsilon^f(\mathcal{K},\mathcal{L})=e}}\ \sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\\ \mathbf{R}^8(\vec{t})=M}} 1-\frac{\gamma_{\mathrm{G}}(e)}{2^{e(e+1)}}\cdot\#T.$$

By our previous results for the narrow 4-rank in the case of Gaussian discriminants (see [21, §12.1.1]), for a pleasant box $\bar{T}$ we have

$$\gamma_{\mathrm{G}}(e)\cdot\#T=\sum_{\substack{(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})\\ e_\varepsilon^f(\mathcal{K},\mathcal{L})=e}}\#T(\mathcal{K},\mathcal{L})+O\Bigl(\frac{\#T}{(\log\log X)^{49/99}}\Bigr)$$

Thus we have

$$A=\sum_{\substack{(\mathcal{K},\mathcal{L})\\ e_\varepsilon^f(\mathcal{K},\mathcal{L})=e}}\Biggl[\sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\\ \mathbf{R}^8(\vec{t})=M}} 1-\frac{\#T(\mathcal{K},\mathcal{L})}{2^{e(e+1)}}\Biggr]+O\Bigl(\frac{\#T}{(\log\log X)^{49/99}}\Bigr),$$

and we can bound the effect of nongeneric $(\mathcal{K},\mathcal{L})$ as

$$\leq\sum_{(\mathcal{K},\mathcal{L})}^{\star}\Biggl|\sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\\ \mathbf{R}^8(\vec{t})=M}} 1-\frac{\#T(\mathcal{K},\mathcal{L})}{2^{e(e+1)}}\Biggr|\leq\sum_{(\mathcal{K},\mathcal{L})}^{\star}\#T(\mathcal{K},\mathcal{L})\ll\frac{\#T}{(\log\log X)^{49/99}},$$

where we used the estimates of Lemmata 6.4.3 and 6.4.6.

[24] One can be more weaselly and get $O\bigl(1/2^{e(e+1)}\sqrt{B}\bigr)+O_\omega\bigl(1/(\log\log X)^\omega\bigr)$ for any $\omega<1/2$ for the relative error; the slim advantages of this are for e between constant multiples of $\sqrt{\log\log\log X}$.

Meanwhile, for generic $(\mathcal{K},\mathcal{L})$ we use character duality and Proposition 10.3.1 for all nontrivial multiplicative characters ψ on $\mathrm{Mat}(e,e+1,\mathbf{F}_2)$ to get

$$\begin{aligned}\sum_{\substack{\vec{t}\in T(\mathcal{K},\mathcal{L})\\ \mathbf{R}^8(\vec{t})=M}} 1 &= \sum_{\psi}\frac{\psi(M)}{2^{e(e+1)}}\sum_{\vec{t}\in T(\mathcal{K},\mathcal{L})}\psi\bigl(\mathbf{R}^8(\vec{t})\bigr)\\ &= \frac{\#T(\mathcal{K},\mathcal{L})}{2^{e(e+1)}}+O\Bigl(\frac{\#T(\mathcal{K},\mathcal{L})}{\sqrt{B}}\Bigr)+O\Bigl(\frac{\#T(\mathcal{K})}{E_2^L(0.03,X)}\Bigr).\end{aligned}$$

We then sum this over generic $(\mathcal{K},\mathcal{L})$, noting that the number of $\mathcal{L}$ is $2^{\binom{\tilde{r}}{2}}$, so is dominated by $E_2^L(0.03,X)$. With $B=\lfloor\sqrt{\log\log X}/999\rfloor$ this gives the Lemma. □

We can then pass from boxes to Gaussian discriminants, upon accounting for boxes that are not very pleasant. Their effect is bounded by $\Phi^{\mathcal{P}}(X)/(\log X)^{\tilde{c}}$, as follows from the analysis in §5.1.3 combined with the (very pleasant) regularity accounting in Lemma 6.2.3. Thus upon summing over very pleasant boxes we get Theorem 1.1.2, namely that for any $M\in\mathrm{Mat}(e,e+1,\mathbf{F}_2)$ we have

$$\frac{\#\{d\le X : d\in\mathcal{G}\mid e_4(d)=e,\mathbf{R}^8(d)=M\}}{\#\{d\le X : d\in\mathcal{G}\}}=\frac{\gamma_{\mathrm{G}}(e)}{2^{e(e+1)}}+O\Bigl(\frac{1}{(\log\log X)^{1/4}}\Bigr).$$

12. Narrow 8-ranks of quadratic class groups

Now we consider the case of the narrow 8-rank of the class group of a quadratic field,[25] and explain the necessary modifications to §6 and the ensuing differences in the computations of §7. The rest of the argument will then follow as before.

The pairing matrix $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ is no longer symmetric in the general quadratic case. Thus the bases for the characters (left kernel) and ideal classes (right kernel) are no longer the same. This is actually perhaps somewhat of an advantage, as we can apply the Rédei symbol relations from hypercube co-ordinates a bit more easily, while the necessary genericity analysis is not too much more difficult.

Also, for imaginary quadratic fields we still have some control of the ideal classes by the relation that $[(\sqrt{D})]$ is trivial. Thus when $f\in\{0,3\}$ the obvious vector will be in the right kernel, and when $f=2$ the quasi-obvious vector $(0,1,1,\dots,1)$ will be (we shall simply refer to both of these as "obvious"). There is no real reason to carry the influence of this vector along to the higher pairing matrices, so in the imaginary case we can take $\mathbf{R}^8$ as an e-by-e matrix. We write $\tilde{e}$ to be the number of columns of $\mathbf{R}^8$, so that this is e when $D<0$ and $(e+1)$ when $D>0$.

12.1. First we describe how we will apply hypercube co-ordinates in this case. We will then show that well-gapped selections of such exist generically.

12.1.1. Let $\psi^\star=\sum_i\sum_j c(i,j)\psi^\star_{i,j}$ be a nontrivial character in $\widehat{\mathrm{Mat}}(e,\tilde{e},\mathbf{F}_2)$, where the $\psi^\star_{i,j}$ are basic characters dual to matrices with exactly one nonzero entry.

Let h^{s}_ψ and h^{z}_ψ be indices such that $c(h^{\mathrm{s}}_\psi,h^{\mathrm{z}}_\psi)=1$. We then take s_ψ to be a co-ordinate such that the (h^{s}_ψ)th vector in the left kernel is the unique nonobvious vector in either kernel that is nonzero at it; and we take z_ψ to be a co-ordinate such that the (h^{z}_ψ)th vector in the right kernel is the unique nonobvious vector in either

[25] Note that Smith only considers imaginary quadratic fields, though at least at the 8-rank level it seems that the generalization to all quadratic fields is rather straightforward.

kernel that is nonzero at it; and z'_ψ to be a co-ordinate where all the nonobvious vectors in either kernel are zero.[26]

We give a pictorial description of this; note that there is no difference in the case where $h^{\mathrm{s}}_\psi = h^{\mathrm{z}}_\psi$, as these are referring to indices for different bases in any event.

$$\begin{array}{r@{\;}c@{}c@{}c@{}c@{}c@{}c@{}c@{}c@{}c@{}c@{}c@{}c}
 & & & 0 & & & & 0 & & & & 0 & \\
 & & & 0 & & & & 0 & & & & 0 & \\
(h^{\mathrm{z}}_\psi)\text{th right kernel basis vector} & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
 & & & 0 & & & & 0 & & & & 0 & \\
(h^{\mathrm{s}}_\psi)\text{th left kernel basis vector} & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
 & & & 0 & & & & 0 & & & & 0 & \\
 & & & \mathrm{z}_\psi & & & & \mathrm{z}'_\psi & & & & \mathrm{s}_\psi &
\end{array}$$

12.2. Let us compute the value of $\partial_{\check{Z}\times S}\bar{\psi}^\star$ in terms of Rédei symbols. As previously, here $\bar{\psi}^\star$ sends $\check{Z} \times S$ to $\mathbf{F}_2$, given by $\bar{\psi}^\star(\check{z}, s) = \psi^\star\big(\mathbf{R}^8(t^\downarrow \times \check{z} \times t^\uparrow \times s)\big)$, where $t^\downarrow$ and $t^\uparrow$ are fixed. As with §7.5, we write $b_k = \prod_{l\notin\mathcal{V}_\psi} t_l^{w_l}$ where $\vec{w}$ is the kth basis vector of the right kernel (for ideals) and (e_l) its components. When we are considering even D we multiply b_k by 2 when the 0th component of the kth right kernel basis vector is 1. Similarly, we write $\underline{b}_k = \prod_{l\notin\mathcal{V}_\psi} \hat{t}_l^{w_l}$ where now $\vec{w}$ is the kth basis vector of the left kernel. Here $\hat{t}_l$ is $t_l \cdot (-1)^{(t_l-1)/2}$, and is thus a discriminantal divisor of D. When D is even and the 0th component of the kth left kernel basis vector is 1, we multiply $\underline{b}_k$ by the one of $\{-4, -8, 8\}$ that is a discriminantal divisor of D. We let u be this product for the left kernel where all $e_l = 1$ (including e_0 if applicable).

Lemma 12.2.1. *Suppose ψ is as above. Then*

$$\partial_{\check{Z}\times S}\bar{\psi}^\star\big((z_1, z'_1, s_1), (z_2, z'_2, s_2)\big) = [s_1 s_2, z'_1 z'_2, z_1 z_2].$$

Proof. Suppose first that (i, j) is a pair with $i \neq h^{\mathrm{s}}_\psi$ and $j \neq h^{\mathrm{z}}_\psi$. In this case we simply have $\bar{\psi}^\star_{i,j}(z, z', s) = \langle \chi_{\underline{b}_i}, b_j \rangle_{\hat{z}\hat{z}'\hat{s}u}$ (where $\hat{z}$ is the discriminantal divisor corresponding to z, etc., noting that the $\mathcal{K}$-condition fixes these modulo 4), so that $\partial\bar{\psi}^\star_{i,j}\big((z_1, z'_1, s_1), (z_2, z'_2, s_2)\big)$ is $\sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{s}}[\underline{b}_i, \hat{\tilde{z}}\hat{\tilde{z}}'\hat{\tilde{s}}\underline{b}_i u, b_j]$ where as with the proof of Lemma 7.7.1 each of the tildes is a co-ordinate to be summed over. By linearity in the second input this is 0.

Suppose next that $j = h^{\mathrm{z}}_\psi$ and $i \neq h^{\mathrm{s}}_\psi$. Then the relevant basis vector in the right kernel is nonzero at the z_ψ-coordinate, and we have $\bar{\psi}^\star_{i,j}(z, z', s) = \langle \chi_{\underline{b}_i}, z b_j \rangle_{\hat{z}\hat{z}'\hat{s}u}$. Here we find that $\partial\bar{\psi}^\star_{i,j}\big((z_1, z'_1, s_1), (z_2, z'_2, s_2)\big)$ is

$$\sum_{\tilde{z}'}\sum_{\tilde{s}}[\underline{b}_i, \hat{z}_1\hat{\tilde{z}}'\hat{\tilde{s}}\underline{b}_i u, z_1 b_j] + \sum_{\tilde{z}'}\sum_{\tilde{s}}[\underline{b}_i, \hat{z}_2\hat{\tilde{z}}'\hat{\tilde{s}}\underline{b}_i u, z_2 b_j] = 0,$$

again by linearity in the second input with both double sums. A similar calculation gives the same answer when $i = h^{\mathrm{s}}_\psi$ and $j \neq h^{\mathrm{z}}_\psi$.

[26] Smith takes the (h^{z}_ψ)th vector in the right kernel to be *nonzero* at z'_ψ (see the first half of Definition 3.4(3) of [15], in particular the third part of the third bullet point, where $w_{b\star}$ refers to the basis of ideals as at the bottom of page 20). On the other hand, the second display in Smith's Theorem 2.8 indicates that each basis vector has at most one variable index at which it is nonzero.

As explained to me by Peter Koymans, the answer to this riddle (at least in the imaginary quadratic case) is that one can multiply by the ideal $(\sqrt{D})$ before applying Theorem 2.8 (cf. Proposition 3.6, particularly the addition with t_b in the proof). It seems more straightforward to select co-ordinates as we do; Smith's setup does only use one fixed index b_i throughout (independent of the character), but this doesn't seem to give any technical gain.

Finally we have the case where $(i,j) = (h^{\mathrm{s}}_{\psi}, h^{\mathrm{z}}_{\psi})$, which will contribute nontrivially. Here $\bar{\psi}^{\star}_{i,j}(z,z',s) = \langle \chi_{\hat{s}\underline{b}_i}, zb_j\rangle_{\hat{z}\hat{z}'\hat{s}u}$, and $\partial\bar{\psi}^{\star}_{i,j}\big((z_1,z_1',s_1),(z_2,z_2',s_2)\big)$ is

$$[\hat{s}_1\underline{b}_i, \hat{z}_1\hat{z}_1'\underline{b}_i u, z_1b_j] + [\hat{s}_1\underline{b}_i, \hat{z}_1\hat{z}_2'\underline{b}_i u, z_1b_j] + [\hat{s}_1\underline{b}_i, \hat{z}_2\hat{z}_1'\underline{b}_i u, z_2b_j] + [\hat{s}_1\underline{b}_i, \hat{z}_2\hat{z}_2'\underline{b}_i u, z_2b_j] +$$
$$+ [\hat{s}_2\underline{b}_i, \hat{z}_1\hat{z}_1'\underline{b}_i u, z_1b_j] + [\hat{s}_2\underline{b}_i, \hat{z}_1\hat{z}_2'\underline{b}_i u, z_1b_j] + [\hat{s}_2\underline{b}_i, \hat{z}_2\hat{z}_1'\underline{b}_i u, z_2b_j] + [\hat{s}_2\underline{b}_i, \hat{z}_2\hat{z}_2'\underline{b}_i u, z_2b_j],$$

which by linearity in the second input is

$$[\hat{s}_1\underline{b}_i, \hat{z}_1'\hat{z}_2', z_1b_j] + [\hat{s}_1\underline{b}_i, \hat{z}_1'\hat{z}_2', z_2b_j] + [\hat{s}_2\underline{b}_i, \hat{z}_1'\hat{z}_2', z_1b_j] + [\hat{s}_2\underline{b}_i, \hat{z}_1'\hat{z}_2', z_2b_j],$$

and then by linearity in the third and first input is

$$[\hat{s}_1\underline{b}_i, \hat{z}_1'\hat{z}_2', z_1z_2] + [\hat{s}_2\underline{b}_i, \hat{z}_1'\hat{z}_2', z_1z_2] = [\hat{s}_1\hat{s}_2, \hat{z}_1'\hat{z}_2', z_1z_2] = [s_1s_2, z_1'z_2', z_1z_2],$$

the last step since $\mathcal{K}$ fixes the classes mod 4, so $s_1 \equiv s_2\,(4)$ and $z_1' \equiv z_2'\,(4)$.

There is thus a unique (i,j)-pair that contributes to the sum, so we get

$$\sum_i\sum_j c_{i,j}\partial\bar{\psi}^{\star}_{i,j}\big((z_1,z_1',s_1),(z_2,z_2',s_2)\big) = [s_1s_2, z_1'z_2', z_1z_2],$$

which is the statement of the Lemma. □

12.3. In order to imitate our previous arguments, we will need to show that for each generic $(\mathcal{K},\mathcal{L})$ there is a well-gapped choice of hypercube co-ordinates.

12.3.1. We will first count how often a given vector in $\mathbf{F}_2^r$ appears in either kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ when aggregated over $(\mathcal{K},\mathcal{L}) \in \mathcal{D}(\tilde{r},\mathcal{P})$, in imitation of Lemma 6.4.5. We also want to count how often a pair $(\vec{v},\vec{w}) \in (\mathbf{F}_2^r)^2$ has $\vec{v}$ in the left kernel and $\vec{w}$ in the right kernel. The case where $\vec{v} = \vec{w}$ will require special accounting.

We consider a fixed $(f,\varepsilon) \in \{0,2,3\} \times \{+,-\}$ and $\tilde{r}$, and have the following.

Lemma 12.3.2. *For every $\mathcal{K}$-specification modulo 8, every nonobvious nonzero vector in $\mathbf{F}_2^r$ is in the left kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ for a proportion $\ll 1/2^r$ of the $\mathcal{L}$. The same holds true for the right kernel.*

One distinction between this and Lemma 6.4.5 is that there we were working with $\mathbf{R}^4$-matrices that were known to be symmetric (perhaps this is best thought of as being part of the $\mathcal{P}$-condition, which here contains all odd primes). However, the proof here could be much the same, even with the extra splitting into $\mathcal{K}$-strata (which is more for later technical convenience than anything else). Yet, we instead give a proof based more on our genericity analysis (following Swinnerton-Dyer) in the 2-Selmer case.

Proof. As in [21, §8.1.2ff] we introduce formal variables as a way of discussing independence of $\mathcal{L}$-entries, though I've decided to omit the dot on $\mathcal{L}$, using $\mathcal{L}'$ instead. Indeed, here we directly take $\mathcal{L}'_{ij} = (\hat{p}_i|\dot{p}_j)^{\star}$ for $1 \le i,j \le \tilde{r}$ with $i \ne j$. When $\varepsilon = +1$ these are independent for $1 \le i < j \le \tilde{r}$, while for $\varepsilon = -1$ they are independent for $1 \le i < j < \tilde{r}$. It is also convenient to notate $\mathcal{L}'_{0j}$ and $\mathcal{L}'_{j0}$ as corresponding to the Kronecker symbols with 2 from $\mathcal{K}_j$ when $f \ne 0$, and moreover $\mathcal{L}'_{jj} = \sum_{i \ne j}\mathcal{L}'_{ij}$ for the column sum (including $i = 0$ when $f \ne 0$). (Thus the $\mathcal{L}'$-matrix is just the Rédei matrix $\mathbf{R}^4$).

We wish to show that for a nonobvious nonzero vector there are at least $r + O(1)$ independent conditions (in the formal variables) that correspond to it being in a kernel. Similar to [21, §8.1.3], a set of linear combinations of formal variables here is *independent* if they are independent in the free $\mathbf{F}_2$-algebra modulo: the relations from the column (and row, where applicable) sums being zero; those arising from quadratic reciprocity (such relations are made precise by the $\mathcal{K}$-conditions modulo 4); and the $\mathcal{L}'_{0j}$ being specified by $\mathcal{K}$.

For instance, taking $\varepsilon = -1$, given a nonzero vector $\vec{v}$ there is an associated nonempty set I of indices at which it is nonzero, and for $\vec{v}$ to be in the left kernel we need $\sum_{i\in I} \mathcal{L}'_{ic} = 0$ for all $1 \leq c < \tilde{r}$. For each column c that has $c \in I$, its condition is the unique one involving $\mathcal{L}'_{cc}$ (and is thus the unique one involving $\mathcal{L}'_{c\tilde{r}}$ if one prefers); this condition is nontrivial in the quotient algebra when $\vec{v}$ is nonobvious, and we thus see that all such conditions are independent of each other and the others. Meanwhile, the condition for each remaining column c has $\mathcal{L}'_{ic}$ appearing for some i (since I is nonempty), whilst $\mathcal{L}'_{ci}$ does not appear (since $c \notin I$). Thus these $(\tilde{r} - 1)$ conditions are all independent.

A similar argument applies for the right kernel. When $\varepsilon = +1$ one can simply ignore (for the right kernel) the condition from the $\tilde{r}$th row, and the argument is then the same. □

12.3.3. Next we consider both kernels simultaneously. A pair of vectors $(\vec{v}, \vec{w})$ is a *kernel pair* for a matrix if $\vec{v}$ is in its left kernel and $\vec{w}$ is in its right kernel.

Lemma 12.3.4. *Suppose $(\vec{v}, \vec{w})$ in $(\mathbf{F}_2^r)^2$ with $\vec{v} \neq \vec{w}$ such that neither $\vec{v}$ nor $\vec{w}$ is zero or obvious. Then $(\vec{v}, \vec{w})$ is a kernel pair for $\mathbf{R}^4(\mathcal{K}, \mathcal{L})$ for a proportion $\ll 1/4^r$ of the $(\tilde{\mathcal{K}}, \mathcal{L})$.*

Proof. As with the previous proof, we employ formal variables. However, as we are now aggregating over $\mathcal{K}$ we take *independent* only to mean they are independent in the free $\mathbf{F}_2$-algebra modulo the relations from the column/row sums being zero and those arising from quadratic reciprocity.

12.3.5. Given a nonzero vector $\vec{v}$ there is an associated nonempty set I of indices at which it is nonzero, and for $\vec{v}$ to be in the left kernel we need $\sum_{i\in I} \mathcal{L}'_{ic} = 0$ for all $1 \leq c < \tilde{r}$. Similarly, there is a nonempty set J of indices at which $\vec{w}$ is nonzero, and for it to be in the right kernel we need $\sum_{j\in J} \mathcal{L}'_{aj} = 0$ for all $1 \leq a < \tilde{r}$ (ignoring the $\tilde{r}$th column regardless of whether $\varepsilon = -1$). This gives a list of $2(\tilde{r}-1)$ conditions. They all are nontrivial in the quotient algebra since $\vec{v}$ and $\vec{w}$ are nonobvious. We wish to show that $2r + O(1)$ of them are independent.

Since $\vec{v} \neq \vec{w}$ we have $I \neq J$, and thus there is some z in their symmetric difference; without loss of generality we will take $z \in I$. Also, since J is nonempty there is some y in it. Finally, we take some arbitrary $x \notin \{0, \tilde{r}, y, z\}$, and remove the conditions from the xth, yth, and zth rows and columns from our list (of course, when $y, z \in \{0, \tilde{r}\}$ they were already not there – we also tacitly assume $\tilde{r} \geq 4$).

In particular, we can note that the collection of $2(\tilde{r} - 2)$ expressions $\mathcal{L}'_{bb}$ and $\mathcal{L}'_{zb}$ for $b \notin \{0, \tilde{r}, x\}$ are all independent (this is why we introduce x), and similarly for the collection with $\mathcal{L}'_{bb}$ and $\mathcal{L}'_{by}$; our strategy to show independence of the members of our list shall essentially be to show the conditions therein can be “triangularized” so as to be reduced to members of the above collections.

12.3.6. For $b \notin (I \cap J) \cup \{0, \tilde{r}, x\}$, we note that the column-condition $\sum_i \mathcal{L}'_{ib} = 0$ depends on $\mathcal{L}'_{zb}$ since $z \in I$. The only other condition that could depend on $\mathcal{L}'_{zb}$ is the row-condition for z; however this dependence does not occur since $b \notin J$ (and moreover we removed the z-row condition anyway). Meanwhile, there are two possibilities for a condition to depend on $\mathcal{L}'_{bz}$, either from the z-column condition or from the b-row condition. The former does not occur because we removed said condition from our list, and the latter has no dependence on $\mathcal{L}'_{bz}$ since $z \notin J$. This implies the b-column condition is independent of all the others. Similarly $\sum_j \mathcal{L}'_{bj} = 0$ depends on $\mathcal{L}'_{by}$ while no other condition depends on $\mathcal{L}'_{by}$ since $b \notin I$, and no remaining[27] condition depends on $\mathcal{L}'_{yb}$ (either since $b \notin J$ or since we removed the y-row condition). Thus the two b-conditions are independent of all the others in this case.

When $b \in I$ but $b \notin J$ (again with $b \notin \{0, \tilde{r}, x\}$), the condition $\sum_i \mathcal{L}'_{ib} = 0$ is the unique one with $\mathcal{L}'_{bb}$, and is thus independent of all the others. Meanwhile the row-condition $\sum_j \mathcal{L}'_{bj} = 0$ depends on $\mathcal{L}'_{by}$, and no other condition depends on this since we removed the y-column condition; moreover, no remaining condition depends on $\mathcal{L}'_{yb}$ since $b \notin J$ (alternatively, since we removed the y-row condition). Thus the two b-conditions are independent of all the others in this case.

When $b \in J$ but $b \notin I$, the condition $\sum_j \mathcal{L}'_{bj} = 0$ is the unique one with $\mathcal{L}'_{bb}$, and is thus independent of all the others. Meanwhile $\sum_i \mathcal{L}'_{ib} = 0$ depends on $\mathcal{L}'_{zb}$, and no other condition depends on this since we removed the z-row condition, while no remaining condition depends on $\mathcal{L}'_{bz}$ due to the removal of the z-column condition (or alternatively that $b \notin I$). Thus the two b-conditions are independent of all the others in this case.

When $b \in I \cup J$ we have one condition as $\sum_i \mathcal{L}'_{ib} = 0$, and then sum the row/column conditions to get $\sum_i \mathcal{L}'_{ib} + \sum_j \mathcal{L}'_{bj} = 0$ for the second condition. This no longer depends on $\mathcal{L}'_{bb}$, so the first condition is the unique one with it, and is thus independent of the others. The second condition depends on $\mathcal{L}'_{zb}$ but not on $\mathcal{L}'_{bz}$, and indeed no unremoved condition depends on $\mathcal{L}'_{bz}$ since $z \notin J$. On the other hand, no remaining condition depends on $\mathcal{L}'_{zb}$ since we removed the row-condition for z. Thus the two b-conditions are independent of all the others in this case.

We conclude that requiring $(\vec{v}, \vec{w})$ to be a kernel pair imposes at least $2(\tilde{r} - 4)$ independent conditions on $\mathcal{L}$, showing the Lemma. □

12.3.7. Now we delve into the case where $\vec{v}$ is in both kernels.

Given a vector $\vec{v} \in \mathbf{F}_2^r$ we can associate to it a product of formal symbols (see §5.3.1) as associated[28] to $\mathbf{R}^4(\mathcal{K}, \mathcal{L})$. This product is given by $\dot{v} = \dot{2}^{v_0} \prod_{i=1}^{\tilde{r}} \dot{p}_i^{v_i}$. Recalling that $\dot{d}$ is $\dot{2}^f \varepsilon$ times the product of the odd $\dot{p}$, the condition that $\vec{v}$ be in the right kernel is that $(\dot{v}, \dot{d})_{\mathbf{Q}}^{\Pi} = 1$ (see §4.3.1). In particular, in the imaginary quadratic case the (quasi-)obvious vector $\vec{o}$ has $\dot{o} = -\dot{d}$ up to squares, and the fact that it is in the right kernel is thus encoded by $(-\dot{d}, \dot{d})_{\mathbf{Q}}^{\Pi} = 1$.

Similarly, for a vector $\vec{v}$ in the left kernel we can associate to it the formal product $\hat{\dot{v}} = \hat{\dot{2}}^{v_0} \prod_{i=1}^{\tilde{r}} \hat{\dot{p}}_i^{v_i}$, where for odd $\dot{p}$ we have $\hat{\dot{p}} = \dot{p} \cdot (-1, \dot{p})_{\dot{p}}$ (thus flipping the sign for $\dot{p}$ associated to a prime that is 3 mod 4), while $\hat{\dot{2}}$ satisfies $\hat{\dot{2}} \prod_i \hat{\dot{p}}_i = \dot{d}$. The condition that $\vec{v}$ be in the left kernel is then that $(\hat{\dot{v}}, -\dot{d})_{\mathbf{Q}}^{\Pi} = 1$, and indeed we see that the obvious vector is in the left kernel via $(\dot{d}, -\dot{d})_{\mathbf{Q}}^{\Pi} = 1$.

[27] We use this adjective to exclude from consideration the b-column condition that we just showed to be independent of all other conditions.

[28] The formal symbols themselves are essentially defined independently of $(\mathcal{K}, \mathcal{L})$, as (in conjunction with (f, ε) that we consider fixed) the residue/Legendre conditions only relevantly specify the number r of formal symbols; moreover, the scope of (f, ε) is mainly to fix the relation with $\dot{d}$.

12.3.8. Thus if $\vec{v}$ is in both kernels of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ we have $(\dot{v},\dot{d})^{\Pi}_{\mathbf{Q}} = (\hat{\dot{v}},-\dot{d})^{\Pi}_{\mathbf{Q}} = 1$. Now $\hat{\dot{v}} = \lambda\dot{v}$ where $\lambda \in \{\pm 1, \pm 2\}$ (up to equivalence by squares) and thus at any odd[29] prime $\dot{p}$ dividing $\dot{d}$ we have $1 = (\dot{v},\dot{d})_{\dot{p}}(\dot{v},-\dot{d})_{\dot{p}}(\lambda,-\dot{d})_{\dot{p}} = (\dot{v},-1)_{\dot{p}}(\lambda,-\dot{d})_{\dot{p}}$.

When $\lambda = 1$ this is a condition just on $\dot{v}$, implying that no prime that occurs in $\dot{v}$ is 3 mod 4; in other words, for primes $\dot{p}$ dividing $\dot{v}$ we have $(-1|\dot{p})^{\star} = 0$ in the $\mathcal{K}$-condition. We will write $\mathcal{S}_4^3(\dot{x})$ for the set of primes (or formal symbols) that are 3 mod 4 that divide a given formal symbol $\dot{x}$. In this $\lambda = 1$ case, the condition that a vector $\vec{v}$ in the left kernel is also in the right kernel is that $\mathcal{S}_4^3(\dot{v})$ is empty.

Of course, we also have to catalogue exactly when $\lambda = 1$ occurs, and handle the other cases. It is convenient to split things based upon (f,ε) and v_0 (the latter is 0 by convention when $f = 0$).

When $v_0 = 0$ we see that $\lambda = 1$ exactly when $\#\mathcal{S}_4^3(\dot{v})$ is even. Otherwise, we have $\lambda = -1$ and can note that $\varepsilon = -1$ (following either from requiring the discriminant to be fundamental or from $(-1,-\dot{d})_\infty = 1$). Since here $(\dot{v},-1)_{\dot{p}}(-1,\dot{p})_{\dot{p}} = 1$ for all odd $\dot{p}$, it is necessary and sufficient for every prime in $\mathcal{S}_4^3(\dot{d})$ to also be in $\mathcal{S}_4^3(\dot{v})$. (One can note that the obvious vector is thus in both kernels when $(f,\varepsilon) = (0,-1)$). This will suffice (in the Lemma below) to restrict the proportion of $(\mathcal{K},\mathcal{L})$ that have $\vec{v}$ in both kernels, essentially since $\mathcal{S}_4^3(\dot{d})$ typically has $r/2$ members.

When $v_0 = 1$ and $f = 2$ we have $\lambda = \pm 2$; when the plus sign occurs we then have $(\dot{v},-1)_{\dot{p}}(2,\dot{p})_{\dot{p}} = 1$ for all odd $\dot{p}$, and thus $\mathcal{S}_8^5(\dot{d})$ is empty, while $\mathcal{S}_8^3(\dot{d}) = \mathcal{S}_8^3(\dot{v})$, and $\mathcal{S}_8^7(\dot{v})$ is empty. The behaviour for the 3 and 7 classes is switched when $\lambda = -2$ (one can also note that $(-2,-\dot{d})_\infty = -1$ in this case, so $\dot{d}$ is negative).

When $v_0 = 1$ and $f = 3$ we have $\lambda = \pm 1$. Indeed from $\hat{\dot{2}}\prod_i \hat{\dot{p}}_i = \dot{d} = \dot{2}^f\varepsilon\prod_i \dot{p}_i$ we have $\hat{\dot{2}} = \dot{2}\varepsilon(-1)^{\#\mathcal{S}_4^3(\dot{d})}$, so that $\lambda = \varepsilon(-1)^{\#\mathcal{S}_4^3(\dot{v})+\#\mathcal{S}_4^3(\dot{d})}$. As above, when $\lambda = 1$ the condition for $\vec{v}$ to be in both kernels (assuming it is in at least one) is that $\mathcal{S}_4^3(\dot{v})$ is empty. Otherwise, when $\lambda = -1$ the condition is that $\mathcal{S}_4^3(\dot{d}) = \mathcal{S}_4^3(\dot{v})$. (Again we can note that the obvious vector is in both kernels when $\varepsilon = -1$).

12.3.9. We sum this up in Table 1. The first column gives the value of v_0, and the second gives any (f,ε) restrictions. (Note that the $\varepsilon = -1$ in the second row is a conclusion from the λ-conditions). The third column then lists the value of λ, and at least[30] for $f \neq 2$ the fourth column lists when such λ occurs (with the $\equiv$-sign being congruence modulo 2). The fifth column then lists the conditions on $\dot{v}$ that are equivalent to $\vec{v}$ being in both kernels of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$ if it is in at least one of them.

12.4. Next we show there are well-gapped choices of hypercube co-ordinates for generic $(\mathcal{K},\mathcal{L})$. We again define $r_{\mathrm{g}} = \lfloor(\alpha_{\mathcal{P}}/2)(\log\log X)\rfloor$ (here $\alpha_{\mathcal{P}} = 1$), and have the integers in $[(5/4)r_{\mathrm{g}}, \tilde{r}]$ as the upper candidate co-ordinates, and the integers in $[r_{\mathrm{g}}/5, r_{\mathrm{g}}/4]$ as the lower candidate co-ordinates. We write ν_{z} and ν_{s} for the numbers of such upper/lower candidate co-ordinates; the former is $\approx (3/4)r_{\mathrm{g}}$ and the latter is $\approx (1/20)r_{\mathrm{g}}$, with both of these $\gg \log\log X$.

[29] By Hilbert reciprocity the number of places at which $(\dot{v},\dot{d})(\dot{v},-\dot{d})(\lambda,-\dot{d})$ is nonzero is even; so we need only check it at all odd primes and the infinite place.

[30] Whereas the distinction between $\mathcal{S}_4^3(\dot{v})$ being empty or equal to $\mathcal{S}_4^3(\dot{d})$ will make a (minor) difference in the analysis, the flipping of the congruence classes mod 8 does not, so we need not be more specific when $f = 2$.

v_0	(f,ε)	λ	λ-conditions	$\vec{v}$-conditions
0	any	1	$\#\mathcal{S}_4^3(\dot{v})$ is even	$\mathcal{S}_4^3(\dot{v})$ is empty
0	$\varepsilon=-1$	-1	$\#\mathcal{S}_4^3(\dot{v})$ is odd	$\mathcal{S}_4^3(\dot{v})=\mathcal{S}_4^3(\dot{d})$
1	$f=2$	2		$\mathcal{S}_8^5(\dot{d})=\mathcal{S}_8^7(\dot{v})=\varnothing,\mathcal{S}_8^3(\dot{v})=\mathcal{S}_8^3(\dot{d})$
1	$f=2$	-2		$\mathcal{S}_8^5(\dot{d})=\mathcal{S}_8^3(\dot{v})=\varnothing,\mathcal{S}_8^7(\dot{v})=\mathcal{S}_8^7(\dot{d})$
1	$\varepsilon=+1$	1	$\#\mathcal{S}_4^3(\dot{v})\equiv\#\mathcal{S}_4^3(\dot{d})$	$\mathcal{S}_4^3(\dot{v})$ is empty
1	$\varepsilon=-1$	1	$\#\mathcal{S}_4^3(\dot{v})\not\equiv\#\mathcal{S}_4^3(\dot{d})$	$\mathcal{S}_4^3(\dot{v})$ is empty
1	$\varepsilon=+1$	-1	$\#\mathcal{S}_4^3(\dot{v})\not\equiv\#\mathcal{S}_4^3(\dot{d})$	$\mathcal{S}_4^3(\dot{v})=\mathcal{S}_4^3(\dot{d})$ (impossible)
1	$\varepsilon=-1$	-1	$\#\mathcal{S}_4^3(\dot{v})\equiv\#\mathcal{S}_4^3(\dot{d})$	$\mathcal{S}_4^3(\dot{v})=\mathcal{S}_4^3(\dot{d})$

TABLE 1. When is a vector in one kernel also in the other?

We then need a slight generalization of the results of §6.4, so as to handle our current situation where we can have two distinct kernel bases. The notion of size-genericity is the same as before (namely $e\le 99\sqrt{\log\log\log X}$), and Lemma 6.4.3 readily gives the desired bound therein.

12.4.1. The situation with pattern-genericity is a bit more articulated. We again have rotten vectors in $\mathbf{F}_2^r$, which fail either to have $\nu_{\rm z}/2+\Theta(\lambda(\log\log X)^{3/4})$ zeros/ones on the $\nu_{\rm z}$ lower candidate co-ordinates, or to have $\nu_{\rm s}/2+\Theta(\lambda(\log\log X)^{3/4})$ zeros/ones on the $\nu_{\rm s}$ upper candidate co-ordinates (where $\lambda=(3+\sqrt{17})/4\approx 1.781$).

We also have *bi-rotten* vector pairs $(\vec{v},\vec{w})$, which are those where $\vec{v}+\vec{w}$ is rotten, trivial, or obvious.[31] Though we will apply this when $\vec{v}$ and $\vec{w}$ are vectors from different kernels, this vector addition makes sense on $\mathbf{F}_2^r$.

We say $(\mathcal{K},\mathcal{L})$ is *pattern-generic* if: there are no rotten vectors in either kernel of $\mathbf{R}^4(\mathcal{K},\mathcal{L})$, and there are also no bi-rotten kernel pairs except when both components are trivial or obvious.

Lemma 12.4.2. *The proportion of $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})$ that are not pattern-generic is $\ll\exp(-\tilde{c}\sqrt{\log\log X})$.*

For a pleasant box $\bar{T}$ the sum of $\#T(\mathcal{K},\mathcal{L})$ over $(\mathcal{K},\mathcal{L})\in\mathcal{D}(\tilde{r},\mathcal{P})$ that are not pattern-generic is $\ll\#T/(\log\log X)^{99}$.

Proof. By Stirling's approximation and tails of the binomial distribution, the number of rotten vectors is $\ll 2^r\exp(-\tilde{c}_{\mathcal{P}}\sqrt{\log\log X})$. Meanwhile, every nontrivial nonobvious vector is in a proportion $\ll 1/2^r$ of the kernels by Lemma 12.3.2, giving an overall proportion that fits into the bound of the Lemma.

Similarly, the number of bi-rotten vector pairs with $\vec{v}$ or $\vec{w}$ as zero or obvious is also $\ll 2^r\exp(-\tilde{c}_{\mathcal{P}}\sqrt{\log\log X})$, and each occurs in a proportion $\ll 1/2^r$ of the kernels to reach the same conclusion.

12.4.3. Next we consider kernel pairs. We first note that we have $4^{\tilde{r}}$ possibilities for $\mathcal{K}$ mod 8, while there are $2^{\binom{\tilde{r}}{2}}$ possibilities for $\mathcal{L}$. We can ignore $\mathcal{K}$ such that $\#\mathcal{S}_4^3(\dot{d})\le 0.49r$ or $\#\mathcal{S}_8^5(\dot{d})\le 0.24r$; indeed, the former contributes a negligible proportion $\ll(1/2)^{0.01r^2/0.50r}\ll 1/(\log X)^{\tilde{c}}$, and the latter similarly contributes a proportion $\ll(3/4)^{0.01r^2/0.25r}$. We call the remaining $\mathcal{K}$ *balanced.*

[31] One can note that $(\vec{v},\vec{v})$ is thus always bi-rotten, and in the case of a symmetric $\mathbf{R}^4$ this would then force (when $e>0$) us into a situation of non-genericity. Hence the analysis for Gaussian discriminants needed to use something more specific to its situation.

We consider bi-rotten vector pairs $(\vec{w}, \vec{w})$ with $\vec{w}$ nonzero and nonobvious. How many of the $(\mathcal{K}, \mathcal{L})$ have some such $(\vec{w}, \vec{w})$ as a kernel pair? The number of such $(\mathcal{K}, \mathcal{L})$ is bounded by $\sum_{\vec{v}} N(\vec{v})$ where $N(\vec{v})$ is the number of $(\mathcal{K}, \mathcal{L})$ that have a specific such $(\vec{v}, \vec{v})$ as a kernel pair. Moreover, this sum is equal to $\sum_{\mathcal{K}} \sum_{\vec{v}} N_{\mathcal{K}}(\vec{v})$ where $N_{\mathcal{K}}(\vec{v})$ counts the number of $\mathcal{L}$ such that $(\mathcal{K}, \mathcal{L})$ has $(\vec{v}, \vec{v})$ a kernel pair. Also, by Lemma 12.3.2 we have the simplistic (yet useful) upper bound $N_{\mathcal{K}}(\vec{v}) \ll 2^{\binom{\tilde{r}}{2}}/2^r$. Referring to Table 1, the condition that $\#\mathcal{S}_8^5(\dot{d}) \geq 0.24r$ on $\mathcal{K}$ already suffices to exclude any kernel pairs $(\vec{v}, \vec{v})$ when $\lambda = \pm 2$. Thus we can ignore the $\mathcal{K}$ with $\lambda = \pm 2$ in the above sum. Otherwise, given $\mathcal{K}$, the proportion of $\vec{v}$ that have $\mathcal{S}_4^3(\dot{v}) = \mathcal{S}_4^3(\dot{d})$ is $1/2^{\#\mathcal{S}_4^3(\dot{d})}$, with the same for the proportion that have $\mathcal{S}_4^3(\dot{v})$ empty. The $\vec{v}$ that do not meet the (fifth column) condition listed in Table 1 cannot have $(\vec{v}, \vec{v})$ as a kernel pair for $\mathcal{K}$, regardless of what $\mathcal{L}$ is. Since $\#\mathcal{S}_4^3(\dot{d}) \geq 0.49r$ for the balanced $\mathcal{K}$ that are under consideration, we thus have $\sum_{\vec{v}} N_{\mathcal{K}}(\vec{v}) \ll (2^r/2^{0.49r}) \cdot (2^{\binom{\tilde{r}}{2}}/2^r)$, and summing over such balanced $\mathcal{K}$ then gives an overall contribution of $1/2^{0.49r} \ll 1/(\log X)^{\tilde{c}}$ to the proportion of $(\mathcal{K}, \mathcal{L})$ that have some nonzero nonobvious kernel pair $(\vec{w}, \vec{w})$.

Finally, the number of bi-rotten vector pairs $(\vec{v}, \vec{w})$ with $\vec{v} \neq \vec{w}$ and neither $\vec{v}$ nor $\vec{w}$ zero or obvious is $\ll 4^r \exp(-\tilde{c}_{\mathcal{P}}\sqrt{\log\log X})$ by tails of the binomial distribution, and every such pair is in a proportion $\ll 1/4^r$ of the kernels by Lemma 12.3.4

So the proportion of nongeneric $(\mathcal{K}, \mathcal{L})$ is as stated.

12.4.4. Then by [21, Lemma 5.5.1] we lose only a factor of $\ll 2^{k_0 k_1}$ when passing to T-sizing, and as the exponent here is $\leq \kappa_0 \eta_0 (\log\log\log X) \cdot 3(\log\log X)^{\eta_1}$, we see that this is dominated by $\exp(-\tilde{c}\sqrt{\log\log X})$. □

Lemma 12.4.5. *If $(\mathcal{K}, \mathcal{L}) \in \mathcal{D}(\tilde{r}, \mathcal{P})$ is generic, then for sufficiently large X there is a selection of well-gapped hypercube co-ordinates for every $\psi^\star \in \widehat{\mathrm{Mat}}(e, \tilde{e}, \mathbf{F}_2)$.*

Proof. Given ψ we have its c-array, and take h_ψ^{z} and h_ψ^{s} so that $c(h_\psi^{\mathrm{s}}, h_\psi^{\mathrm{z}}) = 1$. We then wish to select well-gapped hypercube co-ordinates based upon these distinguished indices.

Since $(\mathcal{K}, \mathcal{L})$ is pattern-generic, every nontrivial nonobvious vector in either kernel has nearly the expected number of zeros/ones on the upper and lower candidate co-ordinates; moreover, the space of dimension $(e + \tilde{e})$ generated by nonobvious basis vectors from the left and right kernels has the same property – there are no rotten (or trivial, or obvious) vectors in it.

We want there to be an upper candidate co-ordinate with the pattern from the nonobvious joint basis vectors having exactly one prescribed nonzero entry. By Lemma 6.3.1, the number of upper candidate co-ordinates that match such a pattern is $\nu_{\mathrm{s}}/2^{e+\tilde{e}} + \Theta\big(1.781^{e+\tilde{e}}(\log\log X)^{3/4}\big)$; since we have $e \leq 99\sqrt{\log\log\log X}$ and $\nu_{\mathrm{s}} \gg \log\log X$, this is positive for sufficiently large X.

The same argument applies to select the lower hypercube co-ordinates. □

12.5. Finally, we can follow §§8-11 above, and get the expected results. (In particular, the choice of fields $L_{\tilde{Z}}$ is as with Type I in §7.7.3).

Theorem 12.5.1. *The asymptotic proportion of negative fundamental discriminants whose class group has 4-rank e_4 and 8-rank e_8 is*

$$\frac{\#\{d \leq X : -d \in \mathcal{F}^- \mid e_4(-d) = e_4, e_8(-d) = e_8\}}{\#\{d \leq X : -d \in \mathcal{F}^-\}} = \gamma_{\mathrm{I}}(e_4)\mathbf{P}_{\mathrm{i}}(e_4, e_8) + O(\tilde{E})$$

where the error $\tilde{E}$ is $\ll 1/(\log\log X)^{1/4}$, while $\mathbf{P}_{\mathrm{i}}(e_4,e_8)$ is the proportion of matrices in $\mathrm{Mat}(e_4,e_4,\mathbf{F}_2)$ *whose kernel has dimension e_8, and*

$$\gamma_{\mathrm{I}}(e)=\frac{1}{2^{e^2}}\prod_{u=1}^{\infty}(1-1/2^u)\prod_{j=1}^{e}(1-1/2^j)^{-1}\prod_{j=1}^{e}(1-1/2^j)^{-1}.$$

Theorem 12.5.2. *The asymptotic proportion of positive fundamental discriminants whose narrow class group has 4-rank e_4 and 8-rank e_8 is*

$$\frac{\#\{d\le X : d\in\mathcal{F}^+\mid e_4(d)=e_4, e_8(d)=e_8\}}{\#\{d\le X : d\in\mathcal{F}^+\}}=\gamma_{\mathrm{R}}(e_4)\mathbf{P}_{\mathrm{r}}(e_4,e_8)+O(\tilde{E})$$

where the error $\tilde{E}$ is $\ll 1/(\log\log X)^{1/4}$, while $\mathbf{P}_{\mathrm{r}}(e_4,e_8)$ is the proportion of matrices in $\mathrm{Mat}(e_4,e_4+1,\mathbf{F}_2)$ *whose left kernel has dimension e_8, and*

$$\gamma_{\mathrm{R}}(e)=\frac{1}{2^{e(e+1)}}\prod_{u=1}^{\infty}(1-1/2^u)\prod_{j=1}^{e}(1-1/2^j)^{-1}\prod_{j=1}^{e+1}(1-1/2^j)^{-1}.$$

One could refine these by splitting into the three cases for $2^f\|D$ if desired.

13. 4-Selmer ranks

Finally we consider the case of 4-Selmer ranks of quadratic twists of a fixed elliptic curve E with full 2-torsion (and no 4-torsion). This was already partially considered by Smith in [14], before his more general result in [15].

13.1. We shall ultimately show the following result, in terms of the numbers

$$\rho_s=\frac{2^s}{\prod_{v=1}^{s}(2^v-1)}\prod_{n=0}^{\infty}\bigl(1-1/2^{2n+1}\bigr)=\frac{1/2^{s(s-1)/2}}{\prod_{v=1}^{s}(1-1/2^v)}\prod_{n=1}^{\infty}\frac{1}{1+1/2^n},$$

for which $\rho_0+\rho_2+\rho_4+\cdots=\rho_1+\rho_3+\rho_5+\cdots=1$ (the n-product is ≈ 0.419422). We assume E is twist-minimal, meaning there are no bad primes p that can be removed by quadratic twisting.

Theorem 13.1.1. *Let $E/\mathbf{Q}$ be a twist-minimal elliptic curve with full rational 2-torsion and no rational 4-torsion point, and consider quadratic twists E_d of E by odd squarefree integers $|d|\le X$ coprime to the product of the bad primes of E. The proportion of such d such that E_d has 2-Selmer rank of (s_2+2) and 4-Selmer rank of (s_4+2) is $\mathbf{P}_{\mathrm{a}}(s_2,s_4)\rho_{s_2}/2+O_E\bigl(1/(\log\log X)^{1/4}\bigr)$, with an effective constant in the error, where $\mathbf{P}_{\mathrm{a}}(s_2,s_4)$ is the proportion of alternating matrices over $\mathbf{F}_2$ of size s_2 that have a kernel of dimension s_4 (here s_2 and s_4 have the same parity).*

In other words, the quotient

$$\frac{\#\{|d|\le X : \mu(d)\ne 0,\ \gcd(d,\Omega)=1\}\mid s_2(E_d)=s_2+2, s_4(E_d)=s_4+2\}}{\#\{|d|\le X : \mu(d)\ne 0,\ \gcd(d,\Omega)=1\}}$$

is asymptotically equal to

$$\frac{\rho_s}{2}\mathbf{P}_{\mathrm{a}}(s_2,s_4)+O_E\Bigl(\frac{1}{(\log\log X)^{1/4}}\Bigr).$$

13.1.2. Similar to [21, §12.1.2] we can count the number of alternating matrices of a given size and rank. We first note that the number $B_n(q)$ of nonsingular alternating matrices of size n over $\mathbf{F}_q$ is

$$q^{\binom{n}{2}} \prod_{\substack{1\le k\le n\\ k \text{ odd}}} (1-1/q^k)$$

for n even (and is 0 for n odd). The number of alternating matrices of size n with rank w (again necessarily even) is then $B_w(q)\left[{n \atop w}\right]_q$ where

$$\left[{n \atop w}\right]_q = \prod_{i=1}^{n}(q^i-1)\Big/\prod_{i=1}^{w}(q^i-1)\prod_{j=1}^{n-w}(q^j-1)$$

is the number of w-dimensional subspaces of a vector space of dimension n, and we indeed have the expected relation $\sum_w B_w(q)\left[{n \atop w}\right]_q = q^{\binom{n}{2}}$.

Using our previous notation of $F_q(u)=\prod_{i=1}^{u}(1-1/q^i)$ we have

$$B_w(q)\left[{n \atop w}\right]_q = q^{\binom{w}{2}}\frac{F_q(w)}{F_{q^2}(w/2)}\cdot\frac{q^{\binom{n}{2}}F_q(n)}{q^{\binom{w}{2}}F_q(w)\cdot q^{\binom{n-w}{2}}F_q(n-w)}.$$

With $(q,s_2,s_4)=(2,n,n-w)$ and dividing out by $q^{\binom{n}{2}}$ we then get

$$\mathbf{P}_{\mathrm{a}}(s_2,s_4)=\frac{1}{2^{\binom{s_4}{2}}}\frac{F_2(s_2)}{F_2(s_4)}\Big/F_4\Big(\frac{s_2-s_4}{2}\Big).$$

Note also that taking $n\to\infty$ with $(n-w)=s$ constant (and w even) gives $\rho_s/2$ as $1/2^{\binom{s}{2}}F_2(s)$ as we previously indicated in the 2-Selmer case.

In Table 2 we give approximations to $\mathbf{P}_{\mathrm{a}}(s_2,s_4)$ for $s_2\le 8$, with the lower-left half corresponding to even parity and the upper-right to odd. For instance, we have that $\mathbf{P}_{\mathrm{a}}(5,3)=155/1024\approx 0.1513$.

$s_2\downarrow$		7	5	3	1	$\leftarrow s_4$
0	1.0000	$5\cdot 10^{-7}$	0.0013	0.1577	0.8410	7
2	0.5000	0.5000	0.0010	0.1513	0.8477	5
4	0.4375	0.5469	0.0156	0.1250	0.8750	3
6	0.4238	0.5563	0.0199	$3\cdot 10^{-5}$	1.0000	1
8	0.4205	0.5585	0.0209	$4\cdot 10^{-5}$	$4\cdot 10^{-9}$	$s_2\uparrow$
$s_4\rightarrow$	0	2	4	6	8	

TABLE 2. Approximations to $\mathbf{P}_{\mathrm{a}}(s_2,s_4)$

13.2. The proof outline is similar to the case of the 8-rank of a quadratic class group. We start with the 2-Selmer pairing matrix $\mathbf{M}_E(\tilde{\mathcal{K}}_\epsilon,\mathcal{L})$; this can be taken to be symmetrical, whereupon the left/right kernels are the same. There is an alternating pairing on kernel vectors that gives the 4-Selmer pairing matrix, and we can ignore the obvious vectors from 2-torsion (which span a 2-dimensional subspace).

We want to show that such pairing matrices are equi-distributed, here in the space of alternating matrices $\mathrm{Alt}(s_2,\mathbf{F}_2)$ where s_2 is the reduced 2-Selmer rank (which is 2 less than the 2-Selmer rank in our case of full 2-torsion). Again this is most easily considered via characters on the matrix space, and the "algebraic input" (analogous to Rédei symbol relations) here gives a sum of character values over varying co-ordinates in terms of a Frobenius element in some field extension. We then need to show that suitable hypercube co-ordinates exist (particularly with respect to genericity), whereupon an argument following §§8-11 will then give the desired result. (Note that the reduced 4-Selmer rank will have the same parity as the reduced 2-Selmer rank, due to the alternating nature of the pairing matrix).

13.3. We first recall the definition of the 2-Selmer pairing matrix ([21, §7.1]), starting with the notation we need for this.

13.3.1. The elliptic curve $E : y^2 = (x-c_1)(x-c_2)(x-c_3)$ has full 2-torsion defined over $\mathbf{Q}$, scaled so that the c_i are integers with no nontrivial common square factor. We write $\delta_{ij} = c_i - c_j$ for $1 \le i \ne j \le 3$, and require that none of the three quantities (with i, j, k distinct) given by $\delta_{ij}\delta_{ik} = (c_i - c_j)(c_i - c_k)$ are square, so that E has no rational 4-torsion points. We write Ω for the set of prime divisors of $\delta_{12}\delta_{13}\delta_{23}$, in particular noting that $2 \in \Omega$. Moreover, we write $\tilde{\Omega}$ for $\{-1\} \cup \Omega$, alternatively interpreted as the set of places in Ω with the infinite place appended.

We will consider quadratic twists of E by squarefree d that are coprime to Ω; these twists are given by $E_d : y^2 = (x-dc_1)(x-dc_2)(x-dc_3)$. As we are interested in twist families, we can assume there is no prime that has the same nonzero valuation for all the δ_{ij}; if there is such a prime p, we can then twist E by it, with the resulting curve then being good at p.

13.3.2. We consider twisting E by $d = \epsilon \prod_i p_i$ where $\epsilon \in \{\pm 1\}$ and the p_i are distinct primes that are not in Ω.

We write $\mathcal{B}$ for the union of $\tilde{\Omega}$ with the primes dividing d. We let $Y_l = \mathbf{Q}_l^\star/(\mathbf{Q}_l^\star)^2$ for $l \in \mathcal{B}$, which is naturally a vector space over $\mathbf{F}_2$, being of dimension 2 for $l \ge 3$ (and dimension 3 for $l = 2$ and 1 for $l = \infty$). We let V_l be the space of 3-tuples $(m_1, m_2, m_3) \in Y_l^3$ with $m_1 m_2 m_3 = 1$, so that this is again naturally a vector space over $\mathbf{F}_2$, of twice the dimension of Y_l. We then let $V_{\mathcal{B}} = \sum_{l \in \mathcal{B}} V_l$.

We then define $U_{\mathcal{B}}$ as the subspace of $V_{\mathcal{B}}$ generated by the diagonally embedded elements $(1, l, l)$ and $(l, l, 1)$ for all $l \in \mathcal{B}$ (with $l = -1$ corresponding to the infinite place). Meanwhile, we define $W_l \subset V_l$ as the image generated by the points on $E_d(\mathbf{Q}_l)$ under the Kummer map, defined away from 2-torsion points as $(X, Y) \to (X - dc_1, X - dc_2, X - dc_3)$, and for 2-torsion points by continuity. With $W_{\mathcal{B}} = \sum_{l \in \mathcal{B}} W_l$, the 2-Selmer group is then the intersection of $U_{\mathcal{B}}$ and $W_{\mathcal{B}}$ (both of these vector spaces have half the dimension of $V_{\mathcal{B}}$).

There is also Tate's pairing-based interpretation. We define e_l on $V_l \times V_l$ by

$$e_l\big((m_1, m_2, m_3) \times (m_1', m_2', m_3')\big) = (m_1, m_1')_l + (m_2, m_2')_l + (m_3, m_3')_l$$

where $(u, v)_l$ is the Hilbert symbol (with values in $\mathbf{F}_2$) defined as 0 if $ux^2 + vy^2$ is soluble in $\mathbf{Q}_l$ and 1 otherwise. We then can extend $\oplus_l e_l$ to $V_{\mathcal{B}} \times V_{\mathcal{B}}$ by additivity. Our desired pairing matrix is then $e_{\mathcal{B}}$ on bases for $U_{\mathcal{B}}$ and $W_{\mathcal{B}}$, and the dimension of the kernel of this matrix is the 2-Selmer rank.

Although it is not immediately important, we can choose bases for $U_{\mathcal{B}}$ and $W_{\mathcal{B}}$ so that the 2-Selmer pairing matrix is symmetric; this is rather nontrivial to achieve, and we only sketched it in [21, §8.2].

In any case, the left kernel of this 2-Selmer pairing matrix then corresponds to the everywhere locally soluble 2-coverings of E_d. Indeed, recall ([21, §1.1]) that a triple $\vec{m} = (m_1, m_2, m_3)$ with $m_1 m_2 m_3$ a nonzero square gives rise to an intersection of three quadric equations $m_i y_i^2 = x - dc_i$ where $m_1 m_2 m_3$ is a nonzero square, and we denote such a curve by $\mathcal{C}(\vec{m})$. The 2-Selmer group is then the subset of $\vec{m} \in V_{\mathcal{B}}$ such that $\mathcal{C}(\vec{m})$ is everywhere locally soluble.

These curves $\mathcal{C}$ each have genus 1, and for each we have a commutative diagram

$$\begin{tikzcd}
E_d \arrow[r, "{[2]}"] & E_d \\
\mathcal{C} \arrow[u, "\iota"] \arrow[ur] &
\end{tikzcd}$$

where [2] is multiplication-by-2 and ι is an isomorphism over $\bar{\mathbf{Q}}$.

13.3.3. As in [21, §8.1] we can re-interpret the above in terms of formal symbols.

In particular, the Hilbert symbols involved in the 2-Selmer pairing matrix are determined (with an obvious identification of ±1 with $\mathbf{F}_2$) by the Legendre symbols $(p_i|p_j)$ for primes p_i and p_j that divide d (so associated to $\mathcal{L}$), by $(q|p_j)$ for $q \in \tilde{\Omega}$, and by the sign of d. The conditions for $(q|p_j)$ are weaker than congruential $\mathcal{K}$-conditions; we refer to these weaker Legendre conditions for bad primes as $\tilde{\mathcal{K}}$-conditions.

Thus, having fixed bases for $U_{\mathcal{B}}$ and $W_{\mathcal{B}}$ appropriately, the entries of the pairing matrix $\mathbf{M}(E_d)$ are the same for all squarefree d that are coprime to Ω and have $\tilde{r}$ prime divisors and meet given $(\tilde{\mathcal{K}}_\varepsilon, \mathcal{L})$-specifications, where ε specifies the sign of d. (The usage of $\mathcal{K}$ and $\mathcal{L}$ here implies a fixity of ordering for the primes dividing d).

With the formal symbols denoted by $\dot{p}_j$ for $1 \le j \le \tilde{r}$, given a specification $(\tilde{\mathcal{K}}_\varepsilon, \mathcal{L})$ we then require that we have $(q|\dot{p}_j)^\star = (q, \dot{p}_j)_{\dot{p}_j} = \mathcal{K}^\star_{qj}$ for $q \in \tilde{\Omega}$ and $(\dot{p}_i|\dot{p}_j)^\star = (\dot{p}_i, \dot{p}_j)_{\dot{p}_j} = \mathcal{L}^\star_{ij}$ for $1 \le i \ne j \le \tilde{r}$, where the $\star$-superscript indicates a result in $\mathbf{F}_2$ rather than $\{\pm1\}$. We can then rephrase the above as involving a 2-Selmer pairing matrix $\mathbf{M}_E(\mathcal{K}_\varepsilon, \mathcal{L})$, applicable to $d = \varepsilon \prod_j p_j$ that meet requisite residue and Legendre conditions when the $\dot{p}_j$ are "evaluated" at its prime divisors.

13.4. Next we wish to discuss the 4-Selmer pairing on members of the 2-Selmer group of E_d. This is known as the Cassels-Tate pairing [2], and is a bit involved to define. For our purposes we will not actually do so precisely, but instead mention the most important facts about it. It is an alternating bilinear pairing on everywhere locally soluble 2-coverings, and an everywhere locally soluble 2-covering $\mathcal{C}$ lifts to an everywhere locally soluble 4-covering $\mathcal{D}$ precisely when it is in the kernel of this 4-Selmer pairing – that is, its pairing with every everywhere locally soluble 2-covering is trivial. For such an everywhere locally soluble 4-covering $\mathcal{D}$ we then have the commutative diagram

$$\begin{tikzcd}
E_d \arrow[r, "{[2]}"] & E_d \arrow[r, "{[2]}"] & E_d \\
\mathcal{D} \arrow[u, "\iota_{\mathcal{D}}"] \arrow[r] & \mathcal{C} \arrow[u, "\iota_{\mathcal{C}}"] \arrow[ur] &
\end{tikzcd}$$

involving multiplication-by-4 on E_d and isomorphisms $\iota_{\mathcal{C}}$ and $\iota_{\mathcal{D}}$ defined over $\bar{\mathbf{Q}}$.

The dimension of the kernel of the 4-Selmer pairing matrix is then the 4-Selmer rank, and the 2-coverings corresponding to the 2-torsion of E_d will always contribute 2 to this dimension.

In fact, the above description with lifts of 2-coverings was later defined by Cassels via a related pairing (in [3]); subsequently Fisher, Schaefer, and Stoll [7] showed that in fact the Cassels pairing and the original Cassels-Tate pairing (given in terms of co-homology) are in fact the same. For computational purposes, the initial method in [3] (or [7, §4]) required difficult tasks such as solving norm equations or computing S-units; this was simplified in the MAGMA implementation (by Donnelly, unpublished notes) to involve only solving norm equations over quadratic extensions of the base field (here $\mathbf{Q}$), which is equivalent to the standard task of solving conics; however, even this necessity has been obviated by Fisher [6], using instead the invariant theory of binary quartics.

In any event, for the case where E has at least one nontrivial 2-torsion point one can additionally exploit the 2-isogenies (rather than the multiplication-by-2 map) to compute the pairing(s), as described by Fisher [5].

13.4.1. For the purposes of Smith's work on the 4-Selmer group, what we need to know is that a sum over a suitable hypercube of 4-Selmer pairings is equal to an Artin symbol involving a quadratic field extension L/K.

Smith already noted in his first preprint [14, Theorem 3.2] that when d_1 and d_2 have the same $(\tilde{\mathcal{K}}_\varepsilon, \mathcal{L})$-specifications one can relate the 4-Selmer pairings $\langle \vec{m}, \vec{m}' \rangle_{d_1}^{\mathrm{CT}}$ and $\langle \vec{m}, \vec{m}' \rangle_{d_2}^{\mathrm{CT}}$, where $\vec{m}, \vec{m}'$ are 2-Selmer elements for both E_{d_1} and E_{d_2}. In particular, he deduces the relation[32]

$$\langle \vec{m}, \vec{m}' \rangle_{d_1}^{\mathrm{CT}} = \langle \vec{m}, \vec{m}' \rangle_{d_2}^{\mathrm{CT}} + \left[\frac{L/K}{d_1 d_2}\right] \tag{3}$$

in terms of an Artin symbol for a field extension L/K.[33]

For the definition of the fields L and K (see [14, §3.2]), given (m_1, m_2, m_3) and (m_1', m_2', m_3') in the 2-Selmer group of $\mathbf{M}_E(\tilde{\mathcal{K}}_\varepsilon, \mathcal{L})$ we have $\prod_i (m_i, m_i')_v = +1$ for all places v, so there is some $b \in \mathbf{Q}$ such that $(m_i, bm_i')_{\mathbf{Q}}^{\Pi} = 1$, and this gives global solutions (x_i, y_i, z_i) to $x_i^2 - m_i y_i^2 = bm_i' z_i^2$ for $1 \le i \le 3$. The basic building block is then the multiquadratic field $K = \mathbf{Q}\big(\sqrt{m_1}, \sqrt{m_2}, \sqrt{m_1'}, \sqrt{m_2'}\big)$ with its extension $L = K(\sqrt{\alpha})$ where $\alpha = \prod_i (x_i + y_i\sqrt{m_i})$. Similar to the Rédei symbol in the 8-class case (§4.5.1), one can scale the global solutions so that L/K is minimally ramified (with adequate care being taken at 2 and ∞).

13.4.2. Let us give a somewhat extended example of this. We let $a = 113 \cdot 193$ and $b = 151 \cdot 167$, and $E : y^2 = x(x+a)(x+b)$. The 2-Selmer group of E has rank 5. Notating the three components of 2-Selmer elements to correspond respectively to the roots $(0, -a, -b)$ of the cubic, a basis for the 2-Selmer group is given by

$$(1,2,2), (-113,-113,1), (151,3,3\cdot 151), (-167,-71,71\cdot 167), (193,3\cdot 71\cdot 193, 3\cdot 71),$$

which we enumerate as $\vec{e}_1$ to $\vec{e}_5$. The 2-torsion points correspond here to $\vec{e}_2 + \vec{e}_5$ and $\vec{e}_3 + \vec{e}_4$ (and their sum). It turns out that the 2-covering associated to $\vec{e}_1 + \vec{e}_5$ lifts to a 4-covering, as do its translates by 2-torsion points; the other non-2-torsion element in the 2-Selmer group do not lift to everywhere locally soluble 4-coverings.

We can then select two of these 2-Selmer elements, and find a twist E_d such that both of them are again everywhere locally soluble. It is perhaps easiest to describe this in terms of matrices (compare [21, §8.3]). Writing M for the 2-Selmer pairing matrix of E, we then have the pairing matrix of E_d as $M_d = \left(\begin{smallmatrix} M & A_d \\ * & * \end{smallmatrix}\right)$, where A_d has

[32] Smith has a condition that $(\vec{m}, \vec{m}')$ not be "alternating", but I think that is simply the happenstance where α is square (so that $L = K$), and in such a case, the Artin symbol is trivial.

[33] Note there is (perhaps curiously) no direct dependence on E in the formula here; indeed it only intervenes indirectly in the requirement that $\vec{m}, \vec{m}'$ be 2-Selmer elements of E_{d_1} and E_{d_2}.

two columns for each prime dividing d (which is positive and coprime to Ω). With $\vec{v}$ a vector in the kernel of M, we see $(\vec{v}\,\vec{0})$ will be in the kernel of M_d when $\vec{v}A_d = 0$, noting here that both $\vec{v}$ and $(\vec{v}\,\vec{0})$ yield the same $\vec{m}$ (as a 3-tuple).

Smith's result then states that for such d where both $\vec{m}$ and $\vec{m}'$ are 2-Selmer elements for E_d (and of course E itself) we have[34]

$$\langle \vec{m}, \vec{m}' \rangle_1^{\mathrm{CT}} = \langle \vec{m}, \vec{m}' \rangle_d^{\mathrm{CT}} + \left[\frac{L/K}{d}\right]$$

where L/K is derived from $\vec{m}$ and $\vec{m}'$ as above.

Let us take $\vec{m} = (-113, -226, 2)$ associated to $\vec{e}_1 + \vec{e}_2$ and $\vec{m}' = (151, 3, 453)$ associated to $\vec{e}_3$, so that $\langle \vec{m}, \vec{m}' \rangle_1^{\mathrm{CT}} = 0$ since $\vec{m}$ lifts to a 4-covering for E.

The field K here is $\mathbf{Q}(\sqrt{-113}, \sqrt{2}, \sqrt{3}, \sqrt{151})$, and we have the nontrivial Hilbert symbol $(-113, 151)_{113} = (-226, 3)_{113} = -1$, with $(-113, 151)_2 = (2, 453)_2 = -1$ too, and finally $(-226, 3)_3 = (2, 453)_3 = -1$; whereupon multiplying $\vec{m}'$ by $b = 3$ then makes these all be $+1$. By solving conics we find a suitable choice of α is $(265 + 74\sqrt{-113})(1 + 0\sqrt{-226})(27 + 17\sqrt{2})$.

We consider $d = p$ to be prime for simplicity, and then find that $\vec{m}$ is a 2-Selmer element for E_p precisely when p has $(-1|p) = (2|p) = (113|p) = +1$, while $\vec{m}'$ is a 2-Selmer element for E_p precisely when $(-1|p) = (2|p) = (3|p) = (151|p) = +1$. The set of such p such that $\left[\frac{L/K}{p}\right]$ is nontrivial (so that p splits no more in L that in K) is $\{241, 313, 1033, 1129, 1777, 2161, 2617, 2857, 3049, 3673, 3793\ldots\}$, which is that same set of such p such that $\langle \vec{m}, \vec{m}' \rangle_p^{\mathrm{CT}} = 1$. In particular, we find that $\vec{m}$ and $\vec{m}'$ do not lift to 4-coverings for E_p. Similarly, the set of such p such that $\left[\frac{L/K}{p}\right]$ is trivial is $\{97, 601, 1009, 2953, 4153, 4177, 4561, \ldots\}$, which is the same as the p with $\langle \vec{m}, \vec{m}' \rangle_p^{\mathrm{CT}} = 0$. Here we cannot yet conclude anything about the lifting of $\vec{m}$ or $\vec{m}'$ to a 4-covering, as the Cassels-Tate pairing with the other 2-coverings of E_p must also be considered.[35]

The Magma code given in Figure 1 demonstrates how to perform some of the relevant computations for this example. Via suitable modifications, the reader can verify a similar computation when starting from $\vec{m}$ and $\vec{m}'$ with $\langle \vec{m}, \vec{m}' \rangle_1^{\mathrm{CT}} = 1$, where the primes with $\left[\frac{L/K}{p}\right]$ nontrivial (with L/K defined from $\vec{m}, \vec{m}'$) are then those that have $\langle \vec{m}, \vec{m}' \rangle_p^{\mathrm{CT}} = 0$.

13.4.3. Smith's formula (3) gives a difference (or sum, as it is in $\mathbf{F}_2$) of two 4-Selmer pairings in terms of an Artin symbol, but we really want is a relation for a sum over eight such pairings; similar to the Rédei case this multi-fold version of the sum will lead to a simplification of the fields that are involved.

The algebraic input that Smith ultimately uses is embodied in Theorem 2.9 and the second part of Proposition 3.6 of [15], though we only need the particular case of the 4-Selmer pairing (rather than general higher 2^k-Selmer pairings).

[34] One can contrast this to the 8-class case, where we have the pairing for d itself in terms of an Artin symbol; essentially therein we don't have the "base point" that here corresponds to the 1st quadratic twist (that is, E itself), which indeed can only be evaluated by other means.

[35] In fact, either by applying `CasselsTatePairing` with a basis for the `TwoSelmerGroup` or using `FourDescent` on the 2-covering we find that: $\vec{m}$ lifts to a 4-covering for $p \in \{97, 601, 2953, \ldots\}$ and does not for $p \in \{1009, 4153, 4177, 4561, \ldots\}$, while $\vec{m}'$ lifts for $p \in \{97, 1009, 4153, \ldots\}$ and does not for $p \in \{601, 2953, 4177, 4561, \ldots\}$.

```
> Qx<x>:=PolynomialRing(Rationals()); KS:=KroneckerSymbol;
> a:=113*193; b:=151*167; E:=EllipticCurve([0,a+b,0,a*b,0]);
> C3d:=func<e,d|TwoCover(quo<Qx|x^3+(a+b)*d*x^2+(a*b)*d^2*x>!e)>;
> Cd:=func<u,d|C3d(CRT([Qx|u[1],u[2],u[3]],[x,x+a*d,x+b*d]),d)>;
> BASIS:=[[1,2],[-113,-113],[151,3],[-167,-71],[193,3*71*193]];
// check that the 5 given BASIS elements are indeed ELS for E
> IsELS:=func<C|IsLocallySoluble(GenusOneModel(C))>;
> assert &and[IsELS(Cd([s[1],s[2],s[1]*s[2]],1)) : s in BASIS];
// now create all the non-trivial 2-covers, compare to TwoDescent
> m:=func<U|[&*[u[1] : u in U],&*[u[2] : u in U],&*[u[3] : u in U]]>;
> VV:=Subsets(Set([[b[1],b[2],b[1]*b[2]] : b in BASIS]));
> COVERS2:=[Cd(m(v),1) : v in VV | #v ne 0]; // (2^5-1) of these
> T:=TwoDescent(E); assert Set(T) eq Set(COVERS2);

//////////////////////////////////////////////////////////

// now do the computation for m1 and m2 and twists by appropriate p
> m1:=[-113,-226,2]; m2:=[151,3,453];
> P:=PrimesUpTo(10000);
> PP:=[p : p in P | KS(-1,p) eq 1 and KS(2,p) eq 1 and KS(3,p) eq 1
                    and KS(151,p) eq 1 and KS(113,p) eq 1];
> assert &and[IsELS(Cd(m1,p)) and IsELS(Cd(m2,p)) : p in PP];
> [p : p in PP | CasselsTatePairing(Cd(m1,p),Cd(m2,p)) eq 0];
// [ 97, 601, 1009, 2953, 4153, 4177, 4561, 5953, 7417, 7489, ...]
> [p : p in PP | CasselsTatePairing(Cd(m1,p),Cd(m2,p)) eq 1];
// [ 241, 313, 1033, 1129, 1777, 2161, 2617, 2857, 3049, 3673, ...]

//////////////////////////////////////////////////////////

// now do the computation with the Artin symbols
> K<sn113,s2,s3,s151>:=NumberField([x^2+113,x^2-2,x^2-3,x^2-151]);
> RationalPoint(Conic([1,-m1[1],-3*m2[1]])); // (265 : -74 : 39)
> RationalPoint(Conic([1,-m1[2],-m2[2]/3])); // (-1 : 0 : 1)
> RationalPoint(Conic([1,-m1[3],-m2[3]/3])); // (27 : -17 : 1)
> alpha:=(265+74*sn113)*(27+17*s2); // signs don't matter
> AK:=AbsoluteField(K); ZK:=Integers(AK);
> AL:=ext<AK|Polynomial([AK!-alpha,0,1])>; ZL:=MaximalOrder(AL);
> Norm(Discriminant(ZL)); // 2^24, which seems minimally ramified
> ANS:=[<p,#Decomposition(ZL,Factorization(p*ZK)[1][1])> : p in PP];
> [a[1] : a in ANS | a[2] eq 1]; // the primes that don't split in L
// [ 241, 313, 1033, 1129, 1777, 2161, 2617, 2857, 3049, 3673, ...]
> [a[1] : a in ANS | a[2] eq 2]; // the primes that do split in L
// [ 97, 601, 1009, 2953, 4153, 4177, 4561, 5953, 7417, 7489, ...]
```

Figure 1. Magma code for Cassels-Tate computations

We shall eventually choose three hypercube co-ordinates in such a way that: given a $(\mathcal{K}_\varepsilon, \mathcal{L})$-specification, a basic character ψ on the space of alternating matrices, and a pair (m, n) at which ψ is nontrivial: the mth basis vector (of the left kernel) is $(10\,00\,00)$ at the three co-ordinates and the nth basis vector is $(00\,00\,01)$, and all the other basis vectors are $(00\,00\,00)$. Then we will be interested in sums such as

$$\sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{s}}\langle(b_i\tilde{z}, c_i, b_i\tilde{z}c_i), (b_j, c_j\tilde{s}, b_jc_j\tilde{s})\rangle^{\mathrm{CT}}_{u\tilde{z}\tilde{z}'\tilde{s}} \tag{4}$$

where $\tilde{z}, \tilde{z}', \tilde{s}$ each range over a set of 2 primes, while the b's and the c's correspond to the components of the basis vectors away from the hypercube co-ordinates, and similarly for u (which is a divisor of d). The above sum in particular corresponds to the case where $(i, j) = (m, n)$ and both $\tilde{z}$ and $\tilde{s}$ vary in the 2-Selmer elements; in other cases we will consider $\sum\sum\sum\langle(b_i\tilde{z}, c_i, b_i\tilde{z}c_i), (b_j, c_j, b_jc_j)\rangle^{\mathrm{CT}}_{u\tilde{z}\tilde{z}'\tilde{s}}$ and its partner (essentially swapping $\tilde{z}$ and $\tilde{s}$), and $\sum\sum\sum\langle(b_i, c_i, b_ic_i), (b_j, c_j, b_jc_j)\rangle^{\mathrm{CT}}_{u\tilde{z}\tilde{z}'\tilde{s}}$.

In all cases, the sum over $\tilde{z}'$ then gives a sum of 4 Artin symbols by using (3). When there is no variation of 2-Selmer elements with $\tilde{z}$ and $\tilde{s}$ these Artin symbols are all the same, so their sum is 0; similarly, when there is variation with only one of the two variables, the four symbols pair off into two copies of two, again summing to 0. On the other hand, the sum given in (4) involves Artin symbols with four (generically) distinct extensions L/K; these each depend on the b's and c's, but when taking the sum over $\tilde{z}$ and $\tilde{s}$ the field extension simplifies, and indeed becomes as in §7.7.3. Omitting the details of this calculation, we end up with L/K as a quadratic extension, with $K = \mathbf{Q}(\sqrt{z_1z_2}, \sqrt{z_1'z_2'})$ and $L = \phi[z_1z_2, z_1'z_2']$ a minimally ramified dihedral octic field, with the sum in (4) then equal to the Artin symbol of s_1s_2 for L/K. (It is thus the Rédei symbol $[z_1z_2, z_1'z_2', s_1s_2]$).

Let us give an example to illustrate this. We let E be $y^2 = x^3 - x$, and consider the twist by $d = 17 \cdot 41 \cdot 73 \cdot 97 \cdot 113 \cdot 193 \cdot 401$ as an exemplar. We will not deal *per se* with this d, but illustrate how to include it in a hypercube.[36]

With the components corresponding to the ordering $(-d, 0, d)$ of the roots of the cubic, we find that $(73\cdot113, 1, 73\cdot113)$ and $(1, 17\cdot73\cdot97, 17\cdot73\cdot97)$ are two examples of 2-Selmer elements. More abstractly with formal symbols, we require d to have 7 prime divisors all of which are 1 mod 8, and fix $(\dot{p}_3, \dot{p}_4, \dot{p}_6, \dot{p}_7) = (73, 97, 193, 401)$. We require the congruence conditions

$$(73|\dot{p}_1) = (97|\dot{p}_1) = (193|\dot{p}_1) = (401|\dot{p}_1) = -1$$
$$(73|\dot{p}_2) = (401|\dot{p}_2) = +1,\ (97|\dot{p}_2) = (193|\dot{p}_2) = -1$$
$$(73|\dot{p}_5) = (193|\dot{p}_5) = -1,\ (97|\dot{p}_5) = (401|\dot{p}_5) = +1$$

and the Legendre conditions $(\dot{p}_1|\dot{p}_2) = -1$, $(\dot{p}_1|\dot{p}_5) = -1$, and $(\dot{p}_2|\dot{p}_5) = +1$.

The above 2-Selmer elements are then given by $(\dot{p}_3\dot{p}_5, 1, \dot{p}_3\dot{p}_5) = (73\dot{p}_5, 1, 73\dot{p}_5)$ and $(1, \dot{p}_1\dot{p}_3\dot{p}_4, \dot{p}_1\dot{p}_3\dot{p}_4) = (1, 73\cdot97\dot{p}_1, 73\cdot97\dot{p}_1)$, with $(\dot{p}_1, \dot{p}_2, \dot{p}_5) = (17, 41, 113)$.

We then want to take another specialization of $(\dot{p}_1, \dot{p}_2, \dot{p}_5)$ that meets the above congruence and Legendre conditions; but we also require that the Legendre conditions are met for all 2^3 specializations of $\dot{p}_1$, $\dot{p}_2$, and $\dot{p}_5$.

Specifically, writing the specializations of $(\dot{p}_1, \dot{p}_2, \dot{p}_5)$ as (z_1, z_1', s_1) and (z_2, z_2', s_2), we require that $(z_i|z_j') = -1$, $(z_i|s_k) = -1$, and $(z_j'|s_k) = +1$ for all $i, j, k \in \{1, 2\}$,

[36] This choice is "non-generic" for various reasons (for instance, in general all the prime divisors of d are not 1 mod 8 as here), but it still demonstrates the formula for the desired 8-fold sum.

which is a total of 12 conditions. This requirement then ensures locally solubility, so that the 8-fold sum of Cassels-Tate pairings over the specializations makes sense; in Table 3 we give some examples, with the final column listing the Artin symbols (as $\mathbf{F}_2$-elements) for s_1 and s_2 in the L/K extension defined via $\phi[z_1z_2, z_1'z_2']/\mathbf{Q}(\sqrt{z_1z_2}, \sqrt{z_1'z_2'})$, while the penultimate column catalogues the 8 Cassels-Tate pairings (again as $\mathbf{F}_2$-elements, and ordered with s_1 corresponding to the first 4 entries). In each row the entries in these two columns have the same $\mathbf{F}_2$-sum, as indeed we have (using the Rédei symbol for notational convenience)

$$\sum_{\tilde z}\sum_{\tilde z'}\sum_{\tilde s}\langle(73\tilde z, 1, 73\tilde z), (1, 73\cdot 97\tilde s, 73\cdot 97\tilde s)\rangle^{\mathrm{CT}}_{73\cdot 97\cdot 193\cdot 401\cdot \tilde z\tilde z'\tilde s} = [z_1z_2, z_1'z_2', s_1s_2].$$

z_1	z_2	z_1'	z_2'	s_1	s_2		
17	98009	41	47969	113	38153	11111010	11
17	98009	41	40177	113	13033	11000001	10
17	46049	41	28393	113	59473	11100101	01
17	12809	41	47969	113	2953	10101001	00
22369	45289	5233	81761	64577	21577	10100100	01
96857	5737	98057	93113	60889	1153	11101111	10
70241	14177	9137	41521	37409	47497	10011000	10
22369	4201	66161	32993	70313	10169	01101000	01
18097	28513	58657	16553	60649	71633	11010001	11
62873	87313	5393	54713	4409	4073	11011001	10
89113	5801	83009	15937	19073	76081	01110011	01
22433	96097	26321	80209	21577	69809	10100000	11

TABLE 3. 8-fold sums of Cassels-Tate pairings, and Artin symbols

13.5. The above calculation with (4) then allows us to set up hypercube co-ordinates. Given a specification $(\mathcal{K}_\varepsilon, \mathcal{L})$ we fix a basis for the kernel of $\mathbf{M}_E(\mathcal{K}_\varepsilon, \mathcal{L})$. In fact, we do so modulo the 2-torsion elements; upon choosing the components at the primes dividing d in an appropriate way, the 2-torsion elements will be the same at each such prime, this sameness being one of 00, 01, 10, or 11. Letting s_2 be the size of the basis of the kernel modulo torsion, we then take a nontrivial multiplicative character ψ on the space of alternating matrices of size s_2 over $\mathbf{F}_2$. There is then some entry $(h^{\mathrm{z}}_\psi, h^{\mathrm{s}}_\psi)$ at which this character is nontrivial. We then take 3 hypercube co-ordinates: the first is z_ψ, at which the (h^{z}_ψ)th basis vector is 10 and all other basis vectors are 00; the second is s_ψ, at which the (h^{s}_ψ)th basis vector is 01 and all others are 00; and the third is z'_ψ where all are 00. Again we give a pictorial representation.[37]

$$\begin{array}{rccccccccccc}
 & & & 00 & & & 00 & & & 00 & & \\
 & & & 00 & & & 00 & & & 00 & & \\
(h^{\mathrm{z}}_\psi)\text{th basis vector} & 10 & 11 & 10 & 11 & 01 & 00 & 10 & 11 & 00 & 00 & 10 \\
 & & & 00 & & & 00 & & & 00 & & \\
(h^{\mathrm{s}}_\psi)\text{th basis vector} & 11 & 01 & 00 & 00 & 10 & 00 & 11 & 10 & 01 & 01 & 11 \\
 & & & 00 & & & 00 & & & 00 & & \\
 & & & \mathrm{z}_\psi & & & \mathrm{z}'_\psi & & & \mathrm{s}_\psi & &
\end{array}$$

An argument as in §6 then shows that for generic cases we can choose such hypercube co-ordinates as desired (including with respect to size constraints).

[37] As with Footnote 26, the conditions on S and T in Smith's Theorem 2.9 compared to his Definition 3.4(3) can be reconciled via a translation by torsion.

We then can bound the effect of non-generic $(\mathcal{K}_\varepsilon, \mathcal{L})$, including Swinnerton-Dyer's definition of this in the mix. There are some technicalities with passing from symmetric $\mathbf{M}_E(\mathcal{K}_\varepsilon, \mathcal{L})$ matrices defined from Legendre (or Hilbert) symbols to a suitable distribution of how often a given vector appears in a kernel, but we should again find that the non-generic cases contribute negligibly.

We then have an analogue of Lemma 7.7.1, namely that

$$\partial_{\check{Z}\times S}\bar{\psi}^\star\big((z_1, z_1', s_1), (z_2, z_2', s_2)\big) = [s_1s_2, z_1'z_2', z_1z_2],$$

and we define $L_{\check{Z}}$ as in §7.7.3 as

$$L_{\check{Z}} = \prod_{k=2}^{B}\prod_{l=2}^{B}\phi[p_1p_k, p_1'p_l']$$

where p_k come from Z and p_l' from Z', and $\phi[a,b]$ is a field from $\mathcal{F}_{a,b}^{\mathrm{mr}}$ as in §4.5.1. We take $K_{\check{Z}}$ as the largest multiquadratic extension of $\mathbf{Q}$ inside $L_{\check{Z}}$, and note that any prime in $T_l(\mathcal{K})[\mathcal{L}|\check{Z}]$ splits completely in $K_{\check{Z}}$.

The rest of the proof of Theorem 13.1.1 is then a rehash of Sections 8-11.

13.6. Although I must admit to not fully understanding the situation, let us make a few preliminary comments about the higher pairings (both for Selmer groups and class groups).

13.6.1. Let us first discuss the Selmer case, starting with geometry of 2^k-coverings.

In general, an m-covering of E is a genus 1 curve that is everywhere locally soluble and isomorphic to E over $\bar{\mathbf{Q}}$, with the isomorphism fitting into a commutative diagram with the multiplication-by-m map. (For $m \geq 4$, it turns out that one can realize an m-covering as an intersection of $m(m-3)/2$ quadrics in $\mathbf{P}^{m-1}$).

In our case of the 2^k-coverings (of E_d) we have maps

$$\begin{tikzcd}
E_d \arrow[r, "{[2]}"] & \cdots \arrow[r, "{[2]}"] & E_d \arrow[r, "{[2]}"] & E_d \arrow[r, "{[2]}"] & E_d \arrow[r, "{[2]}"] & E_d \\
\mathcal{C}_{2^k} \arrow[u, "\iota_{2^k}"] \arrow[r] & \cdots \arrow[r] & \mathcal{C}_8 \arrow[u, "\iota_8"] \arrow[r] & \mathcal{C}_4 \arrow[u, "\iota_4"] \arrow[r] & \mathcal{C}_2 \arrow[u, "\iota_2"] \arrow[ur] &
\end{tikzcd}$$

where the ι are isomorphisms over $\bar{\mathbf{Q}}$.

Swinnerton-Dyer [19] generalizes the Cassels pairing (on everywhere locally soluble 2-coverings) to a pairing for everywhere locally soluble 2^k-coverings. This can be viewed in various ways, for instance as a pairing $\langle\mathcal{C}_{2^k}, \mathcal{C}_{2^k}'\rangle$ directly on such 2^k-coverings; however, as the pairing-value only depends on the underlying 2-coverings, one can notate it as $\langle\mathcal{C}_2, \mathcal{C}_2'\rangle^{\mathrm{CT}_{2^{k+1}}}$ for 2-coverings $\mathcal{C}_2$ and $\mathcal{C}_2'$ that each lift to a 2^k-covering; or as $\langle\mathcal{C}_{2^k}, \mathcal{C}_2'\rangle$ where again $\mathcal{C}_2'$ is a 2-covering that lifts to a 2^k-covering. The import of this pairing is: a 2^k-covering $\mathcal{C}_{2^k}$ lifts to a 2^{k+1}-covering exactly when $\langle\mathcal{C}_{2^k}, \mathcal{C}_2'\rangle$ is trivial for all 2-coverings $\mathcal{C}_2'$ (see [19, Theorem 1, p. 719]).

Similar to (4), one might hope that a sum of 2^{k+2} pairings of 2^k-coverings might be writable in terms of an Artin symbol. For instance, for the 8-Selmer pairing, we might hope to write

$$\sum_{\tilde{z}}\sum_{\tilde{z}'}\sum_{\tilde{z}''}\sum_{\tilde{s}}\langle(b_i\tilde{z}, c_i, b_i\tilde{z}c_i), (b_j, c_j\tilde{s}, b_jc_j\tilde{s})\rangle_{u\tilde{z}\tilde{z}'\tilde{z}''\tilde{s}}^{\mathrm{CT}_8}$$

in terms of an Artin symbol for s_1s_2 in an extension of $\mathbf{Q}(\sqrt{z_1z_2}, \sqrt{z_1'z_2'}, \sqrt{z_1''z_2''})$. (Here $(b_i\tilde{z}, c_i, b_i\tilde{z}c_i)$ and $(b_j, c_j\tilde{s}, b_jc_j\tilde{s})$ are 2-coverings that each lift to a 4-covering, and indeed do so for all 16 values of $u\tilde{z}\tilde{z}'\tilde{z}''\tilde{s}$).

However, as Smith [15] notes after Proposition 4.3 on page 39, there are extra conditions that should be met for such a formula to follow.[38] Koymans and Pagano [9] give a similar discussion with their (8.5), noting that their Theorems 7.7 and 7.8 give such multi-fold sums in terms of Artin symbols, and then after subsequent comments on page 48 (after the proof of Lemma 8.5) refer the matter to their Lemma 13.10 for specifics.

In brief, following the bullet points in Smith's Definition 3.5 there are conditions involving: minimality (defined at the top of page 17), agreement (bottom of page 18), and acceptable ramification (middle of page 33); each of which involve notions of consistency (middle of page 16).[39] An important feature is that all three of these are additive conditions, and one can control the size of the image (as a vector space over $\mathbf{F}_2$) of a map whose kernel contains the desired hypercubes on which the sum can be evaluated as an Artin symbol; moreover, this control over the drop in density from evaluable hypercubes compared to all hypercubes is sufficiently tight to allow grids to be applied as in §10.

13.6.2. The class group case is not too different from the Selmer case, though there is the distinction between bases of the left and right kernels of the pairing matrices.

In either case, once a suitably wide enough class of hypercubes is found where the multi-fold sum can be evaluated in terms of an Artin symbol, the arguments given in §§8-11 then need only comparatively minor modifications to prove an equidistribution result for character values on the matrix spaces of the applicable pairing, whereupon the distribution of the Selmer or narrow class group follows readily.

13.6.3. Let us try to give some idea of what the above conditions on hypercubes mean. I thank Peter Koymans for his comments on this. Firstly, although Smith has written the conditions in terms of co-cycles with Galois co-homology, the nomenclature already hints at a version involving fields.[40] For instance, in the 8-class case we can consider a datum $[\{p_1, p_2\}, \{q_1, q_2\}, s, u]$ where χ_s is a quadratic character that is 2-divisible in each of the four dual class groups of $\mathbf{Q}(\sqrt{p_i q_j s u})$. If we consider the four conics $X^2 = \tilde{p}\tilde{q}uY^2 + sZ^2$ where $\tilde{p}$ ranges over $\{p_1, p_2\}$ and similarly with $\tilde{q}$, we find that if any three of the conics are soluble then all four are; in terms of fields, this is equivalent to the existence of a collection of minimally ramified dihedral extensions (as in §4.5.1) involving $\mathbf{Q}(\sqrt{a}, \sqrt{s})$ for each $a \in \{p_1q_1u, p_1q_2u, p_2q_1u, p_2q_2u\}$, such that the relative degree of the compositum has degree 2^3 over $\mathbf{Q}(\sqrt{p_1q_1u}, \sqrt{p_1q_2u}, \sqrt{p_2q_1u}, \sqrt{p_2q_2u}, \sqrt{s})$, rather than degree 2^4 as would generically be expected.[41] In particular, due to the local-global

[38] One can note that in the 4-Selmer or 8-class cases the additional conditions are met: indeed, in Smith's Definition 3.5 we only consider S with (at most) one element, so the only proper subset is empty, with the condition that $\hat{x}(\emptyset)$ have nonempty intersection with $\overline{Y}^{\circ}_{\emptyset}$ then being immediate.

[39] Meanwhile, Koymans and Pagano cover minimality and agreement in §7.3; acceptable ramification in Lemma 13.10; and consistency in §7.4, though from I can tell they use a somewhat different notion therein.

[40] Note that Koymans and Pagano [9] need to keep track of degree l cyclic extensions for l an odd prime, wherein the quadratic case has a simpler correspondence between extensions and characters. See also their comments in the Remark after Corollary 4.12.

[41] Note here that $\mathbf{Q}(\sqrt{p_1q_1u}, \sqrt{p_1q_2u}, \sqrt{p_2q_1u}, \sqrt{p_2q_2u}, \sqrt{s})$ is not "independent", and only has degree 2^4 over $\mathbf{Q}$; moreover, as Koymans indicates, one could enlarge this multi-quadratic field to $\mathbf{Q}(\sqrt{\hat{p}_1}, \sqrt{\hat{p}_2}, \sqrt{\hat{q}_1}, \sqrt{\hat{q}_2}, \sqrt{s}, \sqrt{\hat{u}})$, with the separation of p's and q's corresponding to including the effect of Stevenhagen's twisting subgroup [18, (43)], which we briefly mentioned in §4.5.1.

principle for conics, we see that in this case Smith's minimality criterion follows immediately from local conditions.

We next translate this into co-cycles. We consider maps from $G_{\mathbf{Q}} = \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ to the 2-power-torsion of modules N_b, which are given by $\mathbf{Q}_2/\mathbf{Z}_2$ twisted by χ_b.[42] In particular (see [15, §2.2]), for any b_1 and b_2 we have $N_{b_1}[2] = N_{b_2}[2]$, and there is an isomorphism $\beta[b_1, b_2]$ between N_{b_1} and N_{b_2} that preserves the Galois structure above $\mathbf{Q}(\sqrt{b_1 b_2})$.[43] Recall that a co-cycle is a map f from a group G to a G-module such that $f(\sigma\tau) = f(\sigma) + \sigma f(\tau)$ for all $\sigma, \tau \in G$; notably, since $1/2 = -1/2$ in $\mathbf{Q}_2/\mathbf{Z}_2$, the co-cycles from $G_{\mathbf{Q}}$ to $N_b[2]$ are precisely the quadratic characters χ_c, where such a character sends $\sigma \in G_{\mathbf{Q}}$ to $1/2$ when it factors through $\mathbf{Q}(\sqrt{c})/\mathbf{Q}$.

Smith then notes (Proposition 2.7) that a 2-torsion element χ_s in the dual class group for $\mathbf{Q}(\sqrt{w})$ lifts to one of order 2^k precisely when there is a co-cycle $\psi_{2^k,w}$ from $\mathrm{Gal}(H_w/\mathbf{Q})$ to $N_w[2^k]$ such that $2^{k-1}\psi_{2^k,w} = \chi_s$ (the multiplication by 2^{k-1} moves the image to $N_w[2]$), where H_w is the Hilbert class field of $\mathbf{Q}(\sqrt{w})$. In the above scenario with the datum $[\{p_1, p_2\}, \{q_1, q_2\}, s, u]$, we take each $\psi_{4,\tilde{p}\tilde{q}su}$ to be a co-cycle from $G_{\mathbf{Q}}$ to $N_{\tilde{p}\tilde{q}su}[4]$ with $2\psi_{4,\tilde{p}\tilde{q}su} = \chi_s$, and minimality is then equivalent to $\sum_{\tilde{p}} \sum_{\tilde{q}} \beta[\tilde{p}\tilde{q}su, p_1q_1su] \circ \psi_{4,\tilde{p}\tilde{q}su}$ being a quadratic character.[44] From this one sees that minimality always holds: for σ with $\chi_s(\sigma) = 1/2$ we have $\psi_{4,\bullet}(\sigma) \in \{1/4, 3/4\}$, and the sum of four such values is in $\{0, 1/2\}$; the same conclusion again holds true when $\chi_s(\sigma) = 0$, implying the image is in $N[2]$ so that the 4-fold co-cycle sum is a quadratic character.

One way Smith uses this minimality condition is apparent in his Proposition 2.5; if we consider a datum $[\{p_1, p_2\}, \{q_1, q_2\}, \{r_1, r_2\}, s, u]$ and selection of co-cycles $\psi_{4,\bullet}$ to $N_{\tilde{p}\tilde{q}\tilde{r}su}[4]$ satisfying $2\psi_{4,\tilde{p}\tilde{q}\tilde{r}su} = \chi_s$ such that the above 4-fold sum is the trivial character for all 6 subcubes (with u notationally absorbing the fixed co-ordinate), then (assuming that χ_s is 4-divisible in each dual class group) upon taking co-cycle lifts $\psi_{8,\bullet}$ such that $2\psi_{8,\tilde{p}\tilde{q}\tilde{r}su} = \psi_{4,\tilde{p}\tilde{q}\tilde{r}su}$, the conclusion of Proposition 2.5 is that the 8-fold sum $\sum_{\tilde{p}} \sum_{\tilde{q}} \sum_{\tilde{r}} \beta[\tilde{p}\tilde{q}\tilde{r}su, p_1q_1r_1su] \circ \psi_{8,\tilde{p}\tilde{q}\tilde{r}su}$ is a quadratic character. (Such a fact about an 8-fold sum would then be used in the 16-class case).

As for how minimality affects the pairing-sum over a suitable hypercube: in a consideration of the 16-class case we want a sum such as

$$\sum_{\tilde{p}} \sum_{\tilde{q}} \sum_{\tilde{r}} \langle \chi_s, I_t \rangle^{\mathrm{Art}_{16}}_{\tilde{p}\tilde{q}\tilde{r}su}$$

to be 0, where I_t is an ideal of norm t that is 4-divisible in each class group (and χ_s is 4-divisible in each dual). The relevant minimality condition for this can be sketched as follows: over each $\mathbf{Q}(\sqrt{\tilde{p}\tilde{q}\tilde{r}su})$ we have a dihedral/$\mathbf{Q}$ extension (with suitable ramification) of relative degree 8 that contains $\mathbf{Q}(\sqrt{s})$; taking the base field K as the multi-quadratic field (of degree 32) containing the $\mathbf{Q}(\sqrt{\tilde{p}\tilde{q}\tilde{r}u}, \sqrt{s})$, these then yield eight extensions $L_{\tilde{p}\tilde{q}\tilde{r}}/K$ with each being cyclic of degree 4 and

[42] The action of $\sigma \in G$ on N_b thus has $\sigma x = -x$ when $\sigma(\sqrt{b}) = -\sqrt{b}$, and $\sigma x = x$ otherwise. (In the Selmer case the modules are twists of $E[2^\infty]$, thus $(\mathbf{Q}_2/\mathbf{Z}_2)^2$ when forgetting the action, but otherwise much the same). Also, the infinite Galois group is a profinite group defined as a inverse limit, and the image under the quotient homomorphism to the extension $\mathbf{Q}(\sqrt{b})/\mathbf{Q}$ determines whether $\sigma(\sqrt{b}) = -\sqrt{b}$ or not.

[43] On the other hand, if one forgets the G-action, this map is just the identity map on $\mathbf{Q}_2/\mathbf{Z}_2$, so the notation could be considered to be a bit superfluous (at least in the class group case).

[44] Again one can be more strict, ensuring that the sum is trivial via translating the ψ_4-lifts by suitable quadratic characters, again corresponding to Stevenhagen's twisting subgroup.

suitably ramified; generically the compositum structure of these eight fields should have no nontrivial containments, whilst Smith's minimality criterion implies that any one of the eight is contained in the compositum of the other seven. (One also requires the same to be true when s is swapped with t).

Unlike the case with 4-class pairings, there is no reason here for minimality to follow simply from local solubility conditions involving the primes in question. Perhaps one could hope to write the obstruction to such a compatible selection of extensions in terms of a Brauer group element, which would then clarify the additive nature of the condition; indeed, we again should stress that a crucial feature of such conditions is that they are additive, and so Smith's combinatorial methods are applicable upon bounding the image of a map to a vector space over $\mathbf{F}_2$.

References

[1] E. Artin, *Zur Theorie der L-Reihen mit allgemeinen Gruppencharakteren.* (German) [On the theory of L-series for general group characters]. Abh. Hamburg **8** (1931), 292–306. http://doi.org/10.1007/BF02941010

[2] J. W. S. Cassels, *Arithmetic on curves of genus 1, IV. Proof of the Hauptvermutung.* J. reine angew. Math. **211** (1962) 95–112. http://eudml.org/doc/150551

[3] ———, *Second descents for elliptic curves. ibid.* **494** (1998) 101–127. http://doi.org/10.1515/crll.1998.001

[4] S. Chan, P. Koymans, D. Milovic, C. Pagano, *On the negative Pell equation.* Preprint, 2019.

[5] T. A. Fisher, *Higher descents on an elliptic curve with a rational 2-torsion point.* Math. Comp. **86** (2017), no. 307, 2493–2518. http://doi.org/10.1090/mcom/3163

[6] ———, *On binary quartics and the Cassels-Tate pairing.* Preprint, 2016. http://www.dpmms.cam.ac.uk/~taf1000/papers/bq-ctp.html

[7] T. Fisher, E. F. Schaefer, M. Stoll, *The yoga of the Cassels-Tate pairing.* LMS J. Comput. Math. **13** (2010), 451–460. http://doi.org/10.1112/S1461157010000185

[8] É. Fouvry, J. Klüners, *On the negative Pell equation.* Ann. Math. **172** (2010), no. 3, 2035–2104. http://doi.org/10.4007/annals.2010.172-3

[9] P. Koymans, C. Pagano, *On the distribution of* $\mathrm{Cl}(K)[l^\infty]$ *for degree l cyclic fields.* Preprint, 2018.

[10] ———, *Higher Rédei reciprocity and integral points on conics.* Preprint, 2020.

[11] G. Landsberg, *Ueber eine Anzahlbestimmung und eine damit zusammenhängende Reihe.* (German) [On a numerical determination and a related series]. J. reine angew. Math. **111** (1893), 87–88. http://eudml.org/doc/148874

[12] L. B. Pierce, C. L. Turnage-Butterbaugh, M. M. Wood, *An effective Chebotarev density theorem for families of number fields, with an application to l-torsion in class groups.* Invent. Math. **219** (2020), 701–778. http://doi.org/10.1007/s00222-019-00915-z

[13] L. Rédei, H. Reichardt, *Die Anzahl der durch 4 teilbaren Invarianten der Klassengruppe eines beliebigen quadratischen Zahlkörpers.* (German) [The number of 4-divisible invariants in the class group in any quadratic field]. J. reine angew. Math. **170** (1934), 69–74. http://eudml.org/doc/149862

[14] A. Smith, *Governing fields and statistics for 4-Selmer and 8-class groups.* Preprint, 2016.

[15] ———, *2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture.* Preprint, 2017.

[16] P. Stevenhagen, *The Number of Real Quadratic Fields Having Units of Negative Norm.* Experiment. Math. **2** (1993), no. 2, 121–136. http://doi.org/10.1080/10586458.1993.10504272

[17] ———, *Rédei-matrices and applications.* In *Number theory (Paris, 1992–1993)*, edited by S. David. London Math. Soc. Lecture Note Ser. **215**, 245–259. http://doi.org/10.1017/CBO9780511661990.015

[18] ———, *Redei reciprocity, governing fields, and negative Pell.* Preprint, 2020.

[19] P. Swinnerton-Dyer, *2^n-descent on elliptic curves for all n.* J. London Math. Soc. **87** (2013), 707–723. http://doi.org/10.1112/jlms/jds063

[20] J. Thorner, A. Zaman, *A zero density estimate for Dedekind zeta functions.* Preprint, 2019.

[21] M. Watkins, *Distribution of the 2-Selmer rank under twisting.* Preprint, 2020.