AVERAGES AND HIGHER MOMENTS FOR THE ℓ-TORSION IN CLASS GROUPS

CHRISTOPHER FREI AND MARTIN WIDMER

ABSTRACT. We prove upper bounds for the average size of the ℓ-torsion $\mathrm{Cl}_K[\ell]$ of the class group of K, as K runs through certain natural families of number fields and ℓ is a positive integer. We refine a key argument, used in almost all results of this type, which links upper bounds for $\mathrm{Cl}_K[\ell]$ to the existence of many primes splitting completely in K that are small compared to the discriminant of K. Our improvements are achieved through the introduction of a new family of specialised invariants of number fields to replace the discriminant in this argument, in conjunction with new counting results for these invariants. This leads to significantly improved upper bounds for the average and sometimes even higher moments of $\mathrm{Cl}_K[\ell]$ for many families of number fields K considered in the literature, for example, for the families of all degree-d-fields for $d \in \{2,3,4,5\}$ (and non-D_4 if $d = 4$). As an application of the case $d = 2$ we obtain the best upper bounds for the number of D_p-fields of bounded discriminant, for primes $p > 3$.

1. INTRODUCTION

In this paper, we provide bounds for the average and higher moments of the size of the ℓ-torsion $\mathrm{Cl}_K[\ell] = \{[\mathfrak{a}] \in \mathrm{Cl}_K \,;\, [\mathfrak{a}]^\ell = [\mathcal{O}_K]\}$ of the ideal class groups of number fields K in certain families, for arbitrary $\ell \in \mathbb{N} = \{1,2,3,\dots\}$. Throughout, we order number fields K by the absolute value D_K of their discriminant. For real-valued maps f and g with common domain we mean by $f(t) \ll_a g(t)$ that there exists a positive constant $C = C(a)$, depending only on a, such that $|f(t)| \le C|g(t)|$ for all t in the domain. Throughout this article we assume $X \ge 2$. To give the reader a quick taste of the results in this paper, here is our first theorem concerning quadratic fields.

Theorem 1.1. *Let $\varepsilon > 0$ and $k \ge 0$ be real numbers and $\ell \in \mathbb{N}$. As K ranges over all quadratic number fields with $D_K \le X$ we have*

$$\sum_K \#\,\mathrm{Cl}_K[\ell]^k \ll_{\ell,k,\varepsilon} X^{\frac{k}{2}+1-\min\left\{1,\ \frac{k}{\ell+2}\right\}+\varepsilon}.$$

We now discuss an application of Theorem 1.1. For a transitive permutation group G of degree d and $X > 0$, let $N(d,G,X)$ be the number of field extensions $K/\mathbb{Q}$ of degree d within a fixed algebraic closure $\overline{\mathbb{Q}}$ with $D_K \le X$ and whose normal closure has Galois group isomorphic to G as a permutation group. Malle's conjecture [Mal02, Mal04] predicts an asymptotic formula for $N(d,G,X)$ as $X \to \infty$. Let p be an odd prime and D_p, $D_p(2p)$

Date: 6 November 2020.

1991 *Mathematics Subject Classification.* Primary 11R29, 11R65, 11R45; Secondary 11G50.

Key words and phrases. ℓ-torsion, class group, moments, Dihedral extensions, counting, small height.

the Dihedral group of order $2p$ and its regular permutation representation respectively. In these cases, Malle's conjecture predicts the formulas

$$N(p, D_p, X) \sim c_p X^{\frac{2}{p-1}} \quad \text{and} \quad N(2p, D_p(2p), X) \sim c_{2p} X^{\frac{1}{p}}$$

with positive constants c_p, c_{2p} (see [Klü06, Example after Conjecture 1.1]). Currently the best upper bounds for $p > 3$ are

$$N(p, D_p, X) \ll_{p,\varepsilon} X^{\frac{3}{p-1} - \frac{1}{p(p-1)} + \varepsilon} \quad \text{and} \quad N(2p, D_p(2p), X) \ll_{p,\varepsilon} X^{\frac{3}{2p} + \varepsilon},$$

the first due to Cohen and Thorne [CT17, Theorem 1.1], the second due to Klüners [Klü06, Theorem 2.7]. As an immediate consequence of Klüners' method and the case $k = 1$ in Theorem 1.1, we can improve both bounds for all primes $p > 3$.

Corollary 1.2. *Let p be an odd prime and $\varepsilon > 0$. Then we have*

$$N(p, D_p, X) \ll_{p,\varepsilon} X^{\frac{3}{p-1} - \frac{2}{(p+2)(p-1)} + \varepsilon} \quad \text{and} \quad N(2p, D_p(2p), X) \ll_{p,\varepsilon} X^{\frac{3}{2p} - \frac{1}{p(p+2)} + \varepsilon}.$$

The special case $p = 5$ was also considered by Larsen and Rolen [LR12]. They suggest to improve Klüners' bound $X^{0.75+\varepsilon}$ [Klü06, Theorem 2.7] by counting integral points on a variety defined by a norm equation. While counting these points seems a difficult matter, their numerical experiments provide evidence that the number of these points is $\ll X^{0.698}$, which, if true, would provide the same bound for $N(5, D_5, X)$. The exponent $0.7 + \varepsilon$ of Cohen and Thorne is just slightly above the latter. Our bound is $X^{0.678\ldots}$, and hence is slightly better than the bound suggested by the numerical experiments in [LR12].

1.1. **Background.** Let us provide here some context for Theorem 1.1 and our further results. Denote the degree of the number field K by d. Landau (see, e.g., [Nar80, Theorem 4.4]) noticed that that the Minkowski bound implies the upper bound

$$\# \operatorname{Cl}_K \ll_{d,\varepsilon} D_K^{\frac{1}{2}+\varepsilon}, \tag{1.1}$$

for arbitrarily small $\varepsilon > 0$. This bound is essentially sharp, and provides the "trivial" upper bound for the ℓ-torsion

$$\# \operatorname{Cl}_K[\ell] \ll_{d,\varepsilon} D_K^{\frac{1}{2}+\varepsilon}. \tag{1.2}$$

However, a standard conjecture asserts that

$$\# \operatorname{Cl}_K[\ell] \ll_{d,\ell,\varepsilon} D_K^{\varepsilon}. \tag{1.3}$$

For some references providing motivation and background for this conjecture, we refer to [PTBW19, Conjecture 1.1] and the discussion thereafter. The conjecture for $d = \ell = 2$ follows from Gauß' genus theory. Since $\# \operatorname{Cl}_K[\ell^t] \leq \# \operatorname{Cl}_K[\ell]^t$ (consider the homomorphism $[\mathfrak{a}] \to [\mathfrak{a}]^\ell$ from $\operatorname{Cl}_K[\ell^t]$ to $\operatorname{Cl}_K[\ell^{t-1}]$) the conjecture also holds true for $(d, \ell) = (2, 2^t)$ and arbitrary $t \in \mathbb{N}$ (see [PTBW19, Section 7.1]). Apart from that the only cases of primes ℓ for which improvements over the trivial bound have been established are $\ell = 3$ for $d \leq 4$ by pioneering work of Pierce, Helfgott, Ellenberg and Venkatesh [Pie05, Pie06, HV06, EV07], and more recently the case $\ell = 2$ for arbitrary d by Bhargava et al. [BST+17]. As noted, again in [PTBW19, Section 7.1], the improvements for $(d, \ell) = (2, 3)$ hold more generally for $(d, \ell) = (2, 3 \cdot 2^t)$ using the fact that $\# \operatorname{Cl}_K[\ell]$ is a multiplicative function (as function of ℓ) and then combining the bounds for $\# \operatorname{Cl}_K[3]$ and $\# \operatorname{Cl}_K[2^t]$. Of course, this argument

also applies to Theorem 1.1 and shows that we could replace ℓ in the exponent on the right hand-side by its maximal odd divisor.

These are all cases (d, ℓ) for which unconditional non-trivial upper bounds for $\#\operatorname{Cl}_K[\ell]$ are known. Assuming the Riemann hypothesis for the Dedekind zeta function of the normal closure of K, Ellenberg and Venkatesh [EV07] proved the bound

$$\#\operatorname{Cl}_K[\ell] \ll_{d,\ell,\varepsilon} D_K^{\frac{1}{2}-\frac{1}{2\ell(d-1)}+\varepsilon} \tag{1.4}$$

for all number fields K. Taking up a key idea of Michel and Soundararajan and generalising it from imaginary quadratic to arbitrary number fields they show in [EV07, Lemma 2.3] that the presence of many small primes splitting completely in K leads to savings over (1.2). Together with the conditional effective version of Chebotarev's density theorem, this leads directly to the bound (1.4). Small splitting primes were also used in [AD03] to lower bound the exponent of the class group of CM-fields.

Subsequently, several papers took the same approach using [EV07, Lemma 2.3], but tried to establish the existence of enough splitting primes unconditionally, at the cost of averaging or having to exclude a zero-density subset of fields in a given family. Number field counting techniques were used in combination with probabilistic methods in [EPW17, FW18], the large sieve in [HBP17], and new effective versions of Chebotarev's density theorem in [PTBW20, An20].

In this paper, we take a different direction by refining the core argument [EV07, Lemma 2.3] itself, see Proposition 2.1. We render the argument in a form from which we then profit by playing two ways of counting number fields, by discriminant and by minimal height of certain generators, against each other. Possible refinements were already proposed in [Ell08], and a first concrete step in this direction was taken by the second author in [Wid18], leading to improvements upon [EPW17] in some cases. Our new technique yields improvements on average in all cases of [EPW17] and [Wid18] (provided ℓ is not too small), as well as on some results in [EV07, PTBW20, An20]. For example, when $\ell > 2$, the case $k = 1$ in Theorem 1.1 improves the case $d = 2$ of [EPW17, Corollary 1.1.1], which gives an upper bound

$$\sum_K \#\operatorname{Cl}_K[\ell] \ll_{\ell,\varepsilon} X^{\frac{3}{2}-\frac{1}{2\ell(d-1)}+\varepsilon}, \tag{1.5}$$

provided $d \in \{2, 3, 4, 5\}$ and $\ell \geq \ell(d)$, where $\ell(2) = \ell(3) = 1$, $\ell(4) = 8$ and $\ell(5) = 25$.

Note that control over averages is often enough for applications, as illustrated by Corollary 1.2. Moreover, having sufficiently good upper bounds for k-th moments with arbitrarily large k would imply (1.3), as shown in [PTBW19, Theorem 1.2]. Here, sufficiently good means with an exponent on X independent of k, and valid for arbitrarily large k.

To our best knowledge, the only published results concerning higher moments are those of Heath-Brown and Pierce [HBP17] on imaginary quadratic fields. One can easily deduce bounds for arbitrary moments from a field count and pointwise results with small exceptional sets, such as those in [EPW17, PTBW20]: for a family S of degree-d-fields we write

$$S(X) = \{K \in S;\ D_K \leq X\}.$$

If all but at most $O_{S,a,b,\ell}(X^a)$ exceptional fields $K \in S(X)$ satisfy $\#\operatorname{Cl}_K[\ell] \ll_{S,a,b,\ell} D_K^{1/2-b}$, then

$$\sum_{K\in S(X)} \#\operatorname{Cl}_K[\ell]^k \ll_{S,a,b,\ell,\varepsilon,k} \#S(X)X^{k(1/2-b)} + X^{k/2+a+\varepsilon}. \tag{1.6}$$

In the following, we call this the *straightforward approach*. In Theorem 1.1 and later results, we give bounds for the k-th moment in cases where the exceptional set is known to be very small. Our bounds are stronger than (1.6) when ℓ is not too small in terms of the other parameters, in particular in terms of k.

Last but not least we should mention that there are very few but spectacular results for the averages of ℓ-torsion in degree-d-fields that provide not only upper bounds but even asymptotics. The case $(d,\ell) = (2,3)$ is due to Davenport-Heilbronn [DH71] (see also the recent improvements [BST13, TT13, Hou16]), and $(3,2)$ due to Bhargava [Bha05]. In particular, these two results show that for $(d,\ell) \in \{(2,3),(3,2)\}$ the conjecture (1.3) holds true on average. Regarding 4-torsion in quadratic fields Fouvry and Klüners [FK07] have established the average value for $\#\operatorname{Cl}_K[4]/\#\operatorname{Cl}_K[2]$. Related results were obtained by Klys [Kly16] for 3-torsion in cyclic cubic fields, and by Milovic [Mil17] for the 16-rank in certain quadratic fields.

1.2. **Further main results.** Let us next consider the other cases of [EPW17], concerning degree-d-fields for $d \in \{3,4,5\}$ (whose normal closure does not have Galois group D_4 in case $d = 4$). In this case, our result is as follows. Define $\delta_0(3) = 2/25$, $\delta_0(4) = 1/48$, and $\delta_0(5) = 1/200$.

Theorem 1.3. *Suppose $d \in \{3,4,5\}$, and $\varepsilon > 0$. As K ranges over number fields of degree d with $D_K \le X$ (and non-D_4 in the case $d = 4$), we have*

$$\sum_K \#\operatorname{Cl}_K[\ell] \ll_{\ell,\varepsilon} X^{\frac{3}{2}-\min\left\{\delta_0(d),\frac{1}{(d-1)\ell+3}\right\}+\varepsilon}.$$

This improves upon Ellenberg, Pierce, and Wood's result mentioned in (1.5) (for large enough ℓ), and moreover upon [Wid18, Corollary 1.5]. Assuming GRH, our method also works for general families S of number fields of fixed degree, but it loses its power if the families are too thin, that is, if $\#S(X) = \#\{K \in S\,;\, D_K \le X\} \ll X^\rho$ for $\rho < 1$ too small compared to the other parameters.

Theorem 1.4. *Let $\varepsilon > 0$, let S be any family of number fields of degree d, and assume that*

(i) *the Dedekind zeta function of the normal closure of each field in S satisfies the Riemann hypothesis,*
(ii) *the numbers $\rho, c_1 > 0$ are such that $\#S(X) \le c_1 X^\rho$ for all $X \ge 2$.*

Then

$$\sum_{K\in S(X)} \#\operatorname{Cl}_K[\ell]^k \ll_{d,\rho,c_1,\ell,k,\varepsilon} X^{\frac{k}{2}+\rho-\min\left\{\rho,\frac{\rho k}{(d-1)\ell+2}\right\}+\varepsilon}.$$

For comparison, an application of the straightforward approach (1.6) with the GRH-bound (1.4) from [EV07] and *no* exceptional fields yields

$$\sum_{K\in S(X)} \#\operatorname{Cl}_K[\ell]^k \ll_{d,\rho,c_1,\ell,k,\varepsilon} \#S(X)X^{\frac{k}{2}-\frac{k}{2\ell(d-1)}+\varepsilon}. \tag{1.7}$$

Taking ρ to be the smallest known value with $\#S(X) \ll_\rho X^\rho$ minimises the bound in Theorem 1.4 as well as the one from (1.7). As long as $\rho > \frac{1}{2} + \frac{1}{\ell(d-1)}$ and $k < 2\ell\rho(d-1)$, our Theorem 1.4 provides a stronger bound than (1.7), thus giving an impression of the density of S that is required for our method to yield improvements.

1.3. **Further results.** In some cases with prescribed Galois groups, our method can also work for families that are thinner than suggested above. For cyclic extensions not covered by Theorem 1.1, we are able to improve upon [FW18, PTBW20] in the case $d = 3$ and, moreover, to cover higher moments using a refinement of the straightforward approach (1.6) based on Proposition 3.2.

Theorem 1.5. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. As K ranges over cubic A_3-extensions of $\mathbb{Q}$ with $D_K \leq X$, we have*

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{\ell,k,\varepsilon} X^{\frac{k+1}{2}-\min\left\{\frac{1}{2},\frac{k}{3\ell+4}\right\}+\varepsilon}.$$

For comparison, the straightforward approach (1.6) applied with the pointwise estimate from [PTBW20, Theorem 7.2] for almost all A_3-fields gives $\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{\ell,k,\varepsilon} X^{\frac{k+1}{2}-\min\{\frac{1}{2},\frac{k}{4\ell}\}+\varepsilon}$ upon which Theorem 1.5 is an improvement as long as $\ell \geq 5$ and $k < 2\ell$.

We can also get improvements in the case of quintic fields whose normal closure has Galois group D_5, the dihedral group of order 10. As already mentioned in the discussion after Theorem 1.1, no asymptotics for the counting function of these fields are known. Moreover, we need to impose the same ramification restrictions as in [PTBW20], since we rely on results from that paper to count small splitting primes. If the rational prime p ramifies tamely in a number field K whose normal closure $\tilde{K}$ has Galois group G then the inertia group $I(\mathfrak{B}) \subset G$ is cyclic for any prime ideal $\mathfrak{B} \subset \mathcal{O}_{\tilde{K}}$ lying above p. For different prime ideals $\mathfrak{B}$ over the same rational prime p these inertia groups are conjugate. Let $n > 2$ be odd and $G = D_n$, the dihedral group of symmetries of a regular n-gon of order $2n$, so that the conjugacy class of a reflection is the set of all reflections. Keeping this in mind we say that the ramification type of a tamely ramified prime p is generated by a reflection if each $I(\mathfrak{B})$ is generated by a reflection.

Theorem 1.6. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. Let S be the family of all quintic D_5-extensions of $\mathbb{Q}$ for which the ramification type of p is generated by a reflection in D_5 for every tamely ramified rational prime p. Suppose moreover that $\rho, c_1 > 0$ are such that*

$$\#S(X) = \#\{K \in S;\ D_K \leq X\} \leq c_1 X^\rho \tag{1.8}$$

holds for all $X \geq 2$. Then, as K ranges over $S(X)$, we have

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{\rho,c_1,\ell,k,\varepsilon} X^{\frac{k}{2}+\rho-\frac{12\rho k}{37\ell+24}+\varepsilon} + X^{\frac{k}{2}+\frac{1}{4}+\varepsilon}.$$

Note that, by [PTBW20, Proposition 2.3], any ρ with (1.8) must satisfy $\rho \geq 1/2$, and Malle's conjecture predicts that $\rho = 1/2$ is indeed the optimal exponent. For comparison, with the conjectured behaviour $\#S(X) \asymp X^{1/2}$, the straightforward approach (1.6) applied to [PTBW20, Theorem 7.2] would yield

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{c_1,\ell,k,\varepsilon} X^{\frac{k+1}{2}-\frac{k}{8\ell}+\varepsilon} + X^{\frac{k}{2}+\frac{1}{4}+\varepsilon}. \tag{1.9}$$

Hence, with Malle's conjectured exponent $\rho = 1/2$, our result provides improvements if $\ell > 2$ and $k < 2\ell$. Taken together, Corollary 1.2 and Theorem 1.6 immediately imply the following unconditional result with $\rho = 19/28 + \varepsilon$.

Corollary 1.7. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. Let S be the family of all quintic D_5-extensions of $\mathbb{Q}$ for which the ramification type of p is generated by a reflection in D_5 for every tamely ramified rational prime p. Then, as K ranges over $S(X)$, we have*

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{\ell,k,\varepsilon} X^{\frac{k}{2}+\frac{19}{28}-\frac{57k}{259\ell+168}+\varepsilon} + X^{\frac{k}{2}+\frac{1}{4}+\varepsilon}.$$

Compared to what one gets from the straightforward approach (1.6) using [PTBW20, Theorem 7.2] and estimating $\#S(X)$ again by Corollary 1.2, this yields an improvement whenever $k < 24\ell/7$ and $\ell \geq 2$.

Moreover, we can get improvements for certain families of quartic D_4-fields studied in very recent work of An [An20]. For distinct and squarefree $a, b \in \mathbb{Z} \smallsetminus \{0, 1\}$, we denote by $S_4(a, b)$ the family of quartic number fields whose normal closure has Galois group D_4 and contains the biquadratic field $\mathbb{Q}(\sqrt{a}, \sqrt{b})$. It is shown in [An20] that the normal closure of every D_4-field contains a unique biquadratic field, and the pairs (a, b) with $S_4(a, b) \neq \emptyset$ are classified in [An20, Condition 1.3].

Theorem 1.8. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. Let $a, b \in \mathbb{Z} \smallsetminus \{0, 1\}$ be distinct and squarefree such that $S_4(a, b) \neq \emptyset$. Suppose moreover that $\rho, c_1 > 0$ are such that*

$$\#\{K \in S_4(a, b);\ D_K \leq X\} \leq c_1 X^\rho \tag{1.10}$$

holds for all $X \geq 2$. Then, as K ranges over the fields in $S_4(a, b)$ with $D_K \leq X$, we have

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{a,b,\rho,c_1,\ell,k,\varepsilon} X^{\frac{k}{2}+\rho-\min\left\{\rho, \frac{3\rho k}{7\ell+6}\right\}+\varepsilon}.$$

By [An20, Theorem 1.2], any ρ with (1.10) must satisfy $\rho \geq 1/2$, and one might expect $\rho = 1/2$ to be the correct order of magnitude. Under the assumption that the expectated order of magnitude $\#\{K \in S_4(a, b);\ D_K \leq X\} \asymp X^{1/2}$ is indeed correct, the straightforward approach (1.6) using [An20, Theorem 1.1] would yield

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{a,b,\ell,k,\varepsilon} X^{\frac{k+1}{2}-\min\left\{\frac{1}{2}, \frac{k}{6\ell}\right\}+\varepsilon},$$

upon which Theorem 1.8 improves whenever $\ell > 3$ and $k < 3\ell$. As one can take the exponent $\rho = 1$ in Theorem 1.8 by [DO02, Corollary 1.4], we immediately obtain the following unconditional result.

Corollary 1.9. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. Let $a, b \in \mathbb{Z} \smallsetminus \{0, 1\}$ be distinct and squarefree such that $S_4(a, b) \neq \emptyset$. Then, as K ranges over the fields in $S_4(a, b)$ with $D_K \leq X$, we have*

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{a,b,\ell,k,\varepsilon} X^{\frac{k}{2}+1-\min\left\{1,\frac{3k}{7\ell+6}\right\}+\varepsilon}.$$

This should be compared to what one gets from [An20, Theorem 1.1] via (1.6), using [DO02, Corollary 1.4] to estimate $\#\{K \in S_4(a, b);\ D_K \leq X\} \ll X$, which yields

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{a,b,\ell,k,\varepsilon} X^{\frac{k}{2}+1-\min\left\{1,\frac{k}{6\ell}\right\}+\varepsilon}.$$

Our techniques can also provide improved average and higher moment bounds for some results that are conditional on open conjectures. In [PTBW20], the assumption of GRH was replaced for certain families of number fields by other assumptions, at the price of introducing certain ramification conditions and allowing a small exceptional set. We can also improve some of these conditional results on average.

Theorem 1.10. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers, and $\ell \in \mathbb{N}$. Let $d \geq 3$ and S be the family of all number fields of degree d with squarefree discriminant, whose normal closure has full Galois group S_d over $\mathbb{Q}$. Suppose that*

(i) *the strong Artin conjecture holds for all irreducible Galois representations over $\mathbb{Q}$ with image S_d,*
(ii) *the numbers $\tau < 1/2 + 1/d$ and c_2 are such that for every integer D, there are at most $c_2 D^\tau$ fields $K \in S$ with $D_K = D$,*
(iii) *the numbers $\rho, c_1 > 0$ are such that $\#\{K \in S;\ D_K \leq X\} \leq c_1 X^\rho$ for all $X \geq 2$.*

Then, as K ranges over all elements of S with $D_K \leq X$, we have

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{d,\rho,c_1,c_2,\ell,k,\tau,\varepsilon} X^{\frac{k}{2}+\rho-\frac{\rho k}{(d-1)\ell+2}+\varepsilon} + X^{\frac{k}{2}+\tau+\varepsilon}.$$

The assumptions (i) and (ii) of Theorem 1.10 are the same as in [PTBW20, Theorem 13] for $d \geq 6$. For a precise formulation of the strong Artin conjecture, see [PTBW20, Conjecture F]. For $d \in \{3, 4, 5\}$, our assumptions can be weakened as in [PTBW20]. If $d = 3, 4$, the result is unconditional if one takes $\rho = 1$ (using [DH71] and [Bha05]) and $\tau = 1/3$ or $\tau = 1/2$, respectively (see Theorem 5.3). If $d = 5$, one still needs (i), but one can take $\rho = 1$ and the upper bound for τ in (ii) can be replaced by 1 (see Theorem 5.3).

Note that Bhargava, Shankar and Wang [BSW16] have shown that $\rho \geq 1/2 + 1/d$, and Bhargava [Bha14] conjectured that (iii) is sharp with $\rho = 1$. On the other hand, it is conjectured that (ii) holds with any $\tau > 0$ (see [EV05]). Assuming these conjectured values for ρ and τ to be the right ones, the straightforward approach (1.6) applied to the bounds from [PTBW20, Theorem 7.2] would yield

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{d,\ell,k,\varepsilon} X^{\frac{k}{2}+1-\min\{1,\frac{k}{2\ell(d-1)}\}+\varepsilon},$$

upon which Theorem 1.10 yields an improvement when $k < 2\ell(d-1)$ and $\ell \geq 2$.

Finally, we can also improve the conditional result of [PTBW20] on A_d-extensions for all $d \geq 5$.

Theorem 1.11. *Let $\varepsilon > 0$ and $k \geq 0$ be real numbers. Let $d \geq 5$ and S be the family of all number fields of degree d, whose normal closure has Galois group A_d over $\mathbb{Q}$. Suppose that*

(i) *the strong Artin conjecture holds for all irreducible Galois representations over $\mathbb{Q}$ with image A_d,*
(ii) *the numbers $\rho, c_1 > 0$ are such that $\#\{K \in S;\ D_K \leq X\} \leq c_1 X^\rho$ for all $X \geq 2$.*

Then, as K ranges over all fields in S with $D_K \leq X$, we have

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{d,\rho,c_1,\ell,k,\varepsilon} X^{\frac{k}{2}+\rho-\min\{\rho,\frac{\rho k}{(d-3/2)\ell+2}\}+\varepsilon}.$$

Here, Malle's conjecture predicts the optimal exponent $\rho = 1/2$. Assuming this conjecture to be correct, we would get from (1.6) applied to [PTBW20, Theorem 7.2] the average bound

$$\sum_K \#\operatorname{Cl}_K[\ell]^k \ll_{d,\ell,k,\varepsilon} X^{\frac{k+1}{2}-\min\{\frac{1}{2},\frac{k}{2\ell(d-1)}\}+\varepsilon}.$$

Theorem 1.11 improves the latter when $\ell > 4$ and $k < (d-1)\ell$.

1.4. **Plan of the paper.** In §2, we introduce invariants $\eta_\ell(K)$ of number fields K and use them to refine the key lemma [EV07, Lemma 2.3] of Ellenberg and Venkatesh. In §3, we prove two general results that use the refined key lemma to deduce average and moment bounds for ℓ-torsion from certain asymptotic counting results. In §4, we provide such counting results for fields K of bounded $\eta_\ell(K)$. In §5, we recall results from the literature that guarantee the existence of enough small split primes. In §6, we deduce all of our theorems, and in §7 we prove Corollary 1.2.

2. A refined key lemma

Let

$$H_K(\alpha) = \prod_{v \in M_K} \max\{1, |\alpha|_v\}^{d_v}$$

be the multiplicative Weil height of $\alpha \in K$ relative to K. Here M_K denotes the set of places of K, and for each place v we choose the unique representative $|\cdot|_v$ that either extends the usual Archimedean absolute value on $\mathbb{Q}$ or a usual p-adic absolute value on $\mathbb{Q}$, and $d_v = [K_v : \mathbb{Q}_v]$ denotes the local degree at v.

For every prime ideal $\mathfrak{p}$ of K lying above a rational prime p, we write $e(\mathfrak{p}) = e(\mathfrak{p}/p)$ for the ramification index and $f(\mathfrak{p}) = f(\mathfrak{p}/p)$ for the inertia degree of $\mathfrak{p}$ over p. For each $\ell \in \mathbb{N}$ we introduce a new invariant of number fields K,

$$\eta_\ell(K) = \inf\left\{ H_K(\alpha)\,;\ \begin{array}{l} \alpha \in K,\ \alpha\mathcal{O}_K = (\mathfrak{p}_1\mathfrak{p}_2{}^{-1})^\ell,\ \text{where } \mathfrak{p}_1 \neq \mathfrak{p}_2 \text{ are prime} \\ \text{ideals of } \mathcal{O}_K \text{ with } e(\mathfrak{p}_i) = f(\mathfrak{p}_i) = 1 \text{ for } i = 1, 2 \end{array} \right\}.$$

We will show in Lemma 4.1 that an element α of this special form necessarily generates K, and moreover its minimal polynomial has a restricted shape. This will allow us to deduce upper bounds for the number of fields K of bounded $\eta_\ell(K)$ which lead to the improved

bounds in our theorems. The following proposition is a refinement of [EV07, Lemma 2.3] and central to all our improvements.

Proposition 2.1. *Let K be a number field of degree d, $\delta < 1/\ell$, and $\varepsilon > 0$. Moreover, suppose that there are M prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$ with norm $N(\mathfrak{p}) \le \eta_\ell(K)^\delta$ that satisfy $e(\mathfrak{p}) = f(\mathfrak{p}) = 1$. If $M > 0$, we have*

$$\#\operatorname{Cl}_K[\ell] \ll_{d,\delta,\varepsilon} D_K^{1/2+\varepsilon} M^{-1}.$$

Proof. We may assume that $\eta_\ell(K) \ge 2$. Write R_K for the regulator of K and set $G := \operatorname{Cl}_K / \operatorname{Cl}_K[\ell]$, so that $\#\operatorname{Cl}_K[\ell] \cdot \#G \cdot R_K = \#\operatorname{Cl}_K R_K \ll_{d,\varepsilon} D_K^{1/2+\varepsilon}$. Hence, we need to show that $\#G \gg_{d,\varepsilon} M/R_K$. Fix a constant $c > 0$ and write $R := \lceil cR_K \rceil$. Our goal is to show that $\#G \ge M/R$, if c was chosen sufficiently large in terms of only d and δ. Since $R_K \gg_d 1$, we may assume that $R \ge 2$. Suppose $\#G < M/R$. Then, by the pigeonhole principle, the classes $[\mathfrak{p}]$ of at least $R+1$ out of our M prime ideals $\mathfrak{p}$ must lie in the same coset in G. We call these prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_{R+1}$ to obtain $[\mathfrak{p}_{R+1}] \operatorname{Cl}_K[\ell] = [\mathfrak{p}_i] \operatorname{Cl}_K[\ell]$ for all $1 \le i \le R$, and thus find $\alpha_i \in K$ with

$$\alpha_i \mathcal{O}_K = (\mathfrak{p}_i \mathfrak{p}_{R+1}^{-1})^\ell.$$

First suppose that K is imaginary quadratic. We choose distinct i and j between 1 and R and conclude

$$H_K(\alpha_i/\alpha_j) \le \max\{N(\mathfrak{p}_i), N(\mathfrak{p}_j)\}^\ell < \eta_\ell(K),$$

which contradicts the minimality assumption in the definition of $\eta_\ell(K)$.

Now suppose that K is not imaginary quadratic. Let $l : K^* \to \mathbb{R}^{q+1}$ be the classical logarithmic embedding, where $q+1$ is the number of Archimedean places of K. After multiplying α_i by a unit we can assume that $l(\alpha_i) = (d_v \log |\alpha_i|_v)_{v|\infty} \in F + (d_v)_{v|\infty}(-\infty, \infty)$, where F is a fundamental cell of the unit lattice $l(\mathcal{O}^*) \subset \mathbb{R}^{q+1}$. We take $F = [0,1)u_1 + \cdots + [0,1)u_q$ where $u_1, \dots, u_q$ is a Minkowski reduced basis of the unit lattice. Write $l(\alpha_i) = v_i + \gamma_i (d_v)_{v|\infty}$, where $v_i \in F$ and $\gamma_i \in (-\infty, \infty)$. We note that the Euclidean length $|u_i| \gg_d 1$, which follows easily from Northcott's Theorem (see, e.g., [Wid10, below (8.2)]). Since F comes from a Minkowski reduced basis we can partition F into at most $R-1$ subcells of diameter $\ll_d (R_K/R)^{1/q} \le c^{-1/q} \le c^{-1/d}$. Again by the pigeonhole principle, we find distinct i and j such that v_i and v_j lie in the same subcell and hence $|(v_i - v_j)_v| \ll_d c^{-1/d}$ for all $v|\infty$. Without loss of generality, we may assume that $\gamma_i \le \gamma_j$. Since $|\alpha_i/\alpha_j|_v = e^{(1/d_v)(v_i - v_j)_v + (\gamma_i - \gamma_j)}$, we conclude that

$$|\alpha_i/\alpha_j|_v = e^{O_d(c^{-1/d}) + (\gamma_i - \gamma_j)} \le e^{O_d(c^{-1/d})} \quad \text{holds for all } v|\infty.$$

Since $\alpha \mathcal{O}_K = (\mathfrak{p}_i \mathfrak{p}_j^{-1})^\ell$, this shows that

$$H_K(\alpha_i/\alpha_j) \le e^{O_d(c^{-1/d})} N(\mathfrak{p}_j)^\ell \le e^{O_d(c^{-1/d})} \eta_\ell(K)^{\ell\delta}.$$

Since $\ell\delta < 1$ and $\eta_\ell(K) \ge 2$, we can choose c large enough in terms of d, δ to ensure that $H_K(\alpha_i/\alpha_j) < \eta_\ell(K)$, contradicting the definition of $\eta_\ell(K)$. Thus, with this choice of c we get $\#G \ge M/R \gg_{d,\delta} M/R_K$. □

3. Framework

Let $d > 1$ be an integer. We set

$$S_{\mathbb{Q},d} = \{K \subset \overline{\mathbb{Q}}\,;\, [K:\mathbb{Q}] = d\} \tag{3.1}$$

for the collection of all number fields of degree d. For a subset $S \subset S_{\mathbb{Q},d}$ we set

$$\begin{aligned} S_X &:= \{K \in S\,;\, X \le D_K < 2X\},\\ \mathscr{B}_S(X;Y,M) &:= \{K \in S_X\,;\ \text{ at most } M \text{ primes } p \le Y \text{ split completely in } K\},\\ N_{\eta_\ell}(S,X) &:= \#\{K \in S\,;\, \eta_\ell(K) < X\},\\ N_D(S,X) &:= \#S_X. \end{aligned}$$

Throughout this section we assume that $\theta, \rho, c_1, c_3 > 0$ are such that for all $X \ge 2$

$$N_D(S,X) \le c_1 X^\rho, \tag{3.2}$$

$$N_{\eta_\ell}(S,X) \le c_3 X^\theta. \tag{3.3}$$

We can now formulate our two main propositions. They differ in their assumption on $\#\mathscr{B}_S(X;X^\delta, cX^\delta/\log X))$. In the first case we have an upper bound that gets worse when δ gets smaller. This situation happens in the work [EPW17] based on probabilistic methods. In the proof we decompose the set of fields in those fields with "small" invariant $\eta_\ell(K)$ compared to the discriminant, those fields with "large" invariant $\eta_\ell(K)$ which are not bad (i.e., they have "sufficiently" many small splitting primes), and those fields with "large" invariant $\eta_\ell(K)$ which are bad. In the first and third case, we use the trivial bound to estimate $\#\operatorname{Cl}_K[\ell]$, and in the second case Proposition 2.1.

Proposition 3.1. *Suppose $S \subset S_{\mathbb{Q},d}$, $\delta_0 > 0$, and that (3.2), (3.3) hold for $\theta, \rho, c_1, c_3 > 0$. Moreover, suppose for every $\delta \in (0,\delta_0]$ and $\varepsilon \in (0,1)$ there are positive $c_4(\delta,\varepsilon)$ and $c_5(\delta,\varepsilon)$ such that*

$$\#\mathscr{B}_S(X;X^\delta, c_4(\delta,\varepsilon)X^\delta/\log X)) \le c_5(\delta,\varepsilon)X^{\rho-\delta+\varepsilon}$$

holds for all $X \ge 2$. Then we have, for all $\varepsilon \in (0,1)$,

$$\sum_{K\in S_X} \#\operatorname{Cl}_K[\ell] \ll_{d,\ell,\theta,\rho,c_1,c_3,\delta_0,c_4(\cdot,\cdot),c_5(\cdot,\cdot),\varepsilon} X^{\frac{1}{2}+\rho-\min\{\delta_0,\frac{\rho}{\ell\theta+1}\}+\varepsilon}.$$

Proof. Let $\varepsilon \in (0,1)$. For sake of clarity, we suppress the dependence of implicit constants in our notation and write $\ll$ instead of $\ll_{d,\ell,\theta,\rho,c_1,c_3,\delta_0,c_4(\cdot,\cdot),c_5(\cdot,\cdot),\varepsilon}$ throughout the proof. We define

$$\gamma_0 := \frac{\rho\ell}{\ell\theta+1}.$$

Hence we have

$$\gamma_0\theta = \rho - \frac{\gamma_0}{\ell}.$$

First let us assume that $\ell \le \frac{1}{\theta}(\frac{\rho}{\delta_0} - 1)$, and thus

$$\gamma_0 \ge \delta_0\ell.$$

We decompose S_X into the three subsets

$$M_0 = \{K \in S_X \,;\, \eta_\ell(K) \le D_K^{\delta_0 \ell}\},$$
$$M_1' = \{K \in S_X \,;\, \eta_\ell(K) > D_K^{\delta_0 \ell}\} \smallsetminus \mathscr{B}_S(X; X^{(1-\varepsilon)\delta_0}, cX^{(1-\varepsilon)\delta_0}/\log X),$$
$$M_1'' = \{K \in S_X \,;\, \eta_\ell(K) > D_K^{\delta_0 \ell}\} \cap \mathscr{B}_S(X; X^{(1-\varepsilon)\delta_0}, cX^{(1-\varepsilon)\delta_0}/\log X),$$

where $c = c_4((1-\varepsilon)\delta_0, \varepsilon)$ comes from the assumptions of the proposition. Using (1.2), we get

$$\sum_{K \in M_0} \#\operatorname{Cl}_K[\ell] \ll \sum_{K \in M_0} D_K^{\frac{1}{2}+\varepsilon} \le \#M_0 \cdot (2X)^{\frac{1}{2}+\varepsilon}.$$

Since $\#M_0 \le N_{\eta_\ell}(S, (2X)^{\delta_0 \ell}) \ll X^{\delta_0 \ell \theta}$ and $\delta_0 \ell \le \gamma_0$ we conclude

$$\sum_{K \in M_0} \#\operatorname{Cl}_K[\ell] \ll X^{\frac{1}{2}+\gamma_0\theta+\varepsilon} \le X^{\frac{1}{2}+\rho-\delta_0+\varepsilon}.$$

Since by assumption $\#M_1'' \ll X^{\rho-(1-\varepsilon)\delta_0+\varepsilon}$, we find similarly

$$\sum_{K \in M_1''} \#\operatorname{Cl}_K[\ell] \ll X^{\frac{1}{2}+\rho-\delta_0+(2+\delta_0)\varepsilon}.$$

For the sum over M_1' we use Proposition 2.1, with the valid choice $M = cX^{(1-\varepsilon)\delta_0}/\log X$, and then bound $\#M_1'$ by (3.2) to conclude that

$$\sum_{K \in M_1'} \#\operatorname{Cl}_K[\ell] \ll \sum_{K \in M_1'} D_K^{\frac{1}{2}-(1-\varepsilon)\delta_0+2\varepsilon} \le \#M_1' \cdot (2X)^{\frac{1}{2}-\delta_0+(2+\delta_0)\varepsilon} \ll X^{\frac{1}{2}+\rho-\delta_0+(2+\delta_0)\varepsilon}.$$

This proves the proposition when $\ell \le \frac{1}{\theta}(\frac{\rho}{\delta_0} - 1)$. Now let us assume that $\ell > \frac{1}{\theta}(\frac{\rho}{\delta_0} - 1)$, and thus

$$\gamma_0 < \delta_0 \ell.$$

We now define M_0, M_1' and M_1'' exactly in the same way but with δ_0 replaced by γ_0/ℓ. Arguing in exactly the same way as in the previous case we get

$$\sum_{K \in S_X} \#\operatorname{Cl}_K[\ell] \ll X^{\frac{1}{2}+\rho-\frac{\gamma_0}{\ell}+(2+\frac{\gamma_0}{\ell})\varepsilon} \le X^{\frac{1}{2}+\rho-\frac{\rho}{\ell\theta+1}+(2+\rho)\varepsilon}.$$

□

Our next main proposition applies when the bound for $\#\mathscr{B}_S(X; X^\delta, cX^\delta/\log X))$ is uniform in δ. For $d = 2$ such a bound can be established by using the large sieve, as shown in [HBP17]. It is a new innovation of the recent work [PTBW20] that such uniform bounds are also available for a much larger class of families S. In this setting it turns out beneficial to use a finer decomposition of the set of fields than just those fields with "small" invariant $\eta_\ell(K)$, and those fields with "large" invariant $\eta_\ell(K)$.

Proposition 3.2. *Suppose $S \subset S_{\mathbb{Q},d}$, $\tau \ge 0$, and that (3.2), (3.3) hold for $\theta, \rho, c_1, c_3 > 0$. Moreover, suppose for every $\delta > 0$ and $\varepsilon \in (0, 1/\ell)$ there are positive $c_4(\delta, \varepsilon)$ and $c_5(\delta, \varepsilon)$ such that*

$$\#\mathscr{B}_S(X; X^\delta, c_4(\delta,\varepsilon)X^\delta/\log X) \le c_5(\delta,\varepsilon)X^{\tau+\varepsilon}$$

holds for all $X \geq 2$. Then we have, for all $k \geq 0$ and $\varepsilon \in (0, 1/\ell)$,

$$\sum_{K \in S_X} \# \operatorname{Cl}_K[\ell]^k \ll_{d,\theta,\rho,c_1,c_3,c_4(\cdot,\cdot),c_5(\cdot,\cdot),\ell,k,\tau,\varepsilon} X^{\frac{k}{2}+\rho-\frac{\rho k}{\ell\theta}+\varepsilon} + X^{\frac{k}{2}+\tau+\varepsilon}.$$

Proof. Let $\varepsilon \in (0, 1/\ell)$. We decompose S_X into $N+2$ subsets M_i, where $N = N(\varepsilon)$ will be chosen later. Let $0 = \gamma_{-1} \leq \gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_N$ and set

$$M_i = \{K \in S_X \,;\, D_K^{\gamma_{i-1}} \leq \eta_\ell(K) < D_K^{\gamma_i}\} \qquad (0 \leq i \leq N),$$
$$M_{N+1} = \{K \in S_X \,;\, D_K^{\gamma_N} \leq \eta_\ell(K)\}.$$

Furthermore, for $1 \leq i \leq N+1$ we decompose M_i into the two sets

$$M_i' = M_i \smallsetminus \mathscr{B}_S(X; X^{\gamma_{i-1}(1/\ell-\varepsilon)}, c_i' X^{\gamma_{i-1}(1/\ell-\varepsilon)}/\log X),$$
$$M_i'' = M_i \cap \mathscr{B}_S(X; X^{\gamma_{i-1}(1/\ell-\varepsilon)}, c_i' X^{\gamma_{i-1}(1/\ell-\varepsilon)}/\log X),$$

where $c_i' = c_4(\gamma_{i-1}(1/\ell-\varepsilon), \varepsilon)$. Hence, we have partitioned S_X into the $1+2(N+1)$ subsets M_0, M_i', M_i'' $(1 \leq i \leq N+1)$. Throughout this proof, we suppress the implicit constants in our notation and write $\ll$ for $\ll_{d,\theta,\rho,c_1,c_3,c_4(\cdot,\cdot),c_5(\cdot,\cdot),\ell,k,\tau,\varepsilon,\gamma_0,\ldots,\gamma_N}$. The values of $\gamma_0, \ldots, \gamma_N$ are fixed later in the proof depending only on the other parameters. Next we record the estimates

$$\begin{aligned}
\#M_0 &\leq N_{\eta_\ell}(S, (2X)^{\gamma_0}) \ll X^{\gamma_0\theta}, \\
\#M_i' &\leq \#M_i \leq N_{\eta_\ell}(S, (2X)^{\gamma_i}) \ll X^{\gamma_i\theta} \qquad (1 \leq i \leq N), \\
\#M_{N+1}' &\leq \#M_{N+1} \leq N_D(S, X) \ll X^{\rho}, \\
\#M_i'' &\ll X^{\tau+\varepsilon} \qquad (1 \leq i \leq N+1).
\end{aligned}$$

We use (1.2) to estimate the sums over M_0 and M_i'' $(1 \leq i \leq N+1)$,

$$\sum_{K \in M_0} \# \operatorname{Cl}_K[\ell]^k \ll \sum_{K \in M_0} D_K^{(\frac{1}{2}+\varepsilon)k} \leq \#M_0 \cdot (2X)^{\frac{k}{2}+k\varepsilon} \ll X^{\frac{k}{2}+\gamma_0\theta+k\varepsilon},$$
$$\sum_{K \in M_i''} \# \operatorname{Cl}_K[\ell]^k \ll \sum_{K \in M_i''} D_K^{(\frac{1}{2}+\varepsilon)k} \leq \#M_i'' \cdot (2X)^{\frac{k}{2}+k\varepsilon} \ll X^{\frac{k}{2}+\tau+(k+1)\varepsilon}.$$

From Proposition 2.1, with the eligible choice $M = c_i' X^{\gamma_{i-1}(1/\ell-\varepsilon)}/\log X$, we conclude for $1 \leq i \leq N$ that

$$\sum_{K \in M_i'} \# \operatorname{Cl}_K[\ell]^k \ll \sum_{K \in M_i'} D_K^{(\frac{1}{2}-\gamma_{i-1}(\frac{1}{\ell}-\varepsilon)+2\varepsilon)k} \ll X^{\frac{k}{2}-\frac{\gamma_{i-1}k}{\ell}+\gamma_i\theta+k(2+\gamma_N)\varepsilon}$$

and similarly

$$\sum_{K \in M_{N+1}'} \# \operatorname{Cl}_K[\ell]^k \ll X^{\frac{k}{2}+\rho-\frac{\gamma_N k}{\ell}+k(2+\gamma_N)\varepsilon}.$$

For $0 \leq i \leq N$, we define $Q_i = \sum_{r=0}^{i} q^r$, where $q = \frac{k}{\ell\theta}$. With these quantities in place, we proceed to choose our γ_i as follows,

$$\gamma_0 = \gamma_0(N) = \frac{\rho\ell}{\ell\theta + kQ_N} \quad \text{and} \quad \gamma_i = \gamma_0 Q_i \leq \frac{\rho\ell}{k} \quad (1 \leq 1 \leq N).$$

Then a quick computation shows that

$$\frac{k}{2}+\gamma_0\theta=\frac{k}{2}-\frac{\gamma_{i-1}k}{\ell}+\gamma_i\theta=\frac{k}{2}+\rho-\frac{\gamma_N k}{\ell},$$

which allows us to estimate

$$\sum_{K\in S_X}\#\operatorname{Cl}_K[\ell]^k\ll X^{\frac{k}{2}+\gamma_0\theta+(2k+\rho\ell)\varepsilon}+X^{\frac{k}{2}+\tau+(k+1)\varepsilon}.$$

The only task left is to choose $N=N(\varepsilon)$. We observe that

$$\tilde{\gamma}_0:=\lim_{N\to\infty}\gamma_0(N)=\begin{cases}\frac{\rho}{\theta}-\frac{\rho k}{\ell\theta^2} & \text{if } q<1,\\ 0 & \text{if } q\geq 1.\end{cases}$$

Hence, choosing $N=N(\varepsilon)$ big enough to ensure $\gamma_0\theta\leq\tilde{\gamma}_0\theta+\varepsilon$, we conclude that

$$\sum_{K\in S_X}\#\operatorname{Cl}_K[\ell]^k\ll X^{\frac{k}{2}+\tilde{\gamma}_0\theta+(1+2k+\rho\ell)\varepsilon}+X^{\frac{k}{2}+\tau+(k+1)\varepsilon},$$

which proves the proposition. □

4. Counting fields of bounded $\eta_\ell(K)$

For $\alpha\in\overline{\mathbb{Q}}$ we write $D_\alpha\in\mathbb{Z}[x]$ for the minimal polynomial of α over $\mathbb{Z}$, i.e., the irreducible polynomial with positive leading coefficient that satisfies $D_\alpha(\alpha)=0$. Our estimates for $N_{\eta_\ell}(S,X)$ hinge upon the following observation.

Lemma 4.1. *Let $\alpha\in K$ be such that $\alpha\mathcal{O}_K=(\mathfrak{p}_1\mathfrak{p}_2{}^{-1})^\ell$, with distinct prime ideals $\mathfrak{p}_1,\mathfrak{p}_2$ of $\mathcal{O}_K$ that satisfy $e(\mathfrak{p}_i)=f(\mathfrak{p}_i)=1$ for $i=1,2$. Then $K=\mathbb{Q}(\alpha)$ and the minimal polynomial D_α has the form*

$$D_\alpha=p^\ell x^d+a_1x^{d-1}+\cdots+a_{d-1}x\pm q^\ell, \tag{4.1}$$

where $a_1,\dots,a_{d-1}\in\mathbb{Z}$ and p,q are the primes below $\mathfrak{p}_2$ and $\mathfrak{p}_1$, respectively.

Proof. First, suppose $\mathbb{Q}(\alpha)=F\subsetneq K$. Let $\mathfrak{q}_1$ be the prime ideal of $\mathcal{O}_F$ below $\mathfrak{p}_1$. Then $e(\mathfrak{p}_1/\mathfrak{q}_1)=f(\mathfrak{p}_1/\mathfrak{q}_1)=1$. Hence, as $[K:F]>1$, there must be another prime ideal $\mathfrak{p}_1'$ of $\mathcal{O}_K$ above $\mathfrak{q}_1$. For the corresponding discrete valuations, we get $v_{\mathfrak{p}_1'}(\alpha)=e(\mathfrak{p}_1'/\mathfrak{q}_1)v_{\mathfrak{q}_1}(\alpha)=e(\mathfrak{p}_1'/\mathfrak{q}_1)v_{\mathfrak{p}_1}(\alpha)=e(\mathfrak{p}_1'/\mathfrak{q}_1)\ell>0$. But there is no other prime ideal of $\mathcal{O}_K$ at which α has positive valuation. Hence, $\mathbb{Q}(\alpha)=K$. The second assertion follows immediately from the well-known formula

$$a_0=\prod_{v\nmid\infty}\max\{1,|\alpha|_v\}^{d_v},$$

where a_0 is the leading coefficient of D_α and the product runs over all non-Archimedean places of $\mathbb{Q}(\alpha)$. The latter formula in turn is essentially a consequence of Gauß' Lemma applied to D_α and each non-Archimedean place of the splitting field of D_α. □

Lemma 4.2. *Suppose $S\subset S_{\mathbb{Q},d}$, and $\theta=d-1+2/\ell$. Then*

$$N_{\eta_\ell}(S,X)\ll_d X^\theta.$$

Proof. Let P_S be the set of all $\alpha \in \overline{\mathbb{Q}}$ such that $\mathbb{Q}(\alpha) \in S$ and $\alpha\mathcal{O}_{\mathbb{Q}(\alpha)} = (\mathfrak{p}_1\mathfrak{p}_2^{-1})^\ell$, for prime ideals $\mathfrak{p}_1 \neq \mathfrak{p}_2$ of $\mathcal{O}_{\mathbb{Q}(\alpha)}$ with $e(\mathfrak{p}_i) = f(\mathfrak{p}_i) = 1$ for $i = 1, 2$. Moreover, let

$$N_H(P_S, X) := \#\{\alpha \in P_S \,;\, H_{\mathbb{Q}(\alpha)}(\alpha) \leq X\}.$$

Using Lemma 4.1, we observe that the image of the map $\alpha \to \mathbb{Q}(\alpha)$ with domain

$$\{\alpha \in P_S \,;\, H_{\mathbb{Q}(\alpha)}(\alpha) \leq X\}$$

covers the set

$$\{K \in S \,;\, \eta_\ell(K) \leq X\}.$$

Hence, we get

$$N_{\eta_\ell}(S, X) \leq N_H(P_S, X).$$

Now if $\alpha \in P_S$ then, as noted in (4.1), the first and last coefficient of its minimal polynomial D_α are, up to sign, $\ell - th$ prime powers. For α to be counted in $N_H(P_S, X)$, we also require $H_{\mathbb{Q}(\alpha)}(\alpha) \leq X$. Now the maximum norm of the coefficient vector of D_α is bounded from above by $2^d H_{\mathbb{Q}(\alpha)}(\alpha)$, and hence by $2^d X$. Thus, we have at most $\ll_d X^{d-1+2/\ell}$ possibilities for these minimal polynomials and thus for α. □

The bound on $N_{\eta_\ell}(S, X)$ from Lemma 4.2 suffices to deduce Theorems 1.1, 1.3, 1.4 and 1.10. Our other theorems involve families of number fields with specified Galois groups $G \subsetneq S_d$. To compensate for the relative thinness of these families, we need to show that families of polynomials of degree d with specified Galois group $G \subsetneq S_d$ are also thin. This was done by Dietmann in [Die12], but his results are not applicable to our situation as they concern monic polynomials with no further restrictions on their coefficients, whereas we have to deal with polynomials of the shape (4.1).

The idea of Dietmann's proof, to detect polynomials with Galois group G through roots of appropriate resolvents, and to control these roots via uniform bounds for integral points on affine surfaces, applies to our situation as well. The following results, culminating in Proposition 4.7 below, modify and refine Dietmann's proofs accordingly. Hence, we keep our notation similar to that of [Die12]. In particular, we will write n instead of d for the degree of our polynomials.

For any field K of characteristic 0 and $n \in \mathbb{N}$, we consider polynomials

$$f = x^n + a_1 x^{n-1} + \cdots + a_n \in K[x]$$

with distinct roots $\alpha_1, \ldots, \alpha_n$ in an algebraic closure of K. Let $G \subset S_n$ be a subgroup, then the Galois resolvent from [Die12, Lemma 5] is defined as

$$\phi(z; a_1, \ldots, a_n) = \prod_{\sigma \in S_n/G} \left(z - \sum_{\tau \in G} \alpha_{\sigma(\tau(1))} \alpha_{\sigma(\tau(2))}^2 \cdots \alpha_{\sigma(\tau(n))}^n \right). \tag{4.2}$$

It is a polynomial in $z, a_1, \ldots, a_n$ with integer coefficients that do not depend on K. It is monic in z of degree $\#(S_n/G)$. It has a root $z \in K$ whenever the Galois group of f, as a subgroup of S_n acting on $\alpha_1, \ldots, \alpha_n$, is contained in G. In case $K = \mathbb{Q}$ and $a_1, \ldots, a_n \in \mathbb{Z}$, this root must clearly lie in $\mathbb{Z}$. Moreover, we denote by $\Delta_\phi(a_1, \ldots, a_n) \in K$ the discriminant of $\phi(z; a_1, \ldots, a_n) \in K[z]$. Again, this discriminant is a polynomial in $a_1, \ldots, a_n$ with integer coefficients independent of K.

Lemma 4.3. *Fix $a_n \in \mathbb{Q}$, $a_n \neq 0$. Then $\Delta_\phi(a_1, \dots, a_{n-1}, a_n)$ is not identically zero as a polynomial in $a_1, \dots, a_{n-1}$.*

Proof. This is a refinement of [Die12, Lemma 7]. Fix $a_n \neq 0$. Then it is enough to find $a_1, \dots, a_{n-1} \in \mathbb{C}$ such that $\Delta_\phi(a_1, \dots, a_n) \neq 0$. For any choice of $a_1, \dots, a_{n-1}$, it is clear from (4.2) that the roots of $\phi(z; a_1, \dots, a_n)$ are the complex numbers

$$\sum_{\tau \in G} \alpha_{\sigma(\tau(1))} \alpha_{\sigma(\tau(2))}^2 \cdots \alpha_{\sigma(\tau(n))}^n, \tag{4.3}$$

where σ ranges over a set of representatives for the cosets in S_n/G. All $\#(S_n/G)$ of these expressions are distinct homogeneous polynomials of degree $n(n+1)/2$ in $\mathbb{Z}[\alpha_1, \dots, \alpha_n]$. Hence, there is a non-empty Zariski-open subset of points $(\alpha_1 : \cdots : \alpha_n) \in \mathbb{P}^{n-1}$ for which all the expressions in (4.3) are distinct. In particular, we find such $(\alpha_1 : \cdots : \alpha_n)$ whose homogeneous coordinates $\alpha_i \in \mathbb{C}$ are all distinct and non-zero. Picking a correctly scaled representative of such a point, we get $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ that satisfy all of the previous conditions and moreover that $(-1)^n \alpha_1 \cdots \alpha_n = a_n$. Let $a_1, \dots, a_{n-1} \in \mathbb{C}$ be the other coefficients of the polynomial $\prod_{i=1}^n (x - \alpha_i)$. Then, by our choice of $\alpha_1, \dots, \alpha_n$, all zeros of $\phi(z; a_1, \dots, a_n)$ are distinct, and hence its discriminant satisfies $\Delta_\phi(a_1, \dots, a_n) \neq 0$. □

Lemma 4.4. *Let $n \geq 3$ and $a_2, \dots, a_{n-2}, a_n \in \mathbb{Z}$ such that $a_n \neq 0$. Then the polynomial*

$$x^n + a_1 x^{n-1} + \cdots + a_{n-2} x^2 + tx + a_n \in \mathbb{Q}(t)[x]$$

has, for all but $\ll_n 1$ values of $a_1 \in \mathbb{Z}$, the full symmetric group S_n as Galois group acting on its roots in an algebraic closure of the rational function field $\mathbb{Q}(t)$.

Proof. This is similar to [Die12, Lemma 2]. By [Her70, Satz 1], the Galois group is S_n for all but finitely many values of $a_1 \in \mathbb{Z}$. As described in [Die12, Lemma 2] and the introduction of [Her72], the proof of [Her70, Satz 1] provides the upper bound n^2 for the number of excluded values of a_1. □

Lemma 4.5. *Let $n \geq 2$ and $a_1, \dots, a_{n-2}, a_n \in \mathbb{Z}$ such that the polynomial*

$$x^n + a_1 x^{n-1} + \cdots + a_{n-2} x^2 + tx + a_n \in \mathbb{Q}(t)[x]$$

has Galois group S_n over the rational function field $\mathbb{Q}(t)$. Moreover, suppose that

$$\Delta_\phi(a_1, \dots, a_{n-2}, t, a_n) \neq 0 \text{ in } \mathbb{Q}(t). \tag{4.4}$$

Then the polynomial $\phi(z; t) = \phi(z; a_1, \dots, a_{n-2}, t, a_n) \in \mathbb{Q}[z, t]$ is irreducible over $\mathbb{Q}$.

Proof. Note that (4.4) states that the roots of $\phi(z; t)$ in an algebraic closure of $\mathbb{Q}(t)$ are all distinct. Hence, we are precisely in the situation of [Die12, Lemma 6], except that we use the variable t for the linear coefficient, whereas Dietmann uses t for the constant coefficient. The proof of [Die12, Lemma 6] is agnostic of this difference and works verbatim in our case. □

The following result is [Die12, Lemma 8], which follows from [BHB05, Theorem 1].

Lemma 4.6. *Let $F \in \mathbb{Z}[x_1, x_2]$ be of degree d and irreducible over $\mathbb{Q}$. Let $P_1, P_2 \geq 1$, and*

$$T = \max_{(e_1, e_2)} \{P_1^{e_1} P_2^{e_2}\},$$

where (e_1, e_2) runs through all pairs for which the monomial $x_1^{e_1}x_2^{e_2}$ appears in F with non-zero coefficient. Then, for $\varepsilon > 0$,

$$\#\{\mathbf{x} \in \mathbb{Z}^2 \,;\, F(\mathbf{x}) = 0 \text{ and } |x_i| \le P_i \text{ for } i = 1, 2\} \ll_{d,\varepsilon} \max\{P_1, P_2\}^{\varepsilon} \exp\left(\frac{\log P_1 \log P_2}{\log T}\right).$$

Note that the implicit constant depends only on the degree, but not on the values of the coefficients of F. This is crucial for our application.

Proposition 4.7. *Let $n \ge 2$, G a transitive subgroup of S_n and $\ell \in \mathbb{N}$. For $B \ge 2$, let $N_{n,G}(B)$ be the number of polynomials $f = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ such that*

1. *$a_0, \dots, a_n \in \mathbb{Z} \cap [-B, B]$,*
2. *a_0, a_n are ℓ-th powers in $\mathbb{Z} \smallsetminus \{0\}$,*
3. *f is irreducible over $\mathbb{Q}$,*
4. *the Galois group of f acts on the roots of f (enumerated in a fixed order) as G.*

Then, for $\varepsilon > 0$, we have the upper bound

$$N_{n,G}(B) \ll_{n,\varepsilon} B^{n-2+2/\ell+\#(S_n/G)^{-1}+\varepsilon}. \tag{4.5}$$

Proof. The result follows from Lemma 4.2 in case $n = 2$, so we assume from now on that $n \ge 3$. Conditions (3) and (4) are invariant under replacing f by

$$a_0^{n-1}f(x/a_0) = x^n + a_1x^{n-1} + \cdots + a_0^{n-3}a_{n-2}x^2 + a_0^{n-2}a_{n-1}x + a_0^{n-1}a_n, \tag{4.6}$$

so we have to bound the number of $a_0, \dots, a_n$ subject to (1) and (2), for which the polynomial in (4.6) satisfies (3) and (4). Lemma 4.4 shows that, for every choice of $a_0, a_2, \dots, a_n$, there are $\ll_n 1$ choices of a_1 for which the polynomial

$$g(x; t) = x^n + a_1x^{n-1} + \cdots + a_0^{n-3}a_{n-2}x^2 + tx + a_0^{n-1}a_n \in \mathbb{Q}(t)[x]$$

does not have full Galois group S_n over the rational function field $\mathbb{Q}(t)$. The total number of $a_0, \dots, a_n$ for which this holds is thus $\ll_n B^{n-2+2/\ell}$. In view of the desired bound (4.5), we may thus restrict our attention to those $a_0, \dots, a_n$ for which

$$g(x; t) \text{ has full Galois group } S_n \text{ over } \mathbb{Q}(t). \tag{4.7}$$

For these polynomials, we consider the corresponding Galois resolvents

$$\phi(z; t) = \phi(z; a_1, \dots, a_0^{n-3}a_{n-2}, t, a_0^{n-1}a_n) \in \mathbb{Z}[z, t],$$

defined in (4.2), and their discriminants $\Delta_\phi(t) = \Delta_\phi(a_1, \dots, a_0^{n-3}a_{n-2}, t, a_0^{n-1}a_n) \in \mathbb{Z}[t]$.

Lemma 4.3 shows that, for any fixed permitted choice of a_0, a_n, the discriminant $\Delta_\phi(t)$ does not vanish identically as a polynomial in $a_1, \dots, a_{n-2}, t$. Hence, there are at most $\ll_n B^{n-3}$ choices of $a_1, \dots, a_{n-2}$ with (1), for which $\Delta_\phi(t) = 0$ in $\mathbb{Q}(t)$. Summing this over all possible choices of a_0, a_{n-1}, a_n with (1) and (2), we obtain a contribution $\ll_n B^{n-2+2/\ell}$ in total, which is negligible when compared to (4.5). Hence, we may assume from now on that $\Delta_\phi(t) \ne 0$ for all our tuples $a_0, \dots, a_n$ under consideration. In this case, together with our previous assumption (4.7), we see from Lemma 4.5 that $\phi(z, t)$ is irreducible over $\mathbb{Q}$ for all choices of $a_0, \dots, a_{n-2}, a_n$. Fixing such a choice, suppose that the polynomial $g(x; a_0^{n-2}a_{n-1})$ from (4.6) satisfies (3) and (4) for some a_{n-1} subject to (1).

Then all complex roots of $g(x; a_0^{n-2}a_{n-1})$ are distinct and moreover the Galois resolvent $\phi(z; a_0^{n-2}a_{n-1})$ has a root $z \in \mathbb{Z}$. Since the roots of a complex polynomial are bounded

polynomially in terms of its coefficients (see, e.g., [Die12, Lemma 1]), this root satisfies $|z| \le B^\alpha$, for some $\alpha > 0$ that depends at most on n. Since the polynomial $\phi(z;t)$, and thus also $\phi(z;a_0^{n-2}t)$, is irreducible over $\mathbb{Q}$, we can apply Lemma 4.6 to bound the number of $(z, a_{n-1}) \in \mathbb{Z}^2$ with $|z| \le P_1 := B^\alpha$ and $|a_{n-1}| \le P_2 := B$ for which $\phi(z; a_0^{n-2}a_{n-1}) = 0$. Since the monomial $z^{\#(S_n/G)}$ appears in $\phi(z;t)$, we get $T \ge B^{\alpha\#(S_n/G)}$, and thus the number of such pairs (z, a_{n-1}) is

$$\ll_{n,\varepsilon} B^\varepsilon \exp\left(\frac{\alpha(\log B)^2}{\alpha\#(S_n/G)\log B}\right) = B^{\#(S_n/G)^{-1}+\varepsilon}.$$

Summing this over all viable choices of $a_0, \dots, a_{n-2}, a_n$ yields the bound (4.5). □

Corollary 4.8. *Suppose $S \subset S_{\mathbb{Q},d}$ consists of all A_d-extensions and $\theta > d - 3/2 + 2/\ell$. Then*

$$N_{\eta_\ell}(S, X) \ll_{d,\theta} X^\theta.$$

Proof. This is analogous to the proof of Lemma 4.2, except that the relevant polynomials are now counted by Proposition 4.7 instead of the trivial argument at the end of that proof.

Let P_S be the set of all $\alpha \in \overline{\mathbb{Q}}$ such that $\mathbb{Q}(\alpha) \in S$ and $\alpha\mathcal{O}_{\mathbb{Q}(\alpha)} = (\mathfrak{p}_1\mathfrak{p}_2^{-1})^\ell$, for prime ideals $\mathfrak{p}_1 \ne \mathfrak{p}_2$ of $\mathcal{O}_{\mathbb{Q}(\alpha)}$ with $e(\mathfrak{p}_i) = f(\mathfrak{p}_i) = 1$ for $i = 1, 2$. By Lemma 4.1, every field counted by $N_{\eta_\ell}(S, X)$ is of the form $\mathbb{Q}(\alpha)$ for some $\alpha \in P_S$ with $H_{\mathbb{Q}(\alpha)}(\alpha) \le X$. By (4.1) and the fact that $\mathbb{Q}(\alpha)$ is an A_d-extension of $\mathbb{Q}$, we see that the minimal polynomial D_α of α is counted by $N_{d,A_d}(2^dX)$. Propostion 4.7 now shows that

$$N_{\eta_\ell}(S, X) \ll_d N_{d,A_d}(2^dX) \ll_{d,\theta} X^\theta.$$

□

Corollary 4.9. *Suppose $S \subset S_{\mathbb{Q},5}$ consists of all D_5-extensions and $\theta > 3 + 1/12 + 2/\ell$. Then*

$$N_{\eta_\ell}(S, X) \ll_\theta X^\theta.$$

Proof. The proof is analogous to Corollary 4.8. Note that $\#(S_5/D_5) = 12$. □

Corollary 4.10. *Suppose $S \subset S_{\mathbb{Q},4}$ consists of all D_4-extensions and $\theta > 2 + 1/3 + 2/\ell$. Then*

$$N_{\eta_\ell}(S, X) \ll_\theta X^\theta.$$

Proof. Again, the proof is analogous to Corollary 4.8. Note that $\#(S_4/D_4) = 3$. □

5. Bounding the number of bad fields

Recall that $d > 1$ is an integer, $S_{\mathbb{Q},d} = \{K \subset \overline{\mathbb{Q}}\,;\, [K : \mathbb{Q}] = d\}$, and for $S \subset S_{\mathbb{Q},d}$ we defined $\mathscr{B}_S(X; Y, M)$ as the set

$$\{K \in S\,;\, X \le D_K < 2X, \text{ at most } M \text{ primes } p \le Y \text{ split completely in } K\}.$$

Lemma 5.1. *Let $d \ge 2$, and let $S \subset S_{\mathbb{Q},d}$ be a family of degree-d-fields. Suppose that the Riemann hypothesis holds for the Dedekind zeta function of the normal closure of each field in S. Then for every $\delta > 0$ there exists $c = c(d, \delta) > 0$ such that*

$$\#\mathscr{B}_S(X; X^\delta, cX^\delta/\log X) \ll_{d,\delta} 1.$$

Proof. This is an immediate consequence of the conditional effective version of Chebotarev's density theorem due to Lagarias and Odlyzko [LO77]. □

Theorem 5.2. *(*[EPW17, Theorem 2.1]*) Let $d \in \{3,4,5\}$, let $S = S_{\mathbb{Q},d}$ if $d \neq 4$ and $S = S^*_{\mathbb{Q},4}$ the family of all quartic non-D_4 fields, if $d = 4$, and let $\varepsilon > 0$. Recall the definition of $\delta_0(d)$ (just before Theorem 1.3), and put*

$$\delta_0 = \delta_0(d).$$

Then for every $0 < \delta \leq \delta_0$ there exists $c = c(\delta) > 0$ such that

$$\#\mathscr{B}_S(X; X^\delta, cX^\delta/\log X) \ll_{\delta,\varepsilon} X^{1-\delta+\varepsilon}.$$

Consider families $S = S(G, \mathscr{I}) \subset S_{\mathbb{Q},d}$ of fields K whose normal closure $\tilde{K}$ has Galois group G, and such that for each rational prime p that is tamely ramified in K, its ramification is of type $\mathscr{I}$, where $\mathscr{I}$ specifies one or more conjugacy classes in G. By this we mean the inertia group $I(\mathfrak{B}) \subset G$ of any prime ideal $\mathfrak{B} \subset \mathcal{O}_{\tilde{K}}$ above p (which is cyclic if p is tamely ramified in K) is generated by an element in the conjugacy class (or classes) specified by $\mathscr{I}$ (see [PTBW20, §1.2.1]). The following result collects some special cases of [PTBW20, Corollary 3.16].

Theorem 5.3 (Pierce, Turnage-Butterbaugh, Wood)**.** *Let $\varepsilon > 0$, let $S = S(G, \mathscr{I}) \subset S_{\mathbb{Q},d}$ be from one of the following five families, and let $\tau = \tau_S$ as below. Then for every $\delta > 0$ there exists $c = c(S, \delta) > 0$ such that*

$$\#\mathscr{B}_S(X; X^\delta, cX^\delta/\log X) \ll_{S,\delta,c_2,\tau,\varepsilon} X^{\tau+\varepsilon}.$$

1. *G is a cyclic group of order $d \geq 2$ with $\mathscr{I}$ comprised of all generators of G (equivalently, every rational prime that is tamely ramified in K is totally ramified), and $\tau = 0$.*
2. *d is an odd prime, and $G = D_d$ the Dihedral group of symmetries of a regular d-gon, with $\mathscr{I}$ being the conjugacy class of reflections and $\tau = 1/(p-1)$.*
3. *$d \geq 5$, $G = A_d$ and $\mathscr{I} = G$ (so no restriction on inertia type), and $\tau = 0$. Moreover, assume that the strong Artin conjecture holds for all irreducible Galois representations over $\mathbb{Q}$ with image A_d.*
4. *$d \in \{3,4\}$, $G = S_d$, with $\mathscr{I}$ being the conjugacy class of transpositions, and $\tau = 1/3$ if $d = 3$ and $\tau = 1/2$ if $d = 4$.*
5. *$d \geq 5$, $G = S_d$, with $\mathscr{I}$ being the conjugacy class of transpositions, and the following two conditions hold:*
   (i) *the strong Artin conjecture holds for all irreducible Galois representations over $\mathbb{Q}$ with image S_d,*
   (ii) *τ and c_2 are numbers such that $\tau < 1$ if $d = 5$ and $\tau < 1/2 + 1/d$ if $d \geq 6$, and for every fixed integer D there are at most $c_2 D^\tau$ fields $K \in S$ with $D_K = D$.*

For the families $S_4(a,b)$ in Theorem 1.8, we have the following bounds, which follow from [An20, Theorem 1.6 and Proposition 6.1].

Theorem 5.4 (An)**.** *Let $\varepsilon > 0$, and let $a, b \in \mathbb{Z} \smallsetminus \{0,1\}$ be distinct squarefree numbers. Then for every $\delta > 0$ there exists $c = c(a,b,\delta) > 0$ such that*

$$\#\mathscr{B}_{S_4(a,b)}(X; X^\delta, cX^\delta/\log X) \ll_{a,b,\delta,\varepsilon} X^\varepsilon.$$

6. Proofs of theorems

Each of our Theorems follows immediately from one of the Propositions 3.1 or 3.2 with suitable parameters, combined with a simple application of dyadic summation.

6.1. **Proof of Theorem 1.1.** Apply Proposition 3.2 with $\theta = 1 + 2/\ell$ (by Lemma 4.2), $\rho = 1$, and $\tau = 0$ (by Theorem 5.3).

6.2. **Proof of Theorem 1.3.** Apply Proposition 3.1 with $\theta = d - 1 + 2/\ell$ (by Lemma 4.2), $\rho = 1$ (by [DH71, Bha05, Bha10]), and $\delta_0 = \delta_0(d)$ (by Theorem 5.2).

6.3. **Proof of Theorem 1.4.** Apply Propostion 3.2 with $\theta = d-1+2/\ell$ (by Lemma 4.2) and $\tau = 0$ (by Lemma 5.1).

6.4. **Proof of Theorem 1.5.** For sufficiently small $\varepsilon' > 0$, we apply Proposition 3.2 with $\theta = 3/2 + 2/\ell + \varepsilon'$ (by Corollary 4.8), $\rho = 1/2$ (by [Wri89]), and $\tau = 0$ (by Theorem 5.3).

6.5. **Proof of Theorem 1.6.** For sufficiently small $\varepsilon' > 0$, we apply Proposition 3.2 with $\theta = 3 + 1/12 + 2/\ell + \varepsilon'$ (by Corollary 4.9) and $\tau = 1/4$ (by Theorem 5.3).

6.6. **Proof of Theorem 1.8.** For sufficiently small $\varepsilon' > 0$, we apply Proposition 3.2 with $\theta = 2 + 1/3 + 2/\ell + \varepsilon'$ (by Corollary 4.10) and $\tau = 0$ (by Theorem 5.4).

6.7. **Proof of Theorem 1.10.** First we note (cf. [PTBW20, Lemma 6.9]) that for each S_d-extension of degree d with squarefree discriminant, the ramification type of each ramified prime p that is tamely ramified is the conjugacy class of transpositions. Now apply Proposition 3.2 with $\theta = d - 1 + 2/\ell$ (by Lemma 4.2) and τ as in the statement of the theorem (by Theorem 5.3).

6.8. **Proof of Theorem 1.11.** For sufficiently small $\varepsilon' > 0$, we apply Propostion 3.2 with $\theta = d - 3/2 + 2/\ell + \varepsilon'$ (by Corollary 4.8) and $\tau = 0$ (by Theorem 5.3).

7. Upper bounds for Dihedral extensions

The aim of this section is to prove Corollary 1.2. In the proof of [Klü06, Theorem 2.5], Klüners has shown the estimates

$$N(p, D_p, X) \leq \sum_{D_K^{(p-1)/2} b^{p-1} \leq X} \frac{p^{\omega(b)+r_K} - 1}{p-1},$$

$$N(2p, D_p(2p), X) \leq \sum_{D_K^{p} b^{2(p-1)} \leq X} \frac{p^{\omega(b)+r_K} - 1}{p-1},$$

where both sums are taken over positive integers b and quadratic fields K with D_K in the indicated range, $\omega(b)$ denotes the number of distinct prime divisors of b, and r_K is the p-rank of Cl_K, so that $p^{r_K} = \#\,\mathrm{Cl}_K[p]$. For the first sum we find

$$N(p, D_p, X) \leq \sum_{D_K^{(p-1)/2} b^{p-1} \leq X} \frac{p^{\omega(b)+r_K} - 1}{p-1} \leq \sum_{b^{p-1} \leq X} p^{\omega(b)} \sum_{D_K \leq X^{2/(p-1)}/b^2} \#\,\mathrm{Cl}_K[p].$$

Plugging in the bound from Theorem 1.1 in case $k = 1$ proves the claim for $N(p, D_p, X)$. The second sum is handled similarly.

Acknowledgments

The authors are grateful to the referee for their careful reading and their valuable comments that significantly improved the exposition of the paper.

References

[AD03] F. Amoroso and R. Dvornicich. Lower bounds for the height and size of the ideal class group in CM-fields. *Monatsh. Math.*, 138, no.2:85–94, 2003.

[An20] C. An. ℓ-torsion in class groups of certain families of D_4-quartic fields. *J. Théor. Nombres Bordeaux*, 32(2):1–23, 2020.

[Bha05] M. Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math.*, 162:1031–1063, 2005.

[Bha10] M. Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math.*, 172:1559–1591, 2010.

[Bha14] M. Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials. *arXiv:1402.0031v1 [math.NT]*, 2014.

[BHB05] T.D. Browning and D.R. Heath-Brown. Plane curves in boxes and equal sums of two powers. *Math. Z.*, 251(2):233–247, 2005.

[BST13] M. Bhargava, A. Shankar, and J. Tsimerman. On the Davenport-Heilbronn theorems and second order terms. *Inventiones mathematicae*, 193 (2):439–499, 2013.

[BST+17] M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y. Zhao. Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves. *arXiv:1701.02458v1*, 2017.

[BSW16] M. Bhargava, A. Shankar, and X. Wang. Squarefree values of polynomial discriminants I. *arXiv:1611.09806v2*, 2016.

[CT17] H. Cohen and F. Thorne. On D_ℓ-extensions of prime degree ℓ. *arXiv:1609.09153*, 2017.

[DH71] H. Davenport and H. Heilbronn. On the density of discriminants of cubic field extensions. II. *Proc. London. Math. Soc.*, 322:405–420, 1971.

[Die12] R. Dietmann. On the distribution of Galois groups. *Mathematika*, 58(1):35–44, 2012.

[DO02] H. Cohen F. Diaz Y Diaz and M. Olivier. Enumerating quartic dihedral extensions of $\mathbb{Q}$. *Comp. Math.*, 133:65–93, 2002.

[Ell08] J. S. Ellenberg. Points of low height on $\mathbb{P}^1$ over number fields and bounds for torsion in class groups. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 45–48. Amer. Math. Soc., Providence, RI, 2008.

[EPW17] J. Ellenberg, L. B. Pierce, and M. M. Wood. On ℓ-torsion in class groups of number fields. *Algebra and Number Theory*, 11-8:1739–1778, 2017.

[EV05] J. S. Ellenberg and A. Venkatesh. *Counting extensions of function fields with bounded discriminant and specified Galois group*, pages 151–168. Birkhäuser Boston, Boston, MA, 2005.

[EV07] J. Ellenberg and A. Venkatesh. Reflection principles and bounds for class group torsion. *Int. Math. Res. Not.*, no.1, Art. ID rnm002, 2007.

[FK07] É. Fouvry and J. Klüners. On the 4-rank of class groups of quadratic number fields. *Invent. Math.*, 167(3):455–513, 2007.

[FW18] C. Frei and M. Widmer. Average bounds for the ℓ-torsion in class groups of cyclic extensions. *Res. Number Theory*, 4:34, 2018.

[HBP17] D. R. Heath-Brown and L. B. Pierce. Averages and moments associated to class numbers of imaginary quadratic fields. *Compositio Math.*, 153:2287–2309, 2017.

[Her70] H. Hering. Seltenheit der Gleichungen mit Affekt bei linearem Parameter. *Math. Ann.*, 186:263–270, 1970.

[Her72] H. Hering. über Koeffizientenbeschränkungen affektloser Gleichungen. *Math. Ann.*, 195:121–136, 1972.
[Hou16] B. Hough. Equidistribution of bounded torsion CM points. *arXiv:1005.1458v3*, 2016.
[HV06] H. A. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.*, 19:527–550, 2006.
[Klü06] J. Klüners. Asymptotics of number fields and the Cohen-Lenstra heuristics. *J. Théor. Nombres Bordeaux*, 18:607–615, 2006.
[Kly16] J. Klys. The distribution of p-torsion in degree p cyclic fields. *arXiv:1610.00226*, 2016.
[LO77] J. C. Lagarias and A. M. Odlyzko. *Effective versions of the Chebotarev density theorem. Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pp. 409-464*. Academic Press Inc., New York, 1977.
[LR12] E. Larsen and L. Rolen. Progress towards counting D_5 quintic fields. *Involve*, 5:1:91–97, 2012.
[Mal02] G. Malle. On the distribution of Galois groups. *J. Number Theory*, 92:315–329, 2002.
[Mal04] G. Malle. On the distribution of Galois groups. II. *Experiment. Math.*, 13(2):129–135, 2004.
[Mil17] D. Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \pmod 4$. *GAFA*, 27 (4):973–1016, 2017.
[Nar80] W. Narkiewicz. *Elementary and Analytic Theory of Algebrac Numbers, edition 2.* Springer, 1980.
[Pie05] L. B. Pierce. 3-part of class numbers of quadratic fields. *J. London Math. Soc.*, 71:579–598, 2005.
[Pie06] L. B. Pierce. A bound for the 3-part of class numbers of quadratic fields by means of the square sieve. *Forum Math.*, 18:677–698, 2006.
[PTBW19] L. B. Pierce, C. Turnage-Butterbaugh, and M. M. Wood. On a conjecture for ℓ-torsion in class groups of number fields: from the perspective of moments. *arXiv:1902.02008*, 2019.
[PTBW20] L. B. Pierce, C. Turnage-Butterbaugh, and 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(2):701–778, 2020.
[TT13] T. Taniguchi and F. Thorne. The secondary term in the counting function for cubic fields. *Duke Math. J.*, 162:2451–2508, 2013.
[Wid10] M. Widmer. Counting primitive points of bounded height. *Trans. Amer. Math. Soc.*, 362:4793–4829, 2010.
[Wid18] M. Widmer. Bounds for the ℓ-torsion in class groups. *Bull. London Math. Soc.*, 50(1):124–131, 2018.
[Wri89] D. J. Wright. Distribution of discriminants of abelian extensions. *Proc. London Math. Soc. (3)*, 58, no. 1:17–50, 1989.

TU Graz, Institute of Analysis and Number Theory, Steyrergasse 30/II, 8010 Graz, Austria
E-mail address: frei@math.tugraz.at

Department of Mathematics, Royal Holloway, University of London, TW20 0EX Egham, UK
E-mail address: martin.widmer@rhul.ac.uk